

Canadian & Cold Regions Rail Research Conference 2021



NOVEMBER 9-10, 2021
Virtual Event

CANADIAN RAIL
RESEARCH LABORATORY

 **UNIVERSITY OF
ALBERTA**

**CANADIAN & COLD REGIONS
RAIL RESEARCH CONFERENCE**

November 9-10, 2021

Virtual Event

Proceedings of the First Canadian & Cold Regions Rail Research Conference
9-10 November 2021, Virtual Event

First Canadian & Cold Regions Rail Research Conference

Edited by

**Dr. Parisa Haji Abdulrazagh, Dr. Michael T. Hendry &
Dr. Renato Macciotta Pulisci**

Canadian Rail Research Laboratory

Copyright ©

All rights reserved. No part of this publication or the information contained herein may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording or otherwise, without written prior permission from the publisher.

Although all care is taken to ensure the integrity and quality of this publication and the information herein, no responsibility is assumed by the publishers nor the author for any damage to property or persons as a result of operation or use of this publication and/or the information contained herein.

Published by: University of Alberta, Dept. of Civil & Environmental Engineering

ISBN 978-1-55195-467-7

Printed in Canada

FORWARD

It is with great pleasure that we present the Proceedings of the inaugural **Canadian & Cold Regions Rail Research Conference 2021 (CCRC 2021)**. Recognizing the need to bring together railroaders, regulators and researchers working in Canada and cold regions in one venue to share their expertise and innovative work, the CCRC 2021 Organizing Committee convened a Technical Committee with representatives from industry, government and academia to create this conference.

The Canadian Rail Research Laboratory (CaRRL) at the University of Alberta is proud to be the host and organizer of CCRC 2021. Established in 2010 by the Railroad Research Advisory Board, Transport Canada, and Canada's railway sector, CaRRL facilitates close collaboration among researchers, industry and government, and identifies emerging technologies for the improvement of safety, reliability, sustainability and economic feasibility of rail transportation. Dr. Michael T. Hendry is the Director of CaRRL and holds the NSERC Industrial Research Chair in Railway Engineering. Dr. Renato Macciotta Pulisci is the Associate Director of CaRRL, and Dr. C. Derek Martin served as inaugural Program Director and concurrently held the NSERC Industrial Research Chair in Railway Geomechanics and the Alberta Innovates – Technology Futures Tier One Industry Chair in Railway Geomechanics.

The aim of CCRC 2021 is to provide opportunities for professionals and researchers involved with railway engineering to present ideas, learn of new developments and technologies, make professional contacts and discuss issues related to railway engineering in Canada and cold regions. The Proceedings of CCRC 2021 represents the current state-of-practice in Canada's rail transportation sector. We are grateful for the authors and presenters who responded enthusiastically to the call for papers; their papers in the Proceedings benefit the Canadian rail sector and the public it serves through providing technical knowledge for current practitioners and educating the next generation of science and engineering students for careers in rail.

We would like to acknowledge the sponsors of CCRC 2021 who recognize the significant contribution of this conference to meeting the challenges of this most important industry—thank you for your support. We want to personally thank members of CaRRL and the Technical Committee for their encouragement and support. The conference would not have been possible without the dedication of Reza Mousapour, Kevin Wallin and especially Vivian Giang who provided so much assistance and leadership.

The importance of rail transportation to both the Canadian economy and the quality of life enjoyed by Canadians is significant. It was with this in mind that the session themes and manuscripts were selected for presentation and inclusion in the Proceedings. May you find further insights to enhance your understanding of the current state-of-practice in the rail industry through CCRC 2021.

Parisa Haji Abdulrazagh, Michael T. Hendry and Renato Macciotta Pulisci
Co-Chairs, CCRC 2021 Organizing Committee

Canadian & Cold Regions Rail Research Conference 2021

Organizing Committee



Michael T. Hendry

*Associate Professor, Faculty of Engineering
University of Alberta*



Renato Macciotta

*Assistant Professor, Faculty of Engineering
University of Alberta*



Parisa Haji Abdulrazagh

*Postdoctoral Fellow, Faculty of Engineering
University of Alberta*



Vivian Giang

*Grant Strategist & Event Coordinator
University of Alberta*

Technical Committee



Daniel Blais

*Senior Program Policy Analyst, Innovation Centre
Transport Canada*



Janet Greene

*Director, Member Engagement
Railway Association of Canada*



Parisa Haji Abdulrazagh

*Postdoctoral Fellow, Faculty of Engineering
University of Alberta*



Michael T. Hendry

*Associate Professor, Faculty of Engineering
University of Alberta*



Jean Hutchinson

*Professor, Faculty of Geological Engineering
Queen's University*



Brooke Jones

*Engineer, Innovation Centre
Transport Canada*



David Lilley

*Assistant Chief, Engineering
Canadian National Railway*



Renato Macciotta

*Assistant Professor, Faculty of Engineering
University of Alberta*



Paul Miller

*Edmonton Community Board Member
Canadian National Railway*



Kyle Mulligan

*Chief Engineer
Canadian Pacific Railway*



Alireza Roghani

*Research Associate
National Research Council Canada*



Albert Wahba

*Program Leader
National Research Council Canada*



Yi Wang

*Mechanical Systems Engineer
Canadian Pacific Railway*

Canadian & Cold Regions Rail Research Conference 2021

Thank you to our sponsors!

PLATINUM



Transports
Canada

Transport
Canada

GOLD



Railway Association
of Canada

Association des chemins
de fer du Canada



SILVER



Endorsed By



THE CANADIAN
GEOTECHNICAL SOCIETY
LA SOCIÉTÉ CANADIENNE
DE GÉOTECHNIQUE

TRANSPORTATION GEOTECHNIQUE COMMITTEE



CANADIAN PERMAFROST ASSOCIATION
ASSOCIATION CANADIENNE DU PERGÉLISOL

TABLE OF CONTENTS

KEYNOTE PRESENTATIONS

TECHNOLOGY TO ENSURE RAILWAY INFRASTRUCTURE RESILIENCE CONSIDERING THE IMPACTS OF CLIMATE CHANGE

C. Moore (Ontario Northland) 2

TECHNOLOGY TO ENSURE RAILWAY INFRASTRUCTURE RESILIENCE CONSIDERING THE IMPACTS OF CLIMATE CHANGE

É. Baril (National Research Council Canada) 3

TECHNOLOGY TO ENSURE RAILWAY INFRASTRUCTURE RESILIENCE CONSIDERING THE IMPACTS OF CLIMATE CHANGE

K.R. Mulligan (Canadian Pacific Railway) 4

SESSION 1: CLIMATE CHANGE AND ENVIRONMENT

TECHNOLOGY TO ENSURE RAILWAY INFRASTRUCTURE RESILIENCE CONSIDERING THE IMPACTS OF CLIMATE CHANGE

T. Sussmann, H. Thompson, S. Wilk and R. Bruzek 6

THE CONSIDERATION OF A CHANGING CLIMATE FOR CONTINUOUSLY WELDED RAIL

M. Zhang and A. Roghani 14

CLIMATE TRIGGER SIGNATURE AND ADVERSE GROUND CONDITIONS RESULTING IN DESTRUCTIVE FLOW EVENTS IN BRITISH COLUMBIA DURING THE FALL 2017

T. Keegan and M. Cormier 22

EFFECT OF CLIMATE CHANGE ON FROST PENETRATION DEPTH IN THE SUBGRADE SOIL BENEATH RAILWAY TRACKS: A CASE STUDY

M. Roustaei, M.T. Hendry and A. Roghani 31

THE CASE FOR ZERO EMISSION PROPULSION FOR THE OKANAGAN VALLEY ELECTRIC REGIONAL PASSENGER RAIL (OVER PR)

T. Boray, T. Goyal, H. Sharma, M. A. Hegazi and G. Lovegrove 39

SESSION 2: COLD REGIONS ENGINEERING

EVALUATION OF CURRENT RAIL TANK CAR TC128B STEEL WELD PERFORMANCE

S. Xu, J. Chen, J. McKinley, J. Liang, L. Yang, A. Laver and M. Spiess 47

RAIL FAILURE ROOT CAUSE ANALYSIS ON NORTH AMERICAN RAILWAY

D. Szablewski and R. Caldwell 56

CONDITION BASED RAIL SURFACE MANAGEMENT USING EMFI TECHNOLOGY

D. Juhaszova-Dubinsky, J. Furlong, A. Milligan and K.R. Mulligan 64

LESSONS LEARNED FROM A LONG-TERM FROST HEAVE MONITORING UNDER A RAILWAY EMBANKMENT	
A. Roghani, R. Caldwell, J. Hiedra-Cobo and P. Charbachi	72
MATERIAL MODEL FOR TC128B AT LOW TEMPERATURE AND APPLICATION IN FULL-SCALE FEA	
B. W. Williams, J. Xue, J. McKinley, S. Xu and M. Spiess	78
INVESTIGATING THE IMPACT OF FROST HEAVE AND THAW SOFTENING ON CHANGING THE VERTICAL FORCE AT WHEEL/RAIL INTERFACE	
A. Roghani, Y. Liu and P. Burgess	86
RELATIONSHIP BETWEEN FRACTURE TOUGHNESS AND AVERAGE GRAIN SIZES OF HIGH STRENGTH RAIL STEELS	
S. Okocha, F. Yu, P.-Y. Ben Jar and M. T. Hendry	92
SESSION 3: INNOVATIONS IN RAILWAY ENGINEERING	
CHARACTERIZING RAIL CORRUGATION THROUGH A DATA-DRIVEN ASSESSMENT OF IMPACTS ON NOISE AND VIBRATION	
A. Lasisi, J. D. Regehr, J. Carneiro, I. Jeffrey, E. Magel, S. Chenier and M. Reimer	101
USING EMERGING TECHNOLOGIES FOR MONITORING SURFACE WATER NEAR RAILWAY TRACKS	
A. Roghani, A. Mammeri and A. J. Siddiqui	106
PROPOSED DYNAMIC LOAD FACTORS DEVELOPED FROM INSTRUMENTED WHEELSET MEASUREMENTS	
D. Behnia, P. Haji Abdulrazagh and M. T. Hendry	115
EVALUATE THE TRACK STIFFNESS VARIATION ALONG THE TRACK USING THE INSTRUMENTED WHEELSET (IWS) MEASUREMENTS	
R. Mousapour, P. Haji Abdulrazagh, M.T. Hendry and M. Gül	122
EVALUATION OF RIDE QUALITY AND RAIL SURFACE ROUGHNESS FROM VIBRATION ANALYSIS OF IN-SERVICE PASSENGER RAIL CARS	
P. Haji Abdulrazagh, M.T. Hendry, A. Roghani, and E. Toma	130
LABORATORY INVESTIGATION OF MECHANICAL PROPERTIES OF USED RUBBER TIRES AS UNDER-SLEEPER PAD IN RAILWAY TRACKS	
M. Esmaeili, A. Shamohammadi, F. Astaraki, M. Movahedi Rad and J. Chalabi	136
EMBANKMENT RECONSTRUCTION USING A WICKING GEOTEXTILE AND ITS EFFECT ON THE DRAINAGE AND STRENGTH OF A LOW PERMEABILITY SUBGRADE	
C. Alvarenga, P. Haji Abdulzaragh and M.T. Hendry	141
SESSION 4: RISK AND RELIABILITY	
FROM PAPER NOTES TO DIGITAL TWINS: THE EVOLUTION OF GEOSPATIAL MODELS ALONG RAILWAY CORRIDORS AND DIGITAL CONVERGENCE	
R. Harrap, P.M. DiFrancesco, D. A. Bonneau, D. J. Hutchinson, Z. Sala, and M. Ondercin.	149

RISK ANALYSIS OF HAZARDOUS MATERIALS RELEASE IN RAILWAY TRANSPORTATION H. Ebrahimi, F. Sattari, L. Lefsrud and R. Macciotta	158
PREDICTION OF EMERGENCY RESPONSE STRATEGIES BASED ON COMBUSTION SIGNATURES FROM FTIR SPECTROSCOPY USING MACHINE LEARNING TECHNIQUES S. Tian, C.G. Lee, D.D.Z. Xi, Y. Ko and N. Elsagan	164
THE USE OF TERRESTRIAL LASER SCANNING FOR THE DEVELOPMENT OF SLOPE PROCESS MODELS D. A. Bonneau, D. J. Hutchinson, P.M. DiFrancesco and R. Harrap	170
EFFECTS OF SEASONAL TRACK GEOMETRY CHANGES ON VEHICLE-TRACK INTERACTION AND POTENTIAL IMPLICATIONS FOR TRAIN MARSHALLING Y. Wang, K.R. Mulligan, K. Oldknow and G. G. Wang	179
DIGITAL ROCKFALL DATABASES: USING LIDAR TO OBSERVE OVER 7-YEARS OF DETAILED ROCKFALL ACTIVITY AT SEVERAL SITES ALONG THE THOMPSON-FRASER RAIL CORRIDOR, BRITISH COLUMBIA, CANADA P.M. DiFrancesco, D. A. Bonneau and D. J. Hutchinson	187
DEVELOPMENT OF A DAILY UPDATED TRAIN DERAILMENT IMPACT MAP AND ANALYTICS SYSTEM FOR THE NATIONAL RAILWAY NETWORK IN CANADA Y. Liu, Chengbi Dai, Z. Schenk and L. Steinginga	197
SESSION 5: ENHANCED TRAIN CONTROL AND OPERATIONS	
ASSESSMENT OF THE LRAIL SYSTEM TO DETECT AND ANALYZE THE CONDITION OF ANGLE BARS ON A 260-MILE RAILROAD L. Faucher, T. Ducheman, M. Jarry, M. Cody and D. Hébert	206
DEVELOPMENT OF RAILWAY TERMINALS WITHIN THE CONTEXT OF PRECISION SCHEDULED RAILROADING PRINCIPLES R. S. Lanyi	211
LITERATURE REVIEW ON COGNITIVE IMPACTS OF CAB WARNING SYSTEMS AND TRAIN CONTROL TECHNOLOGIES M. A. Rad, L. M. Lefsrud and M.T. Hendry	221
SELECTION OF TRACK FORMS – A LIFE CYCLE COST PERSPECTIVE W. Li and R. Dwight	231
SESSION 6: HUMAN FACTORS	
APPLICATIONS OF INNOVATIVE ACCIDENT ANALYSIS METHODS IN RAILWAYS: A REVIEW M.A. Rad, L. M. Lefsrud and M. T. Hendry	241
WITH GREAT POWER COMES GREAT RESPONSIBILITY: ORGANISATIONAL FACTORS INFLUENCING POWER DISPARITIES BETWEEN NETWORK CONTROLLERS AND OTHER OPERATIONAL GROUPS B. Luva and A. Naweed	251

KEYNOTE PRESENTATION ABSTRACTS

High Performance Leadership: Establishing a continuous improvement culture to enable successful technology implementations

Corina Moore
President and CEO, Ontario Northland

ABSTRACT

In order for railways to continue to expand their impact and prosper, a high performance and continuous improvement culture must be established. Many railways and organizations throughout the supply chain are lacking innovative ideas and are notorious for the coined expression “we can’t change...we’ve always done it that way”. This isn’t surprising given the strong and rich history of the rail industry and the engrained command-and-control culture and non-diverse workforce that have historically obstructed progress. A focused program to change culture is not only necessary, it’s critical to ensuring people and technology are running at their full potential. This will lead to exceptional organizational success and financial gains far beyond expectations.

Generally speaking, organizational culture is the shared set of values and beliefs that make the unique social and psychological environment of an organization. It’s what dictates “how things are done” in a company. For an organization to stay relevant and competitive, its culture has to evolve and adapt. People drive culture and therefore investment in leadership and diversity programs are foundational to successful culture shifts, implementing new technology or making processes more effective and efficient.

The presentation will share the steps taken to successfully transform Ontario Northland, a 120-year transportation company, to a culture of continuous improvement with diverse employees who celebrate change and value innovation. Foundational to this success is leadership and how employees at all levels of the organization influence each other. The discussion will dive into different leadership styles and levels, communication and engagement strategies, and provide direction on diversity and inclusion.

Conference attendees should leave the presentation with strategies from a real-life transformation journey and be ready to guide their company’s change journey by creating the environment necessary to develop, implement and evolve technology. Superior customer service and enhanced safety, reliability and productivity are all important consequences of committing to this plan of action.

The NRC's role in supporting railway research and development in Canada

Éric Baril

Vice President of Transportation and Manufacturing (Acting), National Research Council Canada

ABSTRACT

After a quick introduction to the National Research Council of Canada (NRC), Mr. Éric Baril, Acting Vice President, Transportation and Manufacturing for the NRC, presented global and Canadian trends in the transportation sector and in particular the rail industry. Mr. Baril, explained the types of research, development, testing and evaluation activities that the NRC is conducting to support industry, academia and government efforts to have cleaner, safer and more efficient transportation systems. He also presented some examples of recent and current research projects that are being led by NRC experts.

Technology for Cold Weather Railroading

Kyle R Mulligan
Chief Engineer, Canadian Pacific

ABSTRACT

Winter weather railroading conditions in the Canadian climate can profoundly affect railway operations. Developing effective winter plans and strategies is essential to Canadian Pacific (CP) and all other North American railways in maintaining safe and efficient railway operations. These plans and strategies must also be constantly renewed and updated based on lessons learned each year. Technology plays a key role in providing world-class safety and service even in the midst of winter weather railroading conditions. These technologies are essential for maintaining rail and equipment integrity to ensure goods continue to move safely. This presentation will take the audience through key technologies which are currently utilized to combat winter railroading conditions showing some of the talented railroaders who are responsible for their implementation. These technologies include, but are not limited to, the following: broken rail detection for non-signaled (dark) territory, top of rail surface quantification, ultrasonic rail flaw detection, technology driven train inspection using Train Inspection Portal Systems (TIPS) and Remote Safety Inspection (RSI), cold wheel technology using Wheel Temperature Detectors, cracked wheel detection, and Big Data analytics. Investments and Innovations in these spaces are a key part of why has CP led the industry for the 15th consecutive year in 2020, with the lowest Federal Railroad Administration (FRA) reportable train accident frequency (per million train miles) among Class I railroads in North America. The presentation will close with an introduction to CP's next innovations in sustainability including an introduction to North America's first line-haul hydrogen electric locomotive.

Session 1

CLIMATE CHANGE AND ENVIRONMENT

Technology to Ensure Railway Infrastructure Resilience Considering the Impacts of Climate Change

Ted Sussmann,
U.S. DOT Volpe Center and The University of Hartford, Cambridge/MA, USA

Hugh Thompson
U.S. DOT Federal Railroad Administration, Washington/DC, USA

Steven Wilk
Transportation Technology Center Inc., Pueblo/CO, USA

Radim Bruzek
Ensco, Inc., Springfield/VA, USA

ABSTRACT

A majority of railway fixed infrastructure consists of subgrade and fill sections constructed from local materials well before the advent of modern geotechnical engineering and compaction technology. In the ensuing century or more following construction, weather extremes highlighted weak zones for reconstruction following floods, washouts, and soil instability related failures. The effects of climate change have been felt in recent years through increased fire hazards and disasters like hurricanes, floods, and extreme precipitation events. Previous weather that was considered extreme has changed greatly with 100-year storms happening much more frequently and causing engineers to reconsider the statistics of flood interval periods. The open track surface with coarse ballast used for traditional railway track is particularly susceptible to extreme precipitation events since the open track surface allows surface water and precipitation to enter the track, affecting the performance of the track substructure. Technology to assess the structural condition of railway track has been pursued by FRA research for over ten years to help evaluate the gradation of track substructure layers, locate trapped moisture and areas of subsurface saturation, and assess the structural support provided to the superstructure. This research is expected to support improved root cause assessment of track support problems that have resulted in an increasing number of derailments in recent years. The specific systems to be covered in the discussion include ground penetrating radar, visual ballast gradation assessment, magnetic resonance moisture detection, vertical track deflection measurement, and seismic track support stiffness measurements. Together these systems provide track structural condition data that can be used to assess the traffic worthiness of track affected by extreme precipitation through an assessment of the strength and stiffness of the track structure along with the presence of moisture that indicates a risk that the structural condition could deteriorate rapidly under traffic.

1 INTRODUCTION

The construction of railway track represents one of the major technological advancements that has stood the test of time by still providing a cost-effective form of infrastructure. Railway track uses steel rails with wood or concrete crossties to distribute load to the ballast and, ultimately, the subgrade. This design provides:

1. the strength to resist some of the highest repeated loads regularly transported,

2. the flexibility to adjust alignment and grade rapidly,
3. the maintainability to compensate for settlement and movement due to climate and other factors, and
4. the drainage capacity to allow the track structure to function through all but the most serious precipitation events.

This last factor, drainage, is one of the keys to the longevity of railway track. Railway track was originally

sited along rivers and canals including along many flat areas with low elevation where drainage is critical. The use of gravel and eventually aggregate ballast allowed water to drain and any repeated load induced pore water pressure to dissipate. Local track gangs maintained the track to provide local drainage where and when needed and re-established track level to compensate for settlement.

The open track surface that allows surface water into the track substructure must be coupled with adequate internal and external drainage capacity to function effectively. As railways have advanced, the local track gangs have been replaced by production and division maintenance gangs. Railways are much more efficient with improved maintenance machines and processes that provide a uniform track structure in terms of grade, layer thickness, etc. However, drainage is often overlooked in the application of these machines. The ability to surface and tamp settlement zones rapidly allows gangs to overlook local changes in drainage that might affect track performance.

In this paper, we will examine emerging track inspection tools targeting several aspects of track substructure inspection. From identifying localized trapped water in the track structure, to identifying layer thicknesses and ballast fouling conditions these techniques provide an opportunity for highly efficient maintenance operations to provide a more uniform track structure with a longer life and reduced risk of failure.

Changes in North American railway operations are happening rapidly with longer trains operating on more fixed schedules. This move toward Precision Scheduled Railroading will require a resilient and predictable track structure to fulfill the promise of these changes toward a more efficient and competitive railway industry. Local track support problems are not representative of most track miles or maintenance expenses, but these problem areas are often a significant cause of slow orders, unplanned maintenance, and train delays. In addition to the previous measurement systems being developed for substructure inspection, a variety of new systems are being developed to quantify the structural condition and properties of track and roadbed. The techniques range from a system to quantify track deflection under a train that can identify variations in deflection due to changing support conditions to systems that can be used to estimate the Resilient Young's Modulus of Ballast to identify the stage of ballast consolidation and deterioration. Together these systems provide some of the basic data needed to forecast track life in terms of number of axle loads needed to reach a given amount of track settlement or track geometry roughness.

Perhaps the most unpredictable challenge to the rail industry is not related to maintenance planning or track life forecasting, but rather it is related to the effects of severe weather events that can disrupt service over entire corridors. From fires and disasters related to hurricanes and tornados, railroads are affected by the

changes in the environment that affect precipitation and weather. As an open track structure, the railroad is particularly susceptible to heavy rainfall events that can saturate the substructure.

Flooding associated with these events represents a special challenge as the saturation level in railway embankments is difficult to measure. However, the age of many railroad embankments placed construction prior to the advent of geotechnical engineering and any associated developments related to compaction control and drainage. Initially, railway embankments were created from local fill placed in uncontrolled dumping with no compaction at steep slopes. This combination makes the durability of the original embankments somewhat limited and difficult to assess in flood and rapid drawdown conditions. In-place instrumentation is advancing to help monitor high risk locations for signs of saturation and instability.

Taken together, these technologies related to substructure inspection, railway track and roadbed engineering properties, and instrumentation for monitoring performance represent a combination of inspection systems and data sources that can provide important information to railway operations personnel responsible for safe operations in a changing environment where historic performance and safety trends might not be indicative of future conditions. While there is no replacement for the knowledge of the local track gangs of the past, the competitiveness of the rail industry requires some measures to provide data on local changes in track conditions that can affect track performance, operating schedules, and, in some cases, safety.

2 MOTIVATION

One of the driving factors behind this research was the derailment of a train near Cox Landing, WV in 1999 (NTSB, 1999). In this derailment a combination of an unstable sidehill fill, poor drainage conditions, and fouled ballast led to a large amount of maintenance. Despite the best efforts of the track crews, inspectors, and engineers, a train carrying a significant amount of hazardous chemicals derailed necessitating a local evacuation. This accident demonstrated the need for information on the subsurface track conditions and engineering properties of the track support materials to help inform decisions regarding track capacity, longevity, and stability. Research to address the measurement of track deflection and high precision track survey methods were pursued to provide better data on localized track support problems areas and track settlement, respectively.

Increasing train loads and traffic volumes concentrated on main routes in recent decades have increased the stress on each component of the track structure. In the substructure, these high loads stress the ballast and subgrade leading to localized settlement that can trap water or localized ballast deterioration that combines

with other sources of ballast fouling materials to limit ballast drainage. One of the results of this deterioration is manifestation at the track surface as mudspots due to fouled ballast.

Ballast fouling is often thought to proceed from the bottom to the top of the ballast layer. However, the main source of fouling is deterioration of ballast under traffic (Selig & Waters, 1994) and maintenance (tamping) that both originate at or near the track surface. Track vibration and water flow through the track tends to move the fouling materials to the bottom of the layer where the fouling agents accumulate from the bottom to the top. In locations where track drainage normally flows from the center toward the edge of track carrying fouling materials along the drainage path, the fouling can accumulate from the edge of the track toward the center where the fouling material settles out of the water flow or becomes trapped between adjacent particles. Based on this hypothesis for ballast fouling propagation, ground penetrating radar has been developed to detect the amount of fouling material in the ballast (ballast fouling index), the depth of clean ballast below the tie bottom, and the location of trapped fouling and moisture locations.

The development of GPR has proceeded as a topic of research jointly between the industry, suppliers, and the Federal Railroad Administration. In addition to basic information on the track support layers, GPR has a demonstrated ability to detect shallow fouling around the ties that occurs in locations with coal dust fouling that does not settle through the track structure, but rather accumulates near the track surface after blowing and spilling off loaded cars. This type of fouling was associated with a series of derailments in the Powder River Basin in 2005 (DOE, 2007) with subsequent research that emphasized the need for GPR data on ballast condition including coal fouling and the condition of ballast surrounding the ties.

Together, GPR and track deflection technology provide the basic information on track support conditions that reflect both the load-deflection behavior of the track, and the ballast fouling conditions that might cause poor track support conditions associated with deterioration of track geometry. Poor track support conditions increase the stress on track components as the track structure works to distribute loads from poorly supported ties to well supported ties. In locations of poor track support and mudspots, track deflection can lead to increased rail stress due to bending, shear, and/or tension.

A 2012 derailment in Ellicott City, MD (NTSB, 2017) highlighted the role of the level of stress and rail condition on rail life and resistance to fracture. Well-supported track reduces the stress on the rail and other components, extending the track life. In locations where track support is not optimal, the deterioration and failure of track components is possible, and the systems used to measure track support can help minimize risk of track component failure in these situations. A Federal

Railroad Administration (FRA) Report to Congress on the status of track deflection measurement described plans for deployment of this technology to mitigate identified risks. While the measurement of track deflection presents many challenges, advances are ongoing that will make the collection of this data more routine in the future. Concepts for quantifying track support conditions and the application of the data to assure resilience against failure was presented by (Sussmann et. al, 2017).

Applying the concept of resilience to railway track under the threat of severe precipitation related to climate change presents some unique challenges that will be described in this paper. One of the challenges is related to extreme precipitation that can cause localized flooding and saturation and the other is related to large scale flooding that can reduce the capacity of the track structure until the end of saturation and a return of the earthworks to a more routine moisture condition.

The Transportation Research Board hosted a 2021 workshop titled "Measuring and Managing Freight System Resilience" that focused on the movement of freight with one specific focus area involving rail freight related to the 2011 and 2019 flooding of the Missouri River Valley. During this part of the workshop, it became clear that the movement of most freight is dependent on resilient and reliable infrastructure. Some corridors house highway, power, rail, and telecommunications infrastructure and flooding or other disruption to the corridor effects all infrastructure. Rationalization of the nation's infrastructure network has concentrated the bulk of traffic along specific corridors. While the country must be efficient and not be burdened with too many miles of infrastructure, the result is a heavy reliance on the remaining corridors for vital supplies. These lifelines often lack backup routes or corridors. In these situations, extra emphasis on resilience is required to help minimize disruptions. Technology to help monitor the performance of the infrastructure in critical locations is being developed that can provide insight to the degree of saturation, the response of the embankment to traffic, and associated changes from baseline conditions that can indicate structural deterioration. These measurements coupled with remote sensing data can provide insight to the status of infrastructure and provide insight to locations that show signs of deterioration under existing conditions that should be prioritized for upgrade to help minimize impact of more severe conditions expected due to flooding or the severe precipitation events associated with climate change.

3 SYSTEM DESCRIPTIONS

The systems for track substructure inspection, support measurement, and monitoring will be described in this section. Li et al. (2015) presents an overview of some of the systems, while more detail can be found in specific articles referenced in each section.

3.1 Ballast Gradation

One of the most critical elements of substructure problem diagnosis is identification of the ballast gradation including the degree of ballast fouling. The ballast fouling index provides a guide to the relative degree of drainage blocked by fouling material in the void space of ballast. Li et al. (2015) presents the ballast fouling definitions that have been developed but found that the fouling index proposed by Selig and Waters (1994) provided the necessary information most readily. However, the technique requires a gradation that includes the percent passing the number 4 and number 200 sieves. With large sized aggregate, obtaining a representative sample size often requires processing over 0.5 kN of ballast, which makes routine sampling labor intensive. Sussmann and Ruel (2012) investigated the use of the ballast fouling index and compared the results with field derived limits for fouling developed at Canadian National and found that the simplified percent passing the ¾ in. sieve could provide a good indication of the necessary data. With one sieve, a pan, and a scale the measurements can be made in the field in a variety of locations to provide better understanding of variations across and along a site. With proper training, an inspector could also collect important data on the nature of the fouling material, such as plasticity, that can greatly influence ballast performance.

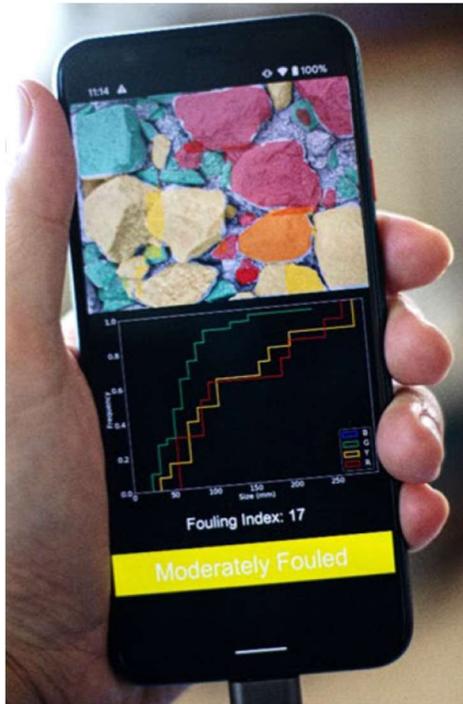


Figure 1. Conceptual Smartphone Application of Optical Ballast Gradation by OceanIT

The FRA has pursued techniques to develop an optical based technique using photos from a smartphone to assess ballast gradation (Figure1). Currently, FRA is seeking techniques to assess the plasticity of the fouling

material to provide insight to variations between sites that can affect the relative severity of the impact of ballast fouling on track performance and safety.

3.2 Substructure Inspection

Ground penetrating radar (GPR) represents the main track substructure inspection technique employed by the rail industry. The ability of GPR to locate zones of fouled ballast, trapped water, and clean ballast provides a strong basis for determining the integrity of the track structure and the durability of track geometry. Loram Technologies HyGround Division and Zetica Ltd are GPR rail industry leaders in North America. Each supplier provides a distinct dataset for its users with various information related to the track subsurface layers, conditions, and variations along and across the track. The interested reader can find more details on this technology in these references (Basye et al, 2020 and Sussmann et al., 2018)). For this paper, the discussion of GPR will be limited to the processed results of the depth of clean ballast, zones of fouled ballast and the ballast fouling index, and zones of increase moisture content which are more thoroughly described in (Bayse et al., 2020).

Identifying locations of trapped water and providing a reliable moisture profile of the track is critical in many of these situations and work by Vista Clara (Walsh and Barret, 2018) is being pursued to provide a routine assessment of the moisture profile with depth at regular intervals.

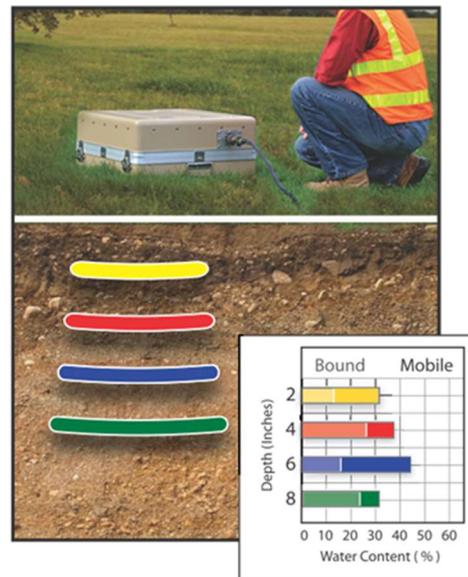


Figure 2. Conceptual Operation of Magnetic Resonance Measurement of Track Moisture Profile by Vista Clara

3.3 Track Deflection

Track deflection is a critical, but complex indicator of track structural condition where, conceptually, softer track structural conditions result in more deflection.

While large deflection by itself is not necessarily an indicator of substandard performance, track deflection can be very useful in diagnosing track support problems.

Measurement of track deflection involves both characterization of the load-deflection curve to capture the non-linearity often associated with deteriorated track conditions and variations in deflection along the track. Measurement of the gap and contact deflection are important elements needed to quantify the non-linearity of the load-deflection curve. The gap deflection refers to the deflection of the track under a light load where little resistance to applied loading is developed (Sussmann et al, 2017) and the contact deflection is the change in deflection from a light load to the test load (often taken as the design wheel load). Variation in the gap, contact, and total deflection along the track can highlight track support variations affecting track structural performance and track geometry deterioration.

Systems that have been used to measure track deflection include the AAR Track Loading Vehicle (TLV) and Harsco Rail's MRail system. The TLV has provided a static load-deflection curve and was configured to provide a moving, continuous measurement of the deflection from lightly loaded to fully loaded. The Mrail System can provide a measure of the deflection and track geometry that tends to indicate variations in track support (Hartsough et al., 2021). Neither system provides a continuous measure of the non-linearity of the load-deflection curve.

Four parameters derived from track deflection data were identified by Sussmann et al. (5) as particularly informative for track support assessment: (a) low track stiffness and large deflection, (b) variable track stiffness and deflection, (c) void deflection, and (d) inconsistent rail deflection. Some relative and approximate values to help bound the ranges expected for these parameters is given in Table 1. Low track stiffness indicates inadequate support from the ballast and/or subgrade. High track stiffness can limit the ability of the track to distribute the applied load and cause excessive load on individual ties. Variable track stiffness highlights locations where variable track support along the track can affect track performance. Inconsistent rail support occurs in locations where differences in support from one rail to the other affects track performance.

Table 1. Track Support Parameters

Parameter	Track Stiffness (kN/mm)	Track Modulus (kN/mm/mm or MPa)
Low Track Stiffness	<43	<19
High Track Stiffness	>71	>65
Variable Track Stiffness	Standard Deviation approaching 50% of the average	
Inconsistent Rail Support	Difference between rails exceeds 25%	

Finally, gap deflection indicates locations where track support non-linearity affects the ability of the ballast to engage and support the track. Selig & Waters noted that gaps more than 1 mm were associated with significant increases in track settlement during lab testing. In the field, 1 mm may be impractical to monitor, but gap deflection should be minimized to avoid excessive dynamic interaction between the tie and ballast. While the gap deflection might be a challenge to measure in the field, locations with total track deflection exceeding 5-10 mm are generally associated with a large gap since contact deflection in well supported track generally does not exceed ~5 mm.

3.4 Ballast Modulus

The ballast layer directly supports the tie and as a large, open graded aggregate, the application of most compaction control equipment is not practical. The degree of ballast compaction is important to ballast performance since loose ballast does not support the track structure properly, allowing the formation of load chains that can increase localized stress levels without providing uniform support, and will result in large vertical settlement, lateral alignment problems, and longitudinal tie restraint issues.

Lack of ballast compaction/consolidation was an issue in the 2002 Crescent City, FL (NTSB, 2003) derailment where multiple track surfacing operations were performed to correct track geometry deviations prior to the derailment. While the derailment was associated with a track buckle, the large number of track surfacing operations in a short time frame indicates that the ballast likely never consolidated properly. The NTSB investigation found that ballast confinement was a problem on the steep embankment. The lack of ballast consolidation appears likely to be a problem in this derailment whereby the track lateral resistance would be reduced and allowed the track to shift as the ballast settled leading to track geometry deviations that could result in lateral track shift and, potentially, a buckle.

While a challenge, data on the degree of compaction (commonly called consolidation in the industry) has not been available for the guidance of construction crews, maintenance operations, and inspectors. To remedy this situation and provide information to the industry on this critical parameter, the FRA has pursued research into field modulus measurement techniques. Important data on ballast compaction/consolidation is expected to be developed from measurements of ballast Young's Modulus, or more specifically, Resilient Modulus. Techniques for measuring modulus investigated to date include:

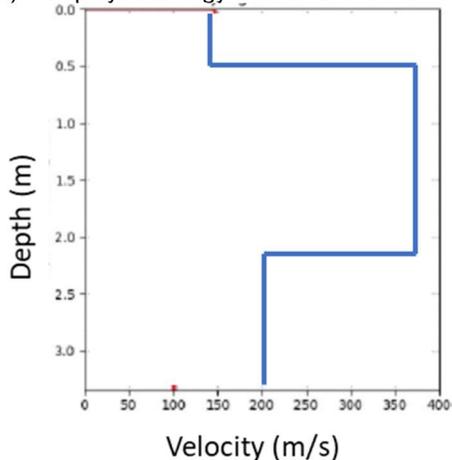
1. falling weight and lightweight deflectometer whereby a weight is dropped on the ballast surface, the impact force is measured along with the resulting deflection to estimate modulus.
2. Spectral analysis of surface waves technology where the shear wave velocity

and surface waves are used to estimate layer thicknesses and moduli.

FRA has pursued more detailed research and sensor system development to advance the use of spectral analysis of surface waves (Figure 3) technology including work to couple the sensors of the seismic energy to the track structure for transmission and monitoring of the shear waves. Current research has focused on automating the data interpretation for common track structure situations to make the system practical for use during routine maintenance operations and suitable for use by inspectors gathering data for planning maintenance.



a) Deployed Energy Source and Sensors



b) Shear Wave Velocity Profile

Figure 3. Ballast Seismic Shear Wave Velocity Profile by Earth Science Systems, Inc.

3.5 Field Monitoring and Remote Sensing

Monitoring of track performance in the field is critical to the assessment of track structural performance including the load path through the track that can be affected by deterioration and the stability of track substructure layers. FRA has pursued a variety of smart sensors to detect track conditions critical to safe operations. The ballast was one of the first targets for smart sensor development with the Smart Rock developed at Penn State Altoona (Liu et al., 2017). This

device has emerged over several developmental generations as a suitable device for monitoring the ballast layer vibration, settlement, rearrangement, and, potentially, stress level. Current research seeks to augment these measurements with pore water pressure to help better understand conditions under which partial saturation could help improve the stability of the ballast and distinguish those field conditions from sites where positive pore water pressure can develop and destabilize the track, perhaps, theoretically, under very few load cycles. These sensors will be used to establish a monitoring network along a section of track where specific locations and monitoring parameters are required. However, due to battery life, sensor life, and other factors these types of smart sensors would not be suitable for large scale or long time-frame investigations of infrastructure performance over annual cycles of environmental condition variations or over several years of flooding intervals for which data is needed to help mitigate climate change concerns.

In cases where monitoring timeframes are long, FRA is pursuing development of autonomous smart sensor systems that require no internal power. These systems consist of passive sensors placed in track coupled with an external interrogator to make the measurement. These systems can be used to monitor track moisture, fouling level, and potentially vibrations. The passive sensors in track would be externally interrogated using electromagnetic energy to make measurements whereby the sensor is basically a reflector of the energy, and that reflected energy contains the information needed for the various measurements.

Remote sensing technology whether from an aircraft, UAV, or satellite can provide data on the position of the track structure. Baseline data on the location of fixed assets is critical to the recovery from disasters such as hurricanes and floods where the position of bridges and elevation of embankments can provide critical insight to the condition of infrastructure. This information can help with interpretation of in-place instrumentation by confirming locations of track settlement that could change baseline reading for in-place sensors.

4 INFORMING CLIMATE CHANGE DECISIONS

Climate change presents many challenges since the risks vary from fires in drought zones, disasters like hurricanes and floods, as well as increasing frequency of severe precipitation. Data from sensors and systems described in this paper can provide insight during both disaster planning and recovery efforts. This section will focus on the impact of extreme precipitation on the structural performance of railway track for planning and recovery from extreme precipitation events.

4.1 Planning

Planning for extreme precipitation events and hardening rail infrastructure against erosion, saturation, and flooding are the main challenges. While increasing the

elevation of rail infrastructure is often the initial concept, the economics of this approach and the impact of these changes on nearby structures are often problematic. Old railway embankments that are still in use have steep side slopes created by uncontrolled dumping and filling in historic trestles that would make elevation increases problematic due to the performance concerns related to the existing embankment and right of way issues. In addition, the industry has been working to lower railway track for clearance of double stack freight below overpasses which is incompatible with a systematic elevation increase of all low-lying areas in all but the most severe of risks. Focusing on low lying areas subject to damage is likely the best approach to identify locations at highest risk for severe damage that could require an increase in elevation to withstand climate change extremes.

Locating drainage problem areas and developing reliable repairs would be necessary for infrastructure to remain operational in extreme weather events. Many drainage structures that provided external track drainage have been lost or deteriorated over the decades. Saturated zones and other signs of these problems can be identified with GPR. Poor internal track drainage can also be identified with GPR along with ballast gradation measurements to identify ballast fouling conditions affecting internal drainage. A well drained subballast that is adequately open graded so that it does not hold water would also be necessary in locations with groundwater flow into the track structure. Finally, maintaining drainage ditches with adequate slope and outlets to carry water away from the tracks are a critical element of the external drainage system needed to ensure proper internal drainage.

Zones of poor track support also must be identified in planning to ensure that the track structure functions properly in areas at highest risk of impact from extreme precipitation events. Track deflection measurements, ballast modulus, and GPR indications of subgrade distress can provide the necessary information to provide a root cause diagnosis of the problems that can be expected to become more critical under extreme precipitation events.

Finally, during implementation of any of these plans, construction specifications and field compaction control for all materials must be considered. These critical locations will often not tolerate the dynamic loading and vibrations induced from track variabilities that develop during ballast compaction/consolidation under traffic. In addition, the use of special railroad subballast gradations will likely be required to provide the necessary filtration properties to prevent migration of subgrade into the ballast while providing adequate drainage capacity. Dense road base should not be used because of the large percentage of fines that help with compaction but that also hold water in the layer keeping the embankment moisture elevated for extended periods once saturated. Open graded bases will likely be

needed along with compaction criteria to ensure proper performance.

4.2 Recovery

In the recovery from extreme precipitation, flooding, and hurricanes, the extent of devastation to the track structure is one of the first questions to be addressed and remote sensing and imaging techniques can provide a rapid assessment. Once areas of track structure damage are identified, an assessment of the condition of the intact infrastructure must be conducted to determine the traffic worthiness of track sections to transport critical aid and equipment to begin the recovery and repair. Ballast modulus and portable GPR systems can provide critical guidance. Ballast modulus can indicate locations where settlement or shifting of the track structure has loosened the ballast or subballast. GPR can identify locations where water is trapped in the track structure and must be drained.

Field instrumentation for monitoring track performance in critical sections can be used to assess performance changes in the track structure. These changes could indicate structural problems that highlight areas where repairs may be required.

Once sections of the corridor have been identified as suitable for traffic, track geometry, GPR, and track deflection can provide critical input to the condition of track in terms of track capacity. Models and design techniques exist that use these datasets to assess track life in terms of number of applied load cycles at a given wheel load to assess track capacity and design life trade-offs to help engineers assess whether temporary weight or traffic volume restrictions might be warranted.

5 CONCLUSIONS

The challenges posed by recent increases in the number of severe precipitation events and the effects of climate change have posed severe challenges to transportation. Planning of infrastructure upgrades to help ensure reliability of infrastructure during extreme weather events and the ability to recover after disasters requires data on the track structure. While the internet of things is providing many sensors and associated monitoring technologies, the expected life of railway track exceeds the design life of most sensors. The railroad industry faces particular challenges related to:

1. Infrastructure age,
2. Open track structure allowing water flow through the structure, and
3. Complex structural interactions of the track superstructure with the substructure.

FRA is pursuing the concept of smart sensor applications which combines critical measurements to be made on track like ballast gradation with sensors on-board trains and inspection cars that provides critical baseline track structural condition data such as track deflection and the

extent and depth of clean ballast from GPR. In some critical situations, in place instrumentation may be necessary to obtain track performance information such as ballast displacement and position using the Penn State SmartRock or pore water pressure measurement technology. In these locations, active sensors with long life that can be measured remotely or passive sensors with no battery life limitations are being pursued. These datasets help quantify the ballast fouling, drainage capacity, load distribution, and load restraint offered by the track structure, all of which are very difficult to reliably assess qualitatively. In addition, this data provides the basis to apply existing track life models to assess load and speed restrictions that might be necessary in extreme situations of flooding and water inundation. Together these datasets provide important data on track structural performance and longevity that has not been available previously. This data can be used to provide information that supports a safe recovery and data that can provide information for infrastructure resilience planning.

The developments highlighted in this paper target improved infrastructure reliability at a time of extreme precipitation and ever-increasing loads and demands for service reliability. Addressing these challenges requires improved understanding of the structural conditions of track that can affect safety related structural performance of track to ensure rail infrastructure is more resilient to severe precipitation and can be rapidly rebuilt to reestablish transportation along these critical lifelines.

6 ACKNOWLEDGEMENT

Research into the characterization of engineering properties of railroad track infrastructure that forms the basis for this research was begun under the direction of Mahmood Fateh at FRA who worked with John Choros at the Volpe Center to define the investigation of the statistical relationships between and interdependencies among track inspection parameters. This work has continued under the direction of Hugh Thompson of FRA along with Gary Carr who has been succeeded by Sean Woody as Chief of FRA Track R&D. Mr. Thompson has worked with Ted Sussmann to expand the initial vision into the current research that has helped to define the failure mechanisms based in the track support system and substructure. This work is accomplished by a motivated group of industry and university researchers whose mission has been focused on advancing this technology to meet the demands of the industry and the emerging risks posed by severe precipitation events that currently test the limits of existing infrastructure systems on a more routine basis.

7 REFERENCES

Bayse, C., Wilk, S., & Y. Gao (2020). Ground Penetrating Radar (GPR) Technology Evaluation and Implementation, Federal Railroad Administration Report DOT/FRA/ORD-20/19.

DOE (2007). Deliveries of Coal from the Powder River Basin: Events and Trends 2005-2007, Infrastructure Security and Energy Restoration, Office of Electricity Delivery and Energy Reliability, U.S. Department of Energy.

Hartsough, C. M., Palese, J. W., and J. P. Kelley (2021). Categorizing Track Mud Spot Risk by Measurement of Vertical Track Deflection, Washington, DC: Federal Railroad Administration, Report No. DOT/FRA/ORD-21/24.

Li, D., Hyslip, J., Sussmann, T., & Chrismer, S. (2015). *Railway Geotechnics* (1st ed.). London: CRC Press. <https://doi.org/10.1201/b18982>.

Liu, S., McHenry, M., Gao, Y. & H. Huang (2017). Evaluating Ballast Migration on Curved Track Using 'SmartRock', American Railroad Engineering and Maintenance of Way Association Conference, Indianapolis, IN.

NTSB (1999). Derailment of CSX Freight Train Q316 and Subsequent Hazardous Material Release at Cox Landing, West Virginia, June 20, 1998, Railroad Accident Report 9901, National Transportation Safety Board, Washington, DC.

Selig, E. & Waters, J. (1994). *Track Geotechnology and Substructure Management*, Thomas Telford, London.

Sussmann, T. R., Ruel, M. & S. M. Chrismer (2012). Source of Ballast Fouling and Influence Considerations for Condition Assessment Criteria, Transportation Research Record 2289, pp. 87-94.

Sussmann, T. R., Stark, T. D., Wilk, S. T., & Thompson, H. B. (2017). Track Support Measurements for Improved Resiliency of Railway Infrastructure. *Transportation Research Record*, 2607(1), 54-61. <https://doi.org/10.3141/2607-08>.

Sussmann, T., Thompson, H., Bayse, C., Atherton, S. & A. Eriksen (2018). Ground Penetrating Radar (GPR) Technology Evaluation and Implementation on an FRA Research Car, AREMA, Chicago, IL.

Walsh, D. O. & B. Barrett (2018). Nuclear Magnetic Resonance Trackbed Moisture Measurement System, Federal Railroad Administration, Report DOT/FRA/ORD-18/03, Washington, DC.

The consideration of a changing climate for continuously welded rail

Merrina Zhang, & Alireza Roghani
National Research Council Canada, Ottawa, Ontario, Canada

ABSTRACT

Canada's climate has been warming at a rate that is approximately double the global average between 1948 and 2016. Climate change is impacting Canada more than the world as a whole, with climate models predicting further warming during all four seasons, ranging from an increase of a low of 1.8 °C to a high of 6.3 °C depending on the emissions trajectory (or representative concentration pathway). Canada's extensive rail network of 46,000 km traverses from coast to coast through different climate zones. Without expansion joints to release thermal stress buildup due to changes in ambient temperature, continuously welded rail (CWR) is vulnerable to climate change. The increase in magnitude and frequency of extreme temperature events projected under future climate will induce additional longitudinal stress in the rail. This paper examines how the projected changes in climate will impact CWR, latest climate data and models, as well as propose a method, along with a case study to incorporate future climate scenarios into CWR thermal stress management practices for specific sites.

1 INTRODUCTION

Canada has one of the most extensive rail networks in the world. Its 46,000 km of track stretches from the Pacific West coast all the way to the Maritimes, in addition to providing vital connections for communities in the far north (Transport Canada, 2020). As a result, the system traverses through some of the harshest climatic zones in the world, including the dry, humid mesothermal, humid microthermal and polar, (Natural Resources Canada, 1957) where temperatures can vary from -63 °C¹ (Government of Canada, 2019) to 49.6 °C².

Continuously welded rail (CWR) is the preferred main track construction in Canada, where long sections of rail steels are welded together to form the running surfaces for trains. Since CWR eliminates the gaps that allows for rail steels to expand or contract with changes in temperature, it is vulnerable to temperature fluctuations and extremes, as large thermal stresses can build up potentially leading to rail buckles in hot weather conditions and rail breaks in cold weather conditions.

Meteorological observations and records show that the climate around the world is changing, and that the change is impacting Canada at a faster rate, which will in turn magnify the CWR vulnerability.

The objective of this paper is to provide information on (1) how the projected changes in climate will impact CWR, (2) climate data and models, (3) a method derived from the latest climate research to incorporate future

climate scenarios into future CWR thermal stress management practices, (4) results from a case study, and (5) recommendations for the future.

2 CURRENT CWR THERMAL STRESS MANAGEMENT METHODS

Track systems generally consist of a superstructure (rails, anchors and ties) and a substructure (ballast, sub-ballast and subgrade).

For new CWR construction, the track system is pre-stressed during initial install such that it experiences no thermal stress at a set temperature point, also known as the rail neutral temperature (RNT). For Canada, RNT for new CWR construction is generally 32 °C, though this value may be adjusted based on the climatic conditions of different locations. Track systems (superstructure and substructure) that is well constrained will resist movements induced by thermal stresses. However, track resistance will weaken over time from wear and tear due to rail operations and maintenance, as well as climatic actions such as freeze-thaw cycles, and temperature fluctuations and extremes, making the system increasingly more vulnerable to failure due to large stress buildup (Zhang et al., 2018). In addition, RNT changes over time due to a variety of reasons including track maintenance activities. These factors make it difficult for track engineers to estimate the thermally induced longitudinal stress in CWR.

¹ Recorded in Snag, Yukon (1947)

² Recorded in Lytton, British Columbia (2021)

Railways employ a variety of methods to manage thermal stress in CWR, including (1) *track inspections (visual and automated)* to identify physical evidence of stress buildup, (2) *speed restrictions*, often imposed as temporary measures during extreme temperature conditions, as well as after maintenance activities such as tamping, (3) *rail de-stressing*, where the rail is cut to release excessive compressive stress buildup, (4) *rail re-stressing*, where rail lengths are adjusted using heaters or coolers, (5) *rail repair*, where plug rails are added to repair rail sections. (Ahmad et al., 2009)

3 THE IMPACT OF AMBIENT TEMPERATURE ON CWR

Thermal stress buildup in CWR is a complex relationship between a variety of factors, including RNT, the temperature of the rail steels and track resistance, all of which are difficult to measure on a network wide basis.

Even though rail temperature can be affected by factors such as sun exposure, wind conditions, shade and topography, it is strongly influenced by ambient temperature conditions, a parameter that is measured and recorded across Canada, and more readily available.

Several railways in North America use equations 1 and 2 to determine rail temperature using ambient measurements in summers and winters, respectively, (Ahmad et al., 2009)

$$T_r = T_a + 16.6 \quad [1]$$

$$T_r = T_a \quad [2]$$

where T_r is the rail temperature, and T_a is ambient temperature, both in °C.

4 THE CONTEXT OF A CHANGING CLIMATE IN CANADA

Official temperature measurements are typically done at a height of 2 m above the ground in Canada. While some weather stations have been operating since the 1840's, few official observation sites' records predate 1948. For this reason, 1948 is most often used as the starting point for climatic analysis in Canada.

Between 1948 and 2016, the average temperature in Canada increased by 1.7°C, approximately double the global average. As shown in Table 1, the warming trend is expected to continue in the future in every region across the country.

Table 1. Projected changes in annual mean surface air temperature, based on Coupled Model Intercomparison Project Phase 5 multi-model ensemble, relative to 1986-2005. (Government of Canada, 2019)

Region	Scenario; Period; Median Temperature Change (25 th , 75 th percentile), °C			
	Low Emission Scenario		High Emission Scenario	
	2031-2050	2081-2100	2031-2050	2081-2100
British Columbia	1.3 (0.8, 1.9)	1.6 (1.1, 2.1)	1.9 (1.4, 2.5)	5.2 (4.3, 6.2)
Prairie	1.5 (1.1, 2.1)	1.9 (1.2, 2.2)	2.3 (1.7, 3.0)	6.5 (5.2, 7.0)
Ontario	1.5 (1.1, 2.1)	1.7 (1.0, 2.1)	2.3 (1.7, 2.9)	6.3 (5.3, 6.9)
Quebec	1.5 (1.0, 2.1)	1.7 (1.0, 2.2)	2.3 (1.7, 2.9)	6.3 (5.3, 6.9)
Atlantic	1.3 (0.9, 1.8)	1.5 (0.9, 2.0)	1.9 (1.5, 2.4)	5.2 (4.5, 6.1)
North	1.8 (1.2, 2.5)	2.1 (1.3, 2.5)	2.7 (2.0, 3.5)	7.8 (6.2, 8.4)
Canada	1.5 (1.0, 2.1)	1.8 (1.1, 2.5)	2.3 (1.7, 2.9)	6.3 (5.6, 7.7)

5 POTENTIAL IMPACTS OF THE PROJECTED CHANGES ON CWR TRACK

Warming has been observed across Canada, however the rate of warming has not been uniform. From 1948 to 2016, Canada's northwest experienced the most rapid rates of warming, where the average temperature of some regions increased by more than 3°C (approximately 3 times the global average). Comparatively, over the same time period, Canada's southeast underwent the slowest rate of warming, where the average temperature in some maritime regions increased by less than 1°C. In most parts of Canada, warming has impacted winter more than other seasons, especially in northern British Columbia (BC) and Alberta, Yukon, Northwest Territories, and western Nunavut, where average temperatures have increased between 4°C to 6°C. This trend is consistent in Ontario and the Prairie provinces, though the warming in these regions has been occurring at a comparatively slower rate. In contrast, Quebec and Atlantic Canada experienced more rapid rates of warming in the summer months than in the winter months. Going forward, accelerated warming is projected to disproportionately impact more northern latitudes. These trend signals the need to incorporate more regional approaches to CWR stresses management based on the local climate.

Warmer winters across Canada have resulted in a country-wide average decrease of more than 15 frost days³ and 10 ice days⁴ from 1948 to 2016. This trend is expected to continue, along with snow shifting to rain in the spring and autumn. In addition, since 1948, the number of extreme hot days⁵ have increased in Canada and the trend is expected to drastically continue. By the end of the century, as seen in Figure 1, under high emissions scenarios, the number of hot days could be increased by 50 per year in regions that currently

³ Daily minimum temperature ≤ 0°C

⁴ Daily maximum temperature ≤ 0°C

⁵ Daily maximum temperature ≥ 30°C

experience them. These trends signal a potential future shift of needing to manage more compressive stresses due to warmer weather than tensile stress resulting from cold weather, including potentially more inspections due to hot days, more hot weather slow orders, and more rail distressing. These factors also signal the need to review RNT values for future new CWR construction, especially in regions projected to undergo more rapid warming.

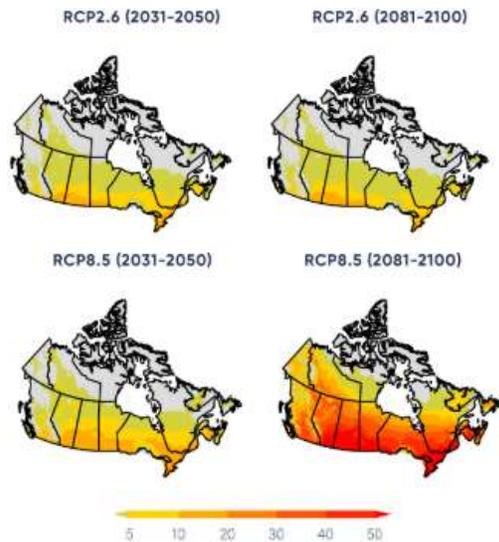


Figure 1. Projected number of annual hot days under different emissions scenarios. (Government of Canada, 2019)

The next sections provide information on modelling of future climate, available data, as well as a method derived from current climate science that could be employed to incorporate climate projections into future CWR thermal stress management practices.

6 CLIMATE MODELS AND DATA

Climate models are simulated by earth systems models, which consists of mathematical representations of the complex physical and biogeochemical processes related to the atmosphere, ocean, land surface, and cryosphere (Government of Canada, 2019). These processes are global in nature, are typically examined at a resolution of 100 km or more, and are used to simulate historical climate as well as make projections of future climate based on expected human activities (or emission scenarios).

All climate models have uncertainties since the mathematical representations used are approximated. To compensate for this, an ensemble of different climate models, each with its own set of assumptions, methodologies and projections, is often used for analytical purposes. The Coupled Model Intercomparison Project (CMIP) established by the World Climate Research Programme sets out standards and experimental protocols that allows for direct comparisons of different models, assumptions and results under a common set of conditions and scenarios.

The climate models and projections developed within the fifth phase of the project (CMIP5), was used in the Intergovernmental Panel on Climate Change's (IPCC) fifth assessment report (released in 2014). The CMIP5 emission scenarios are referred to as *representative concentration pathways* (RCP) with an associated number indicating the change in radiative forcing (W/m^2). RCP2.6 represent a low emission scenario, while RCP8.5 represent a high emission scenario. A new phase of the project (CMIP6) is currently underway with an expanded set of standards and modified emissions scenarios. Projections from the newest suite of climate models are expected to be incorporated into IPCC's 6th assessment reports, scheduled for release from late 2021 through to 2022. The ensemble approach enables researchers to examine how climate models differ, how parameters affect results, and the range of results produced under the same scenarios.

At the time of writing, there are a number of curated future climate projections datasets available for use in the public domain. They are downscaled from ensembles of CMIP5 global climate model projections of 100 km+ resolution to a resolution of 10 km to 15 km for regions within Canada (Zhang et al., 2021). These datasets include, Ouranos (Consortium sur la climatologie régionale et l'adaptation aux changements climatiques, 2021), the Pacific Climate Impacts Consortium (Pacific Climate Impact Consortium, 2021), and ClimateData.ca (Environment and Climate Change Canada et al., 2018). With the release of CMIP6, it is expected that datasets from the new suite of climate models will be available for use in the near future.

Climate records such as daily temperature observations from weather stations are readily available in Canada through the historical data site of the Government of Canada (Government of Canada, 2021a) and ClimateData.ca. As shown in Figure 2, there is a relatively extensive distribution of weather stations along key rail corridors.

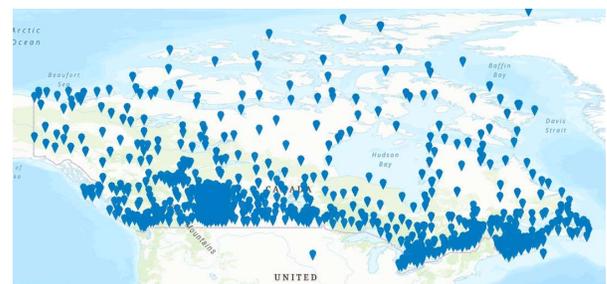


Figure 2. Meteorological stations across Canada. (Government of Canada, 2021b)

7 PROPOSED METHOD TO INCORPORATE FUTURE CLIMATE SCENARIOS INTO CWR THERMAL STRESS MANAGEMENT PRACTICES

The proposed method makes use of data readily available in the public domain, including climate projections and historical temperature records as described in section 6.

Results from climate models typically have a simulated historical climate as shown in Figure 3 in grey, and projections of future climate as shown in Figure 4 in colour, under various emissions scenarios.

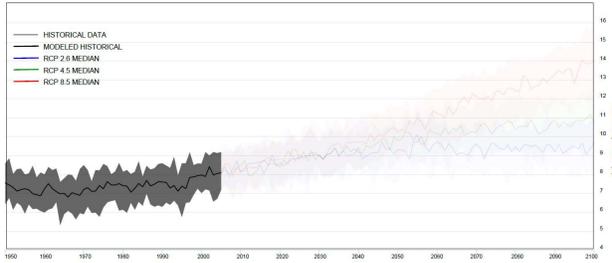


Figure 3. An example of simulated historical climate by climate models. (Environment and Climate Change Canada et al., 2018)

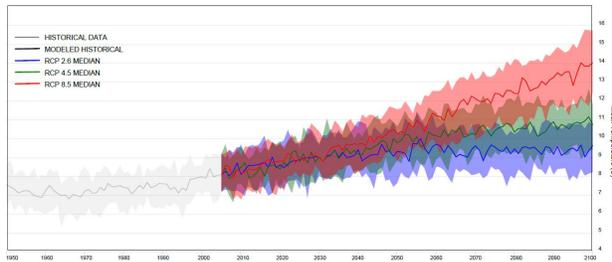


Figure 4. An example of predicted future climate by climate models. (Environment and Climate Change Canada et al., 2018)

The first step is to evaluate how realistic the selected climate model (or ensemble) represents the key earth systems that drive the model. This is done by comparing its simulated past climate with actual historical temperature records (observations) from weather stations as shown in Figure 5, and determining whether a calibration/transfer function is needed to correlate the two datasets.

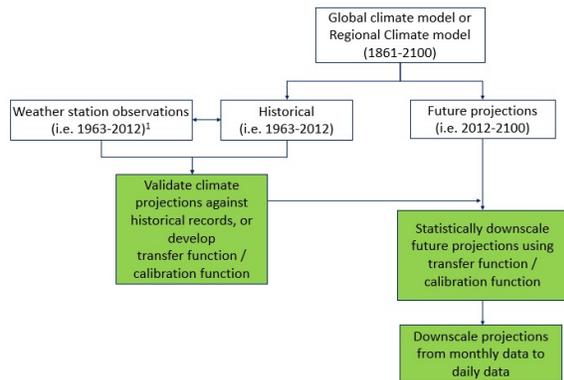


Figure 5. Schematic of the proposed method to derive future data for analytical purposes.

The next step is to integrate the calibration/transfer function into future projections as per Figure 5, to produce a calibrated dataset. This step is also known as statistical downscaling.

Future temperature projections within climate models (Figure 4) are often available in the form of monthly averages. For analysis requiring daily temperature projections, it will be necessary to generate daily temperature datasets using one of many established downscaling methods, such as delta change, enhanced change factors, and synoptic map, etc. (Zhang et al., 2021).

These projected monthly/daily temperatures can be used to compare with historical and present baseline environmental conditions to determine how the climate has changed, and the projected change, along with the rate of projected change within a 5-year, 10-year, 20-year and 30+ year window. This information will be valuable in determining the regions within Canada most vulnerable to climate change, whether there is a need to review current CWR thermal stress management methods, including RNT values.

8 CASE STUDY AND ANALYSIS

Kamloops, BC was selected as a case study for the initial phase of this research for a number of reasons. The Kamloops A weather station, as shown in Figure 6, has continuous observational records dating back to 1951. The region is situated near a key rail corridor where both Canadian Pacific Railway and Canadian National Railway Company operate. The region experienced record heat during the summer of 2021. In addition, as shown in Figure 1, the number of annual hot days for the region is projected to increase under all future emission scenarios.

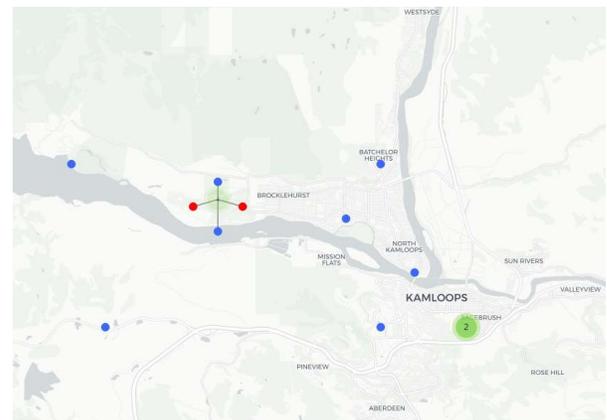


Figure 6. Location of Kamloops A weather station (the red dot on the right stores data from 1951 to 2013, the red dot on the left stores data from 2013 to present). (Environment and Climate Change Canada et al., 2018)

To evaluate how the climate within the region has been changing, as well as the impact of future climate on CWR, the full dataset from the Kamloops A weather station cell (Figure 6, 10 km resolution) was extracted from ClimateData.ca. The extraction included the actual temperature records from the weather station, as well as temperature projections from a 24-model ensemble. The simulated historical temperature by the climate model ensemble is from 1950 to 2005. The projected future

climate is in the form monthly averages under different emission scenarios. For each RCP, low, median, and high are provided to reflect the range of uncertainties within the models. (Environment and Climate Change Canada et al., 2018)

As can be seen in Figure 7, the climate models projected an increase of annual average temperature for the region starting at around 1960. At around 2030, the rate of increase will diverge depending on the emission scenario, ranging from an increase of around 2 °C to more than double by century end (compared to 1960).

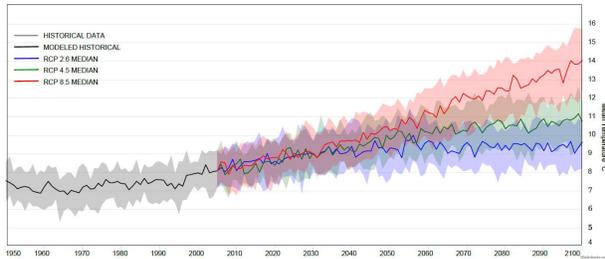


Figure 7. Projected mean annual temperature by climate models. (Environment and Climate Change Canada et al., 2018)

Figure 8 compares the climate projections with actual temperature records from the weather station. Due to data availability, only the median value of each emission scenario is plotted in Figure 8. In general, the temperature records (from 1951 to 2021) indicate that the annual average temperature ranged between 6.5 °C and 11 °C, and is at least 1 °C higher than the projected historical range (Figure 7) by the climate models. If this trend was to continue, it would appear that the future annual average temperature increase would be more in line with a medium or high emission scenario. This set of climate projections only requires minor calibrations to correlate with the Kamloops A weather station data.

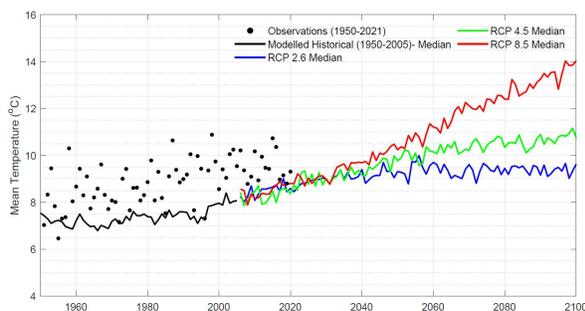


Figure 8. Comparison of mean annual temperature observations versus climate model projections. (Environment and Climate Change Canada et al., 2018)

The coldest annual temperatures projected by climate models (Figure 9) also indicate a warming trend for the region. Though, there appears to be more overlap in the range of projected future coldest annual temperatures under the different emission scenarios, as there is not a clear diverging trend in the rate of change until the 2070's.

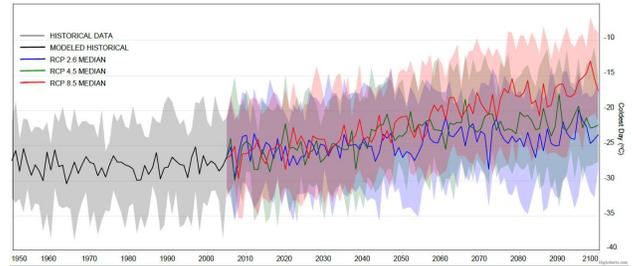


Figure 9. Projected coldest annual temperature by climate models. (Environment and Climate Change Canada et al., 2018)

Figure 10 compares the projections of the coldest annual temperatures with actual temperature records from the weather station. The temperature records (from 1951 to 2021) indicate that the annual coldest temperature ranged between -37.5 °C and -13 °C. Since 1980, rarely has the coldest annual temperature dropped below -29 °C. In general, temperature observations are either within the projected historical range of the climate models, or higher than the projected range (by up to 13 °C). Since there is quite an overlap between the projected future temperature ranges of the different emissions scenarios, the current trend appear to be in line with all three scenarios in the near future. Compared to average annual temperature dataset, this set of climate projections requires more calibrations to correlate with the Kamloops A weather station data.

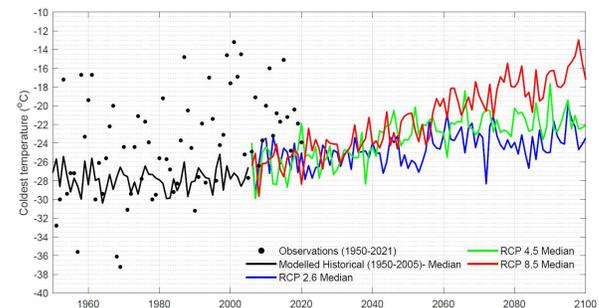


Figure 10. Comparison of coldest annual temperature observations versus climate model projections. (Environment and Climate Change Canada et al., 2018)

The number of annual days with temperature less than -25 °C projected by climate models, as shown in Figure 11, is decreasing for the region. Depending on emission scenario, the region might not experience below -25 °C starting as early as the mid 2030's.

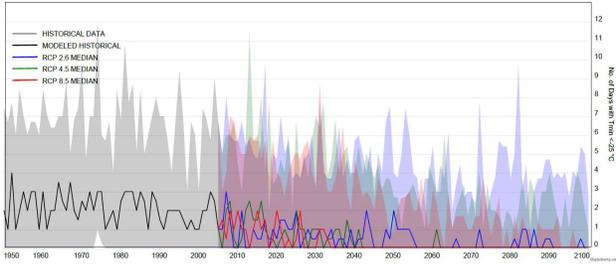


Figure 11. Projected number of annual cold days by climate models. (Environment and Climate Change Canada et al., 2018)

Figure 12 compares the projections of the number of < 25 °C days with actual temperature records from the weather station. In general, the temperature records (from 1951 to 2021) indicate that most years had 6 or less cold days, with many winters without any cold days, which is within the projected historical range by the climate models. This trend appears to be in line with future projections regardless of emission scenario. This set of climate projections only requires minor calibrations to correlate with the Kamloops A weather station data.

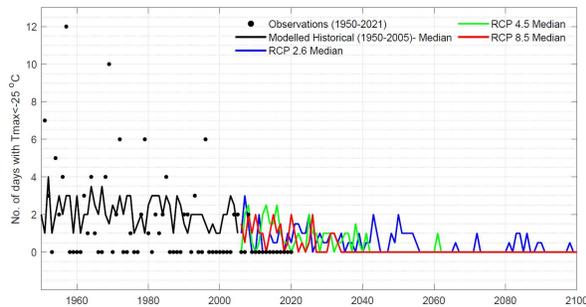


Figure 12. Comparison of the number of annual cold days recorded versus climate model projections. (Environment and Climate Change Canada et al., 2018)

The hottest annual temperatures projected by climate models (Figure 13) also indicate a warming trend for the region starting around the mid 1970's. At around 2045, the rate of increase will diverge depending on the emission scenario, ranging from an increase of around 2 °C to 12 °C by century end (compared to 1975).

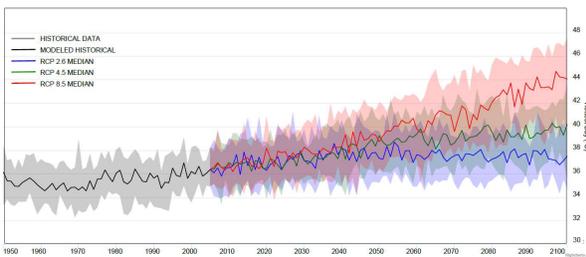


Figure 13. Projections of hottest annual temperature by climate models. (Environment and Climate Change Canada et al., 2018)

Figure 14 compares the projections of the hottest annual temperatures with actual temperature records from the weather station. The temperature records (from 1951 to 2021) indicate that the annual hottest temperature ranged between 32.5 °C and 41 °C, and in general is either within the projected historical range of the climate models, or higher than the projected range by around 2 °C. If this trend was to continue, it would appear that the future hottest annual temperature increase would be more in line with a medium or high emission scenario. This set of climate projections only requires minor calibrations to correlate with the Kamloops A weather station data.

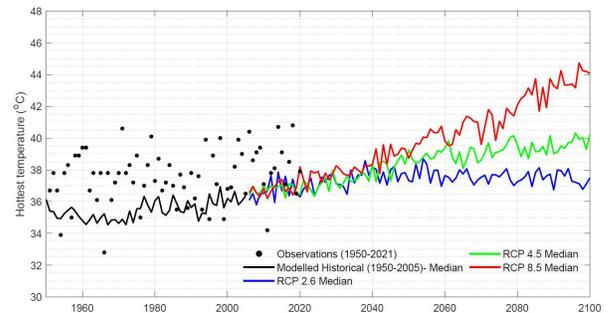


Figure 14. Comparison of hottest annual temperature observations versus climate model projections. (Environment and Climate Change Canada et al., 2018)

The number of annual hot days projected by climate models, as shown in Figure 15, is increasing for the region starting around the 1970's. Depending on emission scenario, the upward range of the number of hot days per year could increase to around 70 days by 2050, and to more than 100 days by century end.

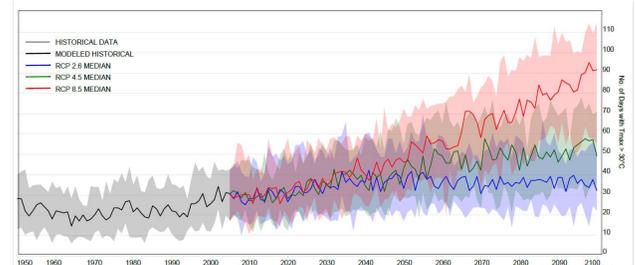


Figure 15. Projections of the number of annual hot days by climate models. (Environment and Climate Change Canada et al., 2018)

Figure 16 compares the projections of the number of hot days with actual temperature records from the weather station. The temperature records (from 1951 to 2021) indicate that the number of hot days within a given year ranged from as few as 4 days, to as many as 67. Quite a number of years had 10 to 15 more hot days than projected by climate models. The trend appears to be more in line with a future medium or high emission scenario. Compared to other datasets, this set of climate projections requires more calibrations to correlate with the Kamloops A weather station data.

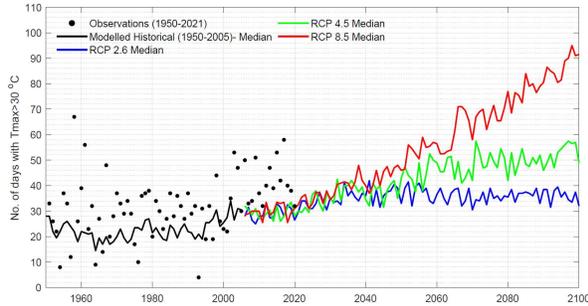


Figure 16. Comparison of the number of annual hot days recorded versus climate model projections. (Environment and Climate Change Canada et al., 2018)

9 CONCLUSIONS

Without expansion joints to release thermal stress buildup due to changes in ambient temperature, CWR is vulnerable to climate change.

This paper explored how projected future changes in ambient temperature will impact CWR, provided information on currently available climate models and data, as well as a method derived from climate research on how to incorporate future climate scenarios into CWR thermal stress management considerations.

A case study was performed on the Kamloops, BC region. Temperature records from 1951 were compared with the projected temperatures from a 24-climate model ensemble with respect to (1) the average annual temperature, (2) the coldest annual temperature, (3) the number of cold days per year, (4) the hottest annual temperature and (5) the number of hot days per year. The simulated historical temperatures by climate models compared well to historical records in 3 of the 5 datasets, with the coldest annual temperature and the number of hot days per year needing additional calibration. The trend indicated that the average annual temperature of the region is on track to increase to a range of 8 °C to 11 °C by 2030 and 2040, 8 °C to 12 °C by 2050 (from a baseline of 6 °C to 11 °C, 1951 to 2021). The coldest annual temperature is on track to increase, potentially significantly from a baseline of -37.5 °C to -13 °C (1951 to 2021). The number of cold days a year is on track to decrease to less than 6 per year by 2030 and 2040 (in line with the baseline, 1951 to 2021), and where many years could potentially experience no cold days. The hottest temperature is on track to increase to a range of 35 °C to 41 °C by 2030, to upwards of 42 °C by 2040 and 2050 (from a baseline of 32.5 °C to 41 °C, 1951 to 2021). The number of hot days per year is also on track to increase, potentially significantly compared to the baseline of 4 to 67 per year from 1951 to 2000, and 20 to 60 per year from 2000 to 2021, indicating a potential need to review future RNT for the region and possible prolonged operation in extreme heat annually.

10 RECOMMENDATIONS

As evident in temperature observations across the country, a changing climate has already impacted every region within Canada, though the rate of change has not been uniform. An important next step is to determine which regions/locations along key rail corridors that are most vulnerable to climate change with respect to the parameters most relevant to CWR thermal stress management, including the rate of warming, projected extreme temperature fluctuations, as well as the number of projected cold days and hot days.

Another recommendation is to examine past climate records, RNT, maintenance records, and operating speed records in those vulnerable locations, to determine whether a changing climate has already impacted CWR thermal stress management.

It is also recommended to use the latest climate models and data to project future climate scenarios similar to that used in the case study to review current construction, operation and maintenance methods to determine whether there will be a need to better adapt them within a 5-year, 10-year, 20-year and 30+ years timeframe.

11 ACKNOWLEDGEMENT

The authors would like to express sincere appreciation for the support of climate resilience related research from Aviva Shiller and Erica Weterings (Crown-Indigenous Relations and Northern Affairs Canada), Paul Barrette, Sylvie Chenier and Derek Morrell (National Research Council Canada).

12 REFERENCES

- Ahmad, S., Mandal, N. K., and Chattopadhyay, G. (2009). *Track Stability Management – Literature Review: Theories and Practices*. Brisbane, Queen's Land: 82.
- Consortium sur la climatologie régionale et l'adaptation aux changements climatiques. (2021). "Climate Portraits." Retrieved September 8, 2021, from <https://www.ouranos.ca/climate-portraits/#/>.
- Environment and Climate Change Canada, The Computer Research Institute of Montréal, Ouranos, Pacific Climate Impacts Consortium, Prairie Climate Centre, and HabitatSeven. (2018). "Climate Data for a Resilient Canada." Retrieved September 9, 2021, from <https://climatedata.ca>.
- Government of Canada (2019). *Canada's Changing Climate Report*. Gatineau.
- Government of Canada. (2021a). "Historical Data." Retrieved September 8, 2021, from https://climate.weather.gc.ca/historical_data/search_historic_data_e.html.

Government of Canada. (2021b). "Meteorological Station Catalogue." Retrieved September 8, 2021, from <https://open.canada.ca/data/en/dataset/9764d6c6-3044-450c-ac5a-383cedbfef17>.

Natural Resources Canada (1957). Atlas of Canada - Climate Zones. Government of Canada, Open Government.

Pacific Climate Impact Consortium. (2021). "Statistically Downscaled GCM Scenarios." Retrieved September 8, 2021, from https://data.pacificclimate.org/portal/downscaled_gcms/map/.

Transport Canada. (2020, July 13). "Rail Transportation." Retrieved September 14, 2021, from <https://tc.canada.ca/en/corporate-services/policies/rail-transportation>.

Zhang, M., Roghani, A., and Barker, C. (2018). Track Research: Longitudinal Rail Stress Management Gap Analysis Research Report (Phase 1). Ottawa, National Research Council Canada.

Zhang, M., Roghani, A., Hill, L., and Barrette, P. (2021). Development of a Winter Road Climate Risk and Vulnerability Review Framework 2020-2021 Update. Ottawa, National Research Council Canada: 89.

Climatic trigger signature and adverse ground conditions resulting in destructive debris flow events in British Columbia 2017 to 2020

Tim Keegan, Matt Cormier, Garry Stevenson
Klohn Crippen Berger, Vancouver, BC, Canada

ABSTRACT

Four debris flow events during Fall 2017 and one in January 2020 in BC were examined: three events near the BC Northwest coast on September 9-10, October 22-23, 2017, and October 27-29, 2020; an event near Lytton, BC November 20-23, 2017; and an event near the BC Southwest coast January 30, 2020. The climatic signature of the five events were similar in terms of setting, rainfall, snow melt, temperature change and trigger causal factors (Climatic Trigger Signature), however differences in factors such as available snowpack in the upper basin and recent wildfires in the basins had a significant effect on the destructiveness of the debris flow/flood event. All debris flows occurred during a climatic system characterized as a landfall atmospheric river event. The Climatic Trigger Signature (CTS), ground conditions and magnitude and duration of run-off resulted in varying degrees of destruction. Minimal snow was present in the upper basins prior to the September 9-10, 2017 event. Run-off for this event was roughly proportional to rainfall over a 12-hour period and resulted in minor damage to road crossings. Preceding the October 22-23, 2017 event, there was an approximate 2.2 m snow pack and near freezing temperatures. The event culminated when temperatures rapidly rose, significant warm precipitation occurred and approximately 2 m of snow melted over an 18-hour period. This resulted in significant run-off and erosion of road crossings and almost 1 km of the road being washed-out. Before the November 20-23 event there was approximately 1.4 m snow in the upper basin and temperatures were slightly below freezing. The event was initiated by rapid temperature rise, significant warm precipitation and approximately 1 m snow melt over a 36-hour duration. The resulting approximate 50,000 m³ debris flow filled the catchment on the uphill side of a railway embankment and eventually overtopped the tracks. A 2015 wildfire contributed to this event by burning part of the upper basin which allowed for increased erosion and debris entrainment. This paper examines the precursor climate trigger signature and ground conditions for the events and suggests a risk estimation model, inspection regime and trigger action response plan matrix to manage them.

1 INTRODUCTION

Climatic Trigger Signatures (CTS) refer to the real-time record of the relevant climate parameters from the time a climatic system can be forecasted until the effects of the system have passed. The CTS that is currently an increasing risk from debris flows in the coastal mountains of Western Canada is the combination of Atmospheric Rivers (rivers of moist and warm air masses from the Southern Pacific, in this case), early high snow accumulation in the upper basins, and recent wildfires in the basins. Relevant climatic parameters, acting on the upper basins, include hourly and preceding warm rainfall; ambient temperature; accumulated snow; and the temperature of that snow. Snow accumulation in the upper basin, combined with warm rain results in flash melting and greatly increases the intensity and duration of surface run-off.

There is increasing evidence in the past 5 to 10 years that this CTS is increasing both in frequency and intensity and that it is associated with global warming.

The intent of this paper is to document the characteristics of five significant runoff events that

occurred in BC, between 2017 and 2020, that resulted from various forms of this CTS and the associated adverse ground conditions, such as topography, recently burnt or logged upstream, accumulated snow, and soil and rock types in the upstream basins. The understandings drawn from this exercise will be used to further develop operational risk management methodologies to identify where the adverse ground conditions exist; runoff paths that threaten elements of risk such as linear or site-specific facility; available forecasting capabilities; real time Trigger Action Response Plans (TARP); and appropriate hazard specific warning and mitigation measures.

The five significant runoff events, examined in this paper, are listed below:

- Three events occurred off the southern slopes of Horetzky Mountain located approximately 120 km inland from the midpoint of the BC Pacific coastline, two during Fall 2017 (Sept. 9-10 and Oct. 22-23) and one during the Fall of 2020 (Sept. 27-29).

- One event occurred off the eastern slopes of the Fraser River Valley approximately 9 km downstream from Lytton, BC during Fall 2017 (Nov. 22-23).
- One event that resulted in several runoff hazard events, including debris flows, earth and rock landslides and hydraulic erosion that occurred between Lytton and Hope in the Fraser Canyon in the Winter of 2020.

The CTS's of the five events had distinctly similar parameters in terms of rainfall, snow melt, and temperature change as all five debris and rock landslide scenarios occurred during a climatic system referred to as an Atmospheric River (AR) (discussed in Section 3). The different magnitude-duration of the CTS's, combined with variable adverse ground conditions resulted in a large variation in the severity of damage to the downstream infrastructure. Through this examination, it was realized that early forecasting of these CTS's combined with identification of the existing adverse ground conditions upslope of elements at risk could be objectively used to manage the risk proactively and operationally in real time.

This paper characterizes Debris Flood/Flow Hazard Scenarios; describes Atmospheric Rivers and their possible correlation to global warming; examines CTS's that led to the five significant runoff events; proposes a Debris flow subjective risk awareness matrix for use in a TARP.

2 DEBRIS FLOOD/FLOW HAZARD SCENARIO

In mountainous terrain, the CTS's described in this paper bring ground responses in the form of damaging rockfalls, debris flows, debris floods, avulsion, gully erosion (washouts), and stream erosion that blocks, erodes and/or damages infrastructure; affect the environment; and threaten lives. The significant ground response resulting from these five CTS's and associated run-off, were caused by a combination of debris flows, debris floods, avulsions, washouts, and stream erosion. This combinations of geohazard events are referred to as ground hazard scenarios (GHS) (Keegan, 2007) and are determined by the initiating process which in this case are debris flood/flows. To manage the risk associated with debris flow scenarios, the entire debris flow hazard scenario needs to be addressed. The debris flow hazard scenario maps the debris flow hazard from initiation (CTS) to failure of the infrastructure and subsequent recovery to full functionality. Figure 1 depicts the failure mode effects analysis (FMEA) for the debris flow hazard scenario developed from review of over 30 case examples in BC.

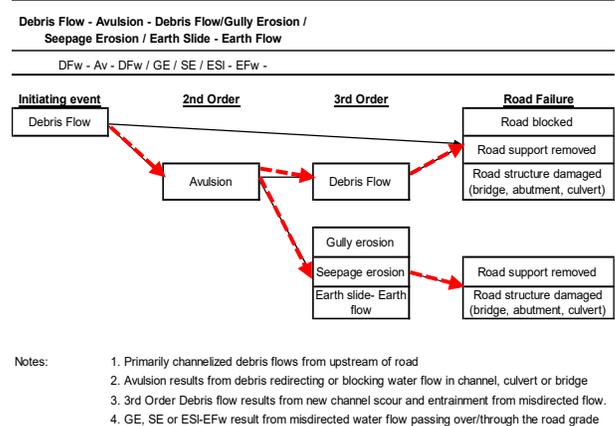


Figure 1. Simplified FMEA for the Debris Flow - Avulsion – Debris Flow / Gully Erosion / Seepage Erosion / Earth Slide - Earth Flow hazard scenario (modified from Keegan (2007)).

3 ATMOSPHERIC RIVER

Atmospheric River (NOAA, 2015) is a non-technical term for a meteorological phenomenon characterized by a strong and persistent flow of warm atmospheric moisture and associated heavy precipitation, in this case, from the waters adjacent to the Hawaiian Islands and extending to any location along the Pacific coast of North America. Atmospheric river refers to such narrow corridors of enhanced water vapor transport at mid-latitudes around the world. The conditions are often created by the Madden-Julian oscillation, an equatorial rainfall pattern which feeds its moisture into this pattern. shows three screen shots from the AR portal sponsored by the Physical Sciences Laboratory of the US National Oceanic & Atmospheric Administration (NOAA) of an actual potential AR system, in this case, forecasted to break up before hitting the BC Coast early morning November 24, 2020.

These forecasting model projects 7.5 days into the future and thus provide a valuable AR forecasting tool that could be effectively used in a TARP. Using the AR Portal, the first Weather Watch (forecast) can be received that conditions are favourable for the development of severe AR in the Pacific and heading towards the element at risk approximately 7 days before it is projected to reach the BC coast. Again, using the AR Portal, a Weather Warning that a fully developed AR will pass over the upper basins above the element at risk can be issued within 48 hours of its arrival.

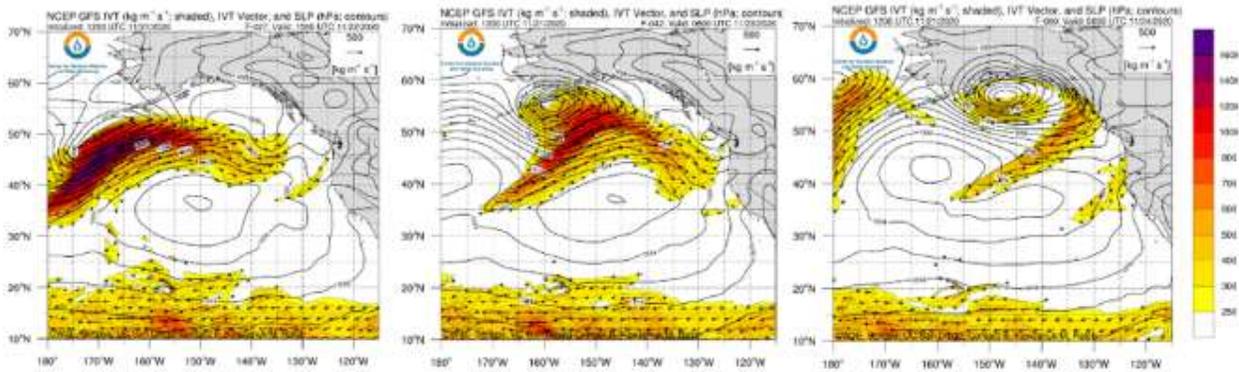


Figure 2. Integrated Water Vapor Transport (IVT) and Integrated Water Vapor (IWV) Analyses and Forecasts 15:00 UTS 11/22/2020 - 00:00 UTS 11/24/2020 (Center for Western Weather and Water Extremes; https://psl.noaa.gov/arportal/ar_forecasts/)

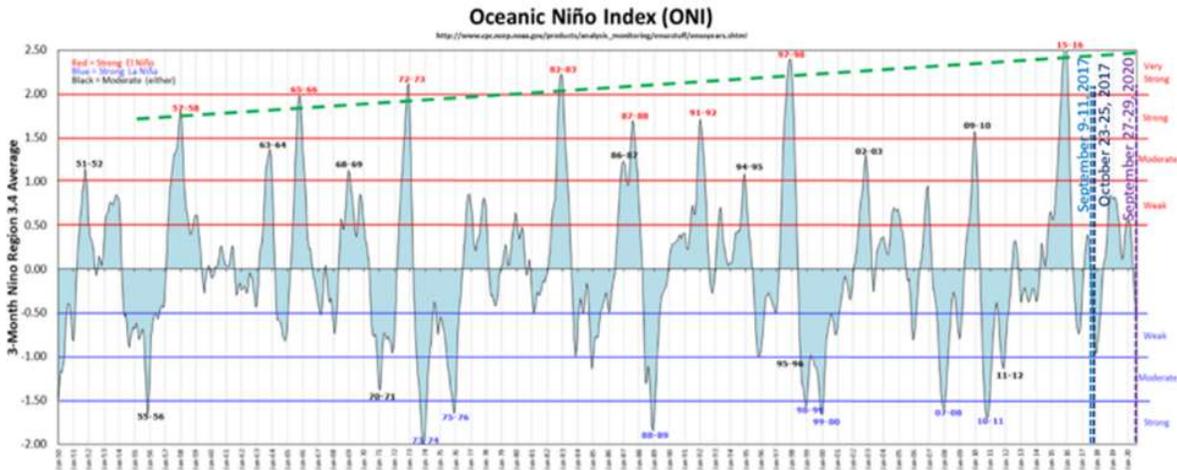


Figure 3. Chronology of El Niño and La Niña 1950-2020

According to the World Meteorological Organization, AR's are slightly more prevalent during years when there is no El Niño (warming) climate phenomena or when there is a La Niña (cooling) climate phenomena.

Perhaps coincidentally, 2017 and 2020 are both years that follow both a very strong and a weak El Niño, respectively, and correspond to weak and moderate La Niñas developing in the latter half of the year (Fall-Winter) (see Figure). Both years also correspond to the occurrence of the AR's that affected the BC Coast in the Falls of 2017 and 2020.

Each El Niño and La Niña climate phenomena lasts 9-12 months, and occurs on average every 2-7 years, but they happen irregularly. While many scientists believe that climate change does impact El Niño, and thus Atmospheric Rivers, they are not yet certain of the mechanism (Pinna Sustainability, 2014).

Also notable from Figure is the apparent trend of increasing intensity of the stronger El Niños, that have occurred 6 times in the last 70 years (on average every

12 years), have shown an increase in intensity from Strong to Very Strong.

4 CASE EXAMPLES

The CTS, ground conditions and magnitude of damage for five debris flow events, which occurred during AR events during Fall 2017 and Fall 2020 in BC were examined. The case histories for each of the five debris flow scenario events presented in the following sections are characterized by their CTS, ground conditions, magnitude-duration of run-off, and relative severity of damage.

4.1 Horetzky Mountain Debris Flows, Fall 2017

From September 9th to 10th and again from October 23rd to 25th, 2017, the south side of Horetzky Mountain, which is located approximately 120 km inland the central coast of BC, was subjected to significant CTS events involving a combination of rapid rise in temperature, intense rain and rapid snow melt that resulted in significant run-off from the upper basins that initiated destructive debris flows that inflicted varying levels of damage to infrastructure at the base of the valley.

The Tahtsa Lake Automated Snow Weather Station (ASWS), used to complete the assessment of the three climatic events, is ideally situated, both in proximity and elevation, to provide representative archival records and real-time indicators of the conditions experienced in the upper basins of Horetzky Mountain.

4.1.1 CTS September 9th to 10th, 2017

The September 9th to 10th CTS, depicted in Figure 2, involved approximately 100 mm of rain in a 19-hour period, which translated into a total order-of-magnitude average run-off of 11m³/sec and total volume of 754,000m³ of water from the 4 main basins. With no accumulated snowpack there was no snow Figure 2. September 9th to 10th, 2017 CTS - BC Central Coast.

melt associated with this run-off. Review of the available record indicates the climatic system and runoff was a 2-year return period event.

The resulting debris flows generally filled the catch basins upstream of the road infrastructure and caused negligible damage.

4.1.2 CTS October 23rd to 25th, 2017

This CTS depicted in Figure 3 involved approximately 100mm of rain and 125mm of vertical snow-water equivalent (SWE) of snow melt, corrected for slope angle as illustrated in , in a 45-hour period which translates to a total order-of-magnitude average run-off of 10.5m³/sec and total volume of 1,696,000m³ of water from the four main basins.

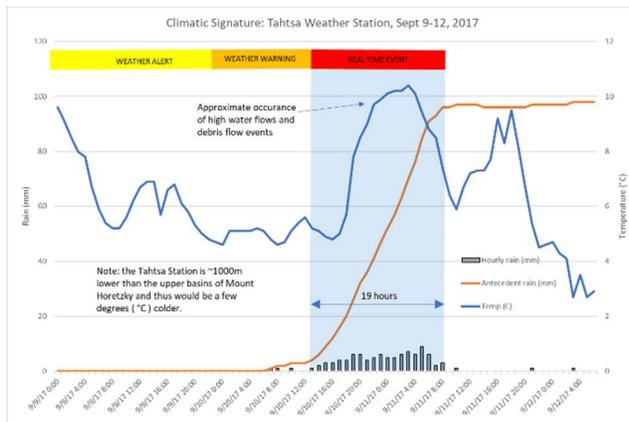


Figure 4. September 9th to 10th, 2017 CTS - BC Central Coast.

This is deemed to have been an anomalous event of greater than a 20-year return period, based on the analysis of the Tahtsa Lake ASWS record and the unfortunate simultaneous occurrence of the following causal factors:

- High early snow accumulations of 400% of average and 150% of the maximum recorded (see Figure .)
- Two days of 0°C prior to the event served to partially melt the snow, priming it for flash snow melt.
- Occurrence of a strong AR which dropped 100 mm of warm rain on the partially melted snow in the upper basin over a 35-hour period.
- Although the 6 to 10°C rise in temperature contributed to the snow melt, the warm (>15°C) rain seems to have significantly intensified the snow melt resulting in an additional 125 mm of SWE.
- The resulting two-day runoff of 225 mm, from combined rain fall and snow melt, is 2.5 times the next highest runoff in the record, supporting the abnormality of this occurrence.

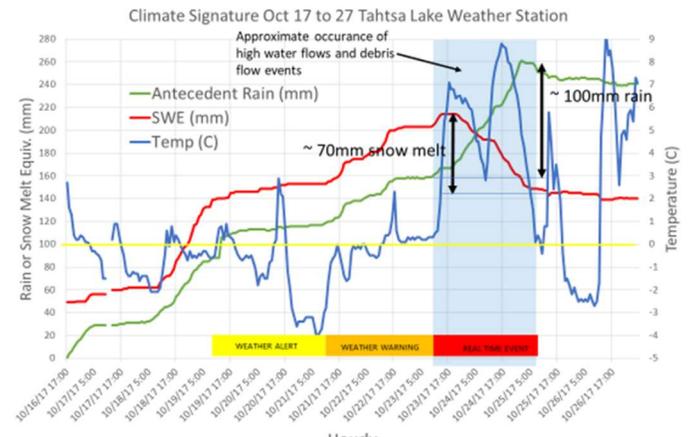


Figure 3. October 23 – 25, 2017 CTS Resulting in Significant Damage to nearby infrastructure.

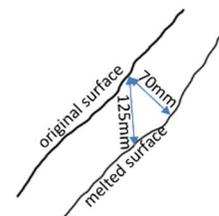


Figure 6. Vertical Equivalent of 70mm of Melt Thickness

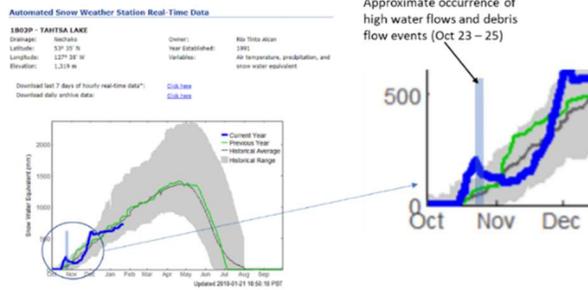


Figure 7. Tahtsa Lake Automated Snow Weather Station Real-Time Record October 2017 to January 2018 imposed on Historical Average and Range.

4.1.3 CTS October 27th to 29th, 2020

The CTS for the September 27-29, 2020 event recorded at the Tahtsa Lake ASWS is presented in Figure . Parameters from Tahtsa Lake ASWS, analyzed for this sequence event, include antecedent rain; accumulated snow measured in Snow Melt Equivalent (SWE); and temperature.

The chronology of preparatory causal factors that developed before the CTS, the trigger causal factors that occurred during the CTS, resulting adverse ground conditions; and the effects of those ground conditions are listed in Table 1. The CTS presented Figure resulted in debris flows which impacted vital infrastructure at the lower reaches of the debris path.

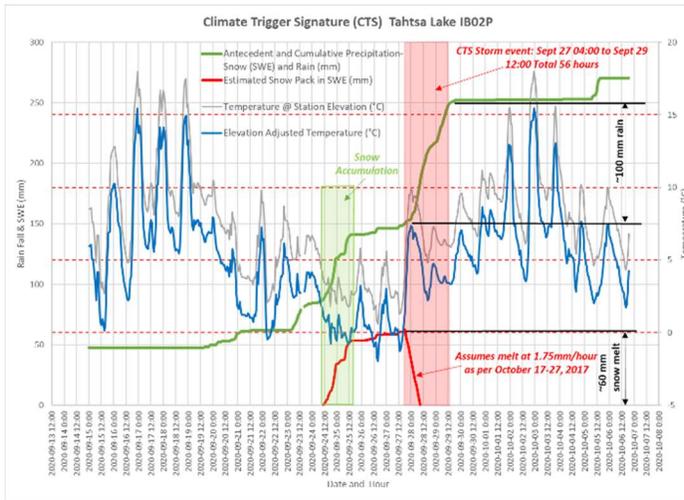


Figure 8. Sep 27 - 29, 2020 CTS from Tahtsa Lake IB02P ASWS (IB02P)

4.2 Fraser Valley Debris Flow near Lytton, November 2017

From November 21st to 23rd, 2017, the North slope of the Fraser River Valley above the Fraser River approximately 9 km downstream of Lytton, BC, was subjected to a significant CTS event involving an AR weather system (see Figure) that tracked directly eastward over the upper catchment basin, bringing warm temperatures and

even warmer intense rain. The combination of rain and rapid snow melt resulted in significant run-off that triggered an estimated 50,000m³ debris flow. Although the runoff was less than most of the other CTS's presented here, the similar magnitude and erosive nature of this debris flow is attributed to the water repelling and erosive characteristics of the soils left by the 2015 wildfire that affected most of the culprit sub-basin.

4.2.1 Blowdown Mid-Mountain RAWS

The Blowdown Mid-Mountain remote automated weather station (RAWS), utilized to assess the November 14th to 25th Debris Flow CTS, is ideally situated, both in proximity (55 km west of the debris flow location) and elevation (1890m ASL), to provide representative archival records and real-time indicators of the conditions experienced in the upper basins (1700 to 2300m ASL).

Additionally, as AR systems predominantly flow due eastward, similar temperature, precipitation, snowpack, and surface runoff can be inferred for the site and can be forecasted with an approximate one-hour lead time before the system reaches the upper basins.

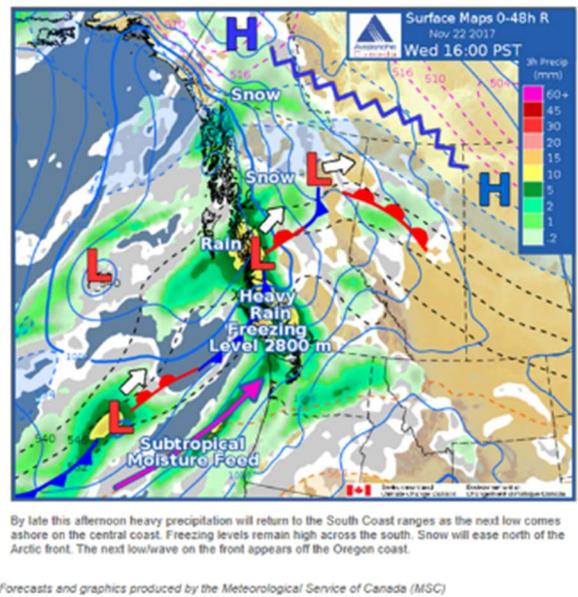


Figure 9. Mountain Weather Forecast for November 22, 2017 (Avalanche Canada, 2017).

Table 1. Sequence of Events Before and During the Sept 13 - Oct 7, 2020 CTS

Dates	Preparatory and Trigger Causal Factors	Resulting Ground Conditions	Effects
September 13th and 24th	Adverse ground conditions start to develop on ~September 20 when the contributing precipitation starts to fall as antecedent rain with the upper basin temperatures remaining ~4°C, by Sept 24 ~25 mm of antecedent rain had accumulated	Likely wetting the surface soils close to saturation.	Potential for high % runoff.
September 24th and 27th	During this period the precipitation at the upper basins is assumed to fall as snow, based on the temperature drop to around 0°C marking the beginning of the snowpack accumulation which continued until when the temperature spiked up to ~7°C.	Total accumulated snow ~60 mm SWE. Snow temperature fluctuates around 0°C.	Snow primed for flash melt.
September 27th to 29th	CTS starts: ~ 12:00 September 27th the temperature abruptly climbed from -2°C to 7°C in 12 hours; then dropped back to 3°C in the next 18 hours; fluctuated around 4°C for the next 18 hours when the rain stopped. The storm lasted a total of ~48 hours.	Ground saturated. Snow close to melting.	100 mm of warm rain (10 to 20°C) fell in the catchment area flash melting 60 mm SWE of snow resulting in near 100% runoff of ~160 mm of SWE water in 48 hours.
September 29 th to October 7 th	CTS ends Rain stops, average daily temperature fluctuates between 5 and 10°C.	Ground dries.	Runoff decreases to pre-storm levels.

4.2.2 CTS November 14th to 25th

The weather preceding the debris flow event is correlated to the historical climatic breakdown of the Blowdown Mid-Mountain RAWS in Figure . This indicates an above average early snowpack started to form October 16th and by November 20th had climbed to 1.25m which was 0.5m above average and near the maximum recorded for these dates. The abrupt change from snow to rainfall, related to the AR, occurred November 21st to 23rd.

The CTS covering November 14th to 25rd for the debris flow event is depicted in Figure . Parameters from the Blowdown Mid-Mountain RAWS analyzed for this event include antecedent rain; accumulated and depleted snow measured in Snow Water Equivalent (SWE); and temperature. During the event, the temperature, which was hovering around -5°C for three days prior, abruptly climbed to 5°C and after 40 hours dropped to around -3°C marking the end of the event. The intense warm rain started falling at the same time the temperature rose and accumulated at an average of 1.6mm/hour, amounting to 63mm by the end of the event. The most startling trend was the rapid drop in the snow cover, at a rate of 0.9 mm/hour of SWE amounting to a total SWE of 35mm. The snow melt started and stopped simultaneously with the intense rain fall and significant temperature fluctuations. This indicates the rapid snow melt was caused primarily by the induction of latent heat from the warm rain drops that fell during the AR event.

The 43mm of vertical SWE was corrected for slope angle as illustrated in Figure , for use in the runoff calculations.

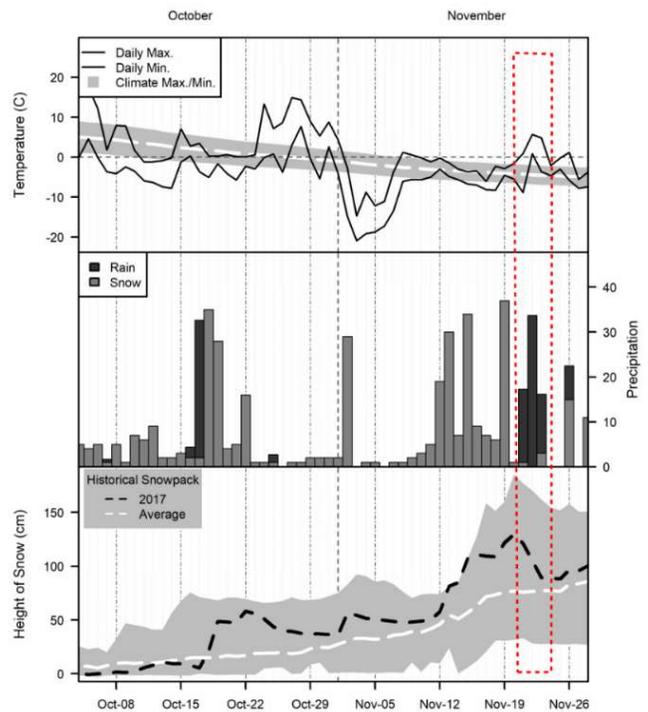


Figure 10. October - November Climate Data for the Blowdown Mid Mountain RAWS. The red dashed line highlights the weather event that contributed to triggering the debris flow.

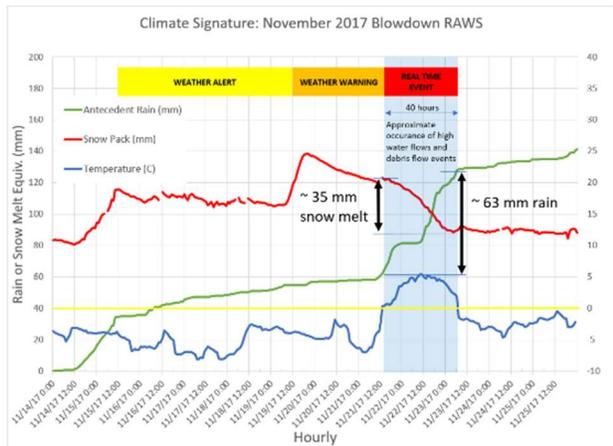


Figure 11. CTS Resulting in 50,000 m³ Debris Flow (Blowdown RAWS Upper Basin Weather Station)

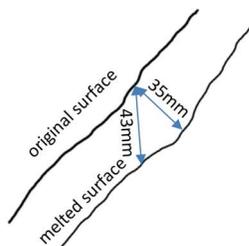


Figure 12. Vertical Equivalent of 35mm of Melt Thickness

4.2.3 Wildfire Effects on Basin Hydrology

Recent wildfire events are known to alter the vegetation and ground characteristics of the watershed. Effects of wildfire include consumption of the soil-binding organic matter and the formation of hydrophobic soils leading to the reduced infiltration and increased erosion potential. The reduced infiltration is caused by the volatilization of organic matter on top of and in the soil, which can bind sand grains together forming an impermeable layer that can persist for 3 to 5 years (Dobson Engineering, 2007). The increased erosion potential is due to the destruction of the binding vegetation root structures and the significant reduction of matrix suction formerly developed by the capillary suction in the tree roots that can result in as much as 10 m of negative head (matrix suction). These effects contribute to a substantial increase in the overland flow of water and movement of soils (Cannon & Gartner, 2005).

4.2.4 Debris Flows Ground Response

Figure provides a comparison of satellite and photogrammetry imagery of the Debris Flow Basin area between August 28, 2017 and May 15, 2018. The comparison shows considerable scour isolated to a narrow band along the pre-existing tributary channel from only the eastern sub-basin. The scour is evident along the entire channel reach from just upstream of vital infrastructure to near the top of the upper sub-basin

inferring that most entrained debris (~50,000m³) that arrived at lower levels of the basin, came from degradation from only the eastern sub-basin channel. It is unclear if the scour initiated at the blockage provided by FSR crossing and gullied up or initiated from the upper sub-basin and gullied down.

Evidently, the combination of high runoff flows; water-repellent soils and erosive soil conditions of the burnt forest ground; and concentration of flows into this anomalously steep tributary channel resulted in an estimated 50,000m³ debris flow/flood event.

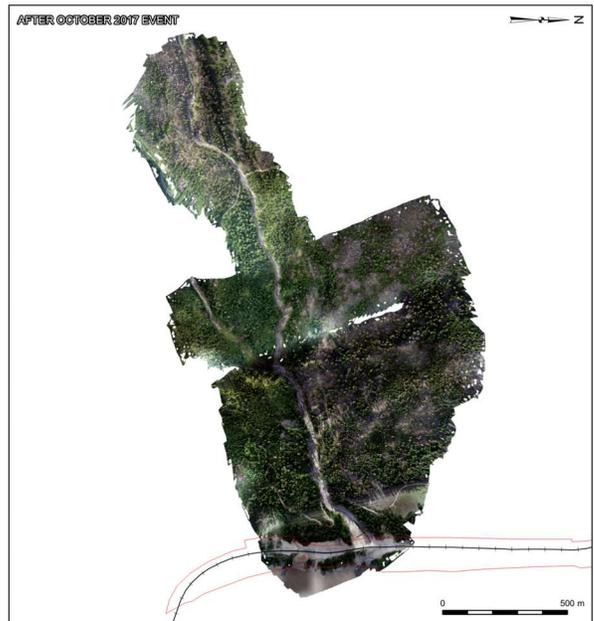
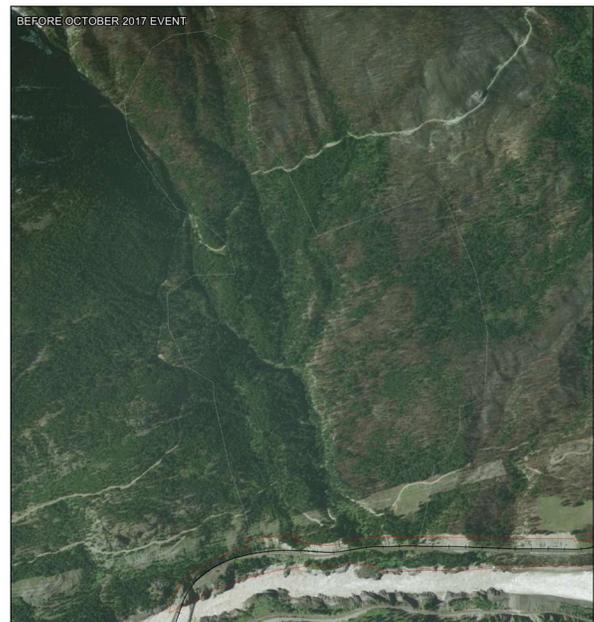


Figure 13. Comparison of Satellite and Photogrammetry Imagery of the Debris Flow Basin between August 28, 2017(upper) and May 15, 2018 (lower).

4.2.5 Basin Assessment

Figure displays an overlay of the area burnt during the 2015 Wildfire on the three basins that drain down to river level. Table 2 summarizes the key basin parameters of the four basins. The basin characteristics present is reviewed, and it is observed the gradient of the upper channel of the eastern most sub-basin is 67% which is 27% higher than the next steepest burnt basin channel.

It is evident that the higher run off percentage and enhanced erodibility of the burnt ground significantly increased the erosion potential of this relatively small sub-basin. Although an estimated 30% of this channel was observed to be scoured to bedrock, ~70% was not, implying a high potential for scour during a future similar CTS.

Also of note is that there is no evidence of previous erosive debris flows in any of the basins noted in Figure . The before and after imagery in Figure clearly shows fully mature forest occupied the stream valley prior to the event and was completely wiped out by the event.



Figure 14. Basins Effected by 2015 Wildfires. West sub-basing in blue, east sub-basin in green, central basin in orange, east basin in light blue, and wildfire area in yellow.

Table 2. Analysis of Basins Effected by 2015 Wildfires

Basin	Slope Area (Ha)	Burnt Area (%)	Upper Basin		Lower Basin	
			Slope Length (km)	Average Gradient (%)	Slope Length (km)	Average Gradient (%)
West sub-Basin	385	1	2.33	51	1.62	26
East sub-basin	161	63	1.52	67	1.52	30
Central Basin	1232	56	2.48	40	3.2	27
East Basin	1070	76	4.23	29	2.82	23

evidently, the combination of high runoff flows; water-repellent soils and erosive soil conditions of the burnt forest ground; and concentration of flows into this anomalously steep tributary channel resulted in an estimated 50,000m³ debris flow/flood event.

4.3 Runoff Event Effecting Fraser Canyon, January 30th to February 2nd, 2020

From January 30th to February 2nd, 2020, the Fraser Valley was impacted by a CTS involving an AR as it moved eastward from the Pacific Ocean. The CTS impacted multiple location in the Fraser Canyon between Lytton and Hope, BC. The Spuzzum Snow Weather Station located at 1,180 m of elevation was used to conduct the CTS analysis as it is analogues to the upper basins where multiple hazards were observed. The CTS began with 12 days of temperatures -3°C and 5°C which began the process of saturating the existing snow pack between January 19th and February 2nd. During the relatively warm weather, an initial approximately 200 mm of rain fell between January 23rd and January 25th, 2020. This resulted in about 200 mm or SWE melt and saturated the snowpack. The already saturated snowpack was then hit with a second burst of rain, approximately the same magnitude, but over a shorter period resulting in higher intensity event. The rain fall resulted in SWE run-off of 620 mm within a 52-hour period, which culminated in a rash of rock landslides, debris flows, earth landslides and hydraulic erosion hazard events along the Fraser Canyon between Lytton and Hope, BC, impacting infrastructure.

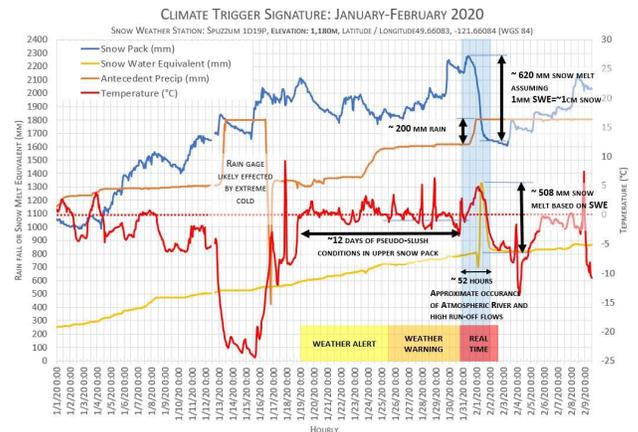


Figure 15. CTS resulting in multiple ground hazard events between January 29th, 2020 and February 2nd, 2020 (Spuzzum Snow Weather Station).

5 CONCLUSIONS

The five case study observations presented show the potential effectiveness for the implementation Climate Trigger Signatures (CTS) in managing risks associated with debris flow hazard scenarios. A summary of the findings for each case is presented in Table 3. The primary CTS and ground risk factor include:

- Unusually warm weather;
- High, saturated snow packs;
- Atmospheric Rivers or other high intensity, long duration precipitation events;
- Steep gradients in the upper basin;

- Available bedload;
- And previous wildfire activity.

Combining the CTS and ground risk factors listed above has the potential to improve risk management through forecasting with automated weather stations and early warning from NOAA AR Portal. Additional observations and analysis should be conducted to determine the effectiveness of implementing CTS monitoring within TARPS or other risk management strategies.

6 REFERENCES

Keegan, T., 2007. Methodology for Risk Analysis of Railway Ground Hazards. Ph. D. Thesis, Department of Civil and Environmental Engineering, University of Alberta, Edmonton.

NOAA, 2015. What are atmospheric rivers?, <https://www.noaa.gov/stories/what-are-atmospheric-rivers> (website accessed on September 28th, 2021)

Pinna Sustainability, 2014. The Future of Atmospheric Rivers and Actions to Reduce Impacts on British Columbia.

Radić, V., Cannon, A. J., Menounos, B., & Gi, N. (2015). Future changes in autumn atmospheric river events in British Columbia, Canada, as projected by CMIP5 global climate models. *Journal of Geophysical Research: Atmospheres*, 120(18), 9279-9302.

Effect of climate change on frost penetration depth in the subgrade soil beneath railway tracks: case study

Roustaei, M., Hendry, M. T.

Department of Civil and Environmental Engineering – University of Alberta, Edmonton, Alberta, Canada

Roghani, A.

Automotive and Surface Transportation Research Centre – National Research Council of Canada, Ottawa, Ontario, Canada

ABSTRACT

Frost heave along Canadian railways is a common issue that can result in track geometry issues and reductions in train speeds and rail traffic capacity. This study was conducted at a location prone to frost heave on VIA Rail Canada's Smith Falls subdivision to determine the maximum frost penetration depth at this site during the last 20 years and explore a correlation between frost depth and weather conditions for future anticipation of frost depth. The results include frost depth fluctuations over the last 20 winters using a thermal analysis conducted by TEMP/W, a regression between weather data and frost penetration depth, and predicted frost depth for the next 75 years based on a high greenhouse gas climate model. The results showed that frost depths estimated using conventional methods and numerical analysis were significantly different and the maximum frost depth tends to decrease in the future but it will be within the frost susceptible layers and could still be problematic at this site. So further monitoring or actions might be needed to prevent frost heave and consequent thaw softening at this location.

1 INTRODUCTION

Frost heave and thaw weakening are two states that result in negative effects of frost action in cold regions such as northern China, Russia, the USA, and Canada (Roghani and Hendry 2016, 2017; Roghani et al. 2015, 2017). Consequential frost heave has been observed in railway embankments along old or recently constructed railways in these areas, limiting train speeds in winter seasons. In theory, frost heave mostly occurs in finer materials or fouled ballast; however, monitored displacements indicate it can also occur in coarse fills considered to be non-frost susceptible materials (Li et al. 2016).

For development of ice lenses and frost heaving to occur, three factors must be present: freezing temperatures, a source of water, and a frost-susceptible soil. The most accepted definition of frost susceptibility is from the Highway Research Board Committee on Frost Heave and Frost Action in Soil (1955) and ISSMFE (1989), wherein frost-susceptible soils are those "*in which significant ice segregation will occur when the requisite moisture and freezing conditions are present*" (Nurmikolu 2010). The frost susceptibility of a material is highly dependent on the content and quality of the fines fraction (Nurmikolu and Kolisoja 2008; Nurmikolu 2010; Nurmikolu and Silvast 2013; Hendry et al. 2016) while the freezing temperature and source of water are site characteristics. Frost depth (or the depth to which soils may freeze and ice lenses develop) can significantly influence the destructive effects of frost heave and thaw softening, and the intensity of these two phenomena depends to a considerable extent on temperature and

precipitation conditions. Therefore, it is crucial to develop a greater understanding of the frost depth and the meaningful effect of any atmospheric weather characteristics.

The present study is part of an extensive project at a section of track in Ontario, Canada that suffers from frost issues. This project aims to monitor the development of frost heave using instrumentation including ShapeArrays, strain gauges, borehole extensometers, thermistors, and piezometers (Roghani et al. 2019). In the first phase of the project, the frost susceptibility of the subgrade soil was investigated using samples from different layers (Roustaei et al. 2019). The experimental results showed medium to very high frost susceptibility classification for the soil layers at this site (Table 1). In addition, the water table was within 3 m (10 ft) of the track surface and thus provides the water source necessary for frost heave (FHWA NHI-05-037 2006). Moreover, the development of a numerical model based on *in situ* data collected during this phase resulted in illustration of the effect of various weather conditions, such as snow cover, on the frost penetration depth at this site. Having the results of the first phase in mind and considering the importance of frost penetration depth as the third factor necessary for the development of frost heave, this second phase of the study aimed to determine the maximum frost penetration depth at this site for the last 20 years and explore a correlation between frost depth and weather conditions for future anticipation of frost depth in this location.

1.1 Frost Depth Calculation

For several decades, frost depth determination and prediction have attracted researchers from academia and industry. Most frost heaving in some parts of Canada seems to occur in the latter half of the winter season, probably when the frost line has reached its maximum penetration (Armstrong and Csathy 1963). As such, predicting the depth to which soils may freeze and thaw can be helpful for guiding engineering designs. Several formulas as well as analytical and semiempirical models—such as Stefan, modified Berggren, and Chisholm and Phang—have been developed to predict the depth of frost penetration. The Stefan equation is derived from the fundamental equations of heat flow and storage (Eqs. 1 and 2; Yoder and Witczak 1975):

$$D = \sqrt{\frac{48kF}{L}}, \quad [1]$$

where D is the depth of frost penetration (ft), k is the thermal conductivity (Btu/ft °F h), F is the freezing index (°F d), and L is volumetric heat of latent fusion (Btu/ft³); and

$$L = 10434 \omega \gamma, \quad [2]$$

where ω is the moisture content and γ is the dry density (lb/ft³). Among meteorological factors such as air temperature, sunshine, precipitation, and wind velocity, air temperature is probably the most significant. The use of 'degree-days of freezing' as a guide for calculating frost depth for a given area illustrates the strong influence of air temperature on soil temperature (Penner 1962).

In addition, the freezing index (also called 'coldness sun' by the U.S. Army Corps of Engineers) is the number of Fahrenheit degree-days above and below 32 °F between the highest and lowest points on the cumulative degree-days time curve for one freezing season. Local maps also exist from which the freezing index or even the frost penetration depth can be directly extracted for different regions of Canada. For example, Figure 1 is a map of southern Ontario published by the Ontario Ministry of Transportation as one of the Ontario Provincial Standards for Roads & Public Works. Although using this map provides a quick estimate of frost depth, which at this site is around 1.6 m, actual field data provide the most accurate information.

Numerical or analytical modeling techniques can also be used to estimate frost depth; however, the required input data are typically unavailable or expensive to collect (Luo 2014). Farrington et al. (2001) presented a frost penetration prediction model using numerical simulation, statistical regression, spatial interpolation, and GIS. Using their methods, they concluded seasonal maximum frost penetration depth can be reliably estimated by its relationship to the actual annual freezing

degree index (AFDI), as long as a pavement-specific relationship is derived using meteorological data that account for region-specific weather dynamics. A regression of maximum seasonal frost penetration depth (derived from dynamic simulations of temperature and moisture flux in a pavement structure using actual climatic data) on AFDI showed a strong positive correlation and was useful for fitting a linear equation to the median and 90% upper prediction limit of maximum frost penetration depth.

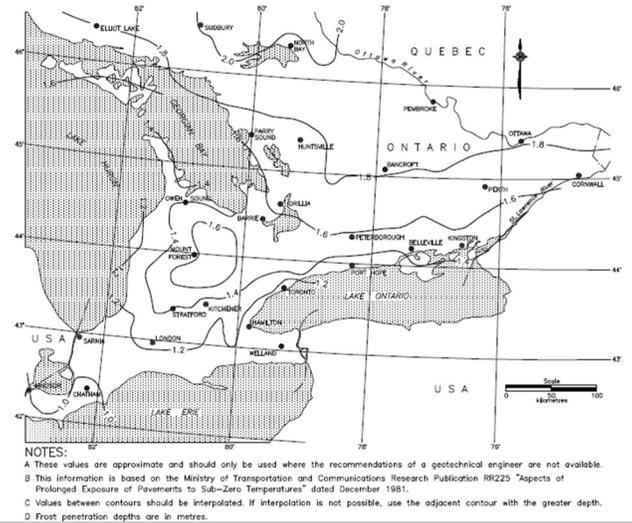


Figure 1. Frost depth map for southern Ontario (Ontario Ministry of Transportation, 2010)

Soliman et al. (2008) introduced a simplified model to predict the frost penetration in Manitoba by developing a relationship with an acceptable accuracy between the depth of frost penetration and the freezing index by minimizing the effect of the other variables. Lee et al. (2013) provided a frost indicator with a methylene blue solution method to measure the frost depth (Luo 2014).

Because the actual frost depth is affected by material type, soil thermal properties, and soil water content as well as climatic conditions such as temperature, wind speed, precipitation, and solar radiation, this study concentrated on finding a simple correlation between air temperature, total snowfall, and frost penetration depth using a numerical model based on the soil properties and climate conditions at this specific site.

1.2 Climate Models

Air temperature and total snowfall are two of the most important requirements for frost depth prediction. Nowadays, various climate models forecast future weather conditions by considering climate change. The Earth's climate, past and future, is not static; it changes in response to both natural and anthropogenic drivers. Human emissions of carbon dioxide (CO₂), methane (CH₄), and other greenhouse gases (GHGs) now overwhelm the influence of natural drivers on the external forcing of the Earth's climate. Over the past 15

to 20 years, the growth rate of atmospheric carbon emissions from human activities has increased from 1.5 to 2 parts per million (ppm) per year due to increasing carbon emissions from human activities, in large part due to growing contributions from developing economies. Beyond the next few decades, the magnitude of future climate change will primarily be a function of future carbon emissions and the response of the climate system to those emissions. Climate projections are typically presented for a range of plausible pathways, scenarios, or targets that capture the relationships between human choices, emissions, concentrations, and temperature change (Hayhoe 2017).

The Representative Concentration Pathways (RCPs) form a set of GHG concentration and emissions pathways designed to support research on the impacts and potential policy responses to climate change (Moss et al. 2010; Van Vuuren et al. 2011). RCP 8.5 corresponds to a high GHG emissions pathway compared to the scenario literature (Fisher et al. 2007; IPCC 2008; Riahi et al. 2011). In this study, RCP 8.5 was used to predict frost depth prediction based on weather conditions for the next 75 years.

2 METHODS AND MATERIALS

2.1 Site Details

The study site is located at a section of VIA Rail's track in eastern Ontario. This subdivision is only used for passenger trains, and consists of 57 kg/m (115 lb/yard) continuously welded rail on ballasted track and wooden ties. It is classified as a class 5 track, which is Transport Canada's highest classification (Transport Canada's Rules

Respecting Track Safety, 2011), and carries 160 km/h trains. The ballast layer and the gravel subballast, both 0.3-m thick, are underlain by old ballast material. Figure 2 shows the different layers of the subgrade with brief descriptions and Table 1 illustrates the frost susceptibility degrees extracted from experimental investigations (Roustaei et al. 2019). Figure 2 also shows the location of instrument installations at this site, including 20 thermistors located below the track surface in different layers and used in this study to monitor changes in temperature of the soil. These thermistors (model 3810, Geokon) are routinely used in permafrost applications to monitor underground soil and water temperatures. These sensors work in a temperature range from -20 to 80 °C with a resolution of 0.1 °C and accuracy of ±0.2 °C. The water table at this site is also within 3 m (10 ft) of the track surface and thus provides the water source necessary for frost heave. So, a frost-susceptible soil and source of water as two main factors for the development of frost heave are present at this site, and the frost depth as the third factor should be further investigated.

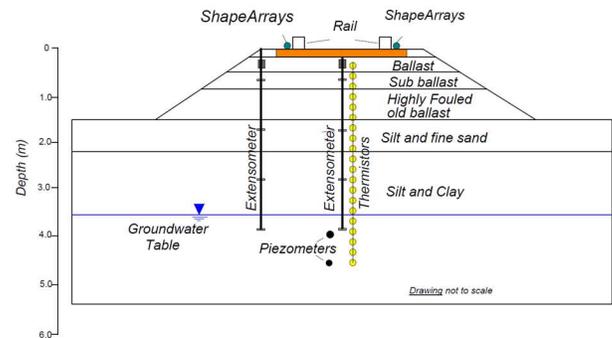


Figure 2. Schematic view of railway embankment and installed instrumentation

A numerical model was first developed to represent the thermal regime of the embankment and the underlying ground at the site using as inputs the data recorded for ambient temperature and from the deepest installed thermistor; it was then calibrated with the data recorded from the other thermistors for the previous winter. To determine the maximum frost penetration depth at this site and explore the correlation between frost depth and weather conditions for future anticipation of frost depth, multiple simulations were conducted based on weather data for the last 20 years (1999-2019) from the Kemptville and Appleton stations (weather stations closest to the site).

Table 1. Degree of frost susceptibility and types of soil samples

Depth (m)	Description	Degree of frost susceptibility
0-0.3	Crushed rock ballast-fouled	-
0.3-0.6	Gravel-grey, uniformly graded, trace of ice	-
0.6-1.2	Highly fouled ballast-the track appears to have been lifted over time and this layer is the ballast from before	Medium
1.2-1.5	Silt-trace fine sand, mostly silt, low plasticity, wet, dark grey	Very High
1.5-1.8	Fine sand and silt-medium plasticity, grey to dark, moist	Very High
1.8-4.2	Clay-trace silt, mostly clay, medium plasticity, moist, dark greenish grey	Very High

2.2 Numerical Model

Modeling was conducted using TEMP/W, a finite element software product for modeling thermal changes in the ground due to environmental changes or due to the construction of facilities, such as buildings or pipelines. TEMP/W requires that functions be specified for thermal conductivity and unfrozen water contents. These functions are dependent on temperature, which means they are only valid for a constant water content soil state, i.e., the water may turn to ice but the overall

amount of water/ice is fixed throughout the analysis. All thermal characteristics of the soil layers shown in Figure 3 were estimated with respect to the type of soil (Table 2) and the TEMP/W manual in phase one of the study, with the model calibrated with temperature data recorded by the thermistors (Roustaei et al. 2019).

The water content of the subgrade layer samples was not available, so the volumetric water content parameter was determined by calibrating the model with the temperature data recorded by the thermistors. Other thermal properties were estimated with respect to the type of soil and the TEMP/W manual. A coupled soil-atmosphere process was used to investigate the effects of climate conditions in terms of snow on the temperature regime of soil layers. Surface Energy Balance (SEB) boundary conditions were considered one of the best solutions for modeling this process in the transient analysis (GEOSLOPE 2010).

In TEMP/W, the user is required to enter the air temperature, wind speed, snow depth, and albedo, all versus time, in addition to the latitude of the site for estimation of solar radiation. Of particular note here is the albedo, which will be influenced by snow cover characteristics, solar zenith angle, and cloud conditions. Snow surface albedo ranges from less than 0.60 for wet and melting snow to greater than 0.85 for fresh snow, can be greater than 0.90 under cloudy sky conditions, and is approximately 0.23 for green vegetation (Wendler and Kelley 1988; Zhang et al. 1996). A high snow surface albedo leads to a reduction in the absorbed solar energy and lowering of the temperature of the snow surface (Zhang 2005).

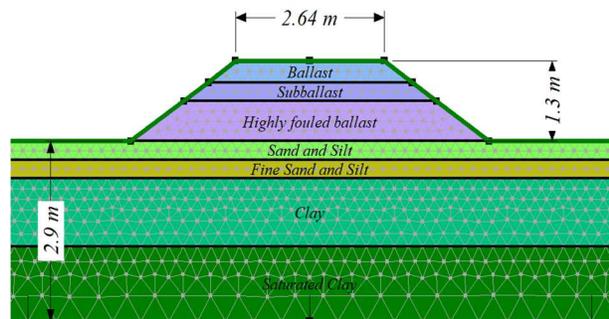


Figure 3. Geometry of the TEMP/W model

3 RESULTS AND DISCUSSION

3.1 Frost Penetration Depth

Frost penetration depth as the depth to which the groundwater in soil and consequently the soil medium are expected to freeze, is one of the main factors for the development of frost heave. It can be affected by different parameters, including ambient temperature, thermal conductivity of soil layers, snow or asphalt cover, and any nearby heat source. The exact depth of the frost line was determined from different thermal conditions.

Table 2. Properties of materials used in the TEMP/W model

Soil layer	Thermal conductivity (kJ/s/m ² /°C)		Volumetric heat capacity (kJ/m ³ /°C)	<i>In situ</i> volumetric water content
	Frozen	Unfrozen		
Ballast	0.00056	0.00055	1800	0.18
Subballast	0.00085	0.0008	1800	0.18
Highly fouled ballast	0.0014	0.0012	1800	0.18
Sand and silt	0.0021	0.0017	2000	0.22
Fine sand	0.0024	0.0019	2000	0.25
Clay	0.0015	0.0013	2000	0.3
Saturated clay	0.00262	0.0013	2000	0.6

The TEMP/W model was calibrated with *in situ* recorded temperature data (2018-2019) in the first phase of this project. The results showed the ground thermal regime was primarily controlled by conductive heat transfer through the snowpack; increased conductivity due to compaction of the snow cover lessened this effect. Thicker snow cover maintained higher mean annual ground temperatures and shallower frost penetration (Roustaei et al. 2019). For a more comprehensive investigation of the effect of different amounts of snowpack as well as ambient temperature on the frost penetration depth, weather data for the last 20 years were used in the current part of the study. Figure 4 shows the ambient temperatures and snow cover depths during these winters, which were used as two of the main important inputs of the model, as well as the outputs in terms of frost penetration depths.

The frost line gradually climbs to its maximum depth starting from the first days of winter through to April, after which it remained deep even when the weather warmed up at the end of April. Although the ambient temperature went up and the weather became warmer, a limited area in the ground remained frozen in early spring and needed more time to thaw (Roustaei et al. 2019). Considering the characteristics of soil layers at this site as well as the ground profile, the frozen area developed in the very fine silty sand layer that has a very high degree of frost susceptibility and thus could result in the development of frost heave.

Additionally, thawing of the upper layers could create free water that is unable to drain through the impermeable frozen zone. This can flow to the freezing front of this zone and feed the growth of ice lenses. This phenomenon corresponds to a saturated substructure due to thaw softening (Li et al. 2016) and can cause extreme strength reduction in thawed soil layers due to the loss of effective stress. The two widest frost zones for the 20 different winters considered are shown in Figure 5, where the temperature of the blue area is below zero while the ambient temperatures are 18.3 and 10.3 °C on April 15 of 2003 and 2015, respectively. These frost zones last until mid-May and do not let moisture

from the upper layers fully drain, which results in a thaw softening phenomenon.

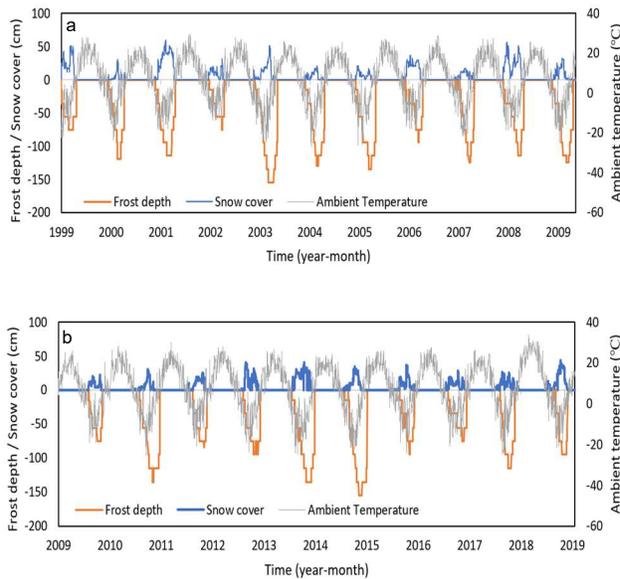


Figure 4. Ambient temperatures vs. snow cover and extracted frost depths from simulation for a) 1999-2009; b) 2010-2019

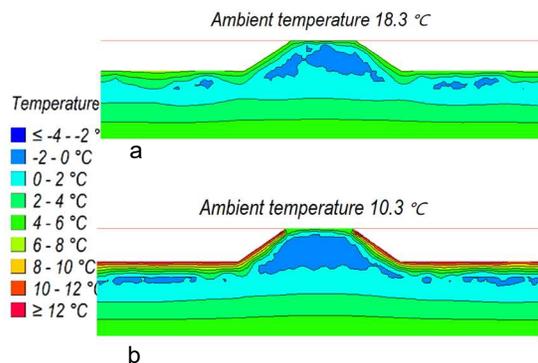


Figure 5. The two widest frost zones determined for the last 20 years: a) 2003-04-15; b) 2015-04-15

3.2 Linear Regression

A linear regression is usually used to find linear relationships between a target variable and one or more predictors. Technically, a regression analysis model is based on the sum of squares, which is a mathematical method to find the dispersion of data points. The goal of a linear regression model is to achieve the smallest possible sum of squares and draw a line that comes closest to the data. To build a regression model for predicting frost depth, frost depth becomes the response variable and two approaches can be taken. The first is to build a model to predict soil temperature at each depth, and then use linear interpolation to calculate frost depth from the predicted soil temperature values at each depth.

The second, which is used in this study, first calculates soil frost depth values as the response variable and then builds the regression model directly to predict frost depth (Luo 2014).

In the first step, the calculated maximum frost depths from the numerical analysis during the previous 20 winters (1999-2019) were regressed against annual freezing degree days (FDDs) and the total annual snowfall using least squares multiple linear regression analysis. Results from this first model are summarized in Table 3 (centre column 'first' model), including an R-squared value indicating an acceptable general fit of our model with real observations and low P values showing meaningful additions of predictors to our model. In addition, the residual plots of our independent variables—annual snowfall and FDD—are randomly dispersed around the horizontal axis and demonstrate the linear regression model is appropriate for the available data (Figure 6a and 6b).

Figure 7 plots the estimated frost penetration depths from the adopted linear regression model for the last 20 years as well as those extracted from the numerical model, with the results showing good consistency between these two types of data. Due to these promising results, the regression equation was then used to predict frost penetration depth, specific to this site and its soil characteristics, using the RCP 8.5 climate model for future years.

Table 3. Summary of multiple linear regression outputs of the two regression models

Regression model	First	Second
Multiple R	0.79	0.72
R-Squared	0.62	0.52
Adjusted R-Squared	0.58	0.49
Standard Error	15.27	16.55
P-value (Variable 1) Snowfall	0.0304	-
P-value (Variable 2) FDD	6.81E-05	0.000306

Notably, many climatic models are based on temperature data and total precipitation while the amount of total annual snowfall is not considered. Although using snowfall data will result in more reliable frost penetration predictions with the first regression model, a second regression model using only the FDD as an independent variable could also be sufficient for frost depth prediction at this site in future years (Table 3, right column 'second' model; Figure 6c).

The freezing degree days extracted from the climate model adopted for this study, RCP 8.5, were then used for frost depth prediction based on weather conditions for the next 75 years (Climate Atlas of Canada 2019). The frost depth penetration threshold shown in Figure 8 illustrates that climate change, reflected in warmer winters and fewer FDDs, will result in shallower frost penetration depths at this site in the future. The maximum frost depth for the next 20 years will be around 100 cm, which is located in the highly fouled old ballast

that contains 18% fine materials and has a medium degree of frost susceptibility at present (which could increase in the future due to further fouling). After 2040, the maximum frost depth tends to decrease and reaches 60-70 cm, which is within the border of the subballast and highly fouled ballast layer.

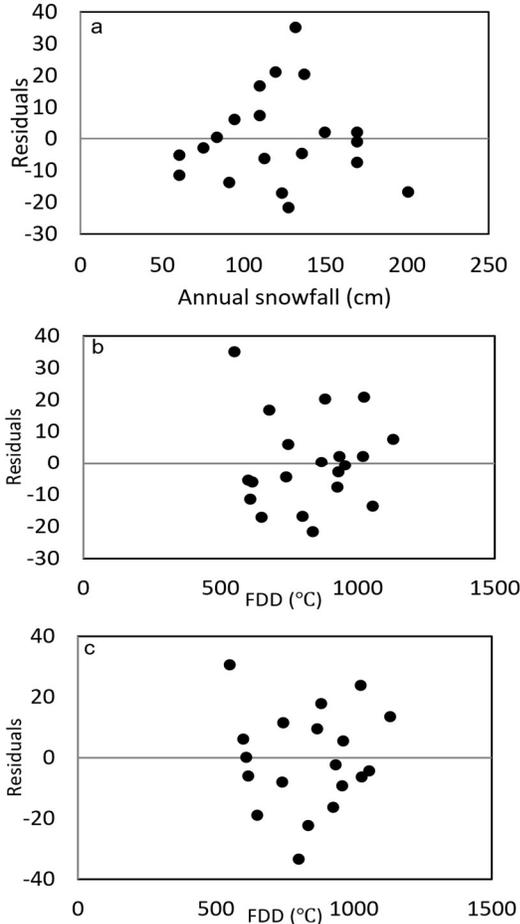


Figure 6. Residual plots of independent variables: a) annual snowfall in the first model; b) FDD in the first model; c) FDD in the second model.

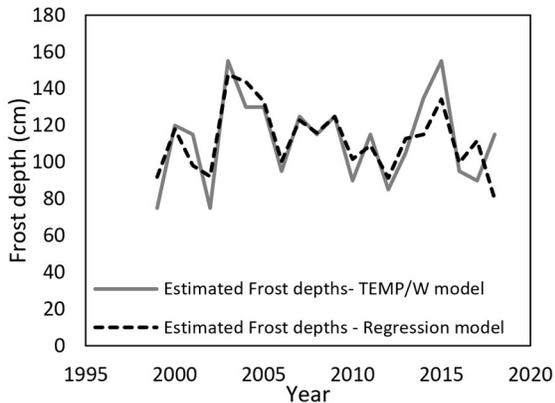


Figure 7. Estimated frost depth for the last 20 years

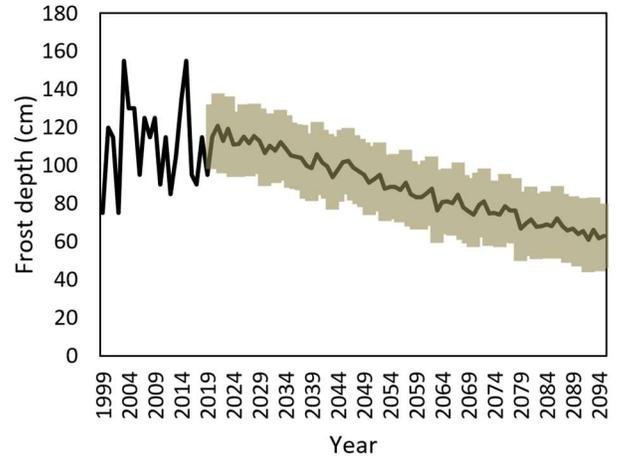


Figure 8. Predicted frost depth based on FDD data from the RCP 8.5 climate model for the next 75 years

4 CONCLUSION

Frost heaving of soils was originally thought to result from cold ambient weather, but present concepts indicate other factors influence the depth of freezing, such as the insulating effect of snow.

This study was undertaken to investigate the correlation of frost penetration depth with ambient temperature and snow cover and thus predict frost penetration depth in the future for a site located on VIA Rail Canada's Smith Falls subdivision.

The high frost susceptibility of soil layers at this site was demonstrated by experimental tests and a thermal model using TEMP/W previously calibrated using data from *in situ* installed thermistors. This model was used to analyze weather data for the last 20 years to determine the maximum frost penetration depth. Conventional maps show the frost penetration depth of this area at around 1.6 m, but the numerical model results for the last 20 years as well as *in situ* data from 2019 show the maximum frost depth at this location varies from 0.75 to 1.55 m and never exceeds this limit.

However, this depth still reaches frost susceptible layers of the subgrade and can result in frost heave. Moreover, a widespread frost zone developed in the embankment under the rail tracks and could last until May, and thus may be the main cause of thaw softening at this location. Two regression analysis models were also developed in this study based on the minimum sum of squares and maximum frost depth data for the last 20 years extracted from the numerical model. These models could be used for future forecasting of frost depth at this site.

Based on one of these regression models and the RCP 8.5 climate model, the maximum frost depth for the next 20 years was predicted to be around 100 cm, which is located in the highly fouled ballast that has a medium degree of frost susceptibility; however, this could change in the future with further fouling. After 2040, the maximum frost depth tends to decrease and reaches 60-70 cm, within the border of the subballast and highly

fouled ballast layer. Therefore, frost heave could still be problematic at this site and further monitoring, or actions might be needed to prevent frost heave and consequent thaw softening at this location.

5 ACKNOWLEDGEMENT

The authors acknowledge VIA Rail Canada for providing access to its track, RailTerm for facilitating the field work, and Dr. Jon Makar and Marianne Armstrong from National Research Council Canada for their support. Funding was provided by Infrastructure Canada, the Natural Sciences and Engineering Research Council of Canada (NSERC-IRC 523369-18), Canadian National Railway, the National Research Council of Canada, and Transport Canada.

6 REFERENCES

Armstrong, M.D. and Csathy, T.I. (1963). Frost Design Practice in Canada. In Highway Research Record 33, HRB, National Research Council, Washington, D.C.

Climate Atlas of Canada. (2019). Ottawa region, https://climateatlas.ca/data/grid/299/fdd_2030_85/line

Farrington, S.P., Gildea, M.L., Dougherty, D. and Rizzo, D. (2001). Frost Penetration Prediction and Mapping. Final Report of Contract # 984024. Agency of Transportation, State of Vermont.

FHWA NHI-05-037. (2006). *Geotechnical Aspects of Pavements*. U.S. Department of Transportation, Federal Highway Administration.

Fisher, B., Nakicenovic, N., Alfsen, K., Corfee Morlot, J., de la Chesnaye, F., Hourcade, J.-C., Jiang, K., Kainuma, M., La Rovere, E., Matysek, A., Rana, A., Riahi, K., Richels, R., Rose, S., van Vuuren, D.P. and Warren, R. (2007). Chapter 3: Issues related to mitigation in the long-term context. In: Climate change 2007. Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, USA.

Foundation, Frost Penetration Depths for Southern Ontario, Volume 3, (2010), Ontario Ministry of Transportation.

GEOSLOPE. 2010. *Thermal Modeling with TEMP/W 2007*, An Engineering Methodology, Fourth Edition, GEOSLOPE International Ltd, Calgary, AB, Canada.

Hayhoe, K., Edmonds, J., Kopp, R.E., LeGrande, A.N., Sanderson, B.M., Wehner, M.F. and Wuebbles, D.J. (2017). Climate Models, Scenarios, and Projections. In: *Climate Science Special Report: Fourth National Climate Assessment*, Volume I.

Hendry, M.T., Onwude, L.U. and Segó, D.C. (2016). A Laboratory Investigation of the Frost Heave Susceptibility of Fine-Grained Soil Generated from the Abrasion of a Diorite Aggregate, *Cold Regions Science and Technology*, 123: 91-98.

Highway Research Board Committee on Frost Heave and Frost Action in Soil. (1995). Highway Research Board Bulletin. No. 111: 107-110.

IPCC (Intergovernmental Panel on Climate Change). (2008). Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies. IPCC Expert Meeting Report on New Scenarios, Noordwijkerhout, Intergovernmental Panel on Climate Change.

ISSMFE (International Society of Soil Mechanics and Foundation Engineering) Technical Committee on Frost, TC-8. (1989). Work Report 1985-1989. *VTT Symposium 94, Frost in Geotechnical Engineering*, Saariselka, Finland, 1: 15-70.

Lee, J., Kim, H.S. and Kim, Y.S. (2013). Estimation of frost depth in south Korea, American Society of Civil Engineers, *Poromechanics V*.

Li, D., Hyslip, J., Sussmann, T. and Chrismer, S. (2016). *Railway Geotechnics*. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Luo, M. (2014). Frost Depth Prediction, Master's Thesis, Graduate Faculty of the North Dakota State University of Agriculture and Applied Sciences, Fargo, ND, USA.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, G.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P. and Wilbanks, T.J. (2010). The Next Generation of Scenarios for Climate Change Research and Assessment. *Nature*, 463: 747-756.

Nurmikolu, A. (2010). Fouling and Frost Susceptibility of Railway Ballast and Subballast, Field and Laboratory Study, VDM VERLAG DR. MÜLLER, Germany; ISBN 978-3-639-23623-1.

Nurmikolu, A. and Kolisoja, P. (2008). The Effect of Fines Content and Quality on Frost Heave Susceptibility of Crushed Rock Aggregates Used in Railway Track Structure. *Proceedings of the 9th International Conf. on Permafrost*. Fairbanks, AK, USA, 2: 1299-1305.

Nurmikolu, A. and Silvast, M. (2013). Causes, Effects and Control of Seasonal Frost Action in Railways. *Sciences in Cold and Arid Regions*, 5(4): 0363-0367.

Penner, E. (1962). Ground Freezing and Frost Heaving, NRC Publications Archive, *Canadian Building Digest*, 02.

- Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N. and Rafaj, P. (2011). RCP 8.5—A Scenario of Comparatively High Greenhouse Gas Emissions. *Climatic Change*, 109: 33-57.
- Roghani, A. and Hendry, M.T. (2016). Continuous Vertical Track Deflection Measurements to Map Subgrade Condition Along a Railway Line: Methodology and Case Studies. *Journal of Transportation Engineering*, 142(12).
- Roghani, A. and Hendry, M.T. (2017). Quantifying the Impact of Subgrade Stiffness on Track Quality and the Development of Geometry Defects. *Journal of Transportation Engineering, Part A: Systems*, 143(7).
- Roghani, A., Hiedra Cobo, J. and Charbachi, P. (2019). Studying the Impact of Freeze Thaw Cycles on Performance of Railway Tracks. *International Heavy Haul Association Conference, IHHA 2019*, Narvik, Norway.
- Roghani, A, Macciotta, R. and Hendry, M.T. (2015). Combining Track Quality and Performance Measures to Assess Track Maintenance Requirements. *2015 Joint Rail Conference*, American Society of Mechanical Engineers, V001T01A009-V001T01A009.
- Roghani, A., Macciotta, R. and Hendry, M.T. (2017). Quantifying the Effectiveness of Methods Used to Improve Railway Track Performance Over Soft Subgrades: Methodology and Case Study. *Journal of Transportation Engineering, Part A: Systems*, 143(9).
- Roustaei, M., Hendry, M.T. and Roghani, A. (2019). Frost Susceptibility of Subgrade Soil Beneath Railway Tracks: Case Study. *72nd Canadian Geotechnical Conference*, St. John's, NF, Canada.
- Soliman, H., Kass, S. and Fleury, N., (2008). A Simplified Model to Predict Frost Penetration for Manitoba Soils. *2008 Annual Conference of the Transportation Association of Canada*, Toronto, ON, Canada.
- Transport Canada (2011). "Rules respecting track safety". <https://www.tc.gc.ca/media/documents/railsafety/track-safety-2012en.pdf>.
- Van Vuuren, D.P., Stehfest, E., den Elzen, M.G.J., Kram, T., van Vliet, J., Deetman, S., Isaac, M., Goldewijk, K.K., Hof, A., Mendoza Beltran, A., Oostenrijk, R. and van Ruijven B. (2011). RCP2.6: Exploring the Possibility to Keep Global Mean Temperature Increase Below 2 °C. *Climatic Change*, 109: 95-116.
- Wendler, G. and Kelley, J. (1988). On the Albedo of Snow in Antarctica: A Contribution to I.A.G.O. *Journal of Glaciology*, 34: 19-25.
- Yoder, E.J. and Witczak, M.W. (1975). *Principles of Pavement Design*, Second Edition, John Wiley & Sons, Inc.
- Zhang, T. (2005). Influence of the Seasonal Snow Cover on the Ground Thermal Regime. *Reviews of Geophysics*, 43: RG4002.
- Zhang, T., Stamnes, K. and Bowling, S.A. (1996). Impact of Clouds on Surface Radiative Fluxes and Snowmelt in the Arctic and Subarctic. *Journal of Climate*, 9: 2110–2123.

The case for zero emission propulsion for the Okanagan valley electric regional passenger rail (OVER PR)

Tye Boray, Tripti Goyal, Harshit Sharma, Mohamed A. Hegazi, Gord Lovegrove
UBC School of Engineering, Kelowna, Canada

ABSTRACT

North American eco-tourism and growth have led to an increase in travel demand between cities, with a commensurate increase in traffic congestion, greenhouse gas (GHG) emissions and crash related injuries. These negative consequences of growth have been acutely felt in the Okanagan Valley, where smaller rural and semi-rural communities are seeing intense growth in eco-tourism-related traffic, essentially doubling each summer, with no inter-regional passenger transit systems that connect to or between the Valley's communities. Therefore, almost all travel into and through the region must be done using the already congested BC provincial Highway 97. Widening the highway has been a very slow and prohibitively expensive process due to mountainous terrain and local opposition. Recent inter-city tram-train innovations in Europe using zero-emission electric Hydrail (Hydrogen fuel cell/battery hybrid) technology have demonstrated an environment-friendly solution that could work in the lower-density, rural Okanagan Valley, similar to Alstom's iLint Coradia regional train that's been operating since 2018 in Germany's Saxony-Holstein Province. This paper presents the results of research on an Okanagan Valley Electric Regional Passenger Rail (OVER PR), including:

1. The geographic and policy context for considering hydrail;
 2. The state and feasibility of hydrail technology, both freight and passenger rail, in Canada;
 3. The economic feasibility for a regional hydrail passenger in the Okanagan Valley, using BC MoTI templates.
- Results suggest OVER PR would cost less than ¼ the \$10 Billion cost of widening highways, and provide significant economic benefits, including: reduced congestion, GHG emissions, and 28 fewer road deaths per year. Over a 30-year project life, a net present value (NPV) is estimated at 6% discount rate of \$20 billion, with Benefit to Cost Ratio (BCR) of 9:1 and 2 year payback period. For reference, the BC MoTI estimates a BCR of only 5:1 to widen Highway 97. The return of passenger rail to the Okanagan Valley holds the key to enabling eco-tourism, one of the fastest-growing tourism sector segments globally.

1 INTRODUCTION

Growing North American (NA) communities and eco-tourism have led to an increase in travel demand across the continent, with a commensurate increase in traffic congestion, greenhouse gas (GHG) emissions and crash-related injuries (RAC, 2020). The UN IPCC has called for drastic action to address climate change, including alternatives to driving, such as regional passenger rail. Rail is the backbone of Canada's economy, and moves over 360 million tons of freight and over 100 million passengers annually, transporting more than \$C 312 billion worth of goods in support of the Canadian economy. While it has a relatively low carbon footprint, even rail significantly impacts the environment, as over 2.26 billion liters of fuel were burned by rail in Canada, including carcinogen particulates PM2.5 and PM10 and GHG gases (Environment and Climate Change Canada, 2019). This is significant in absolute terms, yet low in relative terms, as railways produce only

1% of Canada's total GHG emissions, and 3.5% of Canada's transportation GHG emissions. Clearly, rail is one of the country's greenest transportation modes. Yet to attract more travel demand there is a need to build more passenger rail capacity and infrastructure. As inter-city, regional, and cross-continent eco-tourism passenger rail traffic increases, so will its carbon footprint unless zero-emission (ZE) rail power is used, which presents the opportunity to accelerate Canada's transition to a hydrogen economy using Hydrail (Hydrogen fuel cell/battery hybrid) technology. Hydrail trains utilize fuel cell technology that combines hydrogen and oxygen in a membrane-controlled manner that generates electricity for battery (re-)charging and/or rail vehicle propulsion.

Progress on incorporating hydrail into NA freight locomotives has been slow, likely due to more diverse and demanding duty cycles, lower new vehicle take-up, private ownership, lack of regulatory pressure on emissions, and a risk averse industry. Diesel-electric

locomotives are proven, durable workhorses, economic to run well over 30 years with regular maintenance. Yet, the US EPA and CALTRANS are ramping up Tier 5+ regulations, which will drive hydrail retrofits of these old, reliable chassis to meet zero emission standards before 2050; Canada will soon follow. As a result, research in freight hydrail retrofits is growing, with Southern Rail of BC (SRY) expected to launch the first yard switcher in the next year, and CP Rail announcing a road engine retrofit the year after.

Moreover, this shifting regulatory climate is also accelerating interest in transitioning to Hydrail passenger trains. Hydrail passenger rail has been in commercial service for several years in Europe (e.g. Alstom's iLint Coradia), where zero-emission electric tram-trains using Hydrail technology have been proven to provide a quiet, environment friendly solution, while also enhancing transport equity and access to housing and community services. Building on this momentum, California governments are contracting Stadler to provide the first hydrail regional passenger rail vehicles and service in NA. In Canada, Alstom recently purchased the former Bombardier rail car plant in Quebec, which provides an opportunity to introduce their iLint hydrail passenger tram-trains. To introduce hydrail in NA, in Canada, and specifically in inter-city, regions to provide commuter and/or eco-tourism, it is necessary to increase the understanding of the costs, benefits, and outstanding issues (regulatory and otherwise) that must be addressed to ensure safe, convenient, affordable, yet sustainable zero emission regional passenger rail service. In this paper, the results of a study are presented on the economic feasibility of implementing a hydrail passenger system, in the geographic context of BC's Okanagan Valley connecting the US and Canadian rail networks.

BC's Okanagan Valley is a 250 km long tourism mecca yet very auto dependent, creating transport equity issues and travel barriers for lower income service workers, non-driving youth and seniors, and fly-in tourists in the Valley. Apart from two regional airports, access to, from, and around the Valley is via Highway 97, which runs north-south to connecting Okanagan communities by car like a string of pearls. Moreover, Highway 97 is the longest north-south highway in North America, running from Alaska, through BC, the USA, and into Mexico. No regional, long-distance bus service exists within the Okanagan Valley (e.g., neither Greyhound, nor BC Transit). Freight rail did run in segments near the border, and north of Kelowna, but tracks were abandoned with the advent of Highway 97 construction, which made trucking an economically convenient substitute. Passenger rail in the Okanagan was abandoned even earlier with the advent of BC's Provincial road network, leading to the rise in auto tourism that doubles the Okanagan population each summer, which often creates bumper to bumper, stop and go congestion along Highway 97 throughout Okanagan cities. Seven of the ten deadliest road crash sites in the BC interior are in the Okanagan Valley on Highway 97. Considering these factors - increasing traffic congestion and safety problems - combined with

rising gas prices and the federal commitments to address the global climate crisis using ZE technology and to improve transport equity, there is a case for passenger rail in the Valley. A successful ZE regional passenger rail would promote active, healthy lifestyles; to reduce greenhouse gases and pollution; to improve traffic safety and congestion; and, ultimately, to provide more affordable access for all to housing, jobs, and services. While this study looks only at ZE regional inter-city passenger rail (hydrail) in this case study analysis, this research provides the methodological framework for future research on the feasibility for either freight and/or passenger zero-emission hydrail rail networks in other regions wishing to provide sustainable transportation solutions.

of the work. The author also should seek to answer the question "Why this research matters. Is this research making a contribution to theory and/or practice?"

1.1 UBC Okanagan Hydrail Research Program

Researchers at UBC's Okanagan campus in Kelowna (UBCO), in the School of Engineering's Sustainable Transport Safety (STS) Research Laboratory, have been working on a design to address the problems reported by previous NA attempts to produce a ZE locomotive (Hegazi *et al.*, 2019; Hegazi and Lovegrove, 2020a; Change Energy Services Inc., 2020). The proposed remedies include a more optimal selection and sizing of the different power sources, and a crew friendly design that will require minimal training and maintenance. This will require the use of embedded control systems that will emulate the dynamics of a regular diesel-electric locomotive, automatically deciding on the power split between the power sources, as well as managing the battery bank's charge. The team also plans to do what no other project has done before, which is to write the first railway industry manual on handling hydrogen with a focus on safety. As a part of this research, testing for the design and construction of a standard gauge, low-powered (35 kW), fuel cell/battery hybrid vehicle (Hydrail 1) capable of hauling a single empty freight car on tangent level track was performed (Hegazi and Lovegrove, 2020b). The initial system design and testing was done at the UBCO STS Laboratory with a further testing on a dead track in the New Westminster rail yard. The results of this study have been used to design a fuel cell/battery hybrid retrofit of Southern Railway of British Columbia (SRY) "Green Goat" switcher locomotive (Hegazi and Lovegrove, 2019).

The SRY switcher retrofit project aims to study the impact of hybridization between hydrogen fuel cells (HFCs) and lithium-ion batteries for switching applications. Restoring the Green Goat to working to hydrail vehicle will help to drive hydrail technology forward in Canada and North America. Results from STS research to develop Hydrail 1 and to retrofit the Green Goat retrofit will allow a variety of powertrain architectures and hybridization for many case studies, showcasing how zero-emission technology can help improve the environmental performance of Canada's rail

transportation system. Moreover, this research will increase understanding of how to retrofit all NA diesel locomotives, including SRYS entire fleet of 31 diesel-electric locomotives, to reduce air emission, noise, vibrations, and fuel/operating costs.

1.2 Hydrail Technology

Hydrogen is the most abundant element in the universe, but at atmospheric pressure and temperatures on earth is present in water (i.e. H₂O) molecules. When electrolyzed into its elemental form via one of several processes, including many zero-emission production methods (e.g. tidal, solar, wind, hydro-electric), hydrogen is considered one of the best energy carriers, as when input into an HFC and re-combined with oxygen, generates electricity, water vapor and heat at relatively high efficiencies (Hoffrichter *et al.*, 2015; Hegazi and Lovegrove, 2019). Previously, HFC technology suffered from very low efficiencies; however, their current efficiency runs over 50%, motivating research and development into its use as an alternative to dirty, noisy, vibrating, internal combustion engines. Additionally, the very high energy density of hydrogen gas makes fuel cells an economically wise option in the long term. They have been successfully employed in the automotive sector despite some limitations on HFC responsiveness to dynamic loads, which can be resolved using hybrid powertrains with a more power dense energy source (e.g. batteries).

Applying this technology to retrofit the existing diesel-electric locomotives or to implement in new hydrail vehicles as in the Coradia iLint case in Germany, provides a sustainable transition for the rail industry. In these locomotives, the fuel cells are connected to charge a battery bank (lithium-ion) which will power the vehicles traction motors and auxiliary loads. Charging the battery bank by HFC provides a constant power transfer to the drive system. Recent advancements in HFC and battery technology have made ZE trains economically feasible [9] in North America, where the majority of rail traffic is carried out on non-electrified lines using diesel-electric locomotives.

HFCs can produce electricity as long as fuel and oxygen are supplied, and thus onboard energy storage is required. While the lack of adequate hydrogen supply infrastructure is often cited as a major hindrance to adoption as an energy carrier for road vehicles, railway vehicles run on a fixed guideway and can be operated in a much more controlled manner. This fact reduces the complexity of the hydrogen supply chain and infrastructure design process (Guerra *et al.*, 2020). Researchers at the National Hydrogen Centre in Spain determined the sustainability of implementing hydrogen refueling infrastructure for a fleet of 20 Multiple Unit hydrail trains at 4 tonnes of hydrogen per day. That research, in addition to Alstom's successful iLint Coradia in-service in Germany since 2018, prove that a railway centric approach to the adoption of hydrogen in transportation is a technically and an economically feasible solution, via refueling the stations and refueling durations comparable to current infrastructure. Regarding the feasibility of onboard hydrogen storage, UBC's STS Lab is researching various on-board hydrogen (H₂) storage formats to optimize operating range. Initial results have already shown that hydrail yard switcher locomotives can successfully emulate operating range of conventional diesel-electric switchers. Thus, hydrogen supply, refueling, and storage technology are no longer barriers to implementing ZE rail transport. The issues remaining to widespread adoption of hydrail in North

America relate more to economic feasibility and, perhaps more important, understanding the need to, and the business case for, change in an historically risk-averse industry.

2 ECONOMIC FEASIBILITY OF OVER PR

As noted previously, the Okanagan Valley's auto-centric travel patterns have led to significant safety, congestion, equity, and air quality problems. Therefore, the economic feasibility of implementing an Okanagan Valley Electric Regional Passenger Rail (OVER PR) Hydrail system was researched, running its entire 290-kilometre length connecting Osoyoos and US rail networks at the border with Canadian rail networks via Kamloops in the north, as shown in Figure 1 below.

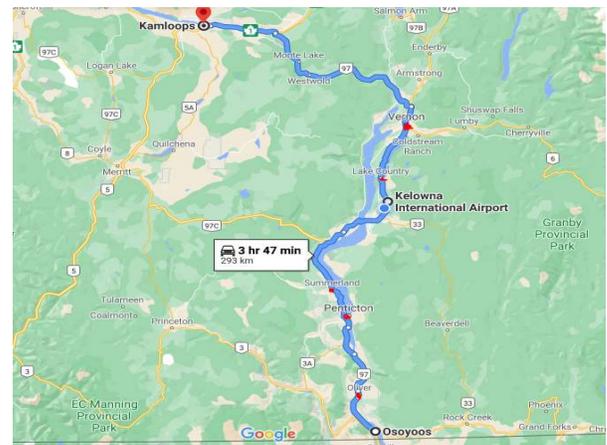


Figure 1. Overview of OVER PR route

Following regional tram-train operations best practises emerging worldwide, the following assumptions were considered reasonable regarding OVER PR:

- OVER PR will be based on tram-train technology operating in Europe (Germany, Netherlands);
- In cities, OVER PR will operate at city speeds, at grade, in imbedded rail shared traffic lanes like a tram;
- Inter-city, it will operate at highway speeds either beside or in the median of Highway 97;
- It will run at peak hour 30-minute headways in major cities, and 60-minute in base and outlying areas; this mean that the train will not always run the full course of the track, and short turns will be instituted based on ongoing demand monitoring
- It will significantly delay if not eliminate future Highway 97 widening via a minimum 30% mode shift; this 30% figure has been based on observed LRT shifts globally, and is considered conservative;
- OVER PR will stimulate rail-based eco-tourism via its connections to US and Canadian rail networks.

In addition to these general assumptions, several other key transport economic and policy assumptions were made, as shown in Table 1 below, following the BC Provincial Business Case template and best practises.

All cost and benefit values also underwent risk analysis, with results presented following.

Table 1. Key Economic & Policy assumptions

Project life	30 years
\$ Value Base Year	2021
Discount rate	6 %
Implementation	All services commence 2025
Fares	\$0.20/km
Operating cost	\$0.20/passenger-km

2.1 Capital & Operating Costs

Table 2 compiles all the costs associated with the proposed project. All cost estimates – construction, operating, and running gear - were based on peer reviewed journals and consultations with railway industry professionals. The construction cost of building the rail is the major contributing factor to the total direct capital which greatly affects the Net Present Value (NPV) of the project. To allow OVER PR tracks to cross Okanagan Lake on Highway 97, the Willian Bennett bridge would require some major modifications. In addition to track construction, other costs include: OVER PR stations (20 stations, each with surrounding area condo/office/retail redevelopment); a hydrogen electrolysis refueling station; railyards, signals, running gear, engineering, and, contingency. Operating costs, including running gear, staffing, fuel and maintenance totalled \$55 million annually. The total capital cost (Direct + Indirect) is estimated conservatively at \$ 1.7 Billion, rising to \$2.5 Billion when the present value of operating costs over 30 years are added in.

Table 2. Project Cost Estimates (rounded, millions)

Capital				
Item	Unit	Cost/Unit	Quantity	Sub-total
Tracks	km	Varies	300	\$ 600
20 Stations	Per station	\$10,000,000	20	\$ 200
Station Redevelopment	Area	One time		\$ 440
H2 Refueling Station		\$ 1,500,000	1	\$ 1.5
Sig & Comm's	Per km	\$ 100,000	290	\$ 29
Yards	Per yard	\$ 10,000,000	1	\$ 10
Tram-trains	Per unit	\$ 10,000,000	16	\$160
Engineering	Lump sum	% of track costs		\$ 90
Contingency	Lump sum	% of track costs		\$ 180
Total Capital Costs				\$ 1.7 Billion

Operating				
Running Gear	Lump Sum	\$ 1,000,000	1	\$ 1.0
Fuel	kg Per Pass-	\$ 5.27	195,582	\$ 1.0
Maintenance	km Per	\$ 0.20	246,440,700	\$ 49.0
Staff	person	\$ 70,000.00	50	\$ 3.5
Total Operating Cost				\$ 55 million/yr
Present Value (Capital & Operating Costs)				\$3 Billion

2.2 Benefits

Table 3 summarizes expected benefits of the OVER PR over its 30-year life, with key considerations including:

- The areas surrounding major stations will be redeveloped into higher activity condo/commercial.
- Conservative occupancy rates (e.g. absentee owners), a pessimistic (5% of condos rented), expected (30%) and optimistic (50%) occupancy.
- Track and running gear salvage values at the end of 30 years are significant, an estimate of 40% for the locomotives and \$ 500,000 per km for track.
- Fare per kilometre reflects local transit fares.
- External benefits (externalities) values draw heavily from those recommended for transportation planning and economic feasibility studies that have been widely sourced and compiled on-line by the Victoria Transport Policy Institute (vtpi.org).

To derive estimates of the various externalities, the number of annual vehicle-kilometres-travelled (VKT) was estimated based on the annual number of vehicles (8,700 vehicles per day on average) removed from along the OVER PR Highway 97 route due to a 30% mode shift to hydrail. Next, probabilistic minimum, expected and maximum values of cost per VKT were used to calculate annual sub-total estimates.

2.2.1 Congestion

As highlighted by Litman (2020a), "Congestion tends to increase travel time, arrival unreliability, fuel consumption, pollution emissions and driver stress, and reduce life satisfaction". It is imperative that congestion, be it recurrent or non-recurrent, must be identified and incorporated in economic evaluations. The expected cost saving for congestion was taken to be \$ 0.168/VKT.

2.2.2 Water Pollution and Hydrologic Impacts

These include oil drips, damage due to de-icing of roads, settlement of air pollutants, and others. But overall, as mentioned by Litman (2015), "runoff from roads and parking lots has a high concentration of toxic metals, suspended solids, and hydrocarbons, which originate largely from automobiles." And "these impacts impose various costs including polluted surface and ground water, contaminated drinking water, increased flooding and flood control costs, wildlife habitat damage, reduced fish stocks, loss of unique natural features, and

aesthetic losses.” The expected cost savings per unit on Water Pollution and Hydrologic Impacts was taken to be \$ 0.0174/VKT.

2.2.3 Waste disposal

As Litman (2009) explained “Waste disposal external costs include damage costs associated with the inappropriate disposal of used tires, batteries, junked cars, oil and other harmful materials resulting from motor vehicle production and maintenance.” The expected value for Waste disposal was taken to be \$0.000621/VKT.

2.2.4 Local Air Pollution & GHG consideration

Tailpipe exhaust emissions (including but are not limited to: particulate matter, sulphates, carbon monoxide, nitrogen oxide, sulphur oxides, lead, CFCs, HCFCs, road dust, methane, VOCs, ozone), lifecycle emissions and non-exhaust emissions damage human health, ecological health, and aesthetic conditions. The benefits of reduced GHG emissions are in addition to those from reduced local air pollutants. After averaging the minimum and maximum ranges for the Cost per VKT estimates on Local Air pollution & GHG emissions, the final value for the savings on their reduction were calculated using a unit value of \$0.00895/VKT & of \$0.1397/VKT respectively. Litman (2020b)

2.2.5 Noise

Litman (2020c) expounded on the ill-effects that noise, defined as unwanted sounds and vibrations, when caused due to engine power changes, and low frequency noise from drivetrains of heavy vehicles can have huge economic implications on the area surrounding any kind of transport corridor. In general, low speed noise comes from engines, whereas at higher speeds aerodynamic and tire/road noise dominate. He concludes that urban traffic noise costs for Rail travel per 1000 kilometres were lower than those for Bus travel and significantly lower than that for cars. The differences of costs on the mode shift transfer then helped in estimating the cost savings on noise for the project, at \$0.0497/VKT.

2.2.6 Safety

Often overlooked in economic considerations, is the loss of life and injuries due to road crashes. Railways, have long been among the safest of all transport modes (Bowen, 2011; Geerrieri, 2018). The final value of road safety benefits was arrived at using an average of two approaches. First, Litman (2018) recommended it be valued at \$ 0.2237/VKT in his VTPI report on Safety and Health Costs. Second, an estimate of avoided injury and fatal death claims due to the expected 30% mode shift was calculated. Knowing that road collision predictions are predominantly VKT based, and the low vehicle occupancy observed in the Okanagan Valley, it was reasonable to extrapolate that a 30% mode shift would remove 30% of traffic crashes, subject to error analysis noted following. As per ICBC, while the average number of annual fatal road crashes along the corridor in consideration are 81, the average number of crashes

involving injury claims are 9,900 per year. The approximate cost of human lives (valued at \$ 8 million each) lost along with the cost of injuries (valued at \$ 300,000 per crash), assuming a 30% mode shift from road to the proposed hydrail. Taking the average of the methods, the benefits on Safety were estimated to be \$570 million per year due to road safety alone!

2.2.7 Tourism

Tourism is one of the main economic drivers for the Okanagan Valley, with eco-tourism expected to rise. The Okanagan offers a host of major tourist destinations, airport connections, golf courses, summer resorts, globally renowned wineries, and more. As tourism increases, the direct expenditure in terms of tourist accommodations such as bed and breakfasts, hotels, motels, resorts, and in terms of miscellaneous expenses (food and transportation) also increases. Historic increases in tourism revenue were researched to calculate OVER PR tourism estimates. In Central Okanagan, the total direct expenditure by visitors in the Greater Kelowna area nearly doubled from \$ 810 million in 2016 (InterVISTAS Consulting Inc., 2017) to \$ 1.4 Billion in 2018 (InterVISTAS Consulting Inc., 2020). In the north, Kamloops, the direct visitor spending increased from \$ 227 million in 2015 to \$ 270 million in 2017 (Tourism Kamloops, 2018). Based on similar rail-based tourism increases, an expect value of between 10% and 30% boost with OVER PR rail connections to American and Canadian networks was reasonable, so that non-driving, non-flying eco-tourists can visit the Valley. Using an average 20% OVER PR boost, and a multiplier of two derived by simply pro-rating Kelowna versus total Okanagan Valley population, an OVER PR annual tourism benefit starting in 2025 at 20% x 2 x \$1.5 Billion of approximately \$600 million/year has been calculated.

2.2.8 Deferred Hwy 97 widening

The lower-cost, high-capacity Hydrail corridor offers a clean, safe and efficient alternative to the higher-cost, lower-capacity Highways 97 widening; thus, there would be an accrued benefit to society (i.e. taxpayers) of in terms of reduced costs to add capacity. It was assumed that the proposed Hydrail project would likely eliminate the need for Highway 97 widening over the 30-year life of the hydrail project, while offering improved transport equity, access to more affordable housing, lowering congestion, and improving travel times, all while a cleaner, reliable, and safer trip. A conservative way to estimate this benefit, in lieu of having equity and housing benefit values, was simply to calculate the present value on interest accrued on the net \$1 Billion estimated savings in costs between building and operating OVER PR versus a widened Highway 97. It has been reported that costs to widen Highway 97 range in the order \$20 million to \$50 million per kilometre due to mountainous terrain, which over its 200-kilometre length would total in the order of \$ 4 to 12 Billion present value, versus a PV of less than \$3 Billion to build and operate OVER PR. A conservative cost value was used of \$5 Billion Hwy 97. The reasonableness of this assumption was further

tested by looking at historic traffic and population growth levels, which in the order of 1% have been observed. When compounded over 30 years would total approximately 35% growth, slightly over 30% expected mode shift (which is a conservative estimate), but well within the upper/lower bounds of the analysis that was also performed (and which follows).

Table 3. Summary of OVER PR Benefits (rounded)*

	Unit	Annual Benefits*
Station revenue from condo sales	One time (OT), PV	\$600 million OT
Salvage value	OT, PV	\$70 million OT
Station office/retail rental revenue	Annuity	\$60 million
Passenger Fares	\$0.20/km	\$50 million
Advertising Revenue	Annuity	\$1 million
Carbon tax savings	Annuity	\$2.5 million
Deferred Hwy 97 construction	OT, PV net of OVER PR	\$3 billion OT
Concession sales/Food sales	Annuity	\$1 million
Light freight (fruit/wines)	Annuity	\$1 million
Congestion	\$0.168/VKT	\$100 million
Water Pollution	\$0.0174/VKT	\$10 million
Local air pollution	\$0.0621/VKT	\$30 million
GHG	\$0.0932/VKT	\$45million
Noise	\$0.0497/VKT	\$27 million
Waste Disposal	\$0.000621/VKT	\$0.3 million
Parking	\$0.087/VKT	\$40 million
User: Car ownership/use savings	\$0.14/VKT	\$85 million
Tourism	Annuity	\$625 million
Travel Time savings	\$0.078/VKT	\$37 million
Safety	Annuity	\$570 million
PV-Benefits over 30 yr project life @ 6%		\$22 Billion

*Note: Benefits accrue annually except as noted (e.g. annuities, one-time benefits).

3 RESULTS AND DISCUSSION

Table 4 shows the expected Net Present Value (NPV) of the project at a 6% discount rate to be nearly \$20 Billion with a Benefit Cost Ratio (BCR) of 9:1, and 2-year payback period (PBP). Increasing the discount rate of future benefits (and costs) to reflect equity partner 15% returns on investments, these results change little due to the long (30 year) project life over which benefits accrue, and, still reflect a better BCR than that forecast from the unsustainable traditional highway widening approach.

Table 4. Economic Feasibility Results

	6% discount rate**	15% discount rate*
PV (Benefits)	\$ 22 Billion	\$12.5 Billion
PV (Costs)	\$ 3 Billion	\$ 2.1 Billion
NPV	\$19 Billion	\$10.4 Billion
BCR	7:1	6:1
Payback period (PBP)	2 years	3 years

*15% reflects closer to expected P3 equity partner RoI, for comparison

**6% reflects cost of borrowing capital to fund by debt

2.3 Risk Analysis

We performed a sensitivity analysis of all variables that would impact the project NPV the most. The results of 50,000 Monte Carlo Analysis simulation (MCS) trials showed in 95% of cases that OVER PR will produce a positive NPV. The results of the MCS at the 95% tail revealed an NPV of C\$6.8 Billion and BCR of 2.8. This MCS result necessarily differs from, and are more conservative than, the originally calculated expected results listed in Table 4, yet still demonstrate the economic feasibility of the OVER PR project with high confidence.

2.4 Limitations

In the absence of a complete engineering design, numerous assumptions had to be made for this academic exercise. A preliminary system engineering design was prepared, rail professionals were consulted, and recognized transport economic methods were followed to minimize errors and confirm reasonable valuation ranges for MCS testing at a 95% level of confidence. Regardless, the significantly positive NPV, BCR, and PBP results warrant a closer look in future research, together with more in-depth engineering and development of a business case. If these positive results continue to hold, the next steps are to explore possible public-private-partnership infrastructure investments to get the project designed, built, and operating. A critical pre-cursor will be a full public consultation process, including engagement of all Indigenous and First Nations communities through which Highway 97 and the future OVER PR corridor runs. OVER PR could also qualify for seed funding as part of federal post-Covid recovery and ZE transport programs (e.g. Build Back Better, Just Transition, Transport 2030).

4 ACKNOWLEDGEMENT

UBCO School of Engineering gratefully acknowledges the generous support of the following organizations in this research:

1. Southern Railway of BC – Mr Laszlo Czihaly, Chief Mechanical Officer, SRY's client representative (and Mr Singh Biln, retired SRY) who have provided in-kind support and on-site facilities use.
2. Government of Canada (NSERC, TC) – Who have graciously supported this research via grants.
3. Hydrogen in Motion – Ms Grace Quan, CEO and Mr

Mark Cannon, CSO, who have graciously supported this research through in-kind and cash contributions via Mitacs grants.

4. Cariboo Central Railway Contractors Ltd – Mr Blair Irwin, President – via in-kind and cash contributions.

5 REFERENCES

Alstom (2020) *Alstom's Hydrogen Train Enters Regular Passenger Service in Austria*, www.alstom.com/press-releases-news/2020/9/alstoms-hydrogen-train-enters-regular-passenger-service-austria.

Bowen, D. J. (2011). *Consensus on crashworthiness: Passenger rail has long been among the safest of all U.S. transport modes. the federal railroad administration, soliciting input from suppliers and others in the industry, seeks to make it even safer in the future*. Railway Age (Bristol), 212(12).

Change Energy Services Inc (2020) *Assessment of the Design, Deployment Characteristics and Requirements of a Hydrogen Fuel Cell Powered Switcher Locomotive Report of the Phase I Work Plan*, Report prepared for Environment & Climate Change Canada – Contract No. 30000704788, Ottawa, ON, 125 pages.

Environment and Climate Change Canada (2019) *Greenhouse gas sources and sink in Canada*. Available online: [En81-4-1-2019-eng.pdf](https://www.ec.gc.ca/en81-4-1-2019-eng.pdf) (publications.gc.ca)

Guerra, Carlos Fúnez, et al. (2020) *Sustainability of Hydrogen Refuelling Stations for Trains Using Electrolysers*, International Journal of Hydrogen Energy, vol. 46, no. 26, 6 Oct. 2022, pp. 13748–13759., doi:10.1016/j.ijhydene.2020.10.044.

Guerrieri M (2018) *Tramways in Urban Areas: An Overview on Safety at Road Intersections*, Urban Rail Transit 4(4): 223-233, Springer, <https://doi.org/10.1007/s40864-018-0093-5>

Haji Akhoundzadeh, Mehran, et al. *A Conceptualized Hydrail Powertrain: A Case Study of the Union Pearson Express Route*. World Electric Vehicle Journal, vol. 10, no. 2, 2019, p. 32., doi:10.3390/wevj10020032.

Hegazi M., Lovegrove G (2019) *Locomotive on-board electric retrofit: Using the UBCO STS research lab's locomotive electric power management expertise to solve Southern Railway of BC (SRY) design problem*, UBCO School of Engineering, Kelowna, BC.

Hegazi M., Markley L., Lovegrove G. (2019) *Examining the influence of battery sizing on hydrogen fuel cell/battery hybrid rail powertrains for regional passenger railway transport using dynamic component models*, Canadian Journal of Civil Engineering, Ottawa, ON.

Hegazi M., Lovegrove G (2020a) *Applying Hydrail as a zero-emission rail power gateway technology to in-service switcher locomotive duty cycles* Proceedings,

Annual Conference - Canadian Society for Civil Engineering. 2019. Montreal.

Hegazi M., Lovegrove G (2020b) *Design Report for a Hydrogen -Hybrid Railway Speeder Conversion Project – Hydrail 1*, Prepared for Transport Canada, Clean Transportation System – Research and Development Program, Ottawa, ON, 23 pages.

Hess, K.S., Miller, A.R., Erickson, T.L., & Dippo, J. (2010). *Demonstration of a Hydrogen Fuel-Cell Locomotive*.

Hoffrichter, A., Fisher, P., Tutcher, J., Hillmansen, S., & Roberts, C. (2014). *Performance evaluation of the hydrogen-powered prototype locomotive 'Hydrogen Pioneer'*. Journal of Power Sources, 250, 120-127. doi: 10.1016/j.jpowsour.2013.10.134

Hoffrichter, Andreas, et al. (2015) *Conceptual Propulsion System Design for a Hydrogen-Powered Regional Train*. IET Electrical Systems in Transportation, vol. 6, no. 2, 11 Apr. 2015, pp. 56–66., doi:10.1049/iet-est.2014.0049.

Litman, T. (2020a). *Transportation Cost and Benefit Analysis II – Congestion Costs*. Victoria Transport Policy Institute (www.vtppi.org). <https://www.vtppi.org/tca/tca0505.pdf>.

Litman, T. (2015). *Transportation Cost and Benefit Analysis II – Water Pollution*, Victoria Transport Policy Institute (www.vtppi.org). <https://www.vtppi.org/tca/tca0515.pdf>.

Litman, T. (2009). *Transportation Cost and Benefit Analysis II – Waste Disposal*, Victoria Transport Policy Institute (www.vtppi.org). <https://www.vtppi.org/tca/tca0516.pdf>.

Litman, T. (2020b). *Transportation Cost and Benefit Analysis II – Air Pollution Costs*. Victoria Transport Policy Institute (www.vtppi.org). <https://www.vtppi.org/tca/tca0510.pdf>.

Litman, T. (2020c). *Transportation Cost and Benefit Analysis II – Noise Costs*. Victoria Transport Policy Institute (www.vtppi.org). <https://www.vtppi.org/tca/tca0511.pdf>.

Litman, T. (2018). *Transportation Cost and Benefit Analysis II – Safety and Health Costs*. Victoria Transport Policy Institute (www.vtppi.org). <https://www.vtppi.org/tca/tca0503.pdf>.

Office of Research, Development and Technology (2017), *Hybrid Locomotive for Energy Savings and Reduced Emissions*, Federal Railroad Administration, Washington, DC, 2017.

Railway Association of Canada (RAC) (2020), *Rail Trends*, Ottawa, ON.

Session 2

COLD REGIONS ENGINEERING

Evaluation of current rail tank car TC128B steel weld performance

S. Xu, J. Chen, J. Mckinley, J. Liang, L. Yang, & A. Laver
Canmet Materials – Natural Resources Canada, Hamilton, Ontario, Canada

Michael Spiess
Transportation of Dangerous Goods - Transport Canada, Ottawa, Ontario, Canada

ABSTRACT

Rail tank cars operate under a wide range of environmental conditions including cold temperatures. In the event of a rail tank car accident, it is of vital importance to ensure adequate fracture toughness to avoid catastrophic fracture and separation, to achieve appropriate puncture resistance during accidents or upon impact loading, and to ensure the continuing safety of tank cars damaged after accidents. TC128B is currently the most common steel used in the manufacturing of tank cars. In this work, weld samples were taken from a DOT-117J tank car for evaluation. The objective of this work was to quantify the performance of TC128B tank car welds at low temperatures in order to assess the potential for brittle behaviour. The experimental work included the determination of (i) chemical composition, microstructure and micro-hardness (ii) tensile strength and cross-weld tensile failure locations in the temperature range of 23 °C to -60 °C, and (iii) Charpy transition curves. The weld metal (WM) and heat-affected zone (HAZ) showed typical microstructures of the weldment made using a double pass procedure. The WM microstructure close to the weld centerline in the second-pass consisted of coarse pro-eutectoid ferrite along WM solidification direction, acicular ferrite and/or bainite, and some pearlite. The hardness values in the WM and HAZ were considerably higher than those of base metal (BM) demonstrating the desired weld strength over-matching (i.e., higher WM hardness and strength than those of BM). Charpy V-notch values (CVN) of WM were significantly lower than those of BM and HAZ specimens at a given temperature. The CVN values of tank car WMs at low temperatures (-34 °C and below) were also significantly lower than at higher temperatures (10 °C and above); at a common reference temperature for pressure tank cars of -34 °C, CVN were approximately 20J which usually represents brittle fracture mode. These results are not unexpected however they clearly demonstrate that there is an increased risk of fracture at weld locations and low temperatures. The large difference of CVN in HAZs was due to the slightly offset two-pass welds and specimen location according to the Association of American Railroads (AAR) Appendix W specification. The Scanning Electron Microscopy (SEM) fractography examinations indicate that the low toughness observed in TC128B WM was not associated with any weld defect. The results raise caution for tank cars operating in low-temperature environments due to the brittle weld behaviour in impact loading situations.

1 INTRODUCTION

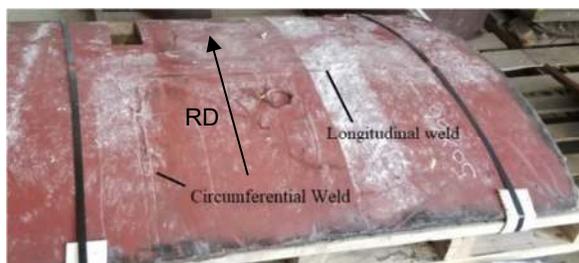
Mechanical properties of rail tank car steel and welds are critical to the integrity of the tank car. Tank cars often operate in cold temperatures where all steels become more brittle. Since WM can be more brittle than the BM it is particularly important to test WM at low temperatures. Tank car shells have longitudinal welds that travel along the length of the cars and circumferential welds. In recent work funded by Transport Canada (Simha & Mckinley, 2017; Xu, Liang, Yang, Laver, & Tyson, 2017; Xu, Mckinley, Chen, Liang, & Laver, 2019), BM, WM, and HAZ samples from a current DOT-117J tank car made of TC128B steel were characterized and evaluated at CanmetMATERIALS.

This paper summarizes the results of the evaluation of current TC128B tank car welds from -80 °C to 23 °C. The experimental work in this report included the determination of (i) chemical composition, microstructure and micro-hardness (ii) tensile strength and failure locations in the temperature range of 23 °C to -60 °C, and (iii) Charpy transition curves. The results of the longitudinal weld were compared to those from a circumferential weld from the same tank car.

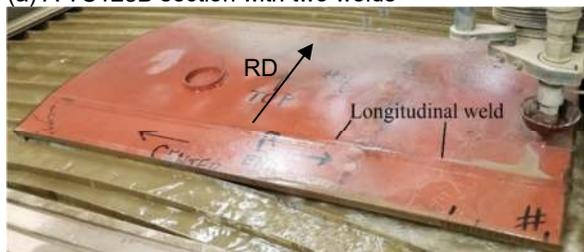
2 MATERIALS AND METHODS

The tank car TC128B weld samples donated by the Federal Railroad Administration (FRA) were taken from a tank car conforming to DOT-117 specifications, and the shell has a nominal thickness between 12.7 mm

(1/2") to 15.9 mm (5/8"). The tank car was fabricated using a submerged arc welding process in 2015 and subjected to a side-impact test in September 2016. The tank car did not see commercial use. The tank car purchase and the impact test was sponsored by the FRA, an agency of the United States Department of Transportation (DOT). The plate rolling direction (RD) is normal to the axis of the tank. The weld sections used in this work are shown in Figure 1. Cross-weld tensile specimens were taken transverse to the weld length and along the RD orientation of the steel plate. The steel thickness was approximately 14.8 mm close to the weld. The compositions of the tank car steel and WM samples were analyzed using optical emission spectroscopy and combustion/fusion analysis technique for oxygen (O) and nitrogen (N), as are given in Table 1. The Carbon (C) content of welds was considered to be high (typical less than 0.08% for production using low-C consumables). The Oxygen (O) of the longitudinal weld (0.00559%) was lower than that of the circumferential weld (0.0826%) and slightly higher than typical submerged arc welding (SAW) (~up to 0.05%) (Amanie, 2011).



(a) A TC128B section with two welds



(b) A TC128B weld section on a water-jet machine
Figure 1. Tank car TC128B I weld samples.

Transverse weld metallographic specimens were cut using a water-jet followed by polishing and etching using a 2% Nital solution for microstructural and micro-hardness characterization.

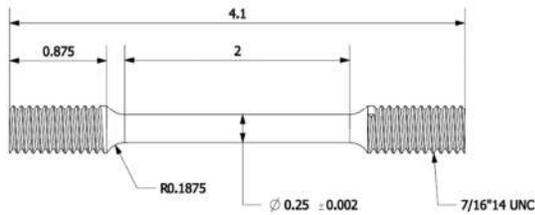
Vickers micro-hardness tests were performed using a diamond pyramid indenter with a load of 300 gram-force and a dwell time of 10 seconds. The hardness values (Hv) were measured across the weld at three through-thickness positions: (i) ~2 mm from the tank outside surface; (ii) at the mid-wall thickness; and (iii) ~2 mm from the tank inside surface.

Table 1. Chemical composition of the BM, circumferential WM, longitudinal WM samples, and maximum allowable by TC128B specification (ASTM International, 2015)

	BM	C - WM	L - WM	TC128B
C	0.19	0.18	0.16	0.26
Mn	1.34	1.54	1.57	1.00-1.70
P	0.0059	0.0182	0.0162	0.025
S	0.0014	0.0117	0.0088	0.009
Si	0.16	0.35	0.41	0.13-0.45
V	0.034	0.016	0.020	0.084
Cu	0.21	0.18	0.19	0.35
Al	0.032	0.021	0.021	0.015-0.060
Nb	<0.001	0.002	-	0.03
Ti	0.0016	0.0129	0.016	0.020
B	0.0002	0.0004	0.0003	0.0005
O	NA	0.0826	0.0559	-
N	0.0069	0.0083	0.0062	0.012
		1	5	
Sn	0.0112	0.006	0.0069	0.020
CE*	0.46	0.53	0.53	0.55
Nb+V+Ti	<0.0366	0.031	0.036	0.11
Cu+N	0.43	0.32	0.36	0.65
i+Cr+				
Mo				
Ti/N	0.23	1.6	2.56	4.0

Cross-weld tensile tests were carried out to evaluate the weld performance at different temperatures. Figure 2 shows the cylindrical cross-weld tensile specimen geometry used with a gauge length of 50.8 mm and diameter of 6.25 mm. Sub-size cross weld specimens were used to allow more specimens for a given length of the weld. Two cylindrical All-Weld-Metal (AWM) specimens with a gauge length of 25.4 mm and diameter of 6.35 mm were taken along the weld close to the middle-thickness to measure WM properties to compare to the specification requirements. Tensile specimens were etched before testing to show weld locations.

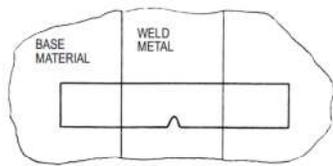
Full-size CVN specimens were prepared from the circumferential and longitudinal welds according to AAR specification (AAR M1002, 2014). Charpy specimen location and orientation for tank car WM and HAZ are shown in Figure 3. Charpy specimen notch was placed through-thickness. Machined Charpy specimens were etched to ensure notch location. Examples of Charpy notch at the longitudinal weld and HAZs are shown in Figure 4. Two types of HAZ notch locations were used, i.e., notched at HAZs of the second-pass weld (Figure 5b) according to [4] and of the first-pass weld (Figure 5c) to illustrate the effect of HAZ location on CVN.



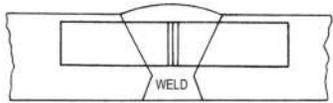
(a) Dimensions of cylindrical tensile specimens



(b) All-weld and cross-weld tensile specimens
Figure 2. TC128B tensile specimens.



(a) WM



(b) HAZ

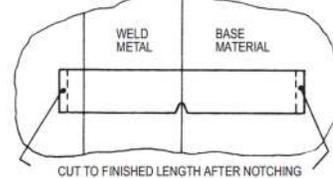


Figure 3. Charpy specimen location and orientation for tank car WM and HAZ according to AAR specification.

3 RESULTS AND DISCUSSION

3.1 Macrography of the Weld

Macrographs of the circumferential and longitudinal welds are shown in Figure 5. The macrography indicates that the welds were produced with a double pass submerged arc weld (SAW) procedure, which is a standardized mechanized welding procedure for TC128B tank cars (Sutton, Abdelmajid, Zhao, Wang, & Hubbard, 2002). The first-pass and second-pass welds of the longitudinal weld were not aligned well as shown

in Figure 5 (b) (Xu et al., 2019). This could have an effect to locate the notch of Charpy specimens to measure the HAZ because the fracture plane could be largely in the BM. There was a doubly tempered HAZ between the first-pass and second-pass WMs in the weld overlap region which is the lightly gray region. This reheated (RH) WM region was, as expected, significantly influenced by the re-heating of the second-pass weld.



(a) Charpy V-notch at the weld

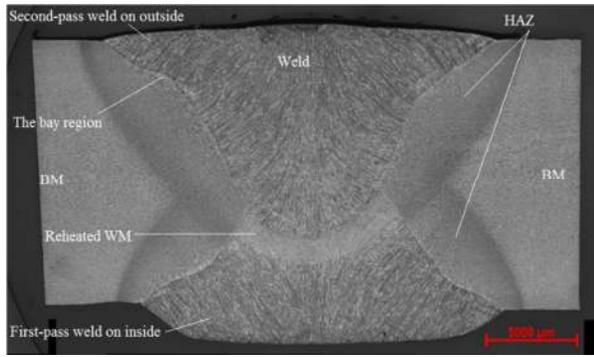


(b) Charpy V-notch at HAZ of the second-pass weld

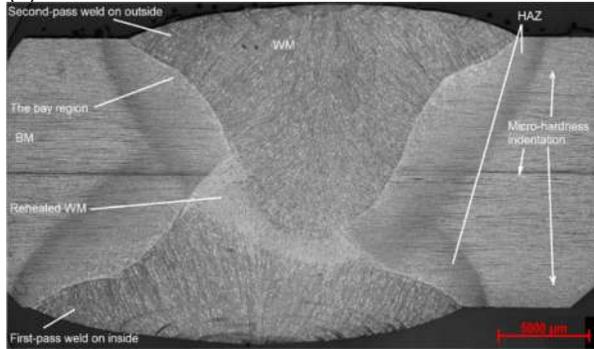


(c) Charpy V-notch at HAZ of the first-pass weld

Figure 4. Photos of Charpy notch location at the weld and HAZ of longitudinal specimens.



(a) Circumferential weld



(b) Longitudinal weld

Figure 5. Macrograph of TC128B circumferential and longitudinal welds.

3.2 Micro-hardness Across the Weld Joint

Micro-hardness measurements were made across the weld joint and through-thickness at the weld centerline. The hardness indentations positions across the weld are indicated in Figure 5 (although many hardness indentations could not be seen well at this magnification). Micro-hardness measurements are shown in Figure 6 and Figure 7. The main observation from the measurements was that the Vickers Pyramid Number (Hv) micro-hardness values measured in the weld and HAZ were generally considerably higher than those of the BM, i.e., demonstrating desired weld strength over-matching. The hardness values (Hv) in the HAZ at 2 mm below the outside surface close to the second-pass weld fusion line (so-called the bay region in Figure 5) were slightly higher than those in the other two thickness positions. It is typical in SAW welding that the bay region close to the fusion line in the final weld pass usually displays a coarse-grained microstructure. The microhardness values (Hv) of the longitudinal weld were comparable to those of the circumferential weld.

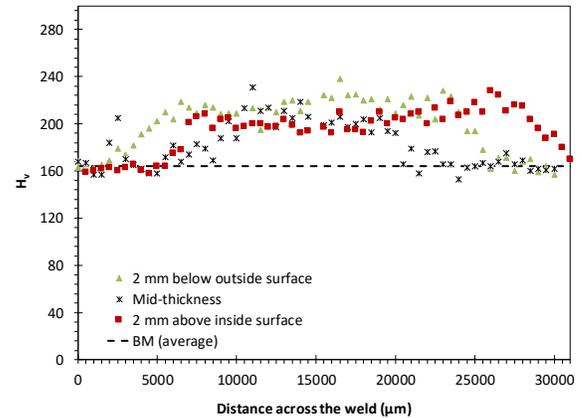


Figure 6. Micro-hardness across the TC128B longitudinal weld at 2 mm below the outside surface, mid-thickness, and 2 mm above the inside surface.

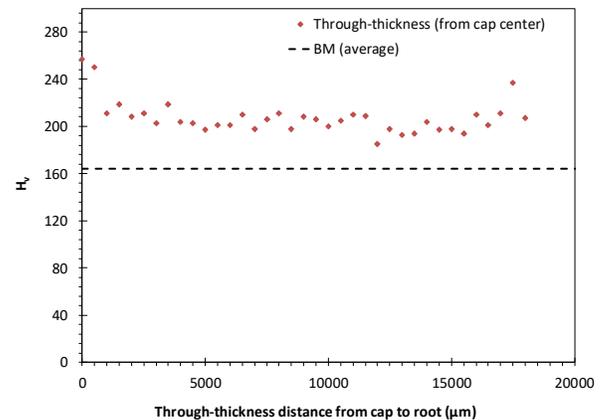


Figure 7. Micro-hardness measured through thickness starting at weld cap centerline of TC128B longitudinal weld.

3.3 Microstructure

The microstructure of the HAZ regions of the longitudinal weld is shown in Figure 8. In the coarse grained HAZ (CGHAZ) region close to the bay region, coarse grain boundary ferrite was present at prior austenite grain boundaries formed during welding, and within austenitic grains, some Widmanstätten ferrite and Acicular Ferrite (AF) and/or tempered bainite and/or tempered martensite were observed. Microstructures of the fine grained HAZ (FGHAZ) region close to the mid-thickness of the weld region consisted of fine equiaxed ferrite grains (white), pearlite (dark) and possibly a small amount of bainite, martensite, and retained austenite (grey). In the transition region of the FGHAZ and intercritical region (ICHAZ), pearlite bands were similar to those observed in the base metal. Microstructure of ICHAZ region close to the mid-thickness of the weld region consisted of fine equiaxed ferrite grains (white) and pearlite (dark). Since 1998, AAR specification (M-128) requires new tank cars to be stress relieved at 649 °C (1200 °F) for one hour after fabrication (as cited in (Hicho, 1993)). The stress-relieving treatment slightly decreases the strength,

increases or maintains the ductility of TC128B and are beneficial to low-temperature fracture properties (Hicho & Harne, 1992). As such, the microstructures presented were taken from the weld and HAZ of TC128B after stress-relieving treatment.

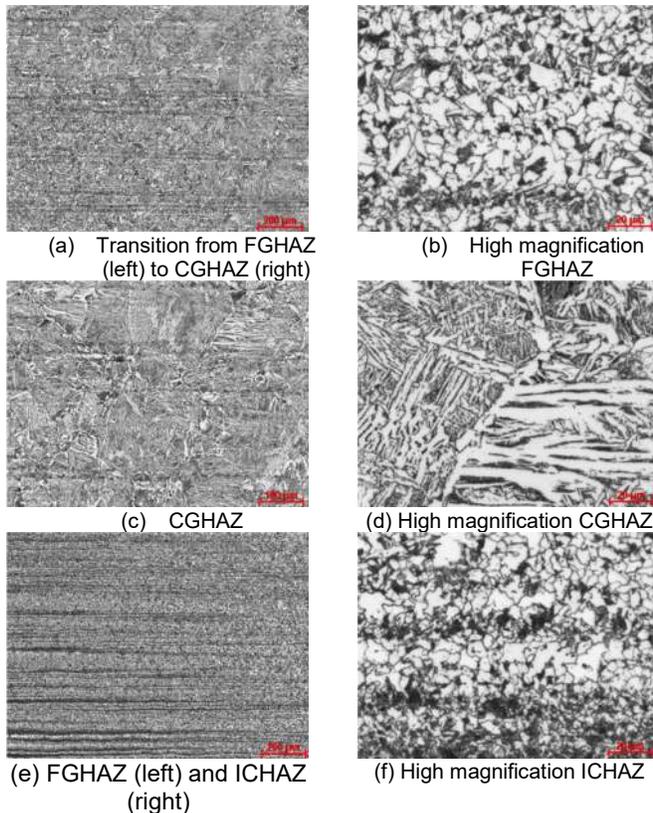


Figure 8. Microstructures of fine (FGHAZ), coarse (CGHAZ), and intercritical region (ICHAZ) of the longitudinal weld (the scales in a, c and e are 200 μm and in b, d and f are 20 μm).

The microstructure in the weld joint is more complex than that of the BM. The WM microstructure close to the middle part of the longitudinal weld centerline in the second-pass is shown in Figure 9. It consisted of coarse pro-eutectoid ferrite (white) along WM solidification direction, acicular ferrite (AF) and/or bainite (grey), and some pearlite (dark). The formation of pearlite in the weld metal indicated that the cooling rate after welding was relatively small. As a result, a very coarse proeutectoid ferrite was formed with a relatively high volume percentage, whereas the volume percentage of AF was small. Low AF volume fraction is usually associated with low toughness.

The microstructure of both WM and HAZ described above would be expected for a C-Mn steel SAW weld (Xu et al., 2019). However, it is worth noting that a very limited volume percentage of AF was formed in the weld metal and very significant austenite grain growth occurred in the CGHAZ region, which indicated that SAW was carried out with relatively high heat input. The volume fraction of AF was found to be important for good fracture toughness of C-N steel SAW welds (Svensson & Grefot, 1990). The welding procedure

specification for manufacturing the tank car needs to be reviewed in order to better understand the relationship between welding parameters, filler alloy composition and microstructure of weld metal.

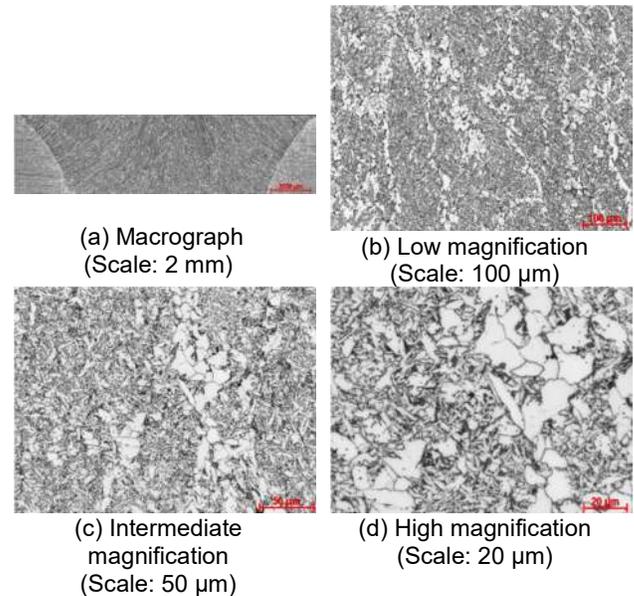


Figure 9. Microstructure of TC128B longitudinal WM at the middle of weld centerline in the second-pass.

3.4 Tensile Properties

The room temperature AWM tensile tests demonstrate that the WM is significantly stronger than the BM while maintaining significant but lower ductility as shown in Figure 10. All of the cross-weld specimens failed in the BM demonstrating strength over-matching which is the desired result as shown in Figure 11. A comparison between the UTS of the BM and the cross-weld tests from -60 °C to 23 °C is shown in Figure 12.

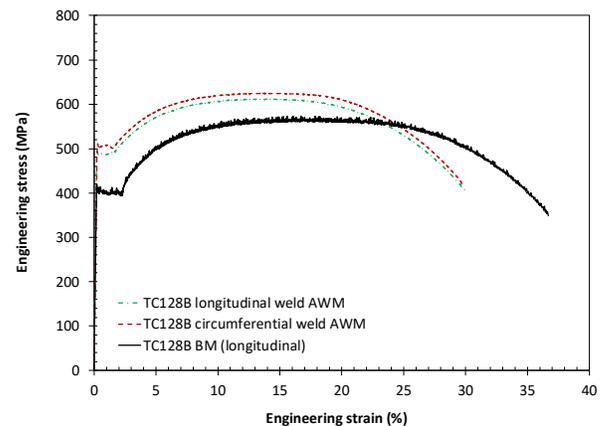


Figure 10. Engineering stress-strain curves of TC128B AWM and BM.

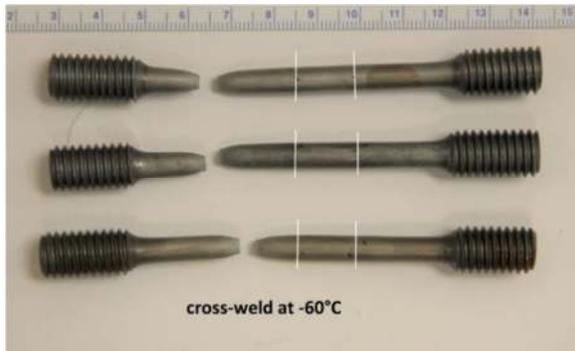


Figure 11. Example of cross-weld specimens showing weld and failure location.

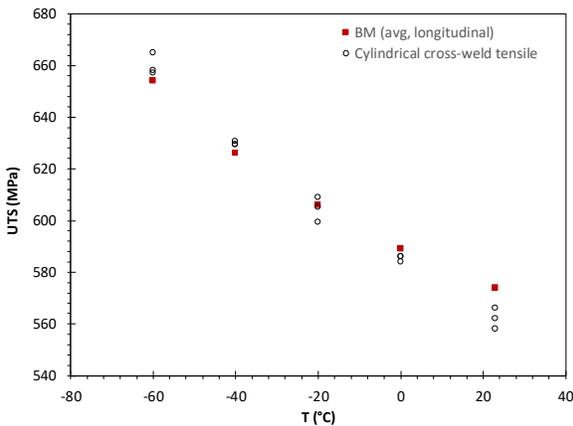


Figure 12. UTS vs. temperature of cross-weld tensile specimens of TC128B longitudinal weld.

3.5 Charpy Impact Properties

Charpy impact toughness is the only toughness requirement included in the AAR tank car steel specifications for pressure tanks. Charpy impact tests are also used to evaluate welding procedures for low-temperature applications (AAR M1002, 2014). This specification requires Charpy testing at -34 and -46 °C for certain types of rail tank cars. The test temperatures were selected to provide enough data points to create a profile of the steel over a wide range of temperatures. The CVN results for circumferential and longitudinal welds are shown in Table 2 and Table 3. The initial Charpy tests of the circumferential weld at -34 °C yielded CVN values of 29 J, 13 J and 24 J. The minimum value of one of the three specimens is less than the benchmark required value (i.e., 13.6 J in Appendix W 8.1.4.1) but met the retest requirements as the average of the three specimens (22 J), which is higher than 13.6 J. According to Appendix W 8.1.4.2, a retest of three additional specimens must be made and each of which must be equal to or exceed the specified minimum AAR required value for the average of the three specimens (i.e., 20.3 J). The retest of three specimens yielded CVN values of 20 J, 19 J and 20 J, each of which were lower than the AAR retest requirement of 20.3 J and hence, the toughness of circumferential weld did not meet the benchmark

specification requirement although the values were close to the requirements in both tests.

The result of a single test series can not be used to conclude that the tank car did not meet the benchmark specification at the time of weld qualification and manufacture. However, the results indicate that Charpy testing is important for rail tank cars in cold climate applications. If this steel sample had come from a pressure tank car or a special commodity non-pressure tank car, the Charpy testing values would have failed at -34 °C (pressure tank car testing temperature) and -46 °C (non-pressure special commodity tank car test temperature). The Charpy tests of TC128B showed that the weld exceeds the AAR standard requirements at moderate temperatures (> -20 °C). The material's performance at low temperatures (-34 °C), particularly in weld regions, can be problematic. Based on these results, future work to research the integrity of tank cars would be well advised to target low temperature weld performance. The Charpy transition curves of the circumferential and longitudinal welds are compared to the BM in Figure 13 and Figure 14. The longitudinal and circumferential welds are compared to each other in Figure 15. BM, WM and HAZ exhibited a gradual transition from upper shelf to lower shelf impact energy. BM and HAZ showed higher Charpy transition temperature than that of WM. For a 47 J transition temperature, it would be -6°C for circumferential WM and approximately -68°C for circumferential HAZ and longitudinal BM based on the fitting curves. For CVN of approximately 18 J or lower, the WM would be brittle.

Table 2. Charpy absorbed energy (CVN) of TC128B circumferential weld: average (and individual)

T (°C)	CVN (J)	
	WM	HAZ
25	80 (102,69,69)	208 (232,145,248)
-20	26 (20,28,30)	109 (159,92,75)
-34	21 (29,13,24,20,19,20)	.
-46	18 (16,24,13)	32 (16,17,64)
-60	9 (10,9,9)	92 (120,142,152,66,16,53)
-80	6 (5,7,5)	21 (25,33,6)

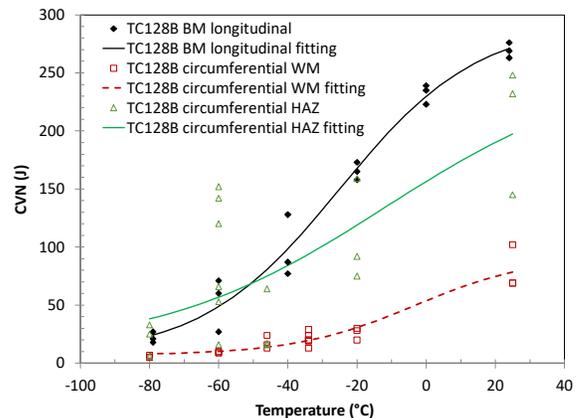


Figure 13. Charpy transition curve of TC128B BM, circumferential WM and HAZ.

Table 3. Charpy absorbed energy (CVN) of TC128B longitudinal weld: average (and individual)

T (°C)	CVN (J)		
	WM	HAZ (notched at the second-pass weld)	HAZ (notched at the first-pass weld)
21	109 (112,120,95)	-	225 (216,241,218)
-20	49 (50,53,43)	165 (172,158)	162 (175,165,145)
-46	16 (31,18,14,9,16,13,12,16,13)	121 (125,123,114)	109 (98,110,118)
-60	11 (9,17,8)	-	88 (69,95,99)
-80	5 (4,4,6)	-	21 (23,19,22)

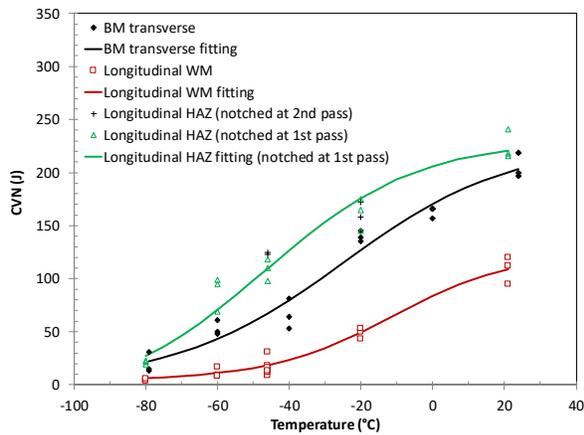


Figure 14. Charpy transition curve of TC128B BM, longitudinal WM and HAZs.

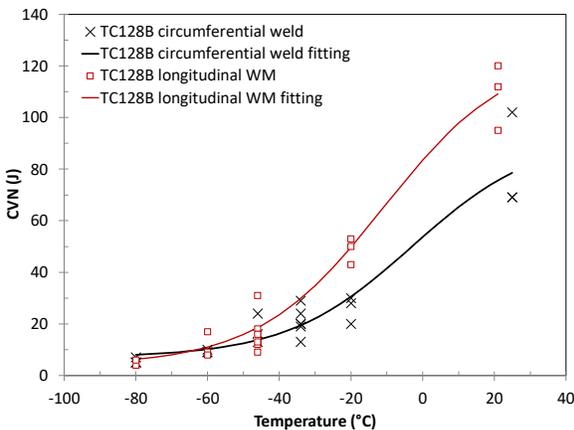
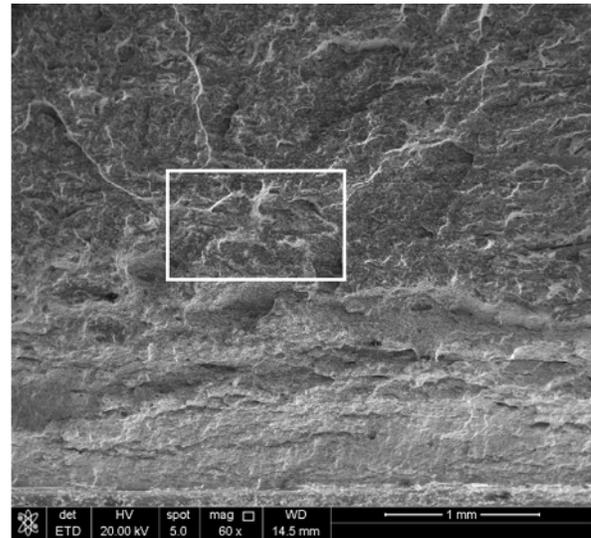


Figure 15. Comparison of Charpy transition curves of TC128B circumferential and longitudinal WM.

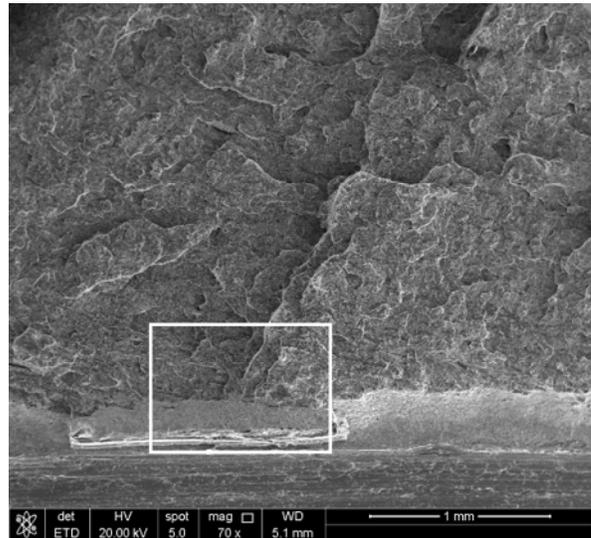
3.6 SEM Fractography

Fracture surfaces of two TC128B circumferential weld metal (WM) Charpy specimens were examined, one

tested at 25°C with CVN=69 J and another tested at -46 °C with CVN=13 J. Both specimens had the lowest CVN at the test temperatures. Both specimens showed cleavage fracture after limited or minimum ductile fracture. The specimen with CVN=69 J showed approximately 35% shear lips (ductile) and 65% brittle (crystallinity) fracture. The specimen with CVN=13 J displayed cleavage fracture (Figure 16). The cleavage initiation regions in the two specimens could be identified and no weld flaws were present (Figure 16 and Figure 17). Many round-shape inclusions were observed in voids in the ductile fracture region. The inclusions are likely oxides although the composition of inclusions was not determined.

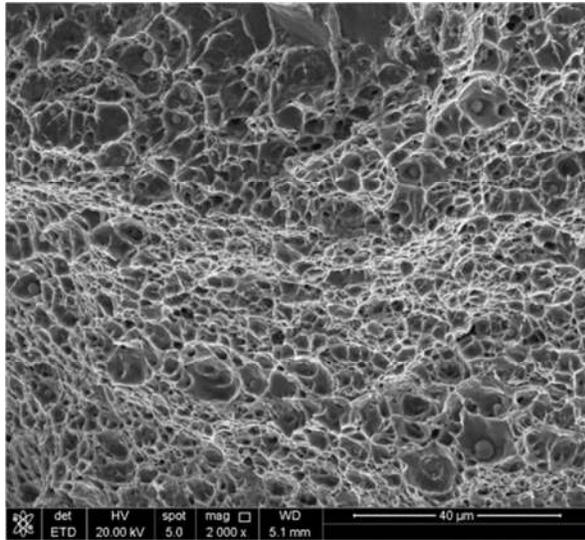


(a) Cleavage fracture after approximately 0.2 mm ductile fracture extension

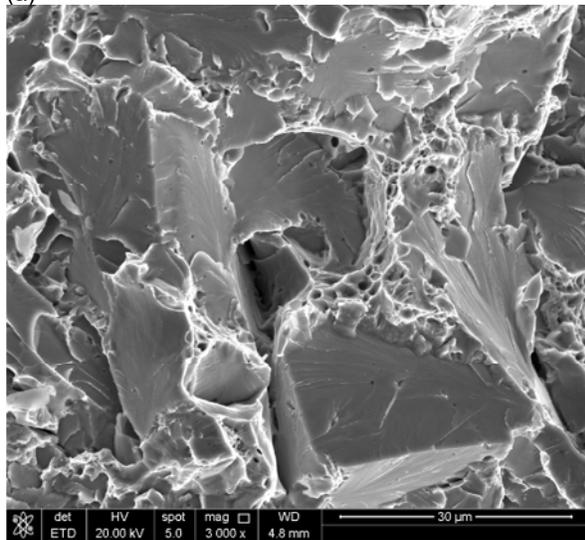


(b) Cleavage fracture after approximately 0.2 mm ductile fracture extension

Figure 16. Cleavage fracture initiation region after approximately 0.2 mm ductile fracture extension (a) and 1.2 mm ductile fracture extension (b).



(a)



(b)

Figure 17. High magnification SEM fractographs of ductile fracture (a) and brittle cleavage initiation (b) in Charpy WM specimens tested at 25 °C and -46 °C.

3.7 Acicular Ferrite Formed in TC128B Weld

The WM was found to consist of large quantities of coarse grain boundary ferrite, some intergranular polygonal ferrite, and ferrite with second phase, as well as a small volume percentage of AF. It has been reported that the formation of AF and its volume percentage are important to the toughness of the weld (Devillers, Keplan, Marandet, Ribes, & Ribond, 1984; Svensson & Grefott, 1990). Acicular ferrite nucleates heterogeneously on the surface of non-metallic inclusions in the weld metal during the austenite-ferrite-transition. As transformation continues, the ferrite grains radiate in various directions, creating a chaotic construction of crystallographically disoriented plates of approximately 5-15 μm in length and 1-3 μm in width. Thereby, the typical fine-grained and interlocking

structure of AF is formed. Compared with other microstructural constituents, it is much more difficult for cleavage cracks to propagate across this chaotically-oriented interlocking structure, which leads to a significant increase in mechanical properties, most notably in toughness.

Quantitative analysis of microstructure showed that the longitudinal WM contained approximately 40% AF, 36% coarse grain boundary ferrite (GBF) as well as 21% intragranular polygonal ferrite (PF(I)) and ferrite with the aligned second phase (FS(A)) (Figure 18). The volume percentage of AF formed in the circumferential WM was estimated to be around 25% by image analysis. As a result, the limited formation of AF may be the main reason leading to low CVN. The formation of AF is mainly affected by weld metal composition, cooling rate, inclusion landscape and austenite grain size. In order to improve the toughness of TC128B weld metal, the current welding procedure needs to be modified or a new welding procedure needs to be developed in order to optimize the formation of AF.

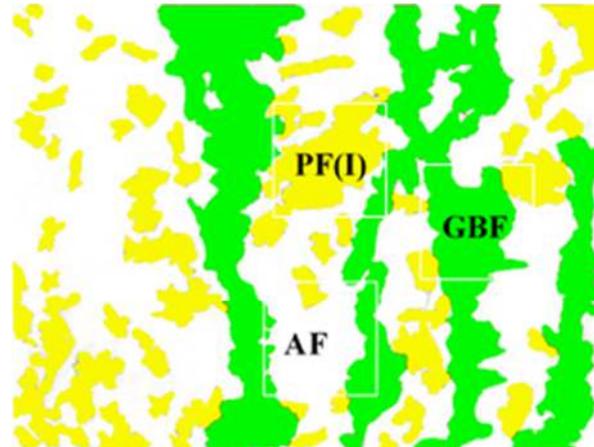


Figure 18. Illustration of areas acicular ferrite (AF) in white, coarse grain boundary ferrite (GBF) in green, and intragranular polygonal ferrite (PF(I)) in yellow for quantitative metallographic analysis.

4 CONCLUSIONS

The tensile performance of the double pass TC128B welds were as expected with sufficient weld strength over-matching.

The Charpy impact energy of the WM was significantly lower than that of the BM. The particularly low CVN of the circumferential weld may be due to high carbon and oxygen contents, and coarse proeutectoid ferrite and low AF (approximately 25%). Pressure tank cars and special commodity non-pressure cars have more different toughness requirements than DOT-117 tank cars. If the weld studied here had occurred in a pressure or special commodity tank car the Charpy testing values would have failed at -34 °C (pressure tank car testing temperature) and -46 °C (non-pressure special commodity tank car test temperature).

During an accident involving impact such as hard coupling or derailment the weld is more likely to fail due to the low toughness. Improving the toughness of the weld metal would reduce the chance of fracture/rupture under all service conditions. Weld toughness can be improved by adjusting the welding process, the filler metal, reducing heat/current, argon shielding, and the post-weld heat treatment.

5 ACKNOWLEDGEMENT

This work was funded by Transport Canada's Transportation of Dangerous Goods Directorate. Mr. Ian Whittal, Mr. Henry Lu and Mr. Shaun Singh are gratefully acknowledged for a very helpful review of the report. We would like to gratefully acknowledge Mr. S. Amey for sample preparation, Ms. R. Zavadil and Ms. P. Liu for performing metallography and microhardness tests.

6 REFERENCES

AAR M1002. (2014). *Manual of standards and recommended practices, section C-III (appendix M, and annex W), "Specifications for tank cars"*,

Amanie, J. (2011). *Effect of submerged arc welding parameters on the microstructure of SA516 and A709 steel welds* Available from Dissertation Abstracts International. Retrieved from <http://www.pqdtcn.com/thesisDetails/76E74E1D074749B62E5F14BFF5F5BD61>

ASTM International. (2015). *ASTM A20/A20M-15, standard specification for general requirements for steel plates for pressure vessels*.

Devillers, L., Keplan, D., Marandet, B., Ribes, A., & Ribond, P. V. (1984). The effect of low level of some elements on the toughness of submerged- arc welded CMn steel welds. *Int. Conf. Proc. the Effects of Residual, Impurity and Microalloying Elements on Weldability and Weld Properties, Abington, the Welding Institute, 31*, 141-194.

Hicho, G. E. (1993). *The mechanical, stress-rupture, and fracture toughness properties of normalized and stress relieved AAR TC128 grade B steel at elevated temperatures*.

Hicho, G. E., & Hame, D. E. (1992). Weld and heat affected zone crack arrest fracture toughness of AAR TC128 grade B steel. *Nistir 5157*,

Simha, C. M. H., & Mckinley, J. (2017). *High-temperature mechanical properties of tank car steel – testing and analysis*. CanmetMATERIALS Report, CMAT-2017-WF 16144032,.

Sutton, M. A., Abdelmajid, I., Zhao, W., Wang, D., & Hubbard, C. (2002). Weld characterization and residual stress measurements for TC-128B steel plate. *Journal of Pressure Vessel Technology, 124*(4), 405-414. doi:10.1115/1.1487878

Svensson, L. E., & Grefot, B. (1990). Microstructure and impact toughness of C-mn weld metals. *Welding Journal, 69* (12), 454S-461S.

Xu, S., Liang, J., Yang, L., Laver, A., & Tyson, W. R. (2017). *Tensile and fracture toughness of a current tank car steel, TC128B*. CanmetMATERIALS Report, CMAT-2017-WF 16144032.

Xu, S., Mckinley, J., Chen, J., Liang, J., & Laver, A. (2019). *Characterization of microstructure, tensile (23 °C to 850 °C) and charpy transition curves of a current tank car steel (TC128B) circumferential weld*. CMAT-2018-WF 34540443.

Rail failure root cause analysis on North American railway

Daniel Szablewski and Robert Caldwell
*National Research Council Canada Automotive and Surface Transportation,
Ottawa, ON, Canada*

ABSTRACT

NRC analyzed a broken rail that occurred on a North American Railway in the springtime. The break took place in 115RE standard rail placed in the high rail position of a 5 degree lubricated curve. Rail inspection focused on verifying mechanical, microstructural and chemistry measurements against current AREMA guidelines for these material properties. In addition, fractography was carried out on the fracture surfaces that led to the critical rail failure.

The rail defect took place in heavily curved track territory. To pinpoint the root cause(s) of this failure NRC performed a site inspection on a 30 mile length of track inspecting 29 curves, observing running surface conditions, and recording rail profiles and eddy current measurements to build an understanding of track conditions that might have contributed to the observed critical rail failure.

The paper describes the methodology undertaken in this investigation and details the outcomes at each investigative step, along with conclusions shedding light on the impact of metrics on the critical rail defect that led to the train derailment. Emphasis is placed on overall running track conditions in the investigated subdivision and on factors affecting the derailment. The paper concludes with a list of recommendations on metrics that need to be monitored with greater scrutiny to prevent future derailments. Improved rail material selection and/or more stringent grinding maintenance practices are also suggested to help prevent rail defect occurrences that might lead to critical track failures in the future.

1 INTRODUCTION

Rail failures are a common occurrence in freight operations (TSBC, 2019). Operating conditions on these lines are such that rail is constantly placed in either the wear regime, where plastic deformation combines with natural wear to result in relatively rapid running surface profile deterioration or in the rolling contact fatigue (RCF) cracking regime, where RCF cracks at the running surface initiate as head checks that then progress to subsurface networks of interconnected cracks. Either of these damage mechanisms is detrimental to the rail life-cycle (Aquib Anis, 2018).

Established solutions address both the removal of pre-existing RCF cracks by grinding the rail at prescribed tonnage intervals to shapes that conform to the predominant wheel profile, and friction management through the gauge face/corner (GFC) lubrication of the high rail and top of rail friction modification (TOR FM) of both rails, to slow down rail wear and RCF progression (Wang et al., 2017; Harmon and Lewis, 2016).

Appropriate combination of these two solutions has been observed to extend the rail life (Tuzik, 2019). The

ultimate goal of their application is to reach the 'magic wear rate' in the rail, wherein RCF crack growth occurs in step with rail wear, thereby minimizing the need for excessive grinding to target RCF removal and instead focus grinding effort on rail profile restoration (Magel et al., 2014).

At times preference is given to only one method (either lubrication or grinding), ignoring the need to apply both practices simultaneously to achieve the desired results. Lubrication is often applied intermittently (i.e. tanks are either not refilled regularly and/or dispensing bars are damaged preventing product from being delivered to the railhead) or grinding cycles are skipped (in favor of longer grinding intervals). Either of these actions creates an environment where rapid rail profile deterioration can take place and/or fatigue cracking in the rail can progress to form critical rail defects leading to derailments (Magel et al., 2016). The time interval to the appearance of critical rail defects usually varies according to track utilization (i.e. traffic frequency, tonnage accumulation), as well as which rail grade is being utilized, among others. Nevertheless, track integrity is decreased over time and often cannot be

reversed by 'catch-up' grinding if RCF progression into the rail is substantial (Sroba, 2003).

To make matters worse, the North American seasonal temperature swings that take place in Canada and northern USA act to increase stress on the rail network (CPR, 2019). In winter time cold temperatures increase rail tension that creates added tensile stress concentrations at the root tips of pre-existing RCF cracks. This condition can lead to rapid crack progression and rail critical failure.

National Research Council Canada (NRC) analyzed two such critical failures in ALGOMA C.C. 115RE 1985 standard rail. This paper addresses the findings from this investigation. In addition, a follow-up track inspection took place on the subdivision where the two critical failures occurred to observe 29 locations where potential future failures could develop. The methodology of selection criteria for these sites is described as well as the findings.

2 MATERIALS AND METHODS

The fractured rails taken out of service for defect root cause analysis had the following characteristics:

- ALGOMA C.C. 115RE 1985 standard rail
- High rail position in a 5.0 deg lubricated curve
- An approximate life-cycle tonnage of 130MGT
- Last grind cycles in 2005 (at 65MGT)
- Penny-sized fatigue cracks originating in high rails at the GFC in multiple locations (see Figure 1)

The laboratory metallurgical investigation consisted of the following analysis steps:

- Fractography of the broken rail section
- Analysis of the rail chemistry (per ASTM E1019-18, E1097-12, E1479-16)
- Analysis of the rail microcleanliness (per ASTM E3-11, E1245-03)
- Analysis of the rail mechanical properties:
 - Tensile testing (per ASTM E8-16a)
 - Fracture toughness testing (per ASTM E399-17) (PTS & MTS, 2021).

In addition, a site visit was arranged to a 30 mile stretch of track which included the locations of the rail fractures to measure and document the following high and low rail conditions:

- Dye penetrant of the rail running surface
- Crack depth measurements from the rail running surface using eddy current (EC)
- Transverse rail profile using Miniprof®
- Track gauge using track gauging bar
- Track superelevation using track gauging bar

The following section outlines the results of this analysis with contributing impacts on root causes.

3 RESULTS AND DISCUSSION

3.1 Fractography of Broken Rail Sections

The broken rail pieces were scanned using portable eddy current (EC) technology to assess the depth of damage. In most locations, the crack depth from the running surface ranged from 4-5 mm. However, at the fractured location the fatigue cracks significantly exceeded those depth limits.

Both rail sections displayed significant wear and RCF cracks with intermittent spalling at both the GFC and TOR running surfaces. The wear was accompanied by significant plastic metal flow at the GFC, as evident by the formation of a plastic lip at the lower GFC. There was also significant field corner plastic flow on the TOR. In addition, there were no visible grinding marks that would indicate recent grinding on the running surface.

The rail sections had two main fracture zones. However, only one of these led to a critical rail failure that resulted in the train derailment. Both defects initiated at the GFC as fatigue cracks (indicated by yellow dotted zones in Figure 1). These fatigue cracks grew relatively slowly over time (as evidenced by fatigue growth rings). Once the remaining rail cross-section could no longer bear the load, critical failure took place. This critical failure was instantaneous, as evidenced by the fast brittle fracture of the remaining rail cross-section.

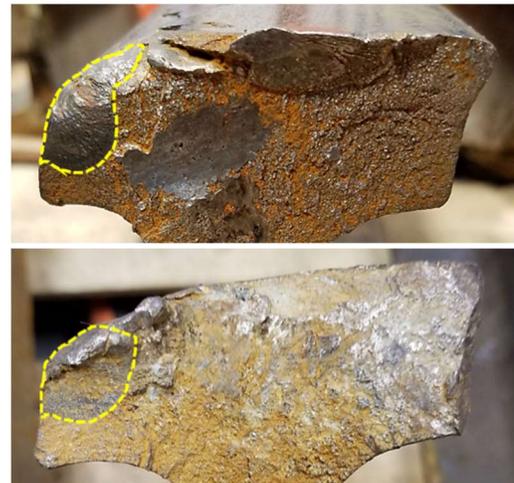


Figure 1. Examples of two (2) slow growth GFC fatigue defects (circled) that grew over time, which led to one of them causing the rail critical brittle failure.

During the track inspection portion of this work NRC took photographs and EC measurements of the scrapped rails which were still on site. This data was compared to the results acquired from the broken rail sections sent to NRC for analysis. The measured crack depths were in the same range (4-5 mm), with GFC damage being very similar to the one investigated at the NRC laboratory (see Figure 2).

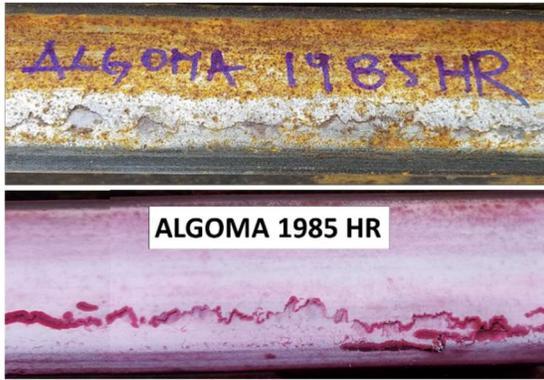


Figure 2. Rail running surface condition of ALGOMA C.C. 115RE 1985 standard rail taken out of service at approximately 130MGT indicating severe GFC damage.

3.2 Rail Chemistry Test Results

Two rail chemistry samples were analyzed (i.e. samples 1C and 2C) to contrast their compositions against AREMA's current recommended standard rail chemistry values (AREMA, 2021). All tested elements were within the current guidelines (see Table 1).

Microstructural analysis of the fractured rail revealed it was nearly entirely pearlitic, with relatively minor amounts of ferrite at the grain boundaries. That was the expected microstructure from standard chemistry rail produced 30 years ago.

Table 1. Rail chemistry test results.

Chemistry Analysis
(per ASTM E1019-18, ASTM E1097-12 (reapproved 2017, modified), and ASTM E1479-16)

Element	Chemical Analysis Weight percent		AREMA recommended values for standard chemistry rail		Product Analysis, Weight Percent Allowance Beyond Limits of Specified Chemical Analysis	
	Chemistry Sample 1C	Chemistry Sample 2C	Minimum	Maximum	Under Minimum	Over Maximum
	Carbon	0.73	0.73	0.74	0.86	0.04
Manganese	0.82	0.83	0.75	1.25	0.06	0.06
Phosphorus	0.015	0.015	-	0.020	-	0.008
Sulphur	0.023	0.023	-	0.020	-	0.008
Silicon	0.29	0.29	0.10	0.60	0.02	0.05
Chromium	0.01	0.01	-	0.30	-	-

3.3 Microcleanliness Test Results

Two rail microcleanliness samples were analyzed (i.e. samples 1MC and 2MC) to contrast their microstructures and reveal the presence of non-metallic inclusions. Both microstructures were found to be free of any voids and had a very low presence of oxides (see Table 2). The rail's microcleanliness values are in line with what is currently considered average in both intermediate and premium grades (see (Szablewski et al., 2011a; Szablewski et al., 2011b)).

Table 2. Rail microcleanliness test results.

Microcleanliness Test
(per ASTM E3-11 (reapproved 2017), and ASTM E1245-03 (reapproved 2016))

Element		Sample 1MC	Sample 2MC
Voids	Mean Volume %	0%	0%
	95% CI	N/A	N/A
	% RA	N/A	N/A
Oxides	Count/mm ²	0/mm ²	0/mm ²
	Mean Volume %	0.0054%	0.0034%
	95% CI	0.00058%	0.00031%
	% RA	11%	9.3%
	Count/mm ²	3/mm ²	2/mm ²
Mean Volume % (Voids + Oxides)		0.0054%	0.0034%
Maximum Volume % (Voids + Oxides)		0.0292%	0.0089%
Sulphides	Mean Volume %	0.040%	0.046%
	95% CI	0.0015%	0.0019%
	% RA	3.7%	4.2%
	Count/mm ²	30/mm ²	34/mm ²
	Maximum Volume %	0.075%	0.080%

3.4 Mechanical Test Results

Analysis of the rail's mechanical properties from two samples (i.e. sample 1T and 2T) indicated that yield strength (YS) and ultimate tensile strength (UTS) were below the current minimum AREMA recommended values (AREMA, 2021), see Table 3.

Table 3. Rail mechanical test results.

Tensile Test
(per ASTM E8-16a)

	Sample 1T	Sample 2T	minimum AREMA recommended values for standard chemistry
0.2% YS [ksi]	65.4	68.2	74.0
UTS [ksi]	129.7	129.4	142.5
Elong [%]	11	12	10

Fracture Toughness Test
(per ASTM E399-17)

	Sample 1F	Sample 2F
K _q [ksi(in) ^{1/2}]	42.0	37.6

Lower YS values indicated that given the same stress level at the running surface, the tested rail was more prone to plastic flow than rail that meets the current minimum AREMA recommended YS values. In addition, lower UTS for the tested rail indicates that the maximum stress the rail had experienced in service before critical failure was lower than that for rail that meets AREMA's current minimum recommended UTS values.

The fracture toughness values (for samples 1F and 2F) were on the upper end of what is currently considered average in both intermediate and premium rail grades (see Table 3 and (Szablewski et al., 2011a; Szablewski et al., 2011b)).

3.5 Track Inspection

A track inspection was carried out to better understand the rail running surface condition in a curve population representative of the entire subdivision. An analysis of the subdivision's track charts was carried out to select the representative curves. This work was done through the site selection criteria.

3.5.1 Curve Site Selection Criteria

During this portion of the work the main focus was to understand the distribution of curves on the subdivision and to develop a better knowledge of the potential problem size. The degree of track curvature and respective length vs. mile post (MP) for the entire subdivision were mapped out. This showed that all curves were in the 1-7 degree range (Figure 3), and that they were mostly below 1,200 feet in length (Figure 4).

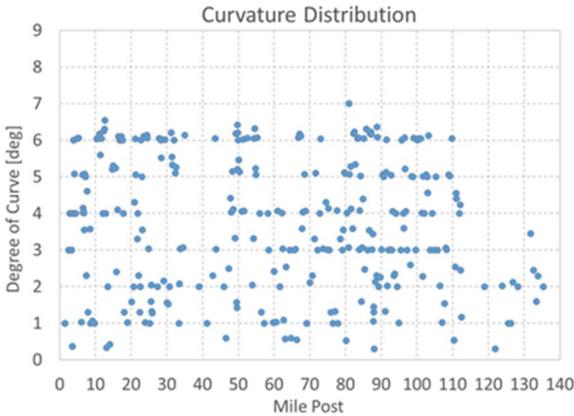


Figure 3. The degree of track curvature vs. mile post on the subdivision.

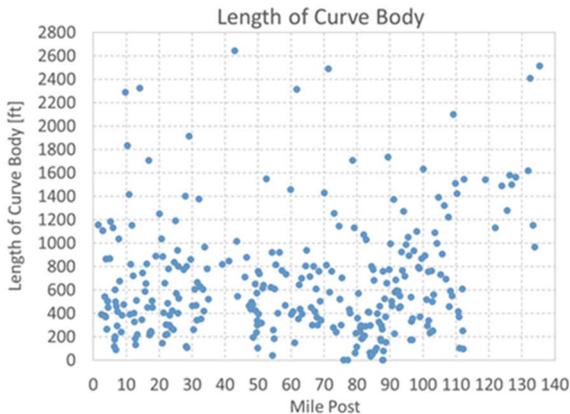


Figure 4. The length of each curve body vs. mile post on the subdivision.

NRC also mapped the posted train speed for all the curve locations (see Figure 5). Based on that data the maximum permissible train speed was 40 mph. However, this speed was the timetable speed of the train. Rail profile measurements would indicate whether trains operated predominantly above or below the balanced speed. The differences between these two conditions are addressed in a later section of this work.

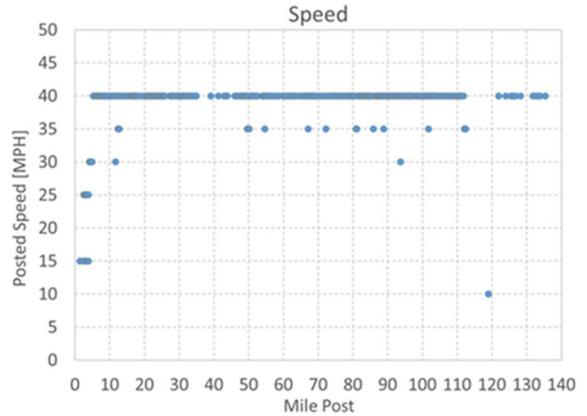


Figure 5. The posted train speed for curves vs. mile post for the subdivision.

The condition of the rail running surface is affected by cant deficiency and whether the track superelevation is adequate for the operating conditions. This has an influence on the amount of loading on the high rail where the rail failure took place. Understanding this relationship across the whole subdivision helped to determine the problem size.

To understand the track geometry condition on the line, NRC plotted the 'required' (based on speed shown in Figure 5) and 'actual' track superelevations vs. MP (see Figure 6).

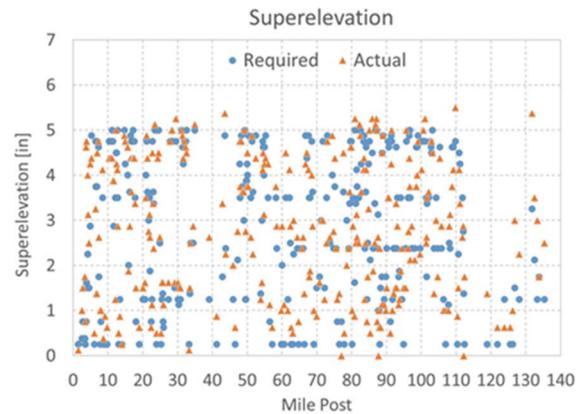


Figure 6. The 'required' and 'actual' track superelevation as a function of MP on the subdivision.

Both of these plots of superelevations vs. MP were essentially point clouds. However, mapping superelevation as a function of track curvature provided a much clearer understanding of its distribution in the subdivision (see Figure 7).

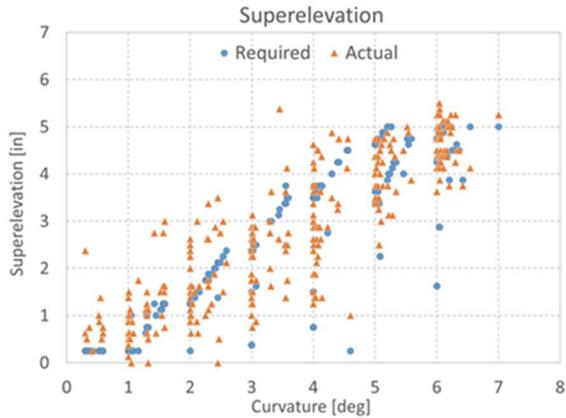


Figure 7. The 'required' and 'actual' track superelevation as a function of track curvature on the subdivision.

Seeing increasing superelevation with increasing curvature was a positive sign that superelevation was being matched to curvature to ensure balanced loading on the high and low rails at all times of train operation on the subdivision. However, as Figure 7 indicates, there was an observable difference between the designed (i.e. 'required') track superelevation and that present in track (i.e. 'actual'). This was an indication that the trend shown in Figure 7 was encouraging, but there were still many curves with incorrect elevation for the typical train speed.

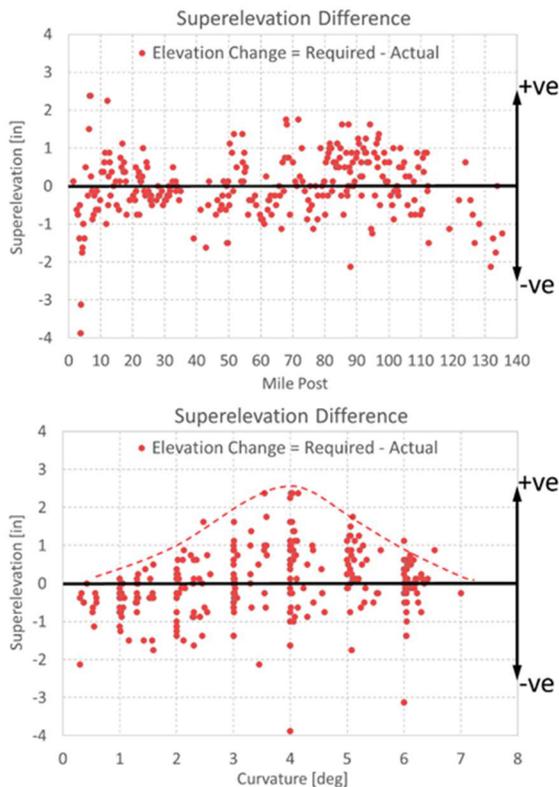


Figure 8. The elevation change (i.e. difference between required and actual superelevations) vs. MP and curvature on the subdivision.

Plotting the cant deficiency (i.e. the difference between the 'required' and 'actual' superelevations) across both MP and curvature (see Figure 8) revealed two things:

1. That superelevation difference varied up to ± 2 inches across the entire subdivision.
2. That there appeared to be a normal distribution to the superelevation difference as a function of curvature for positive differences (i.e. when the high rail was overloaded due to insufficient curve elevation).

The physical significance of the positive and negative superelevation differences are best shown schematically.

At constant train speed in a constant-radius curve the overbalanced condition occurs when the loading on the low rail is elevated (or excessive). This results in excessive wear and plastic flow on the low rail. On the other hand, an underbalanced condition occurs when the loading on the high rail is elevated (or excessive). This results in excessive wear and plastic flow on the high rail. Neither of these conditions is desirable, as it shortens the overall life cycle of the track, as well as increases track maintenance costs.

The ideal condition is to have equal loading on both rails, as this distributes the wear and plastic flow evenly on both rails, thereby maximizing track longevity and reducing overall maintenance costs.

The differences between these three conditions are shown schematically in Figure 9.

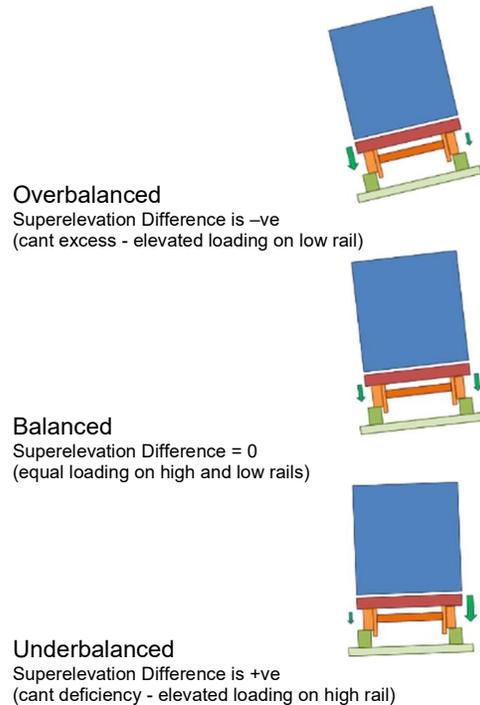


Figure 9. Schematic representation of the overbalanced, balanced, and underbalanced rail loading conditions due to superelevation differences in a curve. High rail is the right hand rail in each diagram.

Therefore a cant deficiency increases loading on the high rail, and this loading increases as the positive difference increases. Curves in the 3-5 degree range are the ones that had the greatest cant deficiency, and therefore suffered the greatest elevated loading on their high rails (see Figure 8).

NRC mapped out the cant deficiency as a function of MP to understand where the greatest density of cant-deficient curves were located on the subdivision. Based on this a 30 mile zone of track was selected for an on-site inspection.

A couple of indicators led to the selection of this 30 mile stretch of track:

1. It had the highest density of curves with cant deficiency (10+ curves where deficiency was greater than 1 inch). Assuming constant train speed, these were the curves where loading on the high rail would be the greatest (as compared to other curves on this subdivision), and therefore the risk of high rail fracture and subsequent train derailment would be most detrimental (assuming the surface damage on these rails was substantial).
2. The curvatures where these occur was in the 1-6 degree range, which was representative of the curve population on the subdivision.

Due to these two conditions the selected 30 mile section of track was deemed to represent the track zone of interest. A total of 29 curves were inspected on the subdivision during a two day visit.

The EC measurements for the RCF cracking at each location were graphed for both high and low rails as a function of track curvature (see two representative locations in Figure 10). These results indicate that maximum low rail cracking was approximately 3.5 mm in depth, whereas in the high rail its maximum depth was approximately 6.1 mm. In all locations, the crack depth on the high rail was greater than that on the low rail.

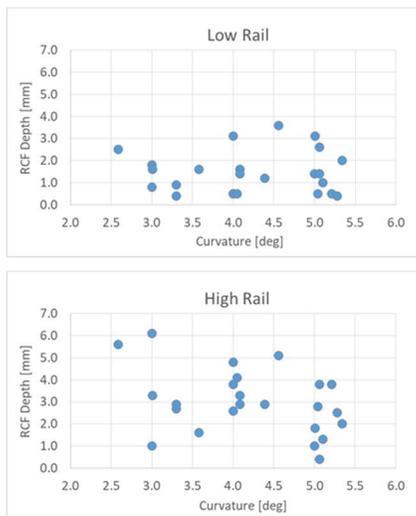
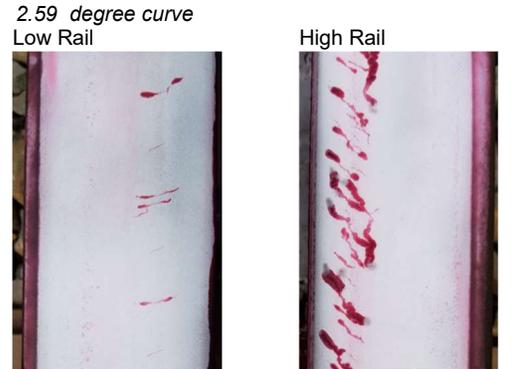
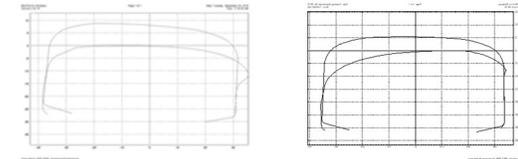


Figure 10. Eddy current RCF crack depth assessment in low and high rails as a function of track curvature in the 30 mile stretch of track.

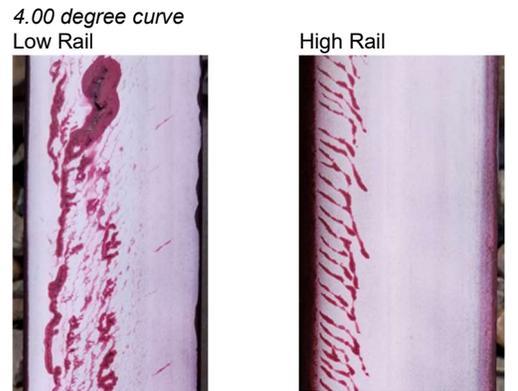
Two representative locations were selected to show typical low and high rail damage observed in curves below 3 degrees of curvature, and in curves of 4 degree curvature and above. Results of these are shown in Figure 11.



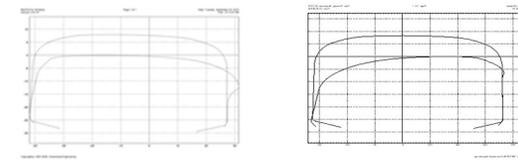
Existing Superelevation (track evaluation car): 2.125 in
 RCF Cracks: 2.5mm deep RCF Cracks: 5.6mm deep
 Track Gauge: 1440.8mm
 Grade: -0.6% (loaded train downhill)
 Rail Type: Algoma 1985
 Approximate Tonnage: 130MGT



W1: 8.28mm W1: 5.88mm
 W2: -0.56mm W2: 0.05mm
 W3: 4.89mm W3: 8.27mm



Existing Superelevation (track evaluation car): 3.625 in
 RCF Cracks: 3.1mm deep RCF Cracks: 4.8mm deep
 Gauge: 1444.6mm
 Grade: -1.2% (loaded train downhill)
 Rail Type: Algoma 1985
 Approximate Tonnage: 130MGT



W1: 8.06mm W1: 8.77mm
 W2: -0.94mm W2: 1.33mm
 W3: 6.08mm W3: 11.82mm

Figure 11. Select track inspection results for a 2.59 and 4.00 degree curvatures.

4 CONCLUSIONS

Two gauge corner fatigue cracks were found in the rail sections delivered to NRC for analysis. Only one of these cracks caused the rail break that led to a derailment. Both defects initiated at approximately the same geometric corner positions on the rail running surface (see Figure 1). They propagated to about the same cross-sectional size and then fractured critically causing multiple rail breaks. It is not possible to say which of these defects broke first. Once the rail broke at one location there would have been enough rail bending to cause a critical failure at the second location, hence the two breaks. However, only one of these defects would have caused the initial critical failure leading to derailment.

It is important to point out that both of the gauge corner fatigue cracks formed and grew in the high rail section independently, but through the same growth mechanism (i.e. metal fatigue due to load reversal that occurred during train travel over these defects). These defects formed over time as a result of inadequate rail grinding. As these defects were relatively deep compared to remaining rail cross-section the suggested course of action was rail replacement. It was also advised to replace any rails that displayed similar gauge corner fatigue defects in track.

Investigation of rail chemistry, microcleanliness and mechanical properties indicated that chemistry and microcleanliness for this 1985 rail grade were within the current AREMA recommended guidelines for standard rail type. Analysis of the mechanical properties indicated that yield strength (YS) and ultimate tensile strength (UTS) values were lower than recommended by the current AREMA guidelines for standard rail grade. These lower values mean the rails were more prone to plastic flow at the gauge corner (leading to earlier onset of fatigue cracking at that location) and critical rail failure (i.e. metal brittle failure at a lower stress) as compared to rails that meet all current AREMA guidelines for standard rail.

The suggested course of action going forward was to focus on more frequent defect inspections of the 1980s rails and prioritizing their replacement in the near future.

The main focus of the track inspection in the subdivision where the derailment took place was to investigate under-elevated curves since these locations would suffer premature high rail damage. Since only freight trains run on the investigated subdivision, changing the curve elevation to be balanced for the typical speed of loaded trains (which will cause more damage than empty trains) should be the main focus going forward.

Knowing the rail balanced loading condition is speed dependant, the first step should be an accurate mapping of the train speed profile in the subdivision. The next step would be to look at train lengths vs. curve and tangent lengths as the speed of loaded trains may change while the train is spanning two or three consecutive curves of different curvatures (i.e. affecting the balanced loading on each curve). This makes the selection of optimal superelevation difficult, since all above mentioned

factors need to be taken into account simultaneously to yield the optimum solution.

To summarize, the following actions were suggested to be taken in this specified sequence:

- Review of subdivision track speed, grade for multiple train passes
- Review of track geometry (curve & tangent lengths) vs. train lengths in the track subdivision
- Review and adjustment of track elevation to fit the two conditions above

Besides adjusting track elevation, the focus should be on replacement of rails based on cracked depth vs. vertical or gauge face wear remaining on the high rail. Changing the elevation in curves with deep cracks and not much rail life remaining may not be as cost effective as just replacing the rail. If the cracks could be removed through rail grinding and the rail still had some wear life remaining, this might make sense. This becomes an economic consideration, and effort should be placed on assigning resources adequately with that in mind.

Concerning rail replacement, premium or head hardened (HH) rails yield longer rail life as compared to standard or intermediate strength rail steels (nearly twice as much in a 5 degree curve location under heavy loading). This is due primarily to their resistance to plastic flow, and if the microstructural grain size is refined, to improved wear resistance. Use of the HH rails should be prioritized for higher degree curves, as these are most prone to premature wear problems (especially in an underbalanced curve condition).

In addition, the application of GF lubrication should also be improved to prevent the high rail lip formation due to plastic flow. Consistent GF lubrication will also improve rail wear performance, but will require periodic light grinding at roughly 60° to 70° on the gauge corner to prevent the formation of deep seated shells. Lack of lubrication causes the transverse rail profile to deteriorate more rapidly than it would with lubrication, which means it needs to be ground more frequently to maintain profile. Preventive grinding and good lubrication results in longer rail life than no lubrication and/or grinding.

TOR FM helps in extending rail life even further but the payback period on low tonnage lines makes its use unattractive due to cost of initial investment. For low tonnage lines this is true from a rail wear perspective. However, use of TOR FM might be useful from an RCF prevention point of view and in terms of increasing grinding interval due to reduced transverse profile deterioration. In either case, both arguments circle back to an economic consideration and whether the initial investment is offset by the payback period.

One of the drawbacks of this analysis was the lack of wheel profile information. During the visit we did not have access to vehicles to acquire wheel profiles (which could then be matched with the rail profiles to perform a W/R contact analysis and to establish its impact on RCF damage). There would be value in acquiring such

information in the future to make the site analysis more complete.

5 ACKNOWLEDGEMENT

The authors would like to express appreciation for the support of the railroad that supplied the rails for analysis and subsequent access to its track for the site investigation.

Special thanks to Alexandre Woelfle (NRC-AST) for contribution to information provided in Figure 9 as well as discussions concerning superelevation vs. rail loading conditions.

6 REFERENCES

Aquib Anis M., Srivastava J.P., Duhan N.R., Sarkar P.K. (2018), *Rolling contact fatigue and wear in rail steels: An overview*, IOP Conference Series: Materials Science and Engineering 377 012098.

AREMA, American Railway Engineering and Maintenance-of-Way Association (2021), Chapter 4 – Rail.

CPR, Canadian Pacific Railroad, *White Paper: Railroading in the Canadian Winter* (2019), <https://www.cpr.ca/en/about-cp-site/Documents/CP-2018-19-WhitePaper.pdf>

Harmon M., Lewis R., (2016), *Review of top of rail friction modifier tribology*, Tribology Materials, Surface and Interfaces, 10 (3).

Magel E., Kalousek J., Sroba P. (2014), *Chasing the magic wear rate*, Proceedings of the Second International Conference on Railway Technology: Research, Development and Maintenance.

Magel E., Mutton P., Ekberg A., Kapoor A. (2016), *Rolling contact fatigue, wear and broken rail derailments*, Wear, <http://dx.doi.org/10.1016/j.wear.2016.06.009>.

PTS & MTS, Physical Testing Standards and Mechanical Testing Standards (2021), ASTM, www.astm.org

Sroba P. (2003), *Rail Grinding Best Practice For Committee 4, Sub-Committee 9*, AREMA

Szablewski D., Kalay S., LoPresti J. (2011a), *Preliminary Evaluation of Premium Rail Steels for Heavy Haul Operations*, AAR Technology Digest, TD-11-031

Szablewski D., Kalay S., LoPresti J. (2011b), *Preliminary Evaluation of Intermediate Hardness Rails for Heavy Haul Operations*, AAR Technology Digest, TD-11-032

TSBC, Transportation Safety Board of Canada, Rail Transportation Occurrences in 2019, <https://www.bst-tsb.gc.ca/eng/stats/rail/2019/sser-ssro-2019.html>.

Tuzik B. (2019), *Getting a Grip on Friction Management*, RT&S, Track Maintenance.

Wang W.J., Lewis R., Evans M.D., Liu Q.Y. (2017), *Influence of Different Application of Lubricants on Wear and Pre-existing Rolling Contact Fatigue Cracks of Rail Materials*, Tribology Letters v. 65, article 58.

Condition based rail surface management using EMFI technology

Dominika Juhaszova-Dubinsky, John Furlong, Adam Milligan and Kyle R Mulligan
Canadian Pacific, Calgary, AB, Canada

ABSTRACT

North American railways use rail wear and operating tonnage data to plan for seasonal rail replacement programs. In addition to using models to forecast and plan for rail maintenance, manual visual and Rail Flaw Detection (RFD) inspections based both on regulated and internal inspection intervals are also used to initiate rail replacements and maintenance. A common RFD method used to detect subsurface metallurgical flaws and fatigue cracks within the rail head uses a series of contact based ultrasonic probes. Tonnage, wear, and seasonal effects result in Rolling Contact Fatigue (RCF) which creates head checking and/or flaking leading to unsmooth rail surface conditions. Due to the contact nature of ultrasonic probes, these conditions may result in masking subsurface defects rendering them more difficult to detect and in more severe cases leading to complete probe signal losses. Unsmooth or rough surface conditions are corrected by railways through rail grinding (more commonly in North America) or rail milling (more commonly in Europe). Similar modeling which is used to initiate rail replacement programs are also used for initiating rail surface conditioning (i.e. tonnage and wear). This paper presents the use of a non-contact rail inspection probe based on Electro-Magnetic Field Inspection (EMFI) technology which is demonstrated to quantify rail surface conditions. The authors propose a technique for initiating condition based rail grinding and rail surface management by aggregating EMFI data into back office railway applications and processes as opposed to relying on traditional railway maintenance models and intervals.

Keywords: Electro-Magnetic Field Inspection (EMFI), Rail Surface Conditions, Rail Surface Management, Rolling Contact Fatigue (RCF), Rail Maintenance, Data Analytics

1 INTRODUCTION

1.1 North American Track Inspection Techniques

North American railways use rail wear and historical tonnage data for trains operating over a given track segment to plan for seasonal rail replacement programs. In addition to using models to forecast and plan for rail maintenance, manual visual and Rail Flaw Detection (RFD) inspections based both on regulated and internal inspection intervals are also used to initiate rail replacements and maintenance.

Track condition must be measured and verified based on internal standards and regulations defined by both Transport Canada (TC) (Transport Canada, 2018) and the Federal Railroad Administration (FRA) (Federal Railroad Administration, 2008) in Canada and the United States respectively. These rules indicate the track measurement: criteria (ex. gauge, alignment, surface), limits, and inspection frequencies. These inspection frequencies are periodic based on the class of track and train tonnages operating across the designated track segments (i.e. subdivisions). To measure track

geometry, the regulations enable the use of technology in the form of Light and Heavy Geometry Inspection Vehicles (LGIV/HGIV). Ensuring a compliant track geometry through track inspection vehicles is critical in avoiding undesirable train dynamics and protecting rail safety (Bonaventura, et al., 2005).

1.2 Ultrasonic Rail Flaw Detection Inspections

In addition to rail wear, railways use Rail Flaw Detection (RFD) vehicles to detect subsurface metallurgical flaws and fatigue cracks within the rail head. These defects can result in a rail break causing a potential derailment. In rail network areas, known as non-signaled territories, where signaling system infrastructure such as Centralized Traffic Control (CTC) does not exist, RFD inspections are the primary derailment risk mitigation for broken rails in addition to periodic visual track inspections to verify complete rail breaks.

Virtually all RFD inspection vehicles, shown in Figure 1, employ ultrasonic probes to inspect the rail for internal defects (Kerchof, 2021). These probes inject and

capture sound waves to and from the rail head using transducers oriented in fixed lateral positions (0°, 45°, 70°). This is such that specific areas of the rail head can be inspected namely the: field side, middle, and gauge

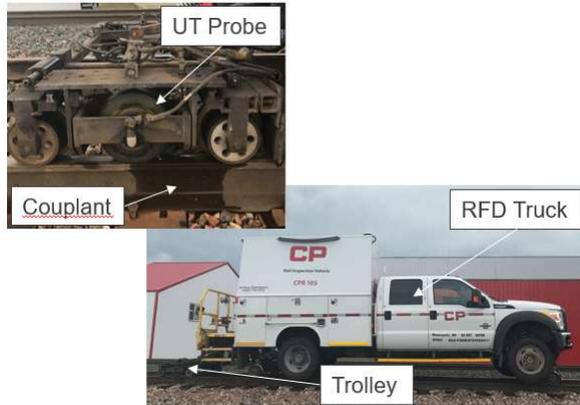


Figure 1. Ultrasonic Rail Flaw Detection (RFD) truck with attached trolley identifying one of the ultrasonic (UT) probes and the couplant residual on the rail.

side of the rail and also particular defect types such as transverse and vertical defects can be found.

Ultrasonic probes require a couplant such as water in warmer weather and washer fluid in cold weather. The purpose of the couplant is to ensure optimal transmission and sensing of the sound waves to and from the rail head through an air-free medium. A proper interface between the ultrasonic probes and the rail head is therefore important.

1.3 Impacts of RCF on Ultrasonic Inspections

Rolling Contact Fatigue (RCF) is a material defect mechanism which is caused by repeated cyclical loading at the contact interface between two components. In rail, RCF initiates at the interface between the wheel and rail where highly localized cyclic stresses over a small volume support the axle loading which enables relative motion of the freight cars or locomotives (Kang, 2013). RCF therefore affects both the wheel and rail contact surfaces.

For the rail, RCF develops on the rail running surface where the wheel and rail contact. This appears in the form of cracks, spalls, and shells which affects the interface between ultrasonic probes and the rail. Varying degrees of RCF severity can be observed visually on the rail head with examples shown in Figure 2 from a North American Class I railway.

RCF defects can cause injected ultrasonic sound waves from ultrasonic probes to deflect without travelling into the rail head. In cases of severe RCF, ultrasonic systems will raise a Shelling, Spalling, and Corrugation (SSC) defect alert indicating an ultrasonic probe signal

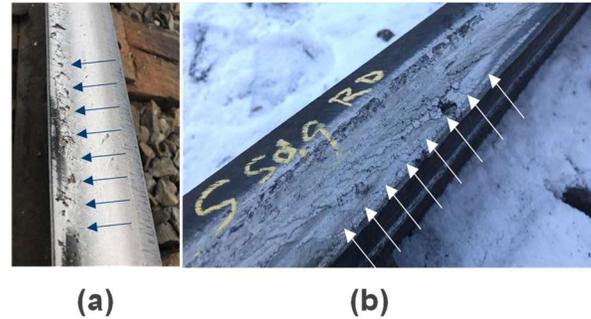


Figure 2. Example of Rolling Contact Fatigue (RCF) on the rail running surface of a North American Class I railway showing increasing severities from medium RCF (a) to severe RCF (b).

loss. When an SSC defect occurs, RFD technicians will need to hand test the affected area. If a hand test is unsuccessful, the rail is marked, condemned, and replaced per railway processes. Due to the contact nature of ultrasonic probes, these conditions may result in masking internal rail defects rendering them more difficult to detect and in more severe cases leading to complete probe signal losses. However, SSC defects generally only apply to the 0° probe and therefore some defects sensitive to the 45° and 70° probes may not be detected.

1.4 Rail Maintenance for RCF

Unsmooth or rough surface conditions (i.e. RCF) are corrected by railways through rail grinding (more commonly in North America) or rail milling (more commonly in Europe). Examples of rail grinding and milling are shown in Figure 3 (a) and (b) respectively. Similar modeling used to initiate rail replacement programs based on tonnage and wear is used to initiate rail surface conditioning through corrective grinding or milling. However, tonnage and wear do not entirely predict areas where RCF may develop and in some cases RCF may develop faster than predicted based on a number of other variables (ex. wheel/rail hardness, relay versus new rail) (Pyke, 2015). Therefore opportunities to utilize a condition based monitoring approach are introduced.



Figure 3. Photo of a North American rail grinder in operation (a) and a photo of a European rail milling head (b). Both (a) and (b) can be used to condition the rail surface (Abbott, 2007) (Tuzik, 2019).

1.5 Inspection Techniques for RCF Quantification

Due to the importance of a robust probe rail interface, technologies to quantify the rail surface have been tested within the domestic and international rail industries. In North America, two dominant technologies are being prominently tested, namely: Eddy Current and Electromagnetic Field Inspection (EMFI).

Both Eddy Current and EMFI methods use an array of probes which electromagnetically excite the rail head, creating an electromagnetic field which forms Eddy Currents on the rail surface. The probes then capture the response of the formed Eddy Currents due to the field. In the presence of SSC and/or cracking along the running surface or slightly below the surface, the propagation path of the eddy currents changes. The change in response is then compared to a baseline (i.e. smooth rail). The response can be overlaid onto the baseline and by leveraging data analytics, crack depth or the area of defect surfaces can be calculated. An example of rail surface measurements using EMFI on a North American Class I railway are shown in Figure 4.



Figure 4. Rail surface quantification using Electromagnetic Field Inspection (EMFI) probes from the rail running surface of a North American Class I railway.

Similar to Ultrasonic probes, the Eddy Current/EMFI arrays can be oriented at 0° , 45° , 70° to assess different angles on the rail running surface. In fact railways are focused on incorporating both UT and Eddy Current/EMFI into track inspection systems (Heckel, et al., 2009).

1.6 Paper Objective

Based on evidence that RCF can mask or prohibit an optimal inspection using ultrasonic probes, it is critical to develop a condition based approach in assessing RCF throughout the rail network. Therefore, this paper presents the use of a non-contact rail inspection probe

based on Electro-Magnetic Field Inspection (EMFI) technology which is demonstrated to quantify rail surface conditions. Using the data acquired from testing on a North American Class I railway. The paper also proposes a condition based monitoring technique for initiating rail grinding and rail surface management. This is performed by aggregating EMFI data into back office railway applications and processes as opposed to relying on traditional railway maintenance models and intervals. This work hopes to reduce risks associated with broken rails by enabling techniques within the industry to ensure an appropriate probe to rail interface is maintained for ultrasonic RFD testing.

2 MATERIALS AND METHODS

The methodology in this paper is described in two parts. The first part assesses laboratory samples from field recovered track service failures (i.e. broken rails) to determine the potential that detection of internal rail defects using UT may be hindered in the presence of heavy rail surface RCF. The second part compares the in-service UT RFD defect detection rates pre and post rail grinding in network areas of a North American railway that have been confirmed with heavy rail surface RCF using EMFI.

2.1 Assessment of Laboratory Rail Samples

North American Class I railways record occurrence data for all in-service track failures (i.e. broken rails). Failure samples are generally recovered for further analysis to determine the root cause of the failure. For this paper, four (4) in-service track failure rail samples have been recovered and are presented.

Assessment of the recovered failures is performed at a North American Class I metallurgical laboratory. Each of the samples are evaluated to determine the root cause of the failure, namely the presence of internal defects.

2.1.1 Metallurgical Analysis Process

Following the evaluation for internal defects, the rail surfaces of each sample are investigated to quantify the level of RCF on the rail surface. This is done initially by visual inspection using Magnetic Particle Inspection (MPI). MPI is a non-destructive technique which detects surface and subsurface cracks in rail samples. By magnetizing the sample and applying magnetic filings onto the rail surface, crack orientations can be highlighted (shown in Figure 5). Ultraviolet light reveals the density and orientation of RCF. MPI inspection confirmed heavy RCF in all four (4) samples.

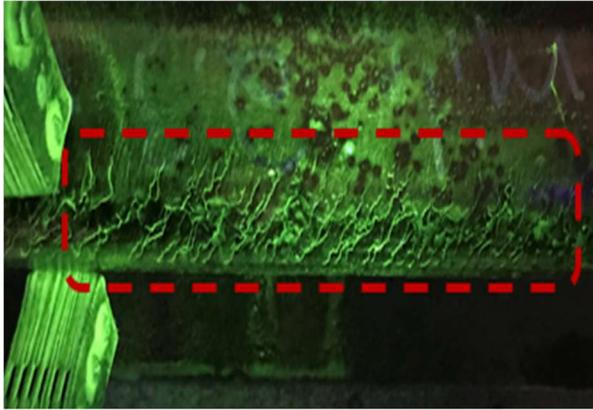


Figure 5. Magnetic Particle Inspection (MPI) reveals the surface and near surface cracks. Magnetic field is induced in the rail, followed by the application of Magnaflux spray containing fluorescent magnetic particles to the rail surface. The particles gather at magnetic flux leakage at the location of discontinuity caused by RCF. Black light illuminates cracks on the rail.

After a visual inspection by MPI, the rails are then sectioned. The cutting direction is chosen depending on the direction of crack propagation as shown in Figure 6. By progressively grinding and polishing the cross-section of each sectioned sample, the crack morphology is revealed (shown in Figure 6). The depth of the exposed cracks (shown in Figure 7) is measured using a microscope.

Detailed fracture is an example of a defect caused by RCF propagation from shelling or spalling on the gauge corner. Impurities such as grease, water or sand are pushed into the microcracks under cycling loading from passing trains. Water has low viscosity and high surface tension and is drawn into cracks by capillary action. Subsequent rapid crack growth occurs under loaded conditions as the crack tip becomes hydraulically pressurized which create large tensile stresses at the tip (Singh, 2011).

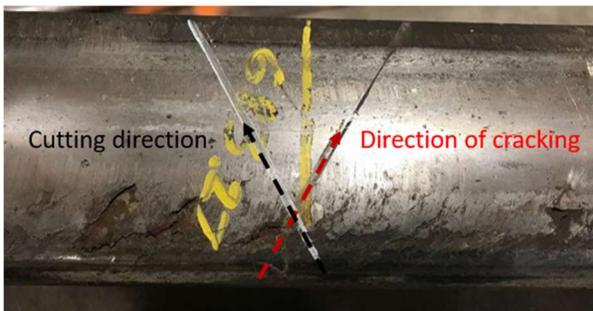


Figure 6. Examination of rail surface and sub surface. The cutting direction is based on the orientation of the crack, to ensure the maximum crack depth is exposed.



Figure 7. Microscope photography of a polished rail sample shows crack propagation from the checking to the rail head

The samples are further evaluated using the chemical etching technique. Application of a 3% nital solution is used to reveal the grain structure of the base rail material and shows any internal defects in grain boundaries. If the inspection determines the cause of the internal defect is an inclusion, the Scanning Electron Microscope (SEM) is used to determine the type of inclusions in rail (Figure 7). An example of an internal rail defect in old rail due to inclusions or impurities in steel during the manufacturing process is a Vertical Split Head as shown in Figure 11.

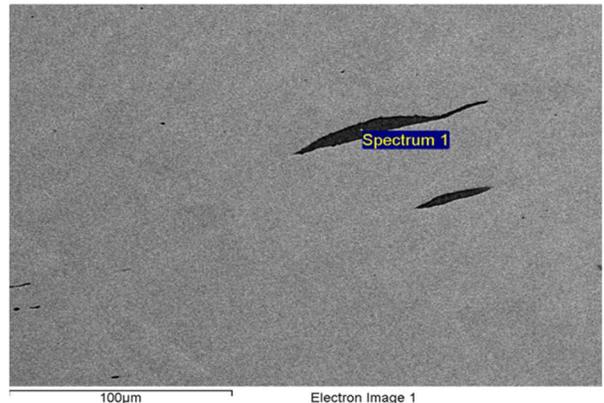


Figure 8. Scanning Electron Microscope inspection reveals inclusions in the rail steel.

Finally, once the metallurgical assessment verifies internal defects in each sample, the last RFD inspection histories are also recovered and compared.

2.2 Assessment of In-Service Defect Detection Rates

Inspections are performed across a 440 mile section of a North American Class I railway using EMFI sensors mounted to a trolley which is towed by a High Rail Truck (shown in Figure 9). The purpose of the EMFI inspections is to quantify levels of RCF on the rail surface. The running surface is visually inspected every 5 miles to verify EMFI measurements. Once validations have been completed, RFD testing is performed along the entire route. The number of defects is recorded electronically and accessible following the run. Any rail defects which are found are addressed by existing railway processes whereby the rails are marked for removal and speed restrictions are applied until the rails are removed from track.



Figure 9. Trolley carrying the Electromagnetic Field Inspection (EMFI) sensors (one per rail) used to perform the rail surface quantification to determine levels of Rolling Contact Fatigue (RCF).

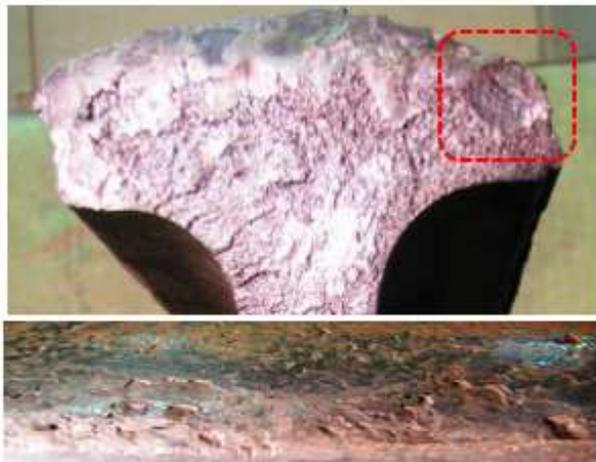


Figure 10. Photo of the detailed fracture found in Sample #1 (shown in the dashed box) in addition to the Heavy RCF on the rail surface identified visually and using Magnetic Particle Inspection. The last RFD inspection occurred 49 days prior to the in-service track failure (i.e. broken rail).

Once the EMFI and RFD inspections are completed, rail grinders are deployed to reduce the levels of RCF (ideally to zero). To verify that sufficient rail grinding has been achieved, a second EMFI inspection is performed. Once validated, after the EMFI inspection a second RFD inspection is performed. Again the number of defects are recorded electronically and are made accessible following the run. Any rail defects which are found are addressed by existing railway processes whereby rails are marked for removal and speed restrictions are applied until the rails are removed from track. Following the procedure used above, the number of defects pre and post rail grinding are plotted and compared.

3 RESULTS AND DISCUSSION

The results of the assessment of rail samples from in-service failures and in-service defect detection rates pre and post grind of areas with heavy RCF of a North American Class I railway are described in the following sections.

3.1 Assessment Results of Laboratory Rail Samples

As described in Section 2.1, four (4) in-service track failure (i.e. broken rails) samples (#) have been recovered for analysis. The samples are characterized in Table 1.

Table 1. Description of the four (4) in-service track failure (i.e. broken rails) samples (#) recovered for analysis including the failure and last RFD inspection dates and a summary of the laboratory findings (i.e. internal defect descriptions and visual/magnetic particle inspection of the rail surfaces).

#	Failure Date (mm/dd/yy)	Last RFD Inspection (mm/dd/yy)	Internal Defect	Heavy RCF
1	12/09/16	10/21/16	Detailed Fracture	Yes
2	10/03/17	09/25/17	Vertical Split Head	Yes
3	12/09/19	11/26/19	Detailed Fracture	Yes
4	02/06/20	11/26/19	Detailed Fracture	Yes

In the Table, all four (4) of the in-service rail samples have been verified to have internal defects leading to the critical failures. The laboratory inspections by visual and magnetic particle have shown an average crack depth greater than 1 mm, which is considered to be heavy RCF, in each sample. The last RFD is also shown for each sample with #1 to #4 at 49, 8, 13, and 67 days respectively prior to the failure date. The average number of days between the failure date and the last RFD test across all samples is 35 with the longest being 67 days. An example of the detailed fracture found in sample 1 and the rail surface is shown in Figure 10. All other samples have been found to exhibit similar levels of RCF as shown in sample #1. Vertical split head is observed along the 5 ft. section of sample #2 removed from the track, shown in Figure 11. A review of the last rail grinding event data for each network section where

the failures occurred suggests the last grinding interval exceeds 2 years for a switch grinder and 1 year for a production grinder in all areas.

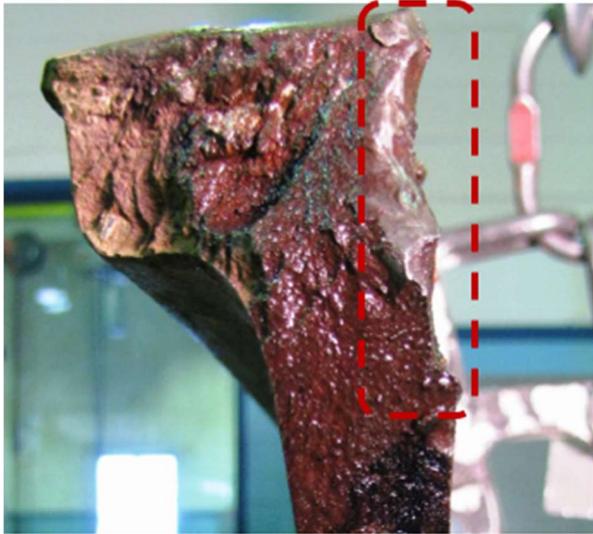


Figure 11. Photo of vertical split head observed along the 5 ft. section of sample #2.

3.2 Assessment Results of In-Service Defect Detection Rates

As described in Section 2.2, an initial inspection using EMFI on a 440 miles section of track across 4 subdivisions of a North American Class I railway is performed to verify heavy levels of RCF. The results of the EMFI inspection indicate crack depths ranging from 0.5 mm to 5 mm with an average overall crack depth of 1 mm over a mile long rail section. This depth, as suggested by the laboratory metallurgical analysis, is interpreted as heavy RCF. An example of the output which is captured for the 440 mile section of track is shown in Figure 13.

The EMFI inspection system provides the operator with a visual output as the inspection is being performed. However, due to the large distance being inspected, the data are stored and plotted following the inspection. Hence, Figure 13 shows the measured crack depth over the testing mileage. Mileage is correlated to crack depth based on odometer and GPS readings. Black lines in Figure 13 represent landmarks such as mile posts, crossings, switches, etc. An example of the RCF and what the operator sees is shown by the red areas in Figure 4.

Average and consistent cracking of greater than 1 mm without any areas showing below 0.5 mm across the entire EMFI run, in addition to regular visual confirmation of areas across the run and laboratory metallurgical assessment, suggest a level of heavy RCF can be confirmed within the test section. Once confirmation is obtained, RFD testing is performed across the test section. The number of defects are gathered and shown

in Figure 12 (Pre-Grind) in 4 separate lines (each representing 1 subdivision of the 440 mile test section).

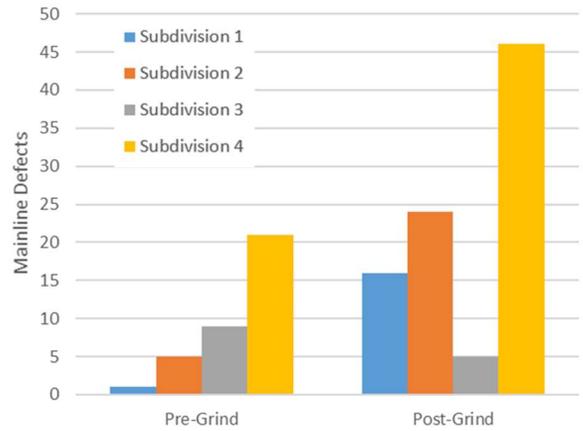


Figure 12. Pre and Post-Grind RFD detection results using ultrasonic testing from the 440 mile test section.

The results of the second EMFI inspection which is performed for the same section of track shown in Figure 13 except the track has now been ground using rail grinding machines are shown in Figure 14. The crack depths of the second inspection post grinding range from 0.1 mm to 5 mm with an average overall crack depth of 0.5 mm which is below the metallurgical quantification of heavy RCF. The frequency of crack depths exceeding 5 mm is also noticeably lower than in Figure 13. Based on the post-grind EMFI data showing that crack depths in many areas approach 0 mm and a lower frequency of cracks exceeding 5 mm overall and previous metallurgical laboratory inputs, it is determined that the rail grinding sufficiently removed enough RCF to warrant a second RFD inspection.

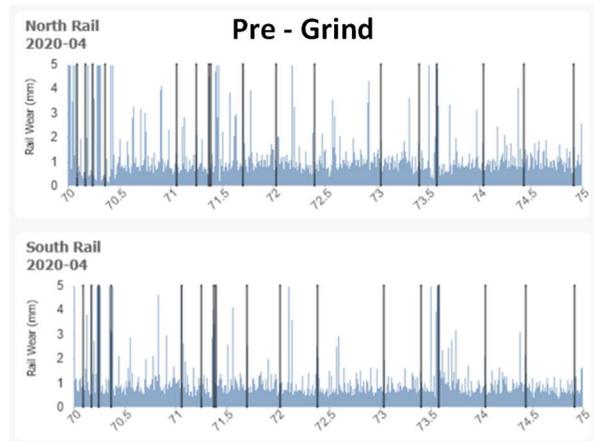


Figure 13. North and South rail crack depth measured by EMFI between Mile Post (MP) 70 and 75 of the 440 mile test section of a North American Class I railway (Pre-Grind).

The results of the second RFD inspection across the 440 mile test section post-grind are presented in Figure 12 next to the pre-grind numbers. The results are also broken down by dividing the 440 miles test section into 4, each representing a respective subdivision of the test section. From the figure, an average of 200% more defects across each section of the 440 test sections is shown. More defects are found on subdivisions with the longest grinding interval. The gray line represents the subdivision with rail replacement between the first and the second RFD inspections.

4 CONCLUSIONS AND DISCUSSION

Rail surface defects can inhibit ultrasound penetration into the rail head and therefore can prevent detection of internal defects. This has been demonstrated in this paper which aims to validate previous mention of this found in literature (Kerchof, 2021), (Transportation Safety Board of Canada, 2019). Based on four (4) separate in-service track failures (i.e. broken rails) exhibiting heavy rail surface RCF, as confirmed by metallurgical analysis by a laboratory, it has been shown that RFD inspections performed within 35 days on average prior to the failure have been unsuccessful at identifying internal rail defects in each case. Rail testing across 4 subdivisions of a North American Class I railway using a non-contact rail inspection probe based on Electro-Magnetic Field Inspection (EMFI) technology to verify conditions of heavy RCF further suggest RFD detection of internal defects increases for smooth rail surfaces. The results also suggest that internal rail defects may be masked in rail surfaces exhibiting heavy RCF. In fact, over 2 times (200%) more defects have been detected using RFD rail testing upon completion of post grinding to restore the rail surfaces within the respective four subdivisions. Maintaining a smooth rail surface using rail grinding or milling equipment therefore increases the Probability of Detection (POD) of ultrasonic based RFD test equipment.

Unfortunately rail grinding/milling programs are scheduled in most North American Class I railways based on measured tonnages across the network along each subdivision. Higher tonnage subdivisions receive more rail maintenance as compared to lower tonnage lines. Lower tonnage lines often do not have the benefits of broken rail detection which is a by-product of signaling systems and therefore rely heavily on detection of internal rail defects through ultrasonic RFD and detection of complete rail failures through visual inspection. The onset of broken rail technology in non-signaled network territories (Stephens, 2020) is increasing within North American Class I railways however, a novel data driven method of quantifying the rail network surface conditions is required for effective ultrasonic RFD testing and ultimately increased internal rail defect detection rates.

Furthermore, standards for rail surface condition and RCF are not well established within the North American Class I industry. Associations such as, but not limited to,

the American Railway Engineering and Maintenance-of-Way Association (AREMA) (American Railway Engineering and Maintenance-of-Way Association, n.d.) are working on developing recommended practices around rail surface condition maintenance. These recommended practices which will become eventual standards are necessary in reducing risks around rail RCF.

Using the ability to integrate data from EMFI systems is another possible avenue for risk reductions due to rail surface conditions exhibiting RCF. Data integration of EMFI systems enables new information sources to North

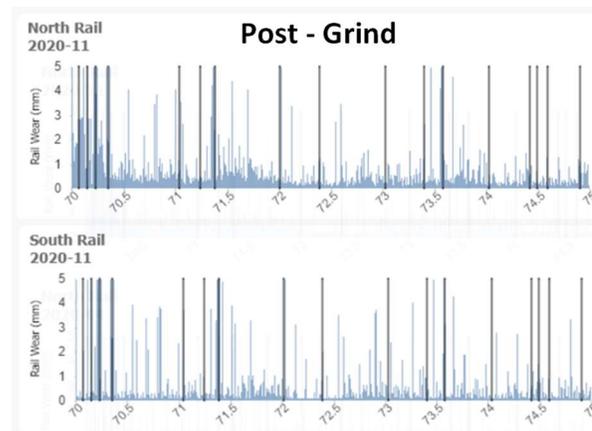


Figure 14. North and South rail crack depth measured by EMFI between Mile Post (MP) 70 and 75 of the 440 mile test section of a North American Class I railway (Post-Grind).

American Class I railways including, but not limited to, a data driven assessment of the rail surface conditions throughout the entire Class I network and subsequent automation. In addition to other data sources such as tonnage data, a condition based rail grinding or milling approach can be established. The data can be used in a number of ways as described in the following.

Firstly, EMFI system data integrated into Class I railway offices can be used to drive targeted rail grinding by identifying network areas exhibiting heavy RCF. Secondly, additional data on the rail used in these areas, the last maintenance actions, and other factors can be further utilized to drill down on the root cause of the accelerated RCF. As an example, the use of a different rail supplier who uses a different manufacturing process whereby the rail life may not follow internal modeling of wear and fatigue may quickly be identified based on the measured condition of the rail. Thirdly, when rail grinding or milling maintenance is performed, the data can be used to ensure that a sufficient amount of material is removed to optimize RFD testing and to audit/validate that grinding or milling has been performed to the specifications outlined by the Class I railway. Data collected from over 440 miles of track with an average crack depth of 0.5 mm show 200% more defects found by RFD post grind in comparison to pre-grind RFD

testing on the track segments with heavy RCF. Targeted grinding leads to more efficient ultrasound testing and higher defect removal rate which increases overall rail safety.

The next steps of this work will focus on analyzing the crack depth levels acquired from the EMFI equipment to determine appropriate thresholds which merit rail maintenance actions. This will include a study comparing defect rates to average crack depths. Further metrics to quantify and trend subdivisions based on crack depth or even crack densities over time to establish areas of rapid RCF growth are also aims of future work.

5 ACKNOWLEDGEMENT

The authors would like to express appreciation for the support of the project sponsors: Scott MacDonald SVP Engineering, Mechanical and Procurement, and Justin Meyer VP Engineering of Canadian Pacific for their continued support in the exploration of knowledge within the North American Heavy Haul railway industry. The authors would also like to thank Lincoln Oree and Rae Castro for the metallurgical support and insights. Finally, the authors appreciate all the support provided by Canadian Pacific field employees in executing all rail related testing.

6 REFERENCES

Abbott, J., 2007. Regular Grinding is the Key to Long Rail Life. *Global Railway Review*, 30 07, Issue 4.

American Railway Engineering and Maintenance-of-Way Association, n.d. *What is Arema?*. [Online] Available at: https://www.arema.org/AREMA_MBRR/About.aspx [Accessed 13 09 2021].

Bonaventura, C. S., Zarembski, A. M. & Palese, J. W., 16-18 March 2005. *TrackSafe: A Track Geometry Car Base Real-Time Dynamics Simulator*. Pueblo, CO, USA, IEEE, pp. 19-30.

Federal Railroad Administration, 2008. *Track Safety Standards Compliance Manual*. [Online] Available at: https://railroads.dot.gov/sites/fra.dot.gov/files/2020-08/2008_Track_Safety_Standards%20%281%29.pdf [Accessed 24 08 2021].

Heckel, T. T., Kreuzbruck, H. M. & Rhe, S., Dec. 2009. *High Speed Non-Destructive Rail Testing with Advanced Ultrasound and Eddy-Current Testing Techniques*. Prague, s.n.

Kang, Y. S., 2013. Rolling Bearing Contact Fatigue. *Encyclopedia of Tribology*.

Kerchof, B., 2021. The Science of Broken Rails - Broken Rails are Common; Derailments Caused by Broken Rails are Not. *Trains Magazine*, 21 09, pp. pp. 18 - 25.

Pyke, D., 2015. *5 Myths About Rails (Part 4) - Hard Rails Generate More RCF*. [Online] Available at: <https://www.linkedin.com/pulse/5-myths-rails-part-4-hard-generate-more-rcf-daniel-pyke> [Accessed 25 08 2021].

Singh, L., 2011. *Rolling Contact Fatigue in Rails*, India: The Institute of Permanent Way Engineers.

Stephens, B., 2020. Canadian Pacific Deploying New Broken-Rail Detection System in Dark Territory. *Trains Magazine*, 27 10.

Transport Canada, 2018. *Rules Respecting Track Safety - Part II - Track Safety Rules - Subpart C - Track Geometry*. [Online] Available at: <https://tc.canada.ca/en/rail-transportation/rules/2011-2012/rules-respecting-track-safety/part-ii-track-safety-rules-subpart-c-track-geometry> [Accessed 24 08 2021].

Transportation Safety Board of Canada, 2019. *Rail transportation safety investigation report R19E0147*. [Online] Available at: <https://www.tsb.gc.ca/eng/rapports-reports/rail/2019/r19e0147/r19e0147.html> [Accessed 10 09 2021].

Tuzik, B., 2019. The Anatomy of a Rail Milling Cutting Head. *RT&S Loram*, 06 03.

Lessons learned from long-term frost heave monitoring under a railway embankment

Alireza Roghani, Robert Caldwell, & Juan Hiedra-Cobo
National Research Council of Canada, Ottawa, Ontario, Canada

Paul Charbachi
VIA Rail Canada, Montreal, Quebec, Canada

ABSTRACT

Frost heave is a major issue for railway tracks constructed in cold regions that degrades track geometry and may affect the safety of railway operations. Canadian railway operators perform frequent maintenance during winter to eliminate track deformation and ensure the safe passage of trains. In early spring, frequent maintenance such as tamping and surfacing are required to alleviate the surface deformation due to thawing. Freeze-thaw cycles are expected to become more frequent under future climate conditions and thus a greater understanding of this phenomenon is essential to develop adequate mitigation measures in the face of a changing climate. National Research Council Canada in collaboration with VIA Rail Canada has conducted a 3-year field investigation to study the mechanism of frost development and its impact on safety and performance of train operations. In this project, a 50 m section of track in eastern Ontario was instrumented with various geotechnical and structural monitoring systems. In addition, measurements from ground penetrating radar and a track geometry car were collected over 90 km of track to map frost-susceptible sections of track and quantify its effect on track geometry degradation. The difference in winter conditions during the monitoring period (which consisted of two freeze-thaw seasons) in terms of temperature and snow on ground, induced different temperature regimes within the track substructure and led to different track responses. This paper summarizes some of the major lessons learned during the field observation period and discusses how the expected future climate may adversely affect the frost heave issues.

1 INTRODUCTION

Frost heave and thaw softening due to freeze-thaw cycles (FTC) are a common problem for Canadian railway tracks. Railway operators must perform frequent maintenance or reduce train speeds to allow safe passage through heaved sections - which increases the cost of operation and reduces the flow of traffic within the network. The ice melting in the early spring leads to non-uniform vertical deformation of railway track, degrades track geometry, shortens the service life of track components, and in extreme cases may lead to unsafe operating conditions. While nearly the entire rail network in Canada is subjected to FTC, there is little research about this phenomenon or its impact on railway safety and performance. It is also anticipated that FTC will become more frequent due to climate change (Henry, 2008; Warren & Lemmen, 2014) and therefore a greater understanding of this phenomenon is essential to develop mitigation / adaptation measures to ensure it does not affect future railway operations.

Under the Pan-Canadian Framework on Clean Growth and Climate Change, Infrastructure Canada has provided funding to the National Research Council Canada (NRC) to deliver the Climate-Resilient Buildings and Core Public Infrastructure (CRBCPI) Initiative over a

5-year period. As part of this project, NRC and VIA Rail Canada conducted a 3-year research project to investigate the effects of FTC on railway embankments. The main goal of this project was to provide a greater understanding about the effect of FTC on the integrity, safety, and performance of railway tracks. The other objectives were to investigate: (i) the mechanism of FT development within the railway embankment under different climatic conditions, (ii) the potential of ground penetrating radar (GPR) to proactively identify the frost-susceptible sections of track, (iii) the potential of various sensors for remotely monitoring heave spots, and (iv) provide recommendations for mitigating the adverse effect of FTC for existing and new tracks.

This paper presents a summary of the major results observed during this project. These observations are divided into two categories: (i) those that resulted from measurements at a short test section and (ii) those that were established from data collected over long distances (subdivision scale).

2 TEST SECTION OBSERVATIONS

A 50 m section of VIA Rail track was instrumented with various geotechnical and structural monitoring systems, including thermistors, borehole extensometers,

ShapeArrays, and piezometers (Figure 1). These instruments measured track response under static (every 10 minutes) and dynamic modes (100 Hz, triggered by each train passage) during the monitoring period which consisted of two freeze-thaw seasons, winter-spring 2018–2019 and winter-spring 2019–2020, (hereafter called winter 2019 and winter 2020, respectively). A more detailed discussion about the installation and types of sensors can be found at Le Borgne et al., 2019. The instrumented section has suffered from frost heave in the past and laboratory tests suggested that the subgrade at this location consisted of highly frost-susceptible material (AREMA, 2012).

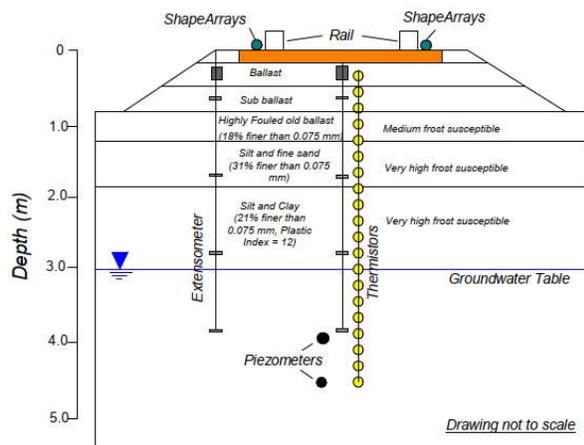


Figure 1. Plot shows embankment composition, degree of frost susceptibility of each layer (as per FHWA NHI-05-037, 2006), and location of instruments (after Roghani et al., 2019).

2.1 Embankment temperature response

The analysis of the weather station data near the test site during the monitoring period indicated that winter 2020 was 2 °C warmer than winter 2019. The freezing index or freezing degree days (FDDs, a common method used to quantify the severity of a given winter) for winter 2020 was reduced by 35% compared to winter 2019 (Figure 2). The total snowfall reduced from 1.74 m in winter 2019 to 1.58 m in winter 2020. Despite this reduction in total snowfall, the average snow on the ground in December and January 2020 (which were the coldest months of winter 2020) was 47.2% and 26.8% higher than in 2019. The insulating effect of thicker snow on the ground is an important factor affecting frost penetration.

The thermistor measurements (installed in the embankment from 0.35 m to 4.15 m below the track surface) showed a distinct difference in the response of embankment temperature to atmospheric conditions during the two winters. Figure 2 presents the variation of frost depth for the two winters. In winter 2019, the frost reached the first thermistor (0.35 m) when the FDDs was approximately 305 degree-days and penetrated down to 1.0 m, while in winter 2020 it took 348 FDDs and never exceeded 0.5 m depth. The higher FDDs to initiate frost in winter 2020 can be attributed to thicker snow cover on

the ground that restricts the loss of heat from the embankment during December and January. In addition, in winter 2019 the track went through one FTC, while in 2020 the track experienced 4 intermittent FTC due to the strong fluctuations of ambient temperature during the freezing season. The increased number of FTC is consistent with the warmer and drier trends in winter 2020 and is expected to be the new normal due to the impact of climate change. In 2019, the track remained in a frozen condition for 65 days whereas in 2020 it was frozen for 38 days. The shorter freezing season was followed by an early thawing stage as indicated in Figure 2.

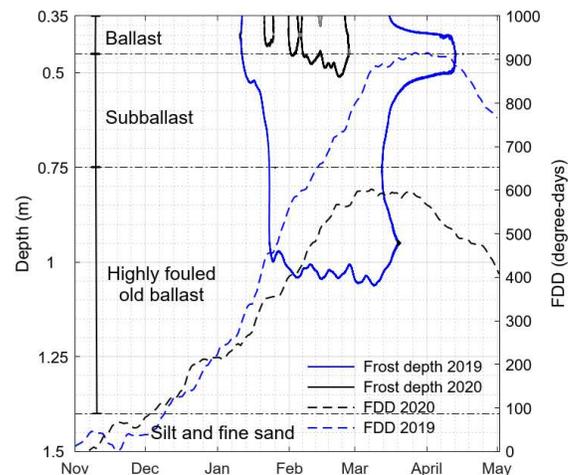


Figure 2. The variation of FDDs and frost depth during winter 2019 and winter 2020 (after Roghani, 2021).

2.2 Track surface deformation

To measure the absolute deformation of the track surface during FTC, the study site was surveyed using a total station. The baseline measurements were collected in June 2018 with subsequent measurements conducted in November 2018, February, and April 2019 for winter 2019 and December 2019, January, February, and March 2020 for winter 2020 track monitoring.

According to the measurements presented in Figure 3, the largest heave occurs at the culvert location, with a maximum of 16 and 17 mm heave on the south and north ends of the culvert in winter 2019 compared to 18 and 21 mm in winter 2020. On average, the track outside of the culvert zone heaved about 6.9 and 2.6 mm during winter 2019 and winter 2020, respectively. The reduction in the heave amount is consistent with lower frost penetration, a shorter freezing period, and thicker snow on the ground in winter 2020. The large differential heave observed in the warmer winter led to the development of crosslevel and twist issues on the track that could have adversely affected the dynamic forces at the wheel/rail interface and led to unsafe operating conditions.

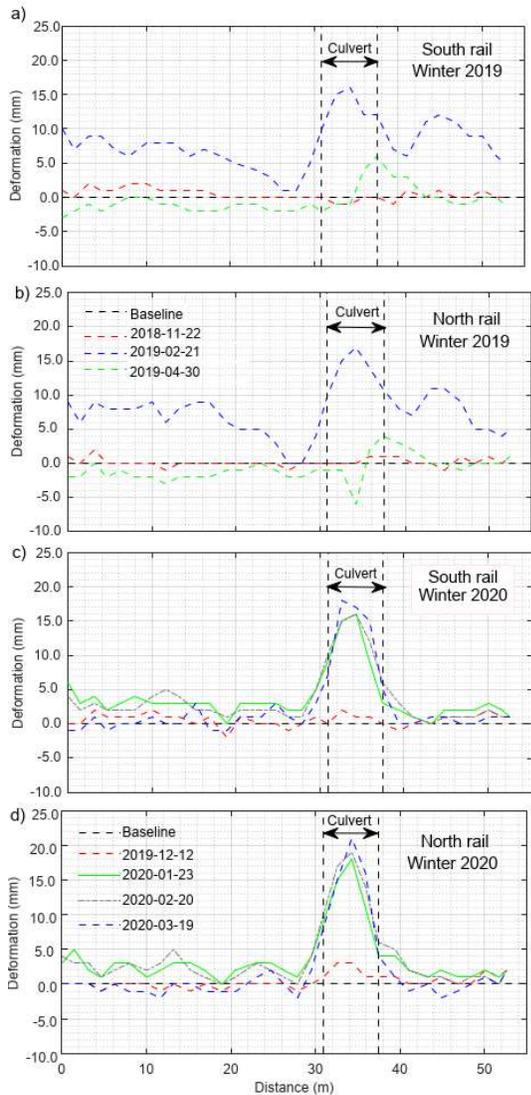


Figure 3. Plots show track surface deformation measured by using total station (after Roghani, 2021).

2.3 Pore water pressure (PWP) response

The measurements from two piezometers, installed 3.7 and 4.3 m below the track surface (hereafter called piezo 1 and piezo 2, respectively), were used to validate the track softening that is known to occur during the thawing season. Figure 4 shows the response of induced PWP to train loading. As this figure suggests, the PWP response is cyclic and consistent with train arrivals. The larger peaks at the beginning of the spectrum is consistent with the fact that locomotives are heavier than coach cars and impose higher stress on the subgrade. Once the train has passed the test section, the PWP returns to its original values. As shown in Figure 4, the maximum PWP measured under train loading can be separated into two components: (1) hydrostatic PWP, and (2) induced PWP caused by train loading. To study the effect of FTC on induced PWP, the hydrostatic pressure was subtracted from the maximum PWP for each train passage;

hereafter it is referred to as ΔPWP . ΔPWP represents the PWP change due to train loading only.

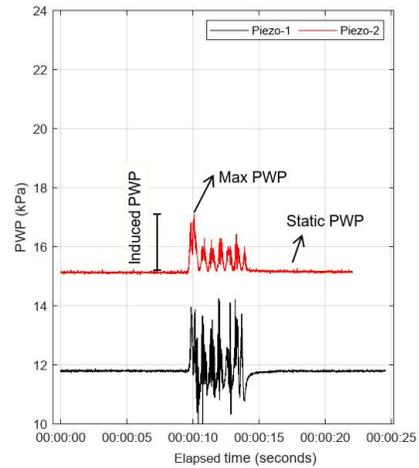


Figure 4. PWP measurements during train passage.

The cumulative distribution function of piezo 1 and 2 during the track's frozen and thawing stages are compared in Figure 5.

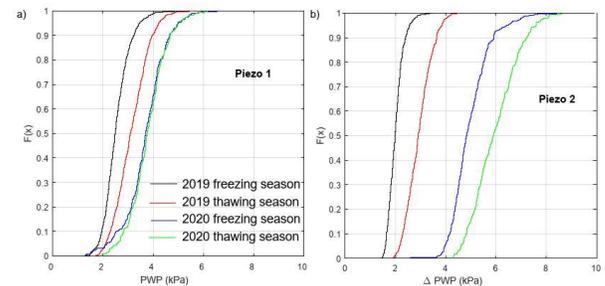


Figure 5. Plots compare cumulative distribution function of ΔPWP during 2019 and 2020 freezing and thawing seasons recorded by piezo 1 and piezo 2.

The analysis suggests that the average ΔPWP recorded by piezo 2 during the thawing stage was 45% and 20% higher than that of the frozen stage for 2019 and 2020 winter-spring period, respectively. Note that this sudden increase in ΔPWP happens while train axle loads remain unchanged, and thus the increase is attributed to a change in track stiffness. This condition, known as track softening, occurs as a result of increasing moisture content during the thawing stage. Another important observation that is evident from Figure 5 is that the average ΔPWP in freezing season 2020 is even higher (about 70%) than thawing season 2019 for piezo 2. This significant increase can be attributed to the mild and short 2020 winter season during which there were 4 subgrade FTC. The higher number of FTC may have caused a higher moisture content and extreme softening of the track foundation and thus created higher ΔPWP . Climate scientists believe that weather conditions similar to what was observed during winter 2020 will be the new normal under the future expected climate and thus more research is required to develop proper countermeasures against the adverse effect of FTC.

3 SUBDIVISION SCALE MEASUREMENTS

This section presents the results of measurements collected over 90 km of VIA Rail track in Ontario. The tested track is only used for passenger trains and consists of 115 lb continuously welded rail and wooden ties.

3.1 Track Geometry measurements

Track geometry measurements are a common inspection method used by the railway industry to ensure that the shape of the track allows for the safe passage of trains at the designated maximum speed of the track [2] [3] [4]. To quantify the effect of FTC on track geometry degradation, three sets of data were analyzed and compared including: (1) measurements collected in May 2019 after the thaw, (2) measurements collected in November 2019 before the freeze, and (3) data collected in April 2020 after the thaw.

For track roughness analysis, a running standard deviation with a 30 m window was used to calculate the surface and alignment roughness, where a higher standard deviation indicates rougher track and infers that it is in poorer condition. Figure 6 presents the cumulative distribution functions for the surface and alignment roughness for the three sets of track geometry measurements and provides a comparison of the overall track condition across these dates. While these figures suggest a very slight improvement in track condition in November 2019 compared to May 2019, there is a significant shift towards higher values in the April 2020 data for both surface and alignment that indicates a significant degradation in track condition.

The surface and alignment roughness at proximity to the culverts (± 15 m) were also processed for the three sets of geometry measurements. The comparison of the roughness values suggests that on average the surface and alignment roughness at the culverts measured after the 2020 thaw were 121% and 94% higher than measurements taken after the 2019 thaw, and 113% and 92% higher than measurements taken before the 2020 freeze.

Analysis of geometry defects, presented in Figure 7 in a normalized format, suggests that there was a considerable increase in the number of surface and alignment defects in spring 2020. It should be noted that only 2% of these defects were urgent defects (which exceed regulatory limits) and the rest were priority defects (which are roughly 65% to 70% of regulatory limits). It is also worth mentioning that the level of maintenance effort put into the track is a key factor affecting track degradation, but this information was not available for this analysis. However, detailed analysis of defect locations for the three runs shows that the defects are at different locations and therefore are new. Therefore it is less likely that the degradation of track geometry is due to lack of maintenance and the subsequent accumulation of defects.

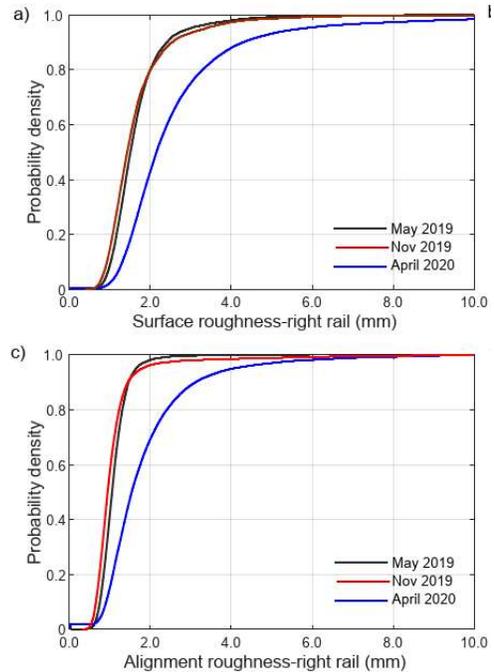


Figure 6. Plots comparing the cumulative distribution function of a) surface roughness and b) alignment roughness (after Roghani 2021).

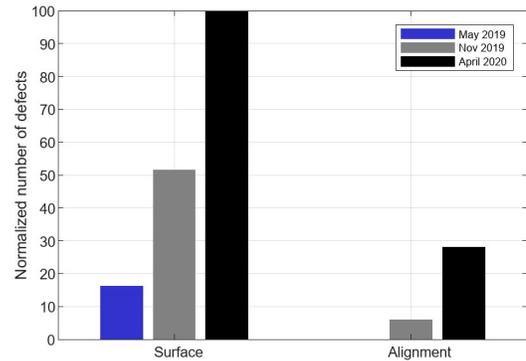


Figure 7. Plot shows the normalized number of surface and alignment defects over the studies track (after Roghani 2021).

3.2 Ground Penetrating Radar surveys

GPR involves transmitting electromagnetic waves into the subsurface and receiving the returning pulses that have reflected off the interfaces between materials with different electromagnetic properties. Water is the dominant factor that determines the dielectric properties of soil and changes the reflection of the GPR pulses at its transition from liquid to ice or vice versa. When the GPR signals travel through the substructure layers, the magnitude of the contrast between the different materials can be monitored to track the interfaces between frozen and unfrozen zones (Li, Hyslip, Sussmann, & Chrismer, 2016; Hyslip, Olhoef, Smith, & Selig, 2005; Silvast, Nurmikolu, & Wiljanen, 2012).

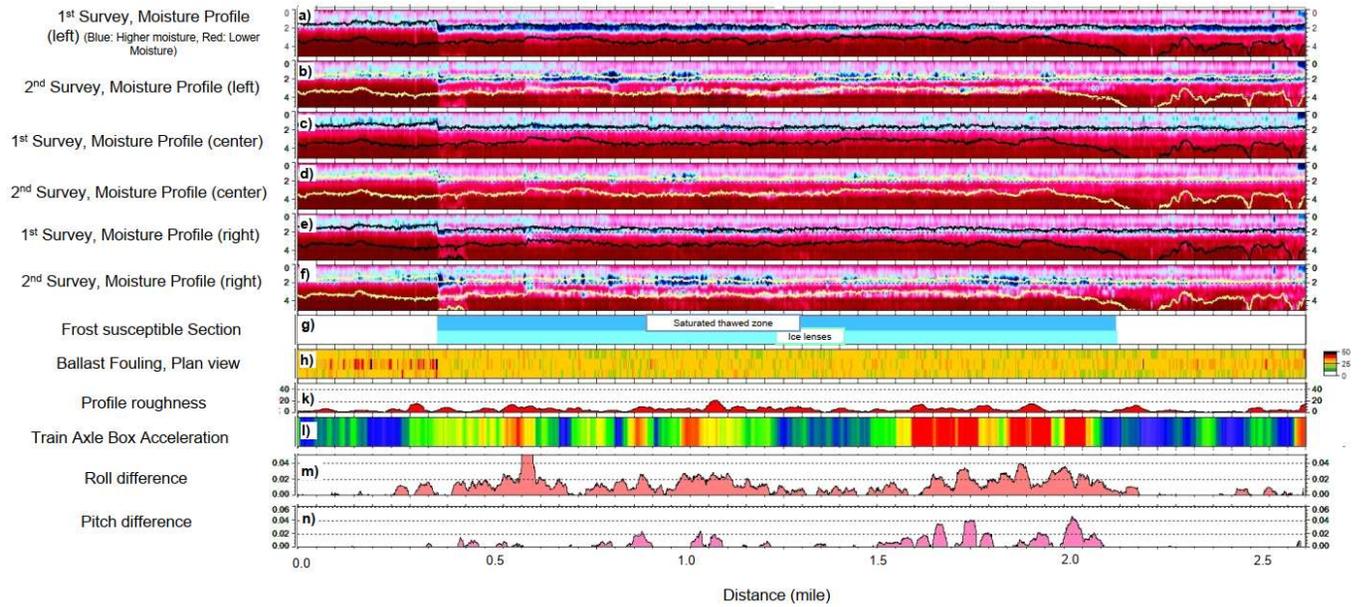


Figure 8. Example of the integrated data. Note: red does not mean unsafe (after Roghani et al., 2021).

Two GPR surveys were conducted to evaluate its potential to identify the frost-susceptible sections of track. The first GPR survey was performed on December 2, 2017, when the track was not yet in a frozen condition. A follow-up survey was performed on February 23, 2018, when most sections of the track were still frozen and some sections were starting to thaw. The difference in the energy of the reflected pulses, as well as the difference in the signal's frequency dispersion between frozen and unfrozen soil, were used to distinguish frozen and unfrozen layers of track. In addition, the axles of the hi-rail truck used for the GPR surveys were equipped with accelerometers. These acceleration measurements were integrated with the other information including processed GPR measurements and track geometry data. Figure 8 shows an example of the results for 2.5 miles of the test track. Figure 8a to Figure 8f present the relative moisture content and the ballast/subballast and the subballast/subgrade interface for left, centre, and right side of the track during the first and second GPR survey. At the centre of this track section, high water content and ice lenses were detected through comparison of the reflected signals between the two runs (Figure 8g), suggesting the presence of frost-susceptible material at this location. The presence of both high moisture at shallow depth and ice lenses below indicates that this track location is in its thawing state. On either side of this section, there is no sign of ice or high moisture content, suggesting the material is not susceptible to frost action.

The acceleration data collected by the hi-rail vehicle were processed to calculate the roll and pitch for each run and the difference of the roll and pitch between the two runs are presented in Figure 8m and n, with positive values indicating that there is an increase in roll and pitch during the second run.

The presence of both frost- and non-frost-susceptible material provides a good contrast to evaluate its impact on track performance. The track surface roughness, calculated from the track geometry measurements collected during the thawing season, is presented in Figure 8k. The results show some higher values of roughness at the frost-susceptible section but not to an extent that is distinguishable from the non-frost-susceptible section. The effect of track softening is evident in the roll and pitch data calculated from the accelerometers mounted on the hi-rail truck, with the second run showing significantly higher values. This confirms the effect of frost action on the performance of the railway track. The GPR assessment was proven to be useful and cost effective when the assessment is needed over long distances. It also helps to identify the source of frost heave, whether it is highly fouled ballast, frost-susceptible subballast, or a problem within the subgrade. The integration of GPR data with other measurements such as track geometry proved useful to not only detect the source of the problem but also identify its impact on track performance.

4 CONCLUSIONS

This paper presented the results of field observations during two winter-spring periods to provide a greater understanding of the process associated with FTC and its effect on the performance of railway track. The field investigation consisted of installing thermistors on a section of frost-susceptible track to measure the response of track temperature to ambient variables, surveying the site periodically to quantify the track surface deformation during FTC, and analyzing GPR and track geometry measurements to quantify track degradation before the freeze and after the thaw seasons.

The limited field observations in this project suggested that even though the warmer and drier winter may result in a shorter freezing period and shallower frost depth, the higher number of intermittent FTC lead to larger differential frost heave in winter and higher track roughness in spring. This may lead to a condition that is more detrimental to track geometry and safe operations than a cold winter. According to predictions by climate scientists, this type of winter, warmer and drier with frequent FTC, is expected to become the new normal, and therefore frequent maintenance may be required to keep railway tracks constructed in seasonally frozen ground within safe limits. Failure to account for these changes may lead to increased costs for infrastructure owners, unforeseen costs for infrastructure users, and considerable negative socioeconomic impacts. The research conducted in this project was limited in scope and was performed over a well-maintained class 5 track in Ontario. The performance of a lower-class track in a colder region of Canada may be different. In the author's opinion, further research and development is required.

5 ACKNOWLEDGEMENT

The authors would like to acknowledge Infrastructure Canada for providing funding for this study and VIA Rail Canada for providing data. The authors also would like to thank Paul Charbachi from VIA rail Canada, Marianne Armstrong, Fiona Hill, Dr. Jon Makar, and Albert Wahba from National Research Council Canada, and Hamed Kashani and James Hyslip from HyGround/Loram for supporting this project.

6 REFERENCES

AREMA. (2012). Manual for Railway Engineering. Lanham, MD: American Railway Engineering and Maintenance of Way Association.

Federal Railroad Administration. (2005). Development of objective track quality Indices. Washington, DC: US Department of Transportation.

Henry, H. A. (2008). Climate change and soil freezing dynamics: historical trends and projected changes. *Climatic Change* (87), 421-434. doi:10.1007/s10584-007-9322-8.

Hyslip, J., Olhoeft, G., Smith, S., & Selig, E. (2005). Ground Penetrating Radar for Railroad Track Substructure Evaluation. Washington, DC: Federal Railroad Administration office of Research and Development, US Dept. of Transportation.

Le Borgne, V, Roghani, A, Hiedra Cobo, J., Charbachi, P. (2019). Design and installation of a geotechnical monitoring system for monitoring freeze-thaw cycles on a railway track. 18th International Conference on Cold Regions Engineering and 8th Canadian Permafrost Conference, ISBN 978-0-7844-8259-9.

Li, D., Hyslip, J., Sussmann, T., & Chrismer, S. (2015). *Railway Geotechnics* (1st ed.). London: CRC Press. <https://doi.org/10.1201/b18982>.

Roghani, A. (2021). Quantifying the effect of freeze-thaw cycles on track surface deformation and degradation of railway track geometry; Case study. *Transportation Geotechnics*. <https://doi.org/10.1016/j.trgeo.2021.100601>.

Roghani A, Pall R, Toma E. (2021) Procedure for combining field measurements and machine learning to quantify impact of different track parameters on ride quality of railway tracks. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*. <https://doi.org/10.1177/09544097211002665>

Roghani, A, Hiedra Cobo, J., Charbachi, P, (2019). Studying the impact of freeze thaw cycles on performance of railway tracks. *International Heavy Haul Association Conference 2019, Narvik, Norway*, ISBN: 9780911382716.

Roustaei, M, Hendry, M, Roghani, A. (2019). Frost susceptibility of subgrade soil beneath railway tracks in different temperature boundary conditions. *Canadian Geotechnical Conference, St. John's*.

Transport Canada. (2011). Rules respecting track safety. <https://tc.canada.ca/en/rail-transportation/rules/2021-2022/rules-respecting-track-safety>.

Silvast, M., Nurmikolu, A., & Wiljanen, B. (2012). Identifying frost-susceptible areas on Finnish railways using the ground penetrating radar technique. *Journal of Rail and Rapid Transit*, 227(1), 3-9. doi:10.1177/0954409712452076.

Warren, F. J., & Lemmen, D. S. (2014). *Canada in a Changing Climate: Sector Perspectives on Impacts and Adaptation*. Ottawa: Government of Canada.

Material model for TC128B at low temperature and application in full-scale FEA

Bruce W. Williams, Jia Xue, Jonathan McKinley, and Su Xu
CanmetMATERIALS, Natural Resources Canada, Hamilton, Ontario, Canada

Michael Spiess
Transportation of Dangerous Goods, Transport Canada, Ottawa, Ontario, Canada

ABSTRACT

Damage mechanics methods are increasingly adopted to describe fracture of metallic alloys, including TC128B steel which is commonly used in rail tank cars. Damage models can accurately capture the response of the material under a wide variety of loading conditions, particularly for stress states that are possible during an impact event such as derailment or hard-coupling. Several mechanical tests are required to generate data for the calibration of the model including the use of tension, notched tension, shear, and fracture specimens. The crack initiation response is measured from small-scale notched tension specimens whereas the crack propagation is measured using specimens that undergo stable tearing, such as Charpy V-Notch specimens. Experimental data was used to calibrate and validate a failure model for TC128B between room temperature and $-40\text{ }^{\circ}\text{C}$. TC128B demonstrated the typical ductile to brittle transition seen in most steels over this temperature range. The model was able to capture the sudden load drop associated with brittle fracture when preceded by ductile crack growth at low temperatures. The material model was applied to large-scale Finite Element Analysis (FEA) of the tank car structure. Bridging the gap between small-scale laboratory size specimens and full-scale tank cars was addressed through a sub-modelling method. Additionally, the low-temperature response was compared to the room-temperature response in the full-scale models. It was shown that the FEA models could accurately predict the response of the large-scale structure subjected to an impact.

1 INTRODUCTION

Finite Element Analysis (FEA) can be used to model full-scale tank car accident scenarios but requires accurate material behaviour. Full-scale simulations have been compared to experimental side impact data of a tank car by Carolan et al. (2018) with good agreement obtained at room temperature ($24\text{ }^{\circ}\text{C}$). At room temperature, steel fails by ductile fracture but at low temperature the steel can fail by brittle, or cleavage fracture, which leads to a sudden drop in load bearing capacity. The focus of this work is to compare a FEA material model that describes failure at both room temperature and low temperature ($-40\text{ }^{\circ}\text{C}$) under a range of stress states that could be expected during tank car impact.

Damage mechanics models have been used in FEA to describe ductile fracture of metals that are subjected to a wide range of loading conditions, from the high stress triaxiality at a crack front to lower triaxiality conditions that are present during crash events. Damage models have recently been adopted and used in FEA to describe the failure response of tank car steel as reported by Kirkpatrick and McKeighan (2018), Przemyslaw et al.

(2019), and Eshraghi and Carolan (2020). These studies focused on the room temperature failure response of TC128B. Typical, damage models comprise two parts: damage initiation and damage evolution. In the first stage, damage increases in the material (or element) without a decrease in strength, until a damage initiation criterion is met. Second, the element is softened according to a damage evolution criterion. When a critical value is met the element is deleted.

The Modified-Mohr-Coulomb (MMC) damage model detailed by Bai and Wierzbicki (2010) was used in the current work. The effective plastic strain to damage is a function of stress triaxiality and Lode angle. Paredes et al. (2018) utilized the MMC damage model to describe the fracture response of a TC128 steel. The damage model was calibrated using experiments, simulations, and optimization techniques. Several mechanical test geometries including notched tension, shear, biaxial and fracture geometries such as Single-Edge-Bending (SEB) and Compact Tension (CT) were considered to produce a wide range of loading conditions described by triaxiality and Lode angle. The model was valid for one

temperature (room temperature) and one strain rate (quasi-static).

Previously, it was shown by Williams et al. (2020) that the well calibrated damage model for TC128 detailed by Paredes et al. (2018) could be adjusted to capture influence of multiple temperatures and strain rates, without the necessity to perform extensive experiment, simulation, and optimization. Only experimental data from quasi-static uniaxial tension tests and dynamic Charpy V-Notch (CVN) fracture tests was considered. This model for TC128B was further assessed in the current effort, particularly for crack initiation type test geometries. For these geometries, once the crack initiates there is no resistance to crack propagation. This is opposed to specimen geometries that undergo stable tearing (for instance, CVN, SEB or CT).

The focus of the current work was to compare the crack initiation response of TC128B between 24 °C and -40 °C. Mechanical test specimens used to calibrate the crack initiation component of the model tend to have lower stress triaxiality than fracture specimens with high stress triaxiality that undergo stable crack propagation. Crack initiation specimens are representative of the stress states that can be present during an impact event. Also of interest was determining whether the brittle fracture response observed in the CVN testing of TC128B at -40 °C was relevant for the lower stress triaxiality geometries.

2 MATERIAL BEHAVIOUR OF TC128B

2.1 Tensile and Charpy Responses

Round tensile specimens were machined from a section of TC128B extracted from a tank car built and qualified in 2015. Quasi-static tensile tests were performed between at room temperature and -40 °C. The tensile specimens had a gauge length of 25.4 mm and diameter of 6.35 mm with the gauge length oriented in the longitudinal (rolling) direction of the steel. The engineering stress-strain curves are shown in Figure 1. Lüders strains were in the range of 2-2.9%. Of note, was that there was no decrease of per cent elongation at -40 °C compared to room temperature.

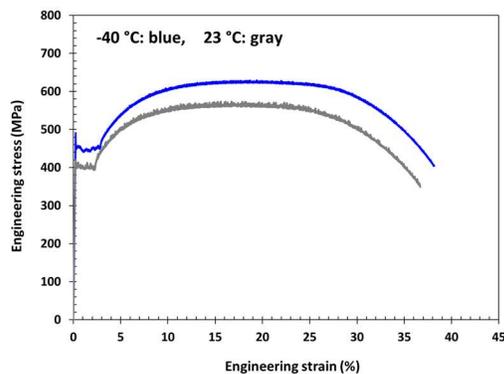


Figure 1. Quasi-Static engineering stress-strain curves for TC128B at room temperature and -40 °C

The hardening response required for FEA, including dynamic behavior, was detailed in Williams et al. (2020). The quasi-static and dynamic stress-strain responses relevant to the current work at 24 ° and -40 °C are shown in Figure 2.

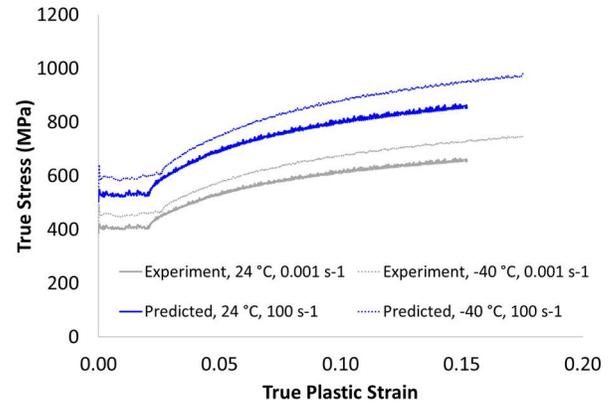


Figure 2: Quasi-static vs. dynamic stress strain curves for TC128B (Williams et al., 2020)

Charpy tests were performed at temperatures of 24 °C and -40 °C at a dynamic loading rate of 5.1 m/s with a computed strain rate of about 100 s⁻¹. The dynamic load versus displacement response is shown in Figure 3. Ductile fracture is dominant at 24 °C whereas brittle, or cleavage, fracture occurred after a small amount of ductile fracture at -40 °C.

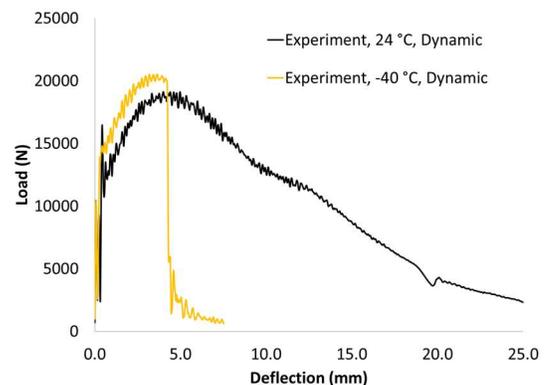


Figure 3. Dynamic load vs. deflection response from Charpy testing of TC128B showing sudden load drop due to brittle fracture at -40 °C

To confirm rate sensitivity of TC128B, a quasi-static Charpy test was performed at 24 °C. The quasi-static Charpy test was performed in a standard load frame at a load rate of 5×10⁻⁵ m/s for a computed quasi-static strain rate of about 0.001 s⁻¹ compared to the dynamic test performed at 100 s⁻¹ using a pendulum machine. All other aspects of the test setup, including the sample size and fixture, were identical. The quasi-static versus dynamic Charpy response is compared in Figure 4. The results show that the energy absorbed (area under the curve) is significantly greater for the dynamic case. Consequently,

it was necessary to account for strain rate effects in the FEA model.

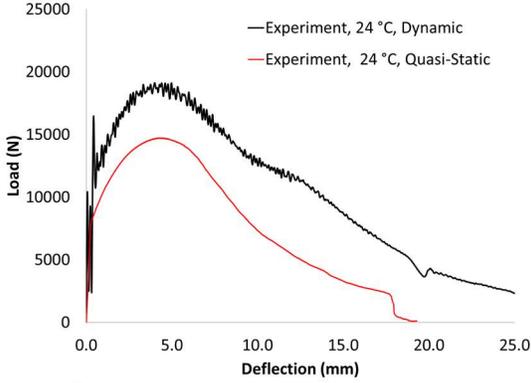


Figure 4. Quasi-static vs. dynamic load vs. deflection response at room temperature from Charpy testing of TC128B

2.2 Ductile Failure Response

The fracture model of TC128B for temperatures ranging from 24 °C to -80 °C was detailed by Williams et al. (2020) and is only briefly outlined in this article. In the MMC model, the strain to initiate fracture, ε_f , depends on stress triaxiality, η , and Lode angle, θ_L . Stress triaxiality is defined by the mean stress divided by the effective stress with the Lode angle defined in Bai and Wierzbicki (2010). The effective strain at damage initiation is given by

$$\varepsilon_f = \left\{ \left[\frac{A}{C_2} \left[C_3 + \frac{\sqrt{3}}{2 - \sqrt{3}} (1 - C_3) \left(\sec\left(\frac{\theta_L \pi}{6}\right) - 1 \right) \right] \right]^{-1/n} \right. \\ \left. \left[\sqrt{\frac{1 + C_1^2}{3} \cos\left(\frac{\theta_L \pi}{6}\right) + C_1 \left(\eta + \frac{1}{3} \sin\left(\frac{\theta_L \pi}{6}\right) \right)} \right] \right\} \quad (1)$$

where C_1 , C_2 , C_3 , A , and n are material constants specific to a single temperature and strain rate. Upon damage initiation at $D_0=1$, damage evolution and element weakening evolve according to

$$D = \int_0^{\varepsilon^p} \frac{1}{\varepsilon_f} d\varepsilon^p \quad \text{and} \quad \varpi = \left(\frac{D_c - D}{D_c - D_0} \right)^m \quad (2).$$

D_c is the critical damage at which to remove the element from the simulation and m is a material constant. The model has eight coefficients that need to be calibrated. The MMC failure model detailed by Paredes et al. (2018) for TC128 was used as a starting point to describe failure of TC128B at various temperatures. The model was adjusted using only tensile and Charpy data (Williams et al., 2020). The model coefficients are given in Table 1.

Table 1: MMC Damage Model Coefficients for TC128B (Williams et al., 2020); $D_0 = 1.0$

Temp.	A	n	C_1	C_2	C_3	D_c	m	Rate
°C	MPa					s ⁻¹		
24	112.87	0.180	0.145	62.94	0.969	2.6	0.05	0.001
24	96.7	0.18	0.150	53.61	0.969	2.6	0.01	100
-40	65.18	0.204	0.168	37.33	0.969	2.0	0.05	0.001
-40	66.35	0.204	0.168	37.33	0.969	2.0	0.05	100

A quasi-static Charpy test was not available at -40 °C therefore the quasi-static MMC coefficients were assumed equivalent to the dynamic values. Figure 5 compares the MMC model (Eqn. 2) between the quasi-static coefficients for 24 °C and -40 °C indicating a small difference in the failure response. Thus, the large difference in the force response observed at 24 °C between the quasi-static and dynamic Charpy tests (Figure 3) can be attributed to the difference in the stress versus strain response (Figure 1) more so than the difference in the fracture response.

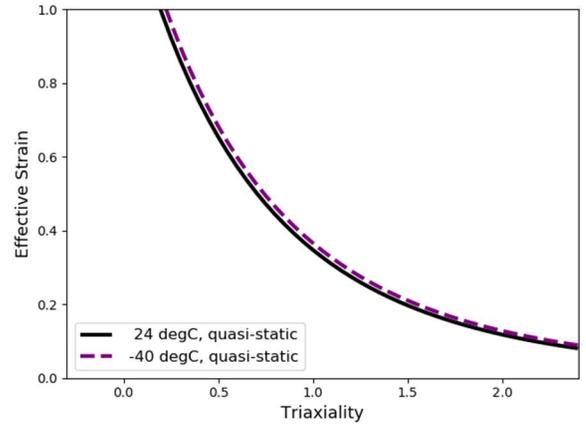


Figure 5. Comparison of failure response for TC128B at 24 °C and -40 °C (normalized Lode angle = 1.0)

2.3 Brittle Failure Response

To capture the sudden load drop associated with brittle fracture when preceded by ductile fracture, a coupled ductile-brittle model was used as detailed in Williams et al. (2020). The measured versus predicted response that was obtained for TC128B at -40 °C using the ductile-brittle model is shown in Figure 6.

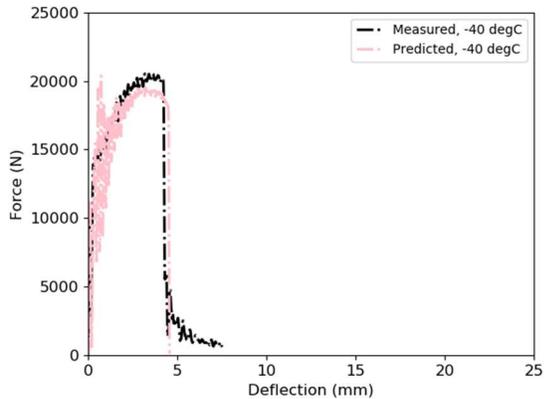


Figure 6. Measured vs. predicted Charpy response for TC128B at -40 °C (Williams et al., 2020)

Brittle fracture of steels usually occurs when a critical stress ahead of a crack tip is reached. At -40 °C, the maximum stress at brittle failure of the TC128B steel was 2200 MPa. This value is typical of the stress state at a crack-tip (high stress triaxiality) but exceeds the stress state expected at lower triaxiality. Charpy V-notched specimens have a machined notch of 2 mm deep and 0.25 mm radius. It was not expected that this model would necessarily predict brittle fracture for the lower triaxiality specimens detailed in the next section.

3 DAMAGE GEOMETRIES FOR CRACK INITIATION

To further validate the model at stress concentrations expected in impact, additional TC128B mechanical testing of notched tension specimens was performed. The data from these tests was used to further validate the damage initiation (Eqn. 2) component of the model. For notched tension geometries, once the crack initiates it propagates very rapidly such that these specimens are not well suited for calibrating and validating the damage evolution component (Eqn. 3) of the failure model.

The five damage geometries shown in Figure 7 were considered. Geometry 1-4 (G1-G4) were notched tension specimens with 20, 15, 10, and 5 mm radius, respectively. Geometry 5 (G5) contained a hole in the centre with a 10 mm radius. Each geometry had a width of 12.7 mm, (gauge) length of 46.5 mm, and a thickness of 3.1 mm. For the notched tension specimens, the cross-sectional width under load was always 7.0 mm. Testing was performed at a quasi-static displacement rate of 1.67×10^{-4} m/s.



Figure 7. Five damage geometries to measure crack initiation response of TC128B

The force versus displacement responses for each geometry is shown in Figure 8 at 24 °C and Figure 9 at -40 °C. Two or three repeat tests were performed for each case and good repeatability was obtained. For brevity, one test per temperature and geometry is presented. For either temperature, G1 has the lowest force response with the highest displacement until final fracture. As the notched radius decreased from G1 to G4, the load slightly increased with a small decrease in the displacement at failure. The centred-hole specimen, G5, had a response similar to G3 but with a slightly lower load.

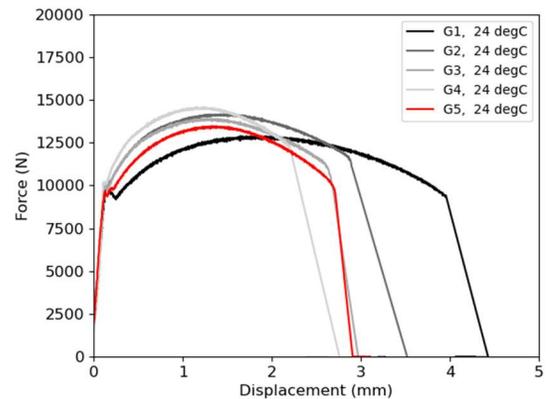


Figure 8. Measured force response of five damage geometries at 24 °C

4 FEA OF LABORATORY-SCALE SPECIMENS

4.1 FEA Model

The elastic modulus was specified as 205 GPa and 207.5 GPa for 24 °C and -40 °C, respectively. This small difference did not significantly influence the predictions. The failure model for TC128B defined in Section 2 was implemented in the explicit FEA analysis software DYNA3D as a user defined subroutine. The model was validated against simulations performed in Abaqus (Williams et al., 2020). The solid element size in the region of fracture was approximately 0.2 mm, transitioning to coarser elements away from the fracture region. The meshes used for G4 and G5 are seen in Figure 11.

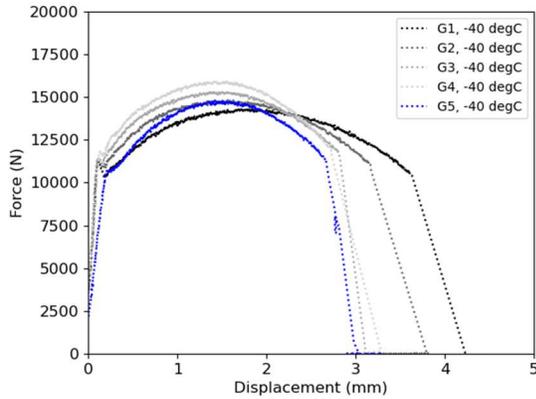


Figure 9. Measured force response of five damage geometries at -40 °C

Figure 10 compares the force versus displacement response of G3 between 24 °C and -40 °C. As expected, the load required for deformation is higher at the lower temperature. Of particular interest was that there was no decrease in the displacement at failure. This was the case for all five geometries, as was also the case in uniaxial tension (Figure 1),

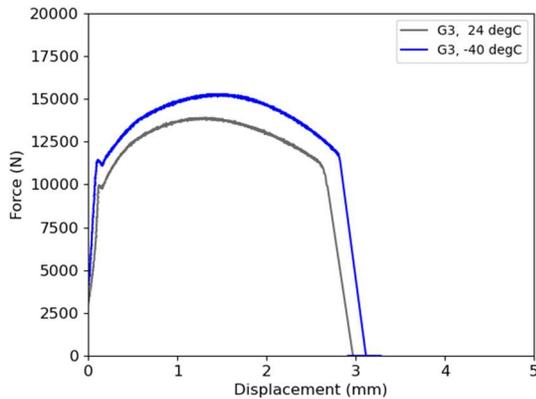


Figure 10. Comparison of measured force response between 24 °C and -40 °C for G3

Unlike the dynamic Charpy test at -40 °C (Figure 3) which displayed the transition from ductile to brittle fracture, none of the five damage geometries displayed this transition. There was no discernable evidence of brittle fracture in any of the tests. This indicates that for TC128B ductile-brittle fracture is active at dynamic strain rates and high stress triaxialities (such as at a crack-tips), whereas ductile fracture is active at quasi-static strain rates at intermediate stress triaxialities (such as at a notch). For impact modelling, this complicates the material model as it becomes necessary to model brittle fracture when preceded by ductile fracture. Though the conditions for brittle fracture might not always be present during impact, from a modelling aspect, it is important that the FEA material model is able to capture the sudden load drop associated with brittle fracture.

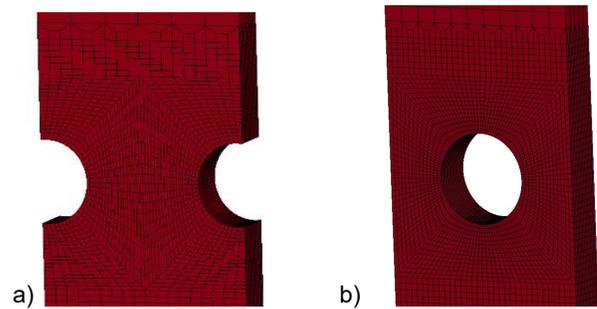


Figure 11. FEA mesh of a) G4 and b) G5

4.2 Predicted vs. Measured Force Response

The predicted versus measured response of G1, G3, and G5 are shown in Figure 12 to Figure 14, respectively. Shown are the responses for 24 °C and -40 °C. The predicted results show good agreement with experiment, which was also the case for G2 and G4 (not shown).

As mentioned in Section 2.3, brittle fracture would be predicted to occur at -40 °C when the maximum principal stress was 2200 MPa. Though this level of stress could be reached at a crack-tip (as with the Charpy specimen), it was below the stress values achieved in any of the five damage geometries. Consequently, brittle fracture was not predicted in any of the geometries. Also, brittle fracture was not readily observed in any of the experiments. However, it is still important to have a ductile-brittle material model in the FEA models that can capture the sudden loss of load bearing capacity associated with brittle fracture.

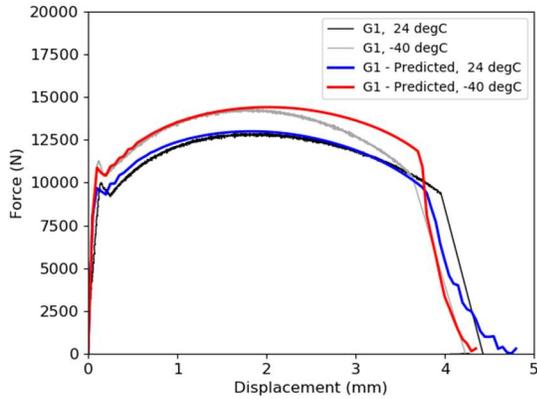


Figure 12. Predicted vs. measured force response at room temperature and -40 °C for G1

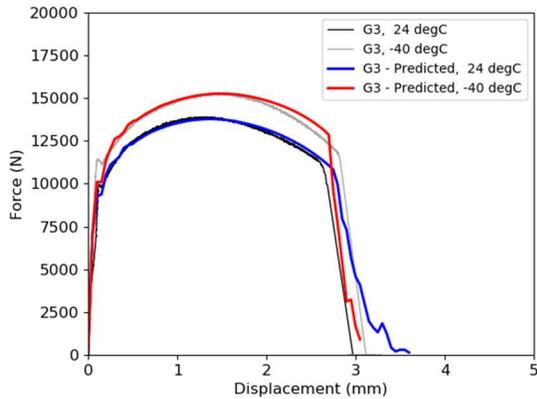


Figure 13. Predicted vs. measured force response at room temperature and -40 °C for G3

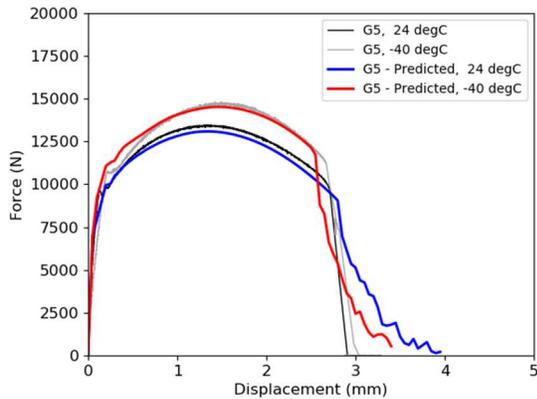


Figure 14. Predicted vs. measured force response at room temperature and -40 °C for G5

5 FULL-SCALE TANK CAR MODELLING – RESULTS AND DISCUSSION

The failure model can be used in large-scale FEA analysis of the tank car structure to simulate impact scenarios such as puncture and crash involving

derailment. In the current work, it was of interest to study hard coupling events. Consequently, the region where the tank is connected to the sill and stub sill was modelled. Bridging the simulation gap between small-scale laboratory size specimens and full-scale tank cars was addressed through a sub-modelling method.

5.1 FEA Model of Full-Scale Tank Car

The structural model of a DOT-117 tank car is shown in Figure 15. The sill (solid region) was connected to the stub sill (shell region) through shell-to-solid coupling. The tank shell and stub sill were connected with tie constraints along weld lines. The solid region of the stub sill was expected to have higher stress and strain than the rest of the stub sill and thus a more refined hex mesh was used for this region. The welds were modeled using the same material properties as the base metal. It is recommended that future work consider the use of a fracture model specifically for TC128B weld material because the weld is expected to have a different fracture response from the steel.

All full-scale tank car simulations were performed in Abaqus/Standard. The MMC fracture model was defined by the damage initiation and evolution routines within Abaqus, with a slight simplification for the damage evolution response compared to the MMC model. Rigid constraint of the tank car in vertical direction (Z) was later found to cause high stress concentration, and it was replaced with elastic constraint to simulate the effect of a tank car suspension. The suspension of the tank car wheel assembly was simulated as an elastic spring element. The bottom surface of the stub sill was tied to one end of the spring element in the vertical direction, and the other end of the spring element was connected to the ground. The spring stiffness was assumed to be 100 kN/mm. A two-step implicit analysis was used. The first step is an implicit static analysis to apply gravity load and the second step is a dynamic implicit analysis to apply coupler forces.

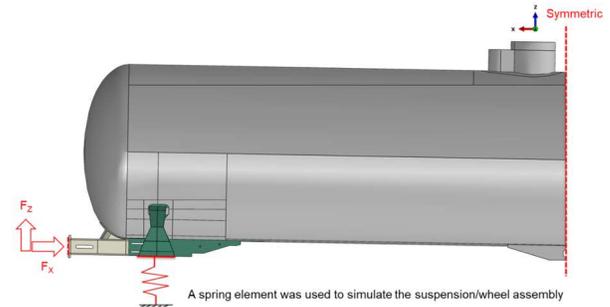


Figure 15. Full-scale tank car model

For balanced efficiency and accuracy, a multi-scale meshing strategy was used. The majority of the tank and stub sill was meshed with 50 mm shell elements which are coarse but computationally efficient. Smaller 10 mm hex elements were used in areas that were expected to have higher stress concentrations. As highlighted in

Figure 16, a high (von Mises) stressed region can be found, typically near the head brace.

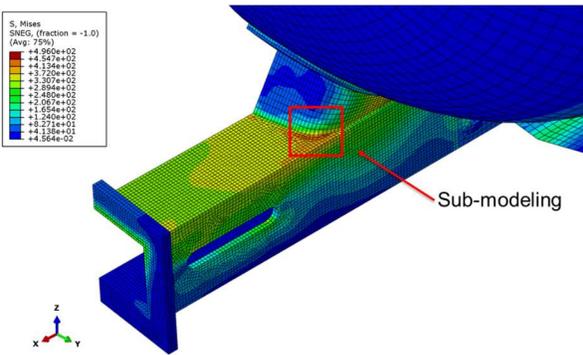


Figure 16: Mesh of the tank car stub sill with 10 mm coarse solid elements (half symmetry shown)

It is important to emphasise that the damage model coefficients are mesh dependent. To circumvent this issue, a sub-modelling approach was adopted where if plastic strain was predicted in the full-scale FEA simulations then the simulation would be repeated in the region of interest with a finer mesh size, as demonstrated in Figure 17.

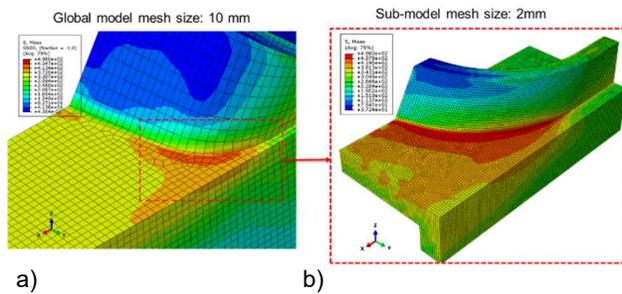


Figure 17: The high stress region in the global model (a) is rebuilt with finer mesh in the sub model (b).

The boundary conditions of the sub-model are read from the simulation results of the global model. The solid elements for the sub model shown have a refined mesh size between around 2 mm but could be further refined to 0.2 mm if damage initiation was expected. The current damage models were calibrated at 0.2 mm mesh size. Sub modeled regions can be re-run without having to run the entire tank car model simulation.

5.2 Hard Coupling Analysis: Results and Discussion

The Federal Railroad Administration (FRA) with the United States Department of Transportation has conducted valuable research in the area of modelling tank car impact scenarios. Much of this work was instrumental to current hard coupling efforts, including details reported by Meymand (2020).

The FRA hard coupling test data was used to validate the finite element model at room temperature. The stub sill location was found to have a significant strain response recorded during the hard coupling test, and thus was chosen to compare with the predicted results from the

FEA analysis. The location of the element in the model corresponding to the location of the strain gauge in the test is highlighted in red in Figure 18.

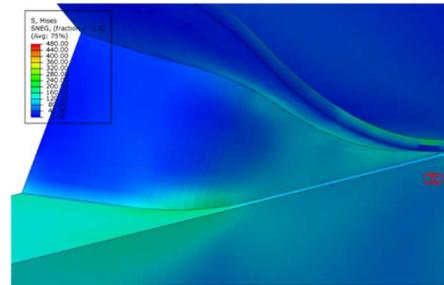


Figure 18: Strain location of interest on the stub sill (red)

Figure 19 compares the strain predicted with experiment for a room temperature hard-coupling event occurring at about 10 km per hour (kph). The maximum load applied in the simulation and experiment was less than 450 ton (1000 kip), the design limit for tank cars. At this load level, the strain induced in the tank car is much less than what is required for damage initiation and might only produce a small amount of plastic strain. Repeating the simulation at $-40\text{ }^{\circ}\text{C}$ resulted in the predicted strain being slightly lower because the structure is slightly stiffer (higher modulus).

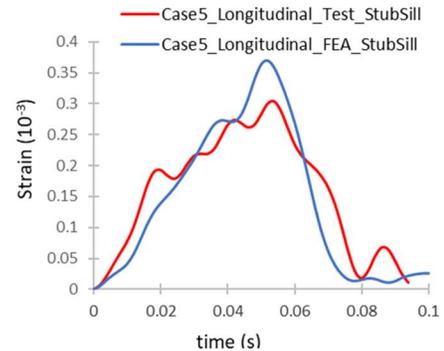


Figure 19: Comparisons of experimental and FEA predicted strain for an impact speed of 10 kph

The results of the full-scale FEA simulations at both $24\text{ }^{\circ}\text{C}$ and $-40\text{ }^{\circ}\text{C}$ did not show failure at the velocities studied (up to 10 kph). To increase the likelihood of failure, a higher stress concentration is required, such as an embedded flaw. Future work should consider study of an embedded flaw in the structure.

6 CONCLUSIONS

The TC128B tank car steel demonstrates the ductile-brittle transition behaviour typical of structural steels, as measured from Charpy V-Notch (CVN) specimens. Charpy specimens contain a V-notch which leads to high stress triaxiality and are tested at a dynamic rate. As these stress states may not be present during impact it was of interest to study lower stress triaxialities. Four

notched tension and one centred-hole specimen were considered at 24 °C and -40 °C at quasi-static strain rates. As expected, there was a slight increase in strength at -40 °C. More interesting was that there was no decrease in ductility at -40 °C compared to tests performed at 24 °C. Unlike the Charpy test at -40 °C, brittle fracture was not readily observed in any of the five geometries with low stress triaxiality and tested at low loading rate. The MMC damage model for the TC128B that was previously developed using only uniaxial tension and Charpy specimens was shown to produce good agreement with experiment for the five geometries further validating the model. The material model was employed in full-scale FEA simulations of hard-coupling events but did not predict fracture. Though fracture was not produced in the current effort, it is still important that the material model be able to capture the sudden load drop associated with brittle fracture at lower temperature, particularly in the presence of a flaw or stress concentration.

7 ACKNOWLEDGEMENT

The authors would like to acknowledge many useful discussions with Mark Gesing and Dr. Bill Tyson of CanmetMATERIALS, and Ian Whittal and Kiran Shoib of Transport Canada. Funding for this work was provided by the Transportation of Dangerous Goods Directorate Scientific Research & Analysis branch at Transport Canada.

8 REFERENCES

Bai, Y. and Wierzbicki, T., 2010. Application of extended Mohr-Coulomb criterion to ductile fracture. *International Journal of Fracture* 161, 1-20.

Carolan, M., Perlman, B., and González III, F., 2018. Validation of puncture simulations of railroad tank cars using full-scale impact test data. In: *Proceedings of the ASME 2018 Verification and Validation Symposium, VVS2018-9322*, May 16-18, Minneapolis, MN, USA.

Eshraghi S. and Carolan M., 2020. Quick calibration of fracture behaviors in TC128 steel for finite element modeling, US Department of Transportation Federal Railroad Association Report #DOT/FRA/ORD-20/46, November 2020.

Kirkpatrick, S.W. and McKeighan, P.C., 2018. Correlating material properties to puncture resistance to enhance the safety and security of tank cars, Final Technical Report for Project TWP-10 – Revision 1, Advanced Tank Car Collaborative Research Program, March 2018.

Meymand S. Impact test data analysis for load environment characterization of tank car stub sill during yard operations. U.S. Department of Transportation, Federal Railroad Administration; 2020.

Paredes, M., Sarzosa, D.F.B., Savioli, R., Wierzbicki, T., and Jeong, D.Y., 2018. Ductile tearing analysis of TC128 tank car steel under mode I loading condition.

Theoretical and Applied Fracture Mechanics 96, 658-675.

Przemyslaw, R., Carolan, M., Gorhum, T., and Eshraghi, S., (2019), Side impact test and analyses of a DOT-117 tank car, US Department of Transportation Federal Railroad Association Report #DOT/FRA/ORD-19/13, May 2019.

Williams, B.W., McKinley J., Xu, S., Xue J., and Spiess, M., 2020. Capturing variability in the fracture response of TC128B steel using damage mechanics, *Procedia Structural Integrity: 1st Virtual European Conference on Fracture* 28, 1024-1038.

Investigating the impact of frost heave and thaw softening on changing the vertical force at wheel/rail interface

Alireza Roghani, Yan Liu, & Paul Burgess
National Research Council Canada, Ottawa, Ontario, Canada

ABSTRACT

The frost heave results in non-uniform deformation and irregularity on railway track. These irregularities increase the dynamic response of the train-track system, resulting in rapid deterioration of track geometry, and poor ride quality. Large surface roughness may cause unloading of wheels and consequently could lead to derailment. This paper presents the results of employing NUCARS® (New and Untried Car Analytic Regime Simulation) software to evaluate the interaction between rail car and the track as it passes through the frost bumps measured over a railway track section during two freeze-thaw monitoring seasons. The vertical wheel/rail forces (force exerted on the wheel by the rail) at each wheel resulting from the simulation is compared against Association of American Railroads (AAR) standard (Chapter 11 of the AAR Manual of Standards and Recommended Practices Section C - Part II) to determine how passing through the frost susceptible sections may affect the safety of train operations. According to AAR Specification M1001 Chapter XI track worthiness limits, the minimum vertical wheel load should be greater than 10% of the static condition. Using the measured track deformation at a study site located on VIA Rail subdivision in eastern Ontario, the minimum vertical force of 72% of the static load was observed as a result of using the track profile measured during thawing season. This value, which is well above the AAR's 10% requirement, occurred at the culvert location where there was large non-uniform deformation. Also, by comparing the results of minimum and maximum vertical force for track in various stages of freezing-thawing cycle, it was observed that the thawing stage is creating the worst combination of the forces.

1 INTRODUCTION

The frost heave results in non-uniform deformation and irregularity on railway track. These irregularities increase the dynamic response of the train-track system, resulting in rapid deterioration of track geometry, and poor ride quality (Cai, et.al, 2019).

To ensure safe passage of trains, railways strive to monitor track conditions during freezing-thawing seasons. During winter, the track inspection is mainly visual and relies on manual measurements while in spring test car is used to measure track geometry parameters over long distances.

According to Transport Canada's Rules Respecting Track Safety (Transport Canada, 2011), the track surface must be maintained to the class 5 level for passenger trains to run at 100 mph. For the class 5 track, the deviation from uniform surface on either rail at the centre of a 62-foot chord may not be more than 1.25 in (or 31.75 mm). Railways normally define a near-urgent defect which is approximately 90% of the TC's threshold values. In the case of the surface bump, the near-urgent limit is 1 in or 25.4 mm.

As measured in winter (Roghani 2021), track surface condition could change considerably in a relatively short period of time caused by freeze-thaw cycles (FTC). Thus changes could come and go without being recorded by track geometry car. This paper aims to evaluate potential safety concern due to the measured surface conditions. The study investigates the effect of track irregularities caused by FTC on the variation of vertical wheel/rail force. The track irregularities have been measured using total station surveys over a 50 m test section during two freeze-thaw seasons. These measurements are imported in a numerical model to estimate the vertical wheel/rail force. These forces are then compared against Association of American Railroads (AAR) standard (Chapter 11 of the AAR Manual of Standards and Recommended Practices Section C - Part II) to determine how passing through the frost susceptible sections may affect the safety of train operations. According to AAR Specification M1001 Chapter XI track worthiness limits, the minimum vertical wheel load should be greater than 10% of the static condition. The other AAR criteria are related to maximum ratio of vertical to lateral force which does not apply to the current study as the test section is located

on a tangent track and as a result has no lateral load generated.

1.1 Study site

A 50 m tangent section of track along one of VIA Rail's subdivisions in eastern Ontario has been used as the test site for this study (Roghani et al., 2019, Le Borgne et al., 2019). This section is only used for passenger trains and consists of 57 kg/m (115 lb/yd) continuously welded rail on ballasted track and wooden ties. Several frost/thaw issues have been reported in the past at this section with the most significant problem occurring towards the east end of the section where three culverts are located. The issues included large frost heave in winter and large dip in track profile during spring. According to track personnel, in some cases in the past they had to put a slow order at this site during winter to ensure the safe passage of the trains.

1.2 Total station survey

The absolute vertical deformation of the test site during the monitoring period was measured using total station. In total, 8 total station surveys were conducted at the test section with the baseline measurements were collected in June 2018. During 2018-2019 freeze–thaw season, the first survey was conducted in November when the track was not started freezing yet. The subsequent surveys of that season conducted in late January and late April when the track was in frozen and thawing conditions respectively. In 2019-2020 monitoring season, 4 surveys were conducted with the first one collected in early December (pre-freezing condition), second and third survey in late January and February (track in frozen condition), and the fourth survey in mid-March (track in thawing condition). Figure 1 presents the results of the total station survey with Figure 1a showing the measurements collected in winter-spring 2018-2019 and Figure 1b showing the results for winter-spring 2019-2020. From these figures, it is evident that the heave not only occurs non-uniformly over each side of the track, but there is also a difference between the deformations of the two sides at each location. The results also demonstrate that the largest heave occurs at the culvert location, with a maximum of 17 mm and 19 mm recorded in winter 2019 and 2020, respectively. On average, the track outside of the culvert zone heaved about 6.9 and 2.6 mm during winter 2019 and winter 2020, respectively (Roghani, 2021). It needs be noted that these numbers should not be compared against the Transport Canada threshold values for profile measurements as the measuring methods for the two datasets are different.

The lower frost heave during 2020 winter is consisted with its warmer ambient temperature, shorter freezing season, and less frost penetration depth compared to winter 2019 (Figure 2). In addition, during 2020 winter, track embankment at the test site experienced 4 intermittent FTC compared to 1 FTC observed during 2019 winter. According to the predictions by climate scientists, winter similar to 2020,

is expected to become the new normal. Table 1 compares the ambient and embankment temperature during the two winter-spring season.

Table 1. Comparing ambient and embankment temperature between two monitoring periods.

Sensor	2018-2019	2019-2020
Start of freezing season ¹	Jan 9, 2019	Jan 18, 2020
Start of thawing season	Mar 15, 2019	Feb 25, 2020
Maximum frost depth	1.0 m	0.5 m
Average temperature ²	-7.5 °C	-4.1 °C

¹ When the thermistor at 0.35 m below track surface recorded 0 °C.

² Average ambient temperature between December and March.

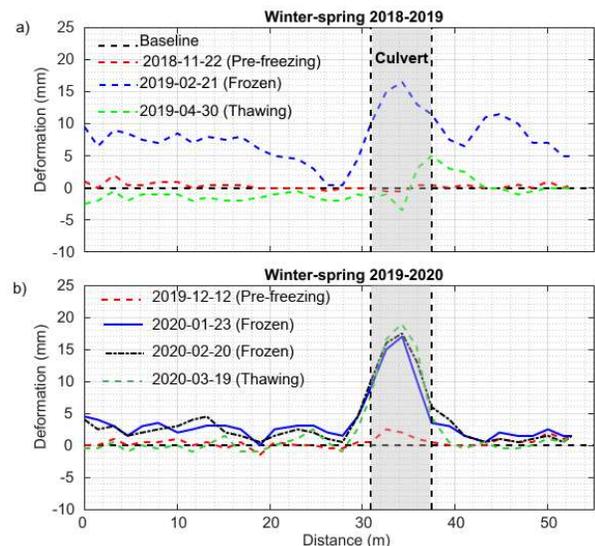


Figure 1. Plots showing track surface deformation during a) winter-spring 2018-2019 and b) 2019-2020 collected by total station surveys (after Roghani, 2021).

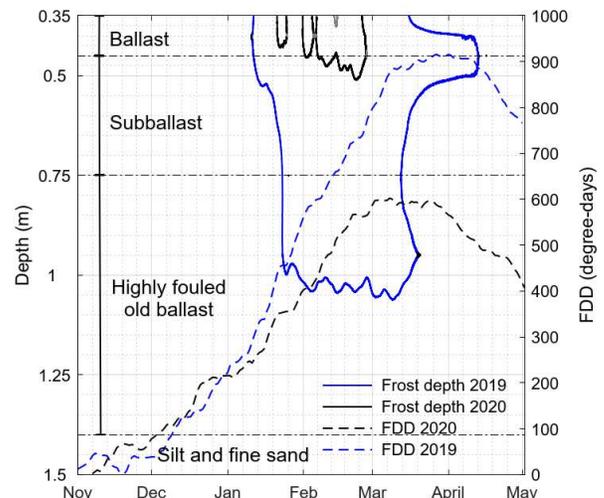


Figure 2. The variation of FDDs and the frost depth during winter 2019 and winter 2020 (after Roghani, 2021).

2 DEVELOPMENT OF NUMERICAL MODEL

This section presents the model developed in NUCARS® (New and Untried Car Analytic Regime Simulation) software to evaluate the interaction between rail car and the track as it passes through the frost bumps measured during each of 2019 and 2020 survey over the test site (Figure 1). NUCARS® is a software developed by Transportation Technology Center, Inc. (TTCI) for modelling rail vehicle transient and steady-state response. This software is an industry-accepted simulation tool used to examine the dynamic performance of rolling stock on track with varying types and levels of perturbation. The following sections discuss the input data into the simulation.

2.1 Rail car model

The rail car model used for this analysis was borrowed from a previous project based on the characteristics of the 4-axle passenger coach car currently used in VIA service. The coach car is modelled as a rigid H-frame bogie equipped with primary and secondary suspensions. The primary suspension is comprised of axle boxes that are connected to the bogie frame through axle box radius rods. There are two coil springs per axle box (four per axle). There are also two primary vertical hydraulic dampers per axle box (four per axle). The primary coil springs are modelled as springs with longitudinal, lateral, vertical, and rotational (roll) stiffness. The secondary suspension uses two air springs per bogie. Each air spring is simplified as a linear spring and a linear series spring-damper.

In the present simulation, the empty coach car with a “worn” suspension was used. The “worn” condition was simulated by reducing damping values by 20%. The gross rail load or the total weight of the simulated coach car is 109,250 lbs. This gives a static wheel load of 13,656 lb (13.7 kips).

2.2 Wheel/rail contact geometry

Since present simulations on the tangent track focuses on the vertical response due to frost heave, the wheel/rail contact details such as tread taper and flange contact angle will have little effect to results. Therefore, a wheel/rail contact file from NUCARS library was used in the simulation (A1b136104.wnt). The wheel back-to-back spacing was set at 53.375 in.

2.3 Track geometry input data

The track geometry parameters imported in this model include: Gauge, Alignment, Crosslevel, and Profile. Figure 3 shows the gauge and alignment data used as import data into the simulation. The gauge measurements shown in Figure 3a represent deviation from 56.5” standard gauge (positive means wider gauge) and Figure 3b represents deviation from a perfectly straight rail line. The profile and crosslevel were calculated using the survey measurements (Figure 1). It

should be mentioned that as there were no geometry measurements available for each survey, the gauge and alignment variations were assumed to be same for all the survey dates. This assumption is expected to have an insignificant effect on the main simulation results. This is because the gauge and alignment only have a very limited impact on vertical dynamics of vehicle track interaction. On the other hand, crosslevel and profile variations provide the main excitation to the vehicle’s vertical response.

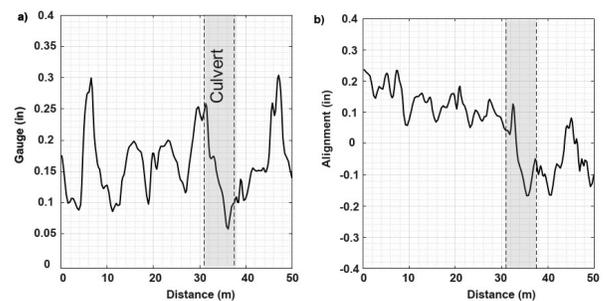


Figure 3. Plots show a) gauge and b) alignment variation over the test section imported in the NUCARS software.

3 RESULTS AND DISCUSSION

The simulation was conducted for each set of profile measurements collected during surveys (Figure 1) and repeated for speed ranging from 40–100 mph, in 20 mph increments. The speed range was selected to define the track resonance behaviour (if any). The variations of vertical force of each wheel on the passenger coach car were then extracted for each simulation. These results are presented in Figure 5, for surface deformations recorded in 2019, and Figure 6, for surface measurements recorded in 2020 (the relative location of axles and wheels are shown in Figure 4). Each plot shows the variation of the vertical force, the magnitude and location of minimum and maximum load, and the static load. It needs to be noted that vertical force for wheels of the first and second axles are presented in these plot. Similar plots were generated for axles 3 and 4.

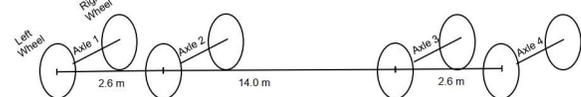


Figure 4. The location of axles and wheels used in simulation.

Figure 5 shows the variation of vertical force for each wheel as the train passes through the track profile measured during 2019 surveys. According to this plot, the minimum force is approximately 10.7 kips and occurs as a result of the track profile measured in 2019-04-30 (track in thawing condition). This value is about 79% of the static load and thus it is well within the safe limits. It is also evident that the minimum load is occurring within the proximity of the culverts. The maximum vertical load

of 17.2 kips also happens during 2019-04-30 survey measurements which is about 25% higher than the static load. Even though AAR does not specify any safety limits for maximum vertical load, the higher forces indicate higher stress on track components and rolling stock that may shorten their service life.

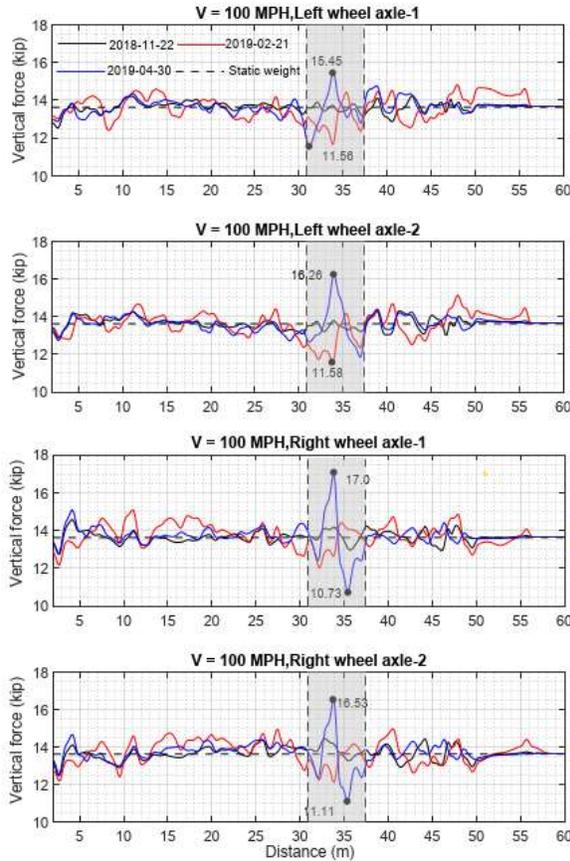


Figure 5. The variation of vertical force along the test section resulted from using 2019 survey measurements in NUCARS simulation. Note: The shaded area is the approximate location of the culverts.

Figure 6 compares the force variation as a result of 2020 profile measurements for a 100 mph operating speed. The minimum and maximum observed force is about 9.9 and 17.4 kips which is approximately equal to 72% and 128% of the static load, respectively. These values occur as a result of 2020-03-19 survey profile measurements in which the track is in a thawing stage.

Similar simulations were conducted for operating speed of 40, 60, and 80 mph and the results presented in

Table 2. Even though all the values are well within the AAR safety limit (greater than 10 percent of static load), they clearly show the effect of frost heave on changing the vertical force at wheel/rail interface. It should also be mentioned that for all these simulations, the worst condition occurred at the culvert location where the maximum frost heave was measured. No resonance effect was observed within the simulated speed range. The train speed plays an important role on the wheel unloading when passing the frost heaved zone. The

higher the speed, the higher the unloading risk. The effect of speed is more clearly demonstrated in Figure 7 where the change in the difference between maximum and minimum force at various operating speed is shown.

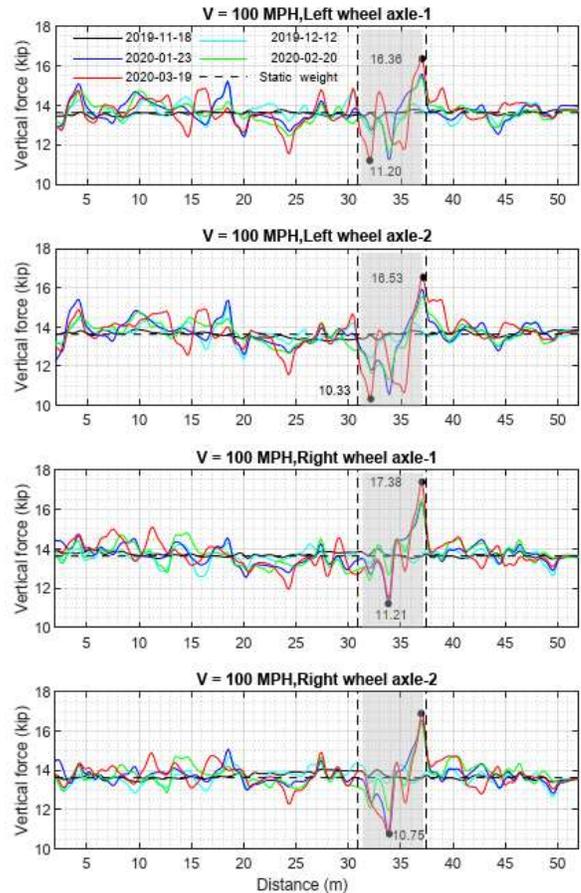


Figure 6. The variation of vertical force along the test section resulted from using 2020 survey measurements in NUCARS simulation.

Table 2. Summary of NUCARS simulation results (units in kips). Note: static vertical force is equal to 13.7 kips.

Track condition	Speed							
	40 mph		60 mph		80 mph		100 mph	
	Min VF	Max VF	Min VF	Max VF	Min VF	Max VF	Min VF	Max VF
2018-11-22 (Pre-freezing)	12.8	14.5	12.9	14.4	12.8	14.4	12.5	14.6
2019-02-21 (Frozen)	12.3	14.8	12.5	14.6	11.9	14.8	11.3	15.2
2019-04-30 (Thawed)	12.2	15.2	11.8	14.9	11.5	16.1	10.7	17.2
2019-12-12 (Pre-freezing)	12.9	14.2	12.9	14.5	12.5	14.8	12.2	15.2
2020-01-23 (Frozen)	12.2	14.9	12.0	15.3	11.2	15.9	10.3	16.6
2020-02-20 (Frozen)	12.0	14.8	12.1	15.23	11.8	15.9	11.2	16.5
2020-03-19 (Thawing)	12.12	15.5	11.4	15.8	10.7	16.5	9.9	17.4

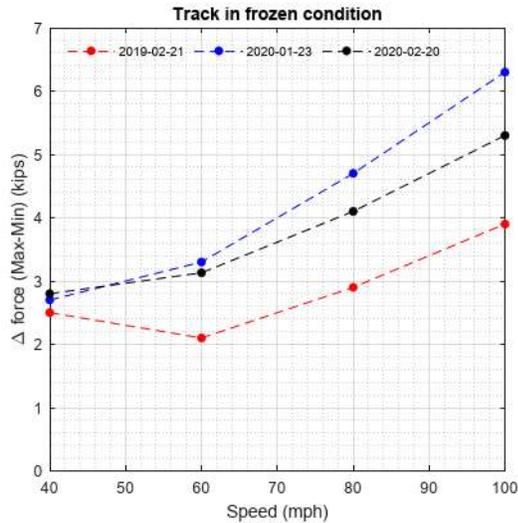


Figure 7. Plot compares the difference between maximum and minimum vertical force while track was in frozen condition.

Figure 8 compares the minimum and maximum vertical force for track in pre-freezing, frozen, thawing, and after thawing conditions. These results suggest that among the various states of track, the thawing stage is creating the worst combination of forces.

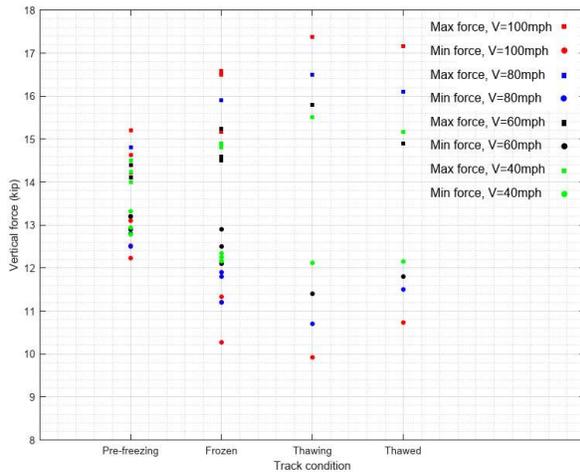


Figure 8. Plot shows the variation of minimum and maximum vertical force for various operating speed and track condition.

4 CONCLUSIONS

A numerical analysis was performed in NUCARS® software to quantify the effect of the non-uniform track deformation measured at the test site on the vertical force at the each wheel of a VIA's coach car. The simulation results were compared against the American Association of Railroad (AAR) specification that requires the minimum vertical load should be greater than 10 percent of static load (to prevent wheel unloading). The worst (minimum) vertical force of 72% of the static load was observed as a result of using the track profile

measured during 2020 thawing season. This value, which is well above the AAR's 10% requirement, occurred at the culvert location. Also, in comparing the results of minimum and maximum vertical force for track in freezing, frozen, thawing, and after thawing condition, it was observed that for these various states of track, the thawing stage is creating the worst combination of the forces.

Overall, the results indicated that 2020 winter created a worse (but still safe) operating condition than 2019. This can be attributed to larger differential deformation resulted from more FTC in winter 2020 and a stronger weakening effect caused by higher moisture content. The simulation results are also consistent with measurements from track geometry car which suggested that track was in poorer conditions in spring 2020 than spring 2019.

The limited numerical modelling conducted in this project suggested that even though the warmer and drier winter (similar to winter 2020) may result in a shorter freezing period and shallower frost depth, the higher number of intermittent FTC lead to larger differential frost heave in winter and higher vertical force at the wheel/rail contact point. This may lead to a condition that is more detrimental for track geometry and safe operations than a cold winter. According to the predictions by climate scientists, this type of winter, warmer and drier with frequent FTC, is expected to become the new normal and therefore, frequent maintenance may be required to keep railway tracks constructed in seasonally frozen ground within the safe limits. The failure to account for these changes may lead to the increased costs to infrastructure owners, unforeseen costs to infrastructure users, and considerable negative socioeconomic impacts. The research conducted in this project was limited in scope and was performed over a well-maintained class 5 track in Ontario. The performance of a lower-class track in a colder region of Canada may be different. In the authors' opinions, further research and development is required.

It also needs to be reminded that the results presented in this paper represent the VIA's coach car response to the limited observations measured at a very well maintained site during winter 2019 and 2020 and thus the results should not be generalized.

5 ACKNOWLEDGEMENT

The author would like to acknowledge Infrastructure Canada for providing funding for this study and VIA Rail Canada for providing data. The author also would like to thank Paul Charbachi from VIA rail Canada, Marianne Armstrong, Robert Caldwell, Fiona Hill, Dr. Jon Makar, Derek Morrell, and Albert Wahba from National Research Council Canada for supporting this project.

6 REFERENCES

Cai, X., Liang, Y., Xin, T., Ma, C., & Wang, H. (2019). Assessing the effects of subgrade frost heave on

vehicle dynamic behaviors on high-speed railway. *Cold Region Science and Technology*, 158, 95-105.

Le Borgne, V., Roghani, A., & Hiedra-Cobo, J., Thivierge, S.-É. T, Charbachi, P. (2019). Design and Installation of a Geotechnical Monitoring System for Monitoring Freeze-Thaw Cycles on a Railway Track. *Cold Region Engineering* 2020:153–60. <https://doi.org/doi:10.1061/9780784482599.018>.

Roghani, A. (2021). Quantifying the effect of freeze-thaw cycles on track surface deformation and degradation of railway track geometry; Case study. *Journal of Transportation Geotechnics*, <https://doi.org/10.1016/j.trgeo.2021.100601>.

Roghani, A., Hiedra Cobo, J., & Charbachi, P. (2019). Studying the impact of freeze-thaw cycles on performance of railway tracks. *International Heavy Haul*. Narvik, Norway.

Transport Canada. (2011). Rules respecting track safety. <https://www.tc.gc.ca/media/documents/railsafety/track-safety-2012en.pdf>.

Relationship between fracture toughness and average grain sizes of high strength rail steels

Stephen Okocha, & P.-Y. Ben Jar

Department of Mechanical Engineering, University of Alberta, Edmonton, Alberta, Canada

Feng Yu,

School of Mechanical Engineering and Mechanics, Ningbo University, PR China

Michael Hendry

Department of Civil and Environmental Engineering, University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

In metals, the microstructure plays a significant role in defining the deformation behaviour and mechanical properties. Sometimes heat treatment is used to enhance mechanical properties by adjusting their microstructural parameters. Since fracture toughness is one of the most important mechanical properties, concerns about how microstructure is related to fracture toughness are considered, especially in areas where cold temperatures and operational conditions can cause unexpected fracture. Hence, various approaches have been studied to use microstructures to predict fracture toughness of metals. In this paper, studies on 9 rail steels were conducted to establish the relationship between the average grain size and critical stress intensity factor (K_{IC}). By applying visual inspection (VI) techniques, the average grain sizes for the 9 rail steels were measured using the mean line intercept method in the rail head regions. Results show that the average grain size gives a linear correlation to the square of K_{IC} , which permits the prediction of K_{IC} using models like the critical fracture strain model that takes into account the rail steels' material properties, ductility and characteristic distance. Emphasis on modifying the model will be demonstrated, offering suggestions on different materials and technologies required for estimating K_{IC} other than the ASTM approach. Also, discussion is provided concerning different heat treatment procedures for hardening of rail heads and their effect on the average grain sizes.

1 INTRODUCTION

The microstructure has been shown over decades to determine the mechanical properties, plastic deformation behaviours and fracture properties of metals (Hahn & Rosenfeld 1973, Bramfitt 1998). With the role of continuum mechanics in defining fracture mechanics, it is expected that estimation of mechanical and fracture properties would present some limitations because of the non-inclusion of microstructural parameters in its assessment (Swable 1977). This can be observed in the scattered results of critical stress intensity factors (K_{IC}) estimated via the American Society for Testing and Materials (ASTM) E399 and E1820 approaches as noted by Sugden & Bhadeshia (1989), who attributed the scattering K_{IC} s to microstructural entropy while Wallin (1984) suggesting cleavage fracture in steels is of a statistical nature. Moreover, the conventional approaches for K_{IC} based on the ASTM E399 and E1820 are time cumbersome in specimen preparation, costly to achieve due to the requirements of several specimens and labour intensive majorly due to the stringent fatigue

pre-cracking requirements. This paper investigates the correlation of microstructural parameters in terms of the average prior austenite grain sizes with the estimated K_{IC} for high strength rail steels according to ASTM approaches and offers analytical models and techniques with the inclusion of microstructural parameters that can be used for K_{IC} estimation. This analytical method for K_{IC} estimation is proposed without the dependency on the conventional pre-fatigue cracking used in ASTM standard. Nine (9) rail steels are investigated microstructurally, finding connections with K_{IC} determined via the ASTM approach and also examining observations based on the heat treatments subjected on the rail heads with possible connections with K_{IC} values.

Hahn & Rosenfeld (1973) explains that the study and measure of fracture toughness, microstructurally, originates from the study of one or more different regions of a fractured surface. These regions involve: the highly distorted lattice (analogous to the dislocation core), region of large plastic stretch, the plastic zone and the elastically strained region. They observed that fractured surfaces of unnotched tensile metallic bars displayed a

fibrous, dimpled appearance which signifies that the failure process followed a sequence of plastic deformation, nucleation of voids from hard particles within the ductile matrix, growth of the voids and the subsequent coalescences of the voids, which transitioned from an internal microcrack to a visually seen macroscopic crack. The K_{IC} was then modelled to contain microstructural parameters like the volume fraction of cracked particles, v_c , and the particle diameter, d^* along with its mechanical properties. In 1975, Hahn & Rosenfeld explained that K_{IC} varies with $v_c^{-1/6}$ for metals with comparable yield strengths. However, by this time, it was already noted that K_{IC} decreases as yield strength increases (Zinkham et al. 1968), which can be an influence either by alloyed metals or heat-treated metals. Yamada (1975) also noted that refinement of the prior austenite grain sizes for fine pearlitic steels transformed at relatively low temperatures improved the ductility of steels. Hence, the interest in the study of microstructural parameters to fracture toughness rose leading many industries especially the rail industry in researching of ways in improving the quality of rail steels. Swable (1977) describes the difficulty on the influence microstructure on fracture toughness and crack propagation due to the several parameters associated with microstructure of metals. Some of these parameters are linked to the influence of microstructural type, alloying elements, base matrix, inclusions etc. Nevertheless, Swable (1977) identified that voids for ductile fracture initiate from inclusions (non-metallic alloying compounds) and the spacing between inclusions which corresponds to the dimple size play a vital role in determining the K_{IC} of metals. Carlson & Ritchie (1977) also studied the effect of prior austenite grain sizes from 30 to 180 μ m which were attained by direct oil quenching followed by a one-hour austenitizing between 870 and 1200°C. They found that increasing the prior austenite grain size increases K_{IC} but further increase causes a drastic drop in K_{IC} values. The increase in grain size also increases the fatigue crack propagation in low strength steels but causes a decrease in high strength steels.

Another study on the cleavage fracture in pearlite eutectoid steels where the prior austenite grain sizes and the pearlite interlamellar spacing were varied by means of appropriate heat treatments (Alexander & Bernstein, 1989) were explored. The study suggests that the effective surface energy for cleavage fracture is independent of both the austenite grain sizes and the pearlite interlamellar spacing when studying the size of the void nucleated at the fracture zone. These outcomes were consistent with some works that suggests of an increase in prior grain size was independent of K_{IC} or in some cases reduced K_{IC} (Wen-Lan Li & Li 1989, Gomez et al 1997, Wang et al 2016) while Qiu et al. (2014) studied the fracture toughness in terms of crack tip opening displacement (CTOD), attributing an increase in ferrite grain size linearly increases CTOD but simultaneously increases the possibility of cleavage fracture. Yu et al. (2017), suggests that fracture toughness is largely controlled by the microplasticity zone/region as suggested by Hahn & Rosenfeld (1973) which is (1-2) the critical CTOD (δ_c), stating that an

increase in prior austenite grain size has little changes with K_{IC} but rather affects the nature of crack propagation depending on the prior austenite grain size and whether it is larger or smaller than the microplasticity zone.

This unclear relationship between prior austenite grain size and fracture toughness (and fracture properties) can arise sometimes from extrapolating the outcomes attained from steels which have little or no pearlite structures with those of fully pearlitic structures as seen in many rail stocks and rail wheel (Lewandowski & Thompson, 1985). Another important factor is the role of inclusions and its density within steel which dictates the inter-particle spacing between voids (Broek 1973, Srivastava et al. 2014) while Ritchie et al. (1979), Rosenfeld (1968) and Rice & Tracy (1969) attribute the ductility killing parameter, stress triaxiality, as a major factor that influences void nucleation and growth which consequentially affects the fracture toughness. Tvergaard (1989) on the other hand, accounts void growth to plastic straining of the surrounding material within the matrix of the microstructure.

The methodology of work presented in this paper requires performing tensile tests of unnotched dog-bone specimens to estimate both mechanical and fracture properties of the high strength rail steels along with 3-point single edged notch bending tests to estimate the K_{IC} in accordance with ASTM E399 procedures, it is to note that this has been performed in previous studies (Yu et al. 2017a, b, Yu et al. 2018). The microstructure of the rail steels was revealed using an optical microscope and scanning electron microscope (SEM) to account for the prior austenite grain sizes and the nature of the microstructure particles respectively for the rail steels. Majority of the tests and analysis are done on the rail head as these are the point of contact from the railway car axle while some are done on the rail foot. Based on the heat treatments done on the rail heads, suggestions based on categorization of grain sizes was correlated with K_{IC} . The results show that K_{IC} has a linear correlation to the square root of average grain diameter. With this in mind, an analytical model was modified from the fracture strain model that can be used for K_{IC} estimation without the need of pre-fatigue cracking. This modified model puts into consideration microstructural parameters along with stress triaxiality effect.

2 MATERIALS AND METHODS

The rail steels used in these experiments are high strength rail steels. Table 1 shows the rail steels provided by the Canadian National (CN) Railway Company and the Transportation Technology Center Inc. (TTCI). A total of nine high-strength rail steels were used in generating the required data for estimating the tensile and fracture properties. The table also shows the different heat treatments included in the manufacturing process for the rail steel.

The mechanical properties of the rail steels were attained in previous studies using ASTM E8/E8M approach, which require dog-bone smooth tensile rods while the fracture toughness is attained using 3-point single-edge notched bending (SENB) specimens based

on ASTM E399 procedures (Yu et al. 2017a, b, Yu et al. 2018). The parameters attained consist of the Modulus (E), yield strength (σ_{ys}) and ultimate tensile strength (UTS), strain hardening exponent (n), fracture strain (ϵ_f) and K_{Ic} .

Table 1- Rail samples and their microstructures

Rail Name	Microstructure	Locations	Company
JP	Deep head hardened	H ¹ F ³	CN
EV	Deep head hardened	H ¹ F ³	CN
CZ	- ₂	H ¹ F ³	CN
Rail #2	Hypereutectoid	H ¹	TTCI
Rail #3	Hypereutectoid	H ¹	TTCI
Rail #4	Head hardened perlite	H ¹	TTCI
Rail #5	Head hardened perlite	H ¹	TTCI
Rail #6	Micro head hardened	H ¹	TTCI
Rail #Control	Micro head hardened	H ¹	TTCI

- 1- Rail Head
- 2- No heat treatment operation
- 3- Foot of rail

In order to reveal the microstructure of the rail steels, firstly, samples were extracted from the rail head (13mm X 7mm) of 13 mm in thickness and firmly held with epoxy-resin mixture with just one location from the foot for JP, EV and CZ rails. The surface was then grounded and polished using SiC paper grits (240-1200mm) and diamond suspension (1mm) respectively until a mirror-like surface is attained. Then, the surface was etched using 2% Nital etchant for an average of 15 seconds via the dipping (immersion) procedure and then washed using distilled water. Fig 1. Shows the images of the microstructure under the optical microscope (OM) and scanning electron microscope (SEM). The optical microscope (OM) images were used to calculate the average prior austenite grain sizes as it was almost impossible to see the perlite colonies within the grain sizes while the SEM images were used to identify the nature of the microstructure which involved: perlite structures, austenite and inclusions within the microstructure. Since the different rail steels were heat treated at different austenitizing temperatures, each large grain demarcated by grain boundaries viewed from the OM was taken to represent the prior average grain sizes.

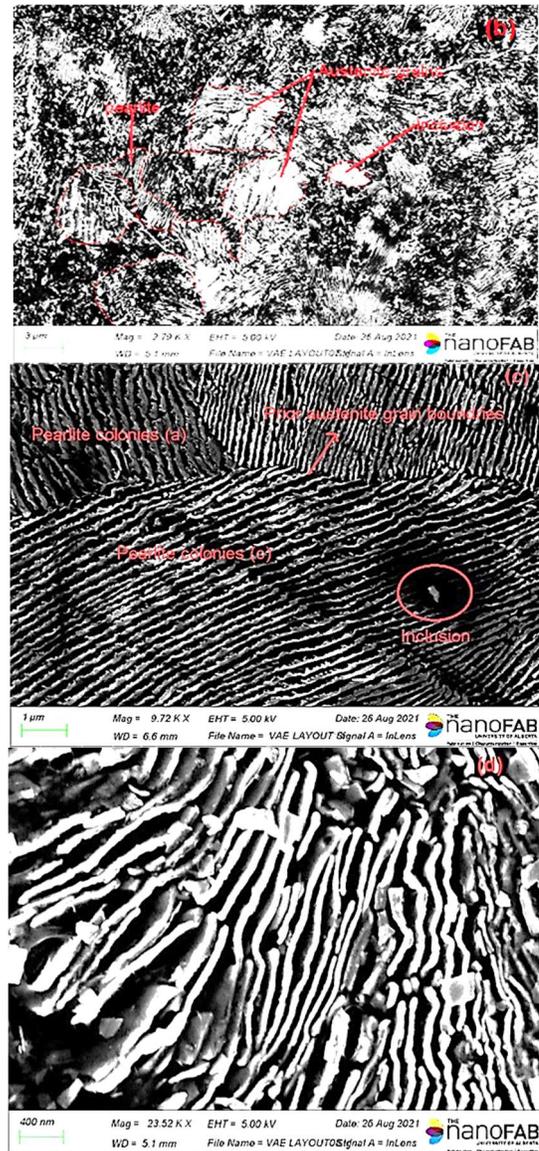


Fig. 1. Microstructure of rail samples: (a) Image viewed via OM showing grain sizes, (b) perlite and austenite sections in a prior austenite grain, (c) prior austenite grain boundaries demarcating perlite colonies with the presence of an inclusion, (d) perlite structures of cementite and ferrite.

Figure 2 shows the test specimen's surface via OM after the rail steels' etching process. Each sample collected from the head of the rail steels were examined using 5 fields of optical images, each provided with 5 random lines of equal lengths to determine the average grain sizes on three different locations per rail sample. Hence a total of 27 samples from 3 locations of the 9 rail steels, 135 optical field images and 675 lines were utilized in total. The prior austenite grain size was calculated using the mean intercept length method (\bar{l}).

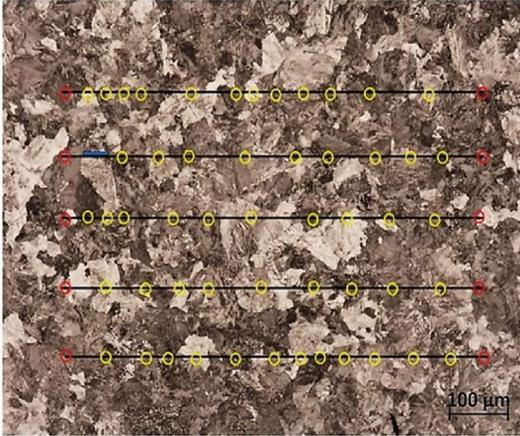


Fig. 2. Average grain diameter determination using Mean intercept line method

Eqn. (1) describes the relationship between \bar{l} and the number of grains intercepted (N) following the ASTM E112 guidelines while Eqn. (2) shows the relationship between the mean intercept length and the ASTM grain Number, G. The ASTM E112 shows the relationship between the average grain diameter, d^* and G from the ASTM number chart, where d^* can be attained via interpolation (see ASTM E112)

$$\bar{l} = \frac{L}{MN} \quad [1]$$

where L is the total test lines and M the magnification of the test field used in each optical image

$$G = (-6.6643858 \log_{10} \bar{l}) - 3.288 \quad \bar{l} = mm \quad [2]$$

3 RESULTS AND DISCUSSION

The summary of the tensile and fracture results is shown in Table 2, which were attained from the tensile and 3-point SENB test. It can be seen that CZ rail has the lowest yield strength, fracture strain (ductility) and K_{IC} because no heat treatment processes were done on this rail steel. Other rails which were involved in heat treatments processes exhibited higher strength and fracture toughness.

Table 2. Mechanical and Fracture properties of the rail steels.

Rail Name	E (GPa)	σ_y (MPa)	n	ϵ_f	K_{IC}
JP	200	820	0.086	0.504	39.50
EV	197	714	0.085	0.301	41.50
CZ	193	632	0.082	0.183	34.10
Rail #2	200	941	0.032	0.260	37.85
Rail #3	200	925	0.039	0.241	35.28
Rail #4	185	943	0.050	0.466	41.16
Rail #5	190	936	0.010	0.638	41.40
Rail #6	220	887	0.054	0.514	42.34
Rail #Control	200	870	0.050	0.381	42.28

Following the procedure for estimating the average grain diameter using the mean intercept length method, the prior-austenite grain sizes were estimated for the 9 rail steels as shown in Table 3. The heat-treated rail steels improved the size of the prior-austenite grain diameter as

compared to CZ rail steel which was not heat treated as seen in Table 3. These also enhanced the K_{IC} values along with the yield and tensile strength. Among the heat-treated rail steels, the hypereutectoid heat treatment (Rail #2 and Rail #3) significantly improved the yield and tensile strength although the ductility was not greatly improved as seen by the ϵ_f when compared to CZ in Table 2. Other heat-treated rails (Head hardened perlite and Micro head hardened) showed significant improvement in strength, ductility and K_{IC} as compared with CZ rail steel.

Fig. 3 shows a relatively linear relationship between the ASTM K_{IC} and square root of d^* .

Table 3. Different microstructural outcome for the 9 rail steels

Rails Steels	Location	Microstructural parameters		
		\bar{l} (mm)	G	d^* (μm)
JP	A (H)	43.692	5.746	58.1
	B (H)	44.037	5.723	58.3
	C(F)	42.105	5.853	58.1
EV	A (H)	69.199	4.419	78.8
	B (H)	69.479	4.407	79.1
	C (F)	67.938	4.471	77.4
CZ	A (H)	51.248	5.285	49.2
	B (H)	51.444	5.277	49.6
	C (F)	51.230	5.290	47.4
Rail #2	A (H)	43.887	5.733	49.9
	B (H)	44.148	5.716	50.2
	C (H)	43.411	5.764	49.3
Rail #3	A (H)	40.812	5.950	45.6
	B (H)	38.345	6.122	43.2
	C (H)	38.151	6.138	42.9
Rail #4	A (H)	52.563	5.213	59.2
	B (H)	51.259	5.285	57.7
	C (H)	52.092	5.238	58.7
Rail #5	A (H)	51.497	5.272	58.4
	B (H)	51.128	5.295	58.0
	C (H)	49.253	5.400	56.1
Rail #6	A (H)	51.267	5.285	57.7
	B (H)	51.943	5.248	58.5
	C (H)	52.536	5.214	59.2
Rail #Control	A (H)	55.586	5.051	63.3
	B (H)	56.336	5.012	64.2
	C (H)	55.842	5.038	63.6

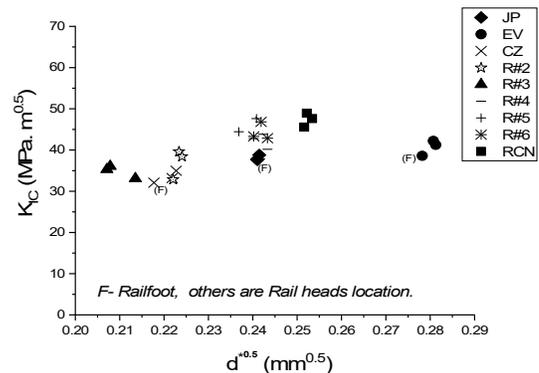


Fig.3 Prior-austenite grain size relationship with ASTM K_{IC} of nine different rail steels

The relationship between K_{IC} and \sqrt{d} suggests that grain size alone cannot estimate K_{IC} but rather suggests a linearity between grain size and K_{IC} as suggested by some

other studies (Carlson & Ritchie 1977, Ritchie et al. 1973, Qiu et al. 2014). However, according to Ravichandran & Vasudevan (2013), an increase in grain size increases K_{IC} for austenitic and ferrite steels but is opposite for aluminium alloys due to coarse grain boundary precipitation. In order to account for other factors, a study of the fracture void coalescence mechanism is needed.

As Hahn & Rosenfeld (1973) pointed out, the different regions of the fractured surface provide different parametric factors that constitutes fracture occurrence. Firstly, the elastically strained region far away from the fractured and plastic zones requires the rigidity of the material which involves E , the extent of stiffness of the material before plastic deformation initiates. The region of large plastic stretch requires that the knowledge of the plastic strain (ϵ_p) be known. As ϵ_p provides ductility limits of the material and relates to the size of the plastic zone. As Irwin proposes, that plastic zone develops at the tip of the crack/flaw which increases in size as more load is added until the crack initiates/grows causing the elastically strained material behind the crack tip to unload (Erdogan 2000). Since the crack tip opening displacements, δ , relates to the energy required to initiate a crack, Shih (1981) established the relationship between the J-integral, J , and δ derived from the near-tip Hutchinson, Rice, Rosengren (HRR) singular crack fields as shown in Eqn. (3)

$$\delta = d_n \frac{J}{\sigma_y} \quad [3]$$

where d_n is a material constant depending on the ratio of E to σ_y for both plane strain and plane stress (Shih 1981). When δ becomes critical, δ tends to $\delta_{c,n}$, and relates to the fracture toughness, K_{IC} under small scale yielding (SSY). The condition which most structural members of low and high strength alloy steels obey. Hence, based on the SSY criteria, J_{IC} can be represented by Eqn. (4) as described by Rice (1968).

$$J_{IC} = \beta \frac{K_{IC}}{E} (1 - \nu^2) \quad [4]$$

where β is a material constant and ν the Poisson ratio. Combining Eqn. (3) on the basis of $\delta \rightarrow \delta_c$ and Eqn. (4), The critical fracture energy model can be deduced as seen in Eqn. (5) (Garrison 1984).

$$K_{IC} = \frac{1}{d_n} \sqrt{E' \sigma_y \delta_c} \quad [5]$$

However, to include microstructural influence, the blunting solution of Rice & Johnson (1970) explains that as the plastic zone develops, ϵ_p increases along with a microstructural distance, x , within the microstructure where voids are initiated. When this microstructural distance, x (distance between voids), becomes critical, it becomes the characteristic distance, l_o^* , which relates to the equivalent plastic fracture strain, ϵ_f^* as well as the δ_c as shown in Eqn. (6). At this critical point, fracture takes place. it is important to note that the characteristic distance, l_o^* , as well as ϵ_f^* independently depends on the stress triaxiality, η (ratio of mean stress to equivalent stress) and defines the extent to which the plastic zone

develops before fracture. i.e, high stress triaxiality amounts to a smaller ϵ_f^*

$$\epsilon_f^* = \alpha_m \frac{\delta_c}{l_o^*} \quad [6]$$

Hence, by combining Eqn. (5) and Eqn. (6), the critical fracture strain model is established.

$$K_{IC} = C_m \sqrt{E' \sigma_y l_o^* \epsilon_f^*} \quad [7]$$

In order to account for triaxiality similar to that attained in an SENB experiment, the fracture strain or plastic strain is extrapolated using a high triaxiality effect similar to that attained while performing SENB experimentation. Models like the Gurson–Tvergaard–Needleman (GTN) (1984) or Bonora’s model (1997) reduce the strain to fracture or plastic strain due to a high triaxiality effect with and without damage necessities, respectively.

Ritchie et al. (1979), described l_o^* as a multiple of the grain sizes (4-7 for ductile fracture), while Rice & Tracy (1969) from a modification of McClintock’s model (1968) relates l_o^* to the growth of voids at the critical state and η . Since voids grow from inclusion as shown in Fig.1c, it is assumed that the distance between two voids, which is the interparticle spacing corresponds to the diameter of the prior austenite grain size. The growth and coalescence of the voids means the linking of the interparticle spacing between voids which results to l_o^* and assumes the relation as a multiple of d^* according to Ritchie. In this study $l_o^* = 6d^*$, the inter particle spacing is taken as ‘6’ only as a fitting parameter. C_m is used as 3 for steels as a constant, which was recommended by Haggag & Nanstad (1989) only on the condition of replacing ϵ_f^* with n^* from Eqn. (7). Hence, by utilizing Bonora’s CDM model with modifications as shown in Eqn. (8), n rather than ϵ_f from uniaxial-testing can be extrapolated to a smaller multi-axial value due to effect of a high η which becomes the equivalent plastic strain, n^*

$$n^* = P_{th} \left(\frac{n}{\epsilon_{th}} \right)^{1/f(\eta)} \quad [8]$$

ϵ_{th} and P_{th} are the damage threshold strains for the uniaxial and triaxial stress states, respectively, which are taken to have values of 0.002 for the rail steels (Yu et al. 2017a, b, Yu et al. 2018) while $f(\eta)$ is taken as

$$f(\eta) = \frac{2}{3}(1 + \nu) + 3(1 - 2\nu)(\eta^2) \quad [9]$$

Therefore, by applying Eqn. (7) and replacing ϵ_f^* with n^* , K_{IC} can be predicted with tensile test for rail steels or

high strength structural steel alloys. Fig. 4 shows a relationship between ASTM K_{IC} and $K_{IC, pred}$.

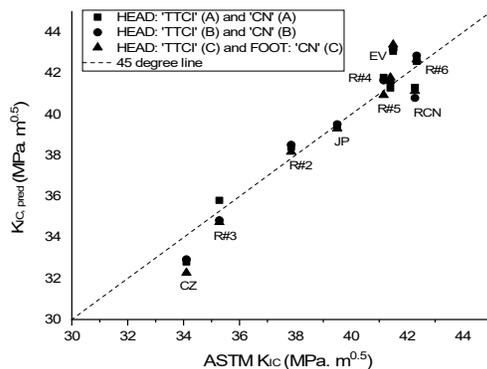


Fig.4. Predicting K_{IC} for nine rail steels in the three different locations

Rails CZ and EV shows little deviation from the 45° equality line while other rails relatively fall within the line. The deviation can be said to occur due to the assumptions used in the Eqn.7. Considering the use of C_m by Haggag & Nanstand (1989), the value can not remain as '3' for all steels and needs further study. Also, the inter-particle spacing should not be a constant for all rail steels. Different heat treatments along with the voids present should alter the values of the inter-particle spacing for different rail steels. In reality, interactions between voids can also play a role in the estimated l_o^* .

With improvements, the prediction of K_{IC} via the modified critical fracture strain model signifies that K_{IC} can be attained using tensile tests or possible non-destructive tests as long as the mechanical and flow properties can be attained, although this needs further work to improve functional parameters like the inter-particle spacing and C_m used by Haggag. From the outcome of K_{IC} values, heat treatment by head hardening (deep and micro) as well as head hardened perlite improved the K_{IC} due to the controlled cooling used in the manufacturing process as compared to increasing the carbon component in hypereutectoid rails as well as rails without any heat treatments. The heat treatments method may not have any way of distinguishing K_{IC} values explicitly but its microstructure can be used for K_{IC} prediction and categorization once other mechanical and flow parameters are known.

4 CONCLUSIONS

Ductile failure initiates from voids present within a microstructure and grows before it finally coalescence, forming the l_o^* . This l_o^* defines the extent at which macroscopic fracture occurs, which permits fracture toughness estimation. Some deductions can be made from the outcome:

1. The role of stress triaxiality is very important in the extent at which l_o^* and ϵ_f^* develops. The high η is used as this is similar to what occurs in plane strain fracture toughness via ASTM E399. Higher η

means lower l_o^* and ϵ_f^* and vice versa. However, the use of high η is for materials that obey SSY criteria, for materials that do not obey SSY criteria like mild steel, low η should be attained and implemented in the modified critical fracture strain model. The degree at which η varies needs to be looked upon.

2. The η is responsible for reducing the energy ($\sqrt{\sigma_y l_o^* \epsilon_f^*}$) required for fracture which is in line with Irwin's postulate against Griffin's fracture theory on the excessive energy required for fracture to occur using Griffin's principle for ductile materials.
3. The value of C_m is likely not a constant as suggested by Haggag (1989) but a factor that depends on the microstructural length, x , during ductile failure process and needs to be investigated. Based on ϵ_p development, C_m should decrease according to the evolution of x , microstructurally.
4. The use of '6' to estimate l_o^* from d^* is only a fitting parameter in this study. It is likely related to the void growth model (VGM) proposed by McClinton (1968), Rice & Tracy (1969) and needs to be further investigated to determine the critical value for interparticle spacing linkage that equates the l_o^* .
5. The modified critical fracture strain model shows a promising approach for estimating K_{IC} for structural members and static equipment. This model even supports the use of tensile experiments and non-destructive approaches like simple punch tests and indentation techniques as long as the mechanical properties and flow properties can be determined.

5 ACKNOWLEDGEMENT

This work was sponsored by the Canadian Rail Research Laboratory (CaRRL). The authors would like to express their appreciation to Kevin Hodder (Ph.D.) in the Department of Chemical and Materials Engineering, University of Alberta for his support and mentorship during the metallographic specimen preparation and etching and to Andre McDonald (Ph.D.) in the Department of Mechanical Engineering, University of Alberta for sharing vital information on the process required for the etching.

6 REFERENCES

- Alexander, D. J., & Bernstein, I. M. (1989). Cleavage fracture in pearlitic eutectoid steel. Metallurgical Transactions A, 20(11), 2321-2335.
- ASTM E112-13, Standard Test Methods for Determining Average Grain Size, ASTM International, West Conshohocken, PA, 2013, www.astm.org

- ASTM E1820-20, Standard Test Method for Measurement of Fracture Toughness, ASTM International, West Conshohocken, PA, 2020, www.astm.org.
- ASTM E399-20a, Standard Test Method for Linear-Elastic Plane-Strain Fracture Toughness of Metallic Materials, ASTM International, West Conshohocken, PA, 2020, www.astm.org.
- Bonora, N. (1997). A nonlinear CDM model for ductile failure. *Engineering fracture mechanics*, 58(1-2), 11-28.
- Bramfitt, B. L. (1998). Structure/Property Relationships in Irons and Steels, *Metals Handbook Desk Edition*, 2nd Ed., 2nd ed., Edited by Joseph R. Davis, ASM International, p 153-173, <https://doi.org/10.31399/asm.hb.mhde2.a0003090>
- Carlson, M. F., & Ritchie, R. O. (1977). On the effect of prior austenite grain size on near-threshold fatigue crack growth. *Scripta Metallurgica*, 11(LBL-6968).
- D. Broek, The role of inclusions in ductile fracture and fracture toughness, *Engineering Fracture Mechanics*, Volume 5, Issue 1, 1973, Pages 55-66, ISSN 0013-7944, [https://doi.org/10.1016/0013-7944\(73\)90007-6](https://doi.org/10.1016/0013-7944(73)90007-6).
- Erdogan, E. (2000). *Fracture Mechanics*, International Journal of Solids and Structures, 37, pp. 171-183.
- Garrison, W. M. (1984). A microstructural interpretation of the fracture strain and characteristic fracture distance. *Scripta Metallurgica*, 18(6), 583-586. [https://doi.org/10.1016/0036-9748\(84\)90345-4](https://doi.org/10.1016/0036-9748(84)90345-4).
- Gomes, M. G., de Almeida, L. H., Gomes, L. C., & Le May, I. (1997). Effects of microstructural parameters on the mechanical properties of eutectoid rail steels. *Materials characterization*, 39(1).
- Haggag, F. M., & Nanstad, R. K. (1989). Estimating fracture toughness using tension or ball indentation tests and a modified critical strain model (No. CONF-890721-24). Oak Ridge National Lab., TN (USA).
- Hahn, G. T., & Rosenfield, A. R. (1973). Relations between microstructure and the fracture toughness of metals. In *International Congress on Fracture*, 3 rd, Munich, West Germany.
- Hahn, G. T., & Rosenfield, A. R. (1975). Metallurgical factors affecting fracture toughness of aluminum alloys. *Metallurgical Transactions A*, 6(4), 653-668.
- J.R. Rice; D.M. Tracey (1969). On the ductile enlargement of voids in triaxial stress fields*. , 17(3), 201-217. [doi:10.1016/0022-5096\(69\)90033-7](https://doi.org/10.1016/0022-5096(69)90033-7)
- Lewandowski, J. J., & Thompson, A. W. (1986). Effects of the prior austenite grain size on the ductility of fully pearlitic eutectoid steel. *Metallurgical transactions A*, 17(3), 461-472. <https://doi.org/10.1007/BF02643953>
- McClintock, F. A. (1968). A Criterion for Ductile Fracture by the Growth of Holes. *Journal of Applied Mechanics*, 35(2), 363-. [doi:10.1115/1.3601204](https://doi.org/10.1115/1.3601204)
- Qiu, H., Hanamura, T., & Torizuka, S. (2014). Influence of grain size on the ductile fracture toughness of ferritic steel. *ISIJ International*, 54(8), 1958-1964.
- Ravichandran, K. S., & Vasudevan, K. (2013). Fracture resistance of structural alloys.
- Rice, J. R. (1968). Mathematical analysis in the mechanics of fracture. *Fracture: an advanced treatise*, 2, 191-311.
- Rice, J. R., & Tracey, D. M. (1969). On the ductile enlargement of voids in triaxial stress fields*. *Journal of the Mechanics and Physics of Solids*, 17(3), 201-217.
- Rice, J.R. and M.A. Johnson (1970), The role of large crack tip geometry changes in plane strain fracture, in: M.E Kanninen, W.E Adler, A.R. Rosenfield and R.I. Jaffee, eds., *Inelastic Behavior of Solids*, McGraw-Hill, New York, p. 641.
- Ritchie, R. O., Server, W. L., & Wullaert, R. A. (1979). Critical fracture stress and fracture strain models for the prediction of lower and upper shelf toughness in nuclear pressure vessel steels. *Metallurgical Transactions A*, 10(10), 1557-1570. [doi: 10.1007/BF02812022](https://doi.org/10.1007/BF02812022).
- Schwalbe, K. H. (1977). On the influence of microstructure on crack propagation mechanisms and fracture toughness of metallic materials. *Engineering Fracture Mechanics*, 9(4), 795-832.
- Shih, C.F. (1981). Relationships between the J-integral and the crack opening displacement for stationary and extending cracks, *Journal of the Mechanics and Physics of Solids*, Volume 29, Issue 4, Pages 305-326, ISSN 0022-5096, [https://doi.org/10.1016/00225096\(81\)90003-X](https://doi.org/10.1016/00225096(81)90003-X).
- Srivastava, A., Ponson, L., Osovski, S., Bouchaud, E., Tvergaard, V., & Needleman, A. (2014). Effect of inclusion density on ductile fracture toughness and roughness. *Journal of the Mechanics and Physics of Solids*, 63, 62-79.
- Sugden, A. A. B., & Bhadeshia, H. K. D. H. (1989). Microstructural entropy and the scatter in toughness.

Recent trends in welding science and technology, 745-748.

Tvergaard, V. (1989). Material failure by void growth to coalescence. *Advances in applied Mechanics*, 27, 83-151., [https://doi.org/10.1016/S0065-2156\(08\)70195-9](https://doi.org/10.1016/S0065-2156(08)70195-9).

Tvergaard, V., & Needleman, A. (1984). Analysis of the cup-cone fracture in a round tensile bar. *Acta metallurgica*, 32(1), 157-169.

Wallin, K. (1984). The scatter in KIC-results. *Engineering Fracture Mechanics*, 19(6), 1085-1093. [https://doi.org/10.1016/0013-7944\(84\)90153-X](https://doi.org/10.1016/0013-7944(84)90153-X).

Wang, J., Enloe, C., Singh, J., & Horvath, C. (2016). Effect of prior austenite grain size on impact toughness of press hardened steel. *SAE International Journal of Materials and Manufacturing*, 9(2), 488-493. doi:10.4271/2016-01-0359.

Yu, F., Jar, P. Y. B., & Hendry, M. (2017). Fracture behaviour at the sharp notch tip of high strength rail steels—Influence of stress triaxiality. *Engineering Fracture Mechanics*, 178, 184-200.

Yu, F., Jar, P. Y. B., Hendry, M. T., Jar, C., & Nishanth, K. (2018). Fracture toughness estimation for high-strength rail steels using indentation test. *Engineering Fracture Mechanics*, 204, 469-481.

Yu, L., Song, X., Yilong, L., Ming, Y., Zemin, W., Hu, X., & Jing, L. (2017). Effect of prior austenite grain size on microstructure and toughness of pearlitic steel. *材料导报*, 31(1), 77-81.

Zinkham, R. E., Dedrick, J. H., & Jackons, J. H. (1968). *Proceedings of 5th International Leichtmetalltagung Leoben Austria, 1968*, Aluminium-Verlag GMBH, Dusseldorf, Germany.

Session 3

INNOVATIONS IN RAILWAY ENGINEERING

Characterizing rail corrugation through a data-driven assessment of impacts on noise and vibration

Ahmed Lasisi, Jonathan D. Regehr, Julian Carneiro, Ian Jeffrey
University of Manitoba, Winnipeg, MB, Canada

Eric Magel, Sylvie Chenier
National Research Council, Canada

Mark Reimer
Advanced Rail Management, Canada

ABSTRACT

This short paper describes an ongoing effort to examine existing measures of rail corrugation using real-world rail performance data with the aim of proffering improvements to the management of rail grinding programs. The paper highlights the development and validation of a novel rail corrugation index (RCI) designed to offer a practical understanding of the impact of corrugation on noise and vibration in rail transit systems. A data-driven approach to match corrugation severity—as measured by the new rail corrugation index—to wayside noise levels is proposed for the corrugation observed within the 30 to 100-mm wavelength. The study attempts to estimate grind effectiveness with observable noise levels tied to corrugation pre- and post-grind, and finds that the proposed index produces intuitively valid results. Specifically, a proportionate sensitivity in noise and corrugation tolerance values shows a corresponding change in the grind effectiveness based on the proposed index. Ongoing work to characterize corrugation with vibration using mapped railcar-mounted accelerometer data is briefly described. Overall, the approach offers a potential low-cost alternative to managing grinding programs using corrugation covariates like noise and vibration.

1 INTRODUCTION

The effective management of rail assets is critical for achieving satisfactory levels of performance—in terms of infrastructure condition, safety, environmental impacts, ride quality, and economic return—within rail transit and freight rail properties (Kalousek 2005; Magel, Kalousek, and Caldwell 2005). Rail grinding programs play a critical role in rail asset management (Magel and Kalousek 2002). There are many reasons for grinding of rail, including the removal of fatigued steel, correcting the transverse rail shape (i.e., rail profile), dressing welds, and eliminating rail corrugation. Through these activities it is possible to extend rail (and wheel) life, minimize noise, improve ride quality, and reduce broken rails. Despite these potential benefits and the increasing availability of relevant data, implementing and managing an effective rail grinding program remains a challenge that confronts both practitioners and researchers in the rail industry.

According to Sroba & Roney (2003), rail corrugation is defined as follows (p. 36):

“Rail corrugation – A pattern of peaks and valleys that form on the running surface of rails at a variety of wavelengths. This condition is self-propagating, and if left unchecked can quickly become severe enough to necessitate changing the rail.”

Rail corrugation is a quasi-sinusoidal type of track wear resulting from wheel-rail contact, where the wavelength is less than about one metre (Grassie 2009). The most common damage mechanism is rail wear, although plastic bending of the rail and plastic flow can also occur. Rolling contact fatigue (RCF), while not a damage mechanism like wear and plastic flow that significantly alters the surface profile, is also exacerbated by rail corrugation (Grassie 2009). Once these damage mechanisms take root on the rail, rail corrugation causes further damage to the track and equipment and also produces noise and vibration (Zerkly and Meehan 2012).

This short paper highlights the development and validation of a novel rail corrugation index (RCI) to support performance-based decision-making within rail maintenance and asset management programs, with an emphasis on real-world rail performance data. In developing and validating the index, the research contributes to improved understanding about the mechanisms that cause rail corrugation, the effectiveness of strategies to mitigate corrugation initiation and development, and the impacts of rail corrugation on noise and vibration.

2 CORRUGATION INDICES IN PRACTICE

Since effective management of rail infrastructure assets is closely tied to overall condition of rail track components (Sadeghi et al. 2020; Soleimanmeigouni, Ahmadi, and Kumar 2018), the measurement and monitoring of various rail and track conditions has gained importance within modern railroading. Routine and reliable measurement of rail corrugation provides the foundation for understanding corrugation development, evaluating the effectiveness of rail grinding as a treatment for rail corrugation, and making performance-based maintenance decisions. However, there are practical questions about how to devise appropriate corrugation monitoring programs and translate corrugation measurements into meaningful indices. Recent work by Magel and Oldknow (2018) proposed a rail corrugation index (denoted as RCI_{2018} in this paper) based on root mean square (RMS) values of corrugation depth (using the 30 to 100-mm band) for short sections or “blocks” of rail. The proposed index is expressed mathematically as in Equation 1:

$$RCI_{2018} = \frac{RMS\ Block - TOL}{6 \times TOL} \times 100 \quad [1].$$

where $RMS\ Block$ is the RMS of the corrugation depth for the defined block of rail and TOL is the tolerance value (herein we use $TOL = 4$ microns). The index is considered applicable for a range of RMS values between 4 (TOL) and 28 microns, inclusive. In an initial attempt to scale the index, Magel and Oldknow (2018) proposed a set of thresholds (and an associated colour scheme) for the range of block RMS values, as shown in **Error! Reference source not found.1**.

With the aim of developing a holistic rail condition index, Magel and Oldknow (2018) also proposed an equivalent grinding index (EGI), defined as a weighted sum of three major components that drive rail grinding decisions: rail corrugation (expressed as RCI_{2018}), a profile quality index (PQI), and a surface damage index (SDI). Mathematically, the EGI was expressed as in Equation 2:

$$EGI = \frac{W_{PQI} \times PQI + W_{SDI} \times SDI + W_{RCI_{2018}} \times RCI_{2018}}{W_{PQI} + W_{SDI} + W_{RCI_{2018}}} \quad [2].$$

where W_{PQI} , W_{SDI} , and $W_{RCI_{2018}}$ are the weights assigned to the PQI , SDI , and RCI_{2018} , respectively.

Table 1. Scale for RCI_{2018} proposed by (Magel and Oldknow 2018).

Block root mean square (RMS) (microns)	RCI_{2018}	Colour
≤ 4	100	Blue
$4 < RMS \leq 8$	83.3	Green
$8 < RMS \leq 12$	66.7	Yellow
$12 < RMS \leq 16$	50	Orange
$16 < RMS \leq 28$	0.0	Red

In addition to the work by Magel and (Magel and Oldknow 2018), there is one known practical application of a rail corrugation index, though its definition differs from the proposed RCI_{2018} . That index, established by Advanced Rail Management (ARM) as a segment-based corrugation quality index (CQI), is defined as the 95th percentile value of the 5-m block RMS value (using the 30 to 100-mm band) within a specified track segment length (0.2 miles in this case).

The CQI can be calculated as follows:

1. Divide a 0.2-mile segment of track into 5-m blocks.
2. Calculate the RMS value of the corrugation measurement for each block.
3. Sort the 5-m block RMS values in the 0.2-mile segment in ascending order.
4. Multiply 0.95 by n (the number of 5-m block RMS values in the segment). The product is called the index.
5. If the index is not a whole number, round it up to a whole number.
6. Count the sorted RMS values in the 0.2-mile segment up to the number indicated by Step 5. The corresponding value is the 95th percentile, which is equivalent to the CQI.

Like the proposed RCI_{2018} , the CQI is based fundamentally on block RMS values. However, since the CQI is oriented towards practical grinding applications, it is expressed in terms of a segment length (usually 0.2 miles) rather than at the block level. While characterizing rail corrugation at a more precise spatial scale (e.g., a 5-m block) may be helpful to understand particular causes of rail corrugation, the treatment of corrugation through grinding can only practically occur for some meaningful length of track.

3 PROPOSED RAIL CORRUGATION INDEX

Based on current practice, this short paper proposes a new corrugation index designed to offer a practical understanding of the impact of corrugation on noise and vibration. The approach used here is to find the absolute or relative difference between the corrugation or noise measurements before and after grinding. In practice, an increase of 5 dB in noise level is sometimes considered to be a trigger for grinding action. This absolute difference does not necessarily account for corrugation severity without considering the level of corrugation/noise before/after grinding. Since grinding can be regarded as artificial wear, it is important to assess grinding effectiveness against the original

corrugation depth (Tyfour 2008). Therefore, grinding effectiveness could be evaluated as a fraction or percentage of the corrugation removed by grinding, relative to the desired corrugation level post-grind.

Mathematically, let the deformed layer, D_l , equal the difference between the corrugation depth and some acceptable corrugation tolerance (assume TOL of 4 microns). Similarly, let the noise nuisance, N_n , equal the difference between the noise level and some tolerable noise limit (assume 86 dB (HealthLink 2019)). Then, Equation 3 gives the absolute difference in corrugation severity, $|\Delta|_c$:

$$|\Delta|_c = RCI_{x[BG]} - RCI_{x[AG]} \quad [3].$$

where $RCI_{x[BG]}$ is the RCI_x before grinding, $RCI_{x[AG]}$ is the RCI_x after grinding, and x is the user-specified percentile (e.g., 95th or 65th). And, Equation 4 gives the difference in the noise level, Δ_n :

$$\Delta_n = LA_{eqBG} - LA_{eqAG} \quad [4].$$

where LA_{eqBG} and LA_{eqAG} are the A-weighted equivalent noise levels before and after grinding, respectively (e.g., 95th percentile, 65th percentile, or mean).

Then, the effectiveness of grinding with respect to corrugation removal, ϵ , may be expressed as a percentage as in Equation 5:

$$\epsilon = \frac{|\Delta|_c}{D_l} \times 100\% \quad [5].$$

Analogously, the effectiveness of grinding with respect to the reduction of noise caused by corrugation, ϵ_{noise} , may be expressed as in Equation 6:

$$\epsilon_{noise} = \frac{|\Delta|_n}{N_n} \times 100\% \quad [6].$$

The effectiveness will be sensitive to the choice of desired limit which will vary from one property to another.

3 PRELIMINARY ANALYSIS RESULTS

This section presents initial results of attempts to apply the proposed rail corrugation index (using corrugation and wayside noise data), and briefly describes initial work understand the impact of corrugation on vibration.

3.1 Application of the Proposed Corrugation Index

As an initial attempt to apply the proposed rail corrugation index, corrugation and noise data were collected from a Canadian rail transit system on a 0.2-mile (320-m) portion of track (test site from STA 277-288), comprising tangent, curve, and spiral segments. Figure 1 shows a snippet from the track chart that describes the studied section of track. In the figure, the entire section of track (STA 277-288) is classified simply as tangent and non-tangent. The non-tangent section is further broken down into spiral-1, spiral-2, and curve

(right hand). The direction of travel is from STA 277 to STA 288; thus, the left rail is the hi-rail and the right rail is the low-rail.

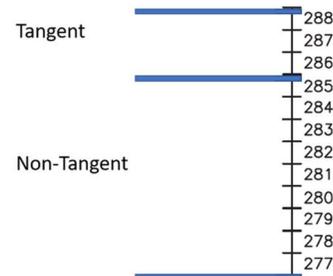


Figure 1. Track chart snippet (STA 277-288) used for preliminary analysis.

Table 2 presents some preliminary results obtained from applying the corrugation index to assess grind effectiveness as it pertains to both corrugation removal and noise reduction. The table shows RCI values (95th and 65th percentiles) for the left rail only and equivalent noise levels (95th and 65th percentiles, and mean).

Table 2: Grind effectiveness for tangent section

	Corrugation index		A-weighted equivalent noise		
	RCI ₉₅ (left rail)	RCI ₆₅ (left rail)	LAeq (95)	LAeq (65)	LAeq (mean)
Before grind ¹	15.76	7.03	100.66	99.75	99
After grind ²	6.08	2.08	89.04	85.61	86
Effectiveness [4,86] ³	82%	100%	79%	100%	100%
Effectiveness [5,87] ³	90%	100%	85%	100%	100%
Effectiveness [3,85] ³	76%	100%	74%	96%	93%

¹ Before grind measurements occurred on 07/02/2020.

² After grind measurements occurred on 10/03/2020.

³ The effectiveness percentages are calculated using corrugation and noise tolerances corresponding to the values in the brackets. As a calculation example, for the RCI₉₅ (left rail) column, $D_l = 11.76$, $|\Delta|_c = 9.67$, so $\epsilon = 82\%$.

From the table, it is evident that the effectiveness percentages are at or close to 100% when using the 65th percentile or mean values. In contrast, when examining the 95th percentile values, the index shows sensitivity when calculating effectiveness. While ϵ , ϵ_{noise} are sensitive to the tolerance values selected for RCI and LAeq, it can be observed that the 95th percentile statistic for corrugation and noise led to a numerically close grind effectiveness values for the different scenarios. Specifically, when the corrugation and noise limits are:

- [4,86], the RCI₉₅ and LAeq (95th percentile) effectiveness values are approximately 80%.
- [5,87], the RCI₉₅ and LAeq (95th percentile) effectiveness values are approximately 90%; and
- [3,85], the RCI₉₅ and LAeq (95th percentile) effectiveness values are approximately 75%.

It is observed that as the tolerable limits are lowered, the grind effectiveness values are also lower and vice versa. This behaviour is quite intuitive and indicates the reasonableness of the proposed index.

3.2 Rail Corrugation and Vibration

As corrugation is known to cause track and vehicle vibration, vibration may also be considered as a tool with which corrugation can be identified through passive observation (i.e., using trains to make observations during normal operations). By fixing accelerometers close to the wheelset and examining the resulting vibration data collected, it is potentially possible to identify and monitor corrugation from a revenue or test vehicle and thereby obviate the need for on-track measurements with a corrugation data collection tool.

For this study, the accelerometer and associated logging device were deployed on an American rail transit system. These devices were fixed to the side frame, since there was no physical space for mounting this system directly on the axle box. The accelerometer sampled at 3225 Hz and logged the acceleration experienced in three dimensions – longitudinal (x-dimension), lateral (y-dimension), and vertical (z-dimension), as shown in Figure 2. The train operated through the full system a total of eight times, and the acceleration was logged for the entire day's run. The data was then split in station-to-station segments.

The goal is to get a summary of the acceleration experienced by the train in this station-to-station segment by removing as much run-to-run variance as possible. This run-to-run variance may arise because of operational differences such as passenger load, train speed, driver action as well as car wheel and suspension condition. This means that each station-to-station segment needs to have a representative summary of the different runs over it in order to minimize the variance present in each run's sample of that segment. The objective is to perform an ensemble average of all runs between two stations, where the acceleration at each point is averaged and the full run is each point's average value. We then aim to relate these values (on the z-axis) to the corresponding values of the rail corrugation index. This work is currently ongoing.

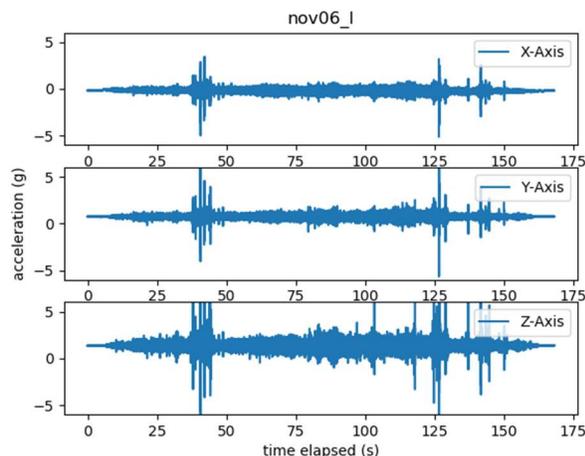


Figure 2: Sample accelerometer data.

4 CONCLUSION

This short paper describes a new proposed rail corrugation index, and helps to validate the index. The index attempts to support characterization of rail corrugation and its impact on noise and vibration. The results demonstrate promising proximity for grinding effectiveness values less than 100%. A proportionate sensitivity in noise and corrugation tolerance values shows a corresponding change in the grind effectiveness based on the proposed index. Ongoing work aims to relate the index to vibration impacts, as measured by onboard accelerometers.

5 ACKNOWLEDGEMENT

The authors acknowledge the support of the National Research Council of Canada and Advanced Rail Management (ARM).

6 REFERENCES

- Grassie, S. L. 2009. "Rail Corrugation: Characteristics, Causes, and Treatments." *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 223(6):581–96.
- HealthLink. 2019. "Harmful Noise Levels | HealthLink BC." *Harmful Noise Levels*. Retrieved February 12, 2021 (<https://www.healthlinkbc.ca/health-topics/ff4173>).
- Kalousek, Joseph. 2005. "Wheel/Rail Damage and Its Relationship to Track Curvature." *Wear* 258(7–8):1330–35.
- Magel, E. and Kevin Oldknow. 2018. "Quality Indices for Managing Rail through Grinding." *Proceedings of the 11th International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, CM 2018* 658–67.
- Magel, Eric, Joe Kalousek, and Robert Caldwell. 2005. "A Numerical Simulation of Wheel Wear." *Wear* 258(7–8):1245–54.
- Magel, Eric and Joseph Kalousek. 2002. "The Application of Contact Mechanics to Rail Profile Design and Rail Grinding." *Wear* 253(1–2):308–16.
- Sadeghi, Javad, Yousef Rahimizadeh, Amin Khajehdezfuly, Mohammad Rezaee, and Esmail Rajaei Najafabadi. 2020. "Development of Rail-Condition Assessment Model Using Ultrasonic Technique." *Journal of Transportation Engineering Part A: Systems* 146(8).
- Soleimanmeigouni, Iman, Alireza Ahmadi, and Uday Kumar. 2018. "Track Geometry Degradation and Maintenance Modelling: A Review." *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 232(1):73–102.
- Sroba, Peter and Mike Roney. 2003. "Rail Grinding Best Practices." *Proceedings of the 2003 American Railway*

Engineering and Maintenance of Way Association Annual Conference.

Tyfour, W. R. 2008. "Predicting the Effect of Grinding Corrugated Rail Surface on the Wear Behavior of Pearlitic Rail Steel." *Tribology Letters* 29(3):229–34.

Zerkly, Omar and Paul Meehan. 2012. *A Literature Survey of Rail Corrugation Measurement Techniques.*

Using emerging technologies for monitoring surface water near railway tracks

Alireza Roghani; Abdelhamid Mammeri; Abdul Jabbar Siddiqui
National Research Council Canada, Ottawa, Ontario, Canada

ABSTRACT

The level of water near the railway track is a major factor affecting the safety of train passage. Prolonged periods of heavy rainfall, rapid snowmelt, flash flooding, river flooding, beaver dams or blockage of a culvert result in a rise of water levels. This situation has been the major cause of many derailments in Canada and resulted in fatalities and serious injuries, damage to environment, loss of property, and service disruption. Railway companies strive to identify the development of problematic water levels in the area surrounding the track. This includes visual inspection performed by qualified track inspectors to visually identify waterway blockage and levels issues and air reconnaissance patrols that take place once or twice each year. These inspections rely on the inspectors' judgment and experience regarding the water level, have a limited range of coverage, and do not provide visibility on the water issue in the areas that are out of the vision range but still close enough to affect the track. The air reconnaissance patrol covers a larger area and provides a bird's eye view of all the waterways and identifies blockage of waterway but they are not as frequent. Recent advances in satellite-based remote sensors and tremendous development in unmanned aerial vehicle (UAV) have promoted the field of sensing surface water to a new era. National Research Council Canada and Transport Canada undertook a collaborative research project to evaluate the feasibility of using satellite imagery (including synthetic aperture radar and optical images) and UAV-based RGB images to detect water near railway tracks using data from two test sites in Canada. In addition, Transportation Safety Board (TSB) Rail Occurrence database and TSB's investigation reports were analyzed to identify the root causes of water-related derailment within Canada's rail network in the last few decades. 4 Canadian railway operators were also interviewed to better understand their main water related issues. The results of this project suggested that even though these technologies cannot entirely replace the current methods of water inspection, they offer an additional and inexpensive method to provide trackside water information to track inspectors. It was also indicated that further investigations and testing of technologies over same section of track would be required for drawing a definitive conclusion.

1 INTRODUCTION

The level of water near a railway track can be a major factor affecting the safety of train operations. Prolonged periods of heavy rainfall, rapid snowmelt, flash flooding, river flooding, beaver dams, blocked and under sized culverts or poor drainage design can result in a rise in water levels. The problems associated with high water levels include washout (a sudden release of a large quantity of water which quickly damages the trackbed), massive shear failure caused by submersion of track substructure, pier scour and undermining, and bridge foundation erosion that may occur with or without train loading. Such failures in a Centralized Traffic Control (CTC) territory often leave the track superstructure intact but unsupported. As a result, the washout is not visible until the train is a short distance away, when the train cannot be stopped or significantly slowed before reaching the washed-out track sections. This situation

has been a major cause of many derailments in Canada and resulted in fatalities and serious injuries, damage to environment, and loss of property. Occasionally the track washout can be caused by another washout or culvert failure on a nearby road.

Railways strive to identify the development of problematic water levels in the area surrounding the track. This includes visual inspection performed by qualified track inspectors to visually identify waterway blockage and levels issues and air reconnaissance patrols that take place once or twice each year. When rainfall occurs over multiple days and a sustained accumulation is identified, the railways increase the number of special inspections ahead of train traffic. Also, when a certain area is expected to have a large amount of continuous rain or when a severe weather advisory is issued, the railway companies implement additional track inspections to monitor drainage along the railway right-of-way. These inspections rely on the inspectors' judgment and experience regarding the water level, have

a limited range of coverage, and do not provide visibility on the water issue in the areas that are out of visual range but still close enough to affect the track. The air reconnaissance patrols cover a larger area but they are not as frequent.

The historical weather data shows that precipitation has increased in many parts of Canada, and there has been a shift toward less snowfall and more rainfall (Bush and Flato 2019). In the future, climate change is expected to exacerbate flooding issues (Khaliq & Attar, 2017) and extreme precipitation is projected to increase in Canada under both low and high emission scenarios. An extreme precipitation event that now occurs every 50, 20, 10 years is expected to occur every 10, 5, and 3 years under high-emission scenarios (Bush & Flato, 2019), a significant change that will affect infrastructure in the future. It is also expected that warmer winters and earlier snowmelt will combine to produce higher winter stream flows. Therefore, to assist in the identification of future impacts from expected climate changes, it is imperative to develop new ways for water inspection in the vicinity of a railway corridor, particularly through the use of new and emerging technologies.

Recent advances in satellite-based remote sensing and the tremendous developments in unmanned aerial vehicle (UAV) technologies have promoted the field of remote sensing to a new level. These technologies can be adopted by railways (some have already started implementing them) to improve/supplement their current waterway inspection procedures for areas beyond the railway corridor as well as the extreme weather policies. National Research Council Canada (NRC) and Transport Canada undertook a collaborative research project to evaluate the potential of satellite imagery and UAV-based imagery to detect water bodies near railway tracks. The main goals of this project were to 1) investigate the major causes of water-related issues around railway tracks by analyzing historical occurrences within Canada's rail network and interviewing the railway operators in Canada and 2) evaluate the potential of UAV and satellite imagery to map water bodies near railway tracks by testing both technologies over two study sites located within Canadian railway network.

2 HISTORICAL OCCURENCES

NRC reviewed eight Transportation Safety Board (TSB) of Canada investigation reports in which washout and high water level were the main attributed causes of incidents (summarized in

Table 1). It was found that in all of these occurrences, a special weather condition, such as heavy rainfall over 24 hours, above average rainfall over several days prior to the occurrence, above average daily temperature, etc. was involved. All these factors are anticipated to be more frequent under future expected climate. This review also indicated that seven of these occurrences happened over a section of track without any history of problems. In three such occurrences, an undetected beaver dam located out of the visual range was the major cause. It was also found that the time it takes for

a water-related issue to develop and lead to a derailment is so variable and it may range between a month to a few hours.

NRC in collaboration with TSB created a subset of the Railway Occurrence Database (RODS) in which one of the major causes was water-related (e.g. washout, culvert failure, beaver dams, high-water level, subgrade failure, etc.). Washout was found to be the most frequent primary cause of the occurrences in this dataset followed by subgrade failure (Figure 1a). The analysis also suggested that the main secondary factors were heavy rainfall followed by beaver dams. It also showed that despite the negative effect of climate change, there has been a declining trend in the number of water-related occurrences in the last 20 years. This can be attributed to more frequent and stringent water monitoring and management methods adopted by the railways. The analysis also suggests that June and July are the most critical months of the year in terms of water-related occurrences (Figure 1b).

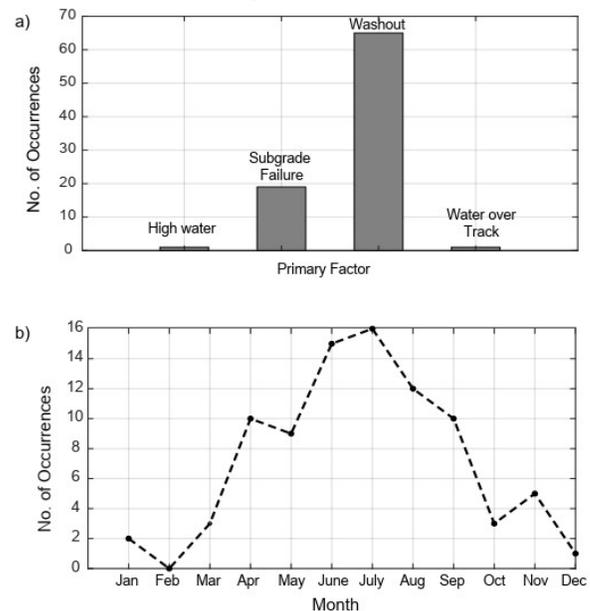


Figure 1. Plots show a) number of water-related occurrences vs primary cause/factor attributed and b) monthly distribution of water-related derailments.

3 INTERVIEWING RAILWAYS

NRC interviewed four Canadian railways (two class-I and two short lines) to obtain their perspectives regarding the major causes of water issues on their networks and to learn more about their current water inspection procedures. These interviews indicated that the water-related issues vary depending on the territories; however beaver dams and culvert failures seem to be the most common issues. The interviewees also indicated that, in their opinion, water issues have become more frequent in recent years due to the effects of climate change. They also indicated that in recent years, they have increased the size of the culverts to the next standard size whenever there is a need for a replacement to allow for higher expected water flow in the future. The information

collected during these interviews also suggested that the railway water inspections are undertaken more frequently than what is required by the regulators. Out of the four, one company indicated that it has evaluated the risk of extreme weather events on its network and identified areas prone to washout, flooding, and drainage

issues. This risk assessment is also supplemented with a criterion used to define and trigger a special weather alert that automatically notifies the inspectors to increase their inspection frequency until the situation is eliminated.

Table 1. Summary of TSB's water related investigation reports

Occurrence	Involved beaver dam	Special environmental condition prior to the occurrence	Impeded external drainage	History of problem
Ottawa Valley; TSB report# R09H0006	Yes	Week before the occurrence, above average rainfall	Accumulated water exceeded the capacity of culvert	No
Canadian Pacific, TSB report# R13E0069	No	Abnormally warm temperature increasing seasonal run-off water from melting snow	Ice in the ditch restricted flow to the culvert	No
VIA Rail, TSB Report# R13W0124	No	5 days prior to occurrence, rapid melt of snowpack	Culvert plugged with ice	No
Huron Central, TSB report# R15H0092	No	Above average monthly precipitation	Culvert had either collapsed, sunk, and/or was plugged	Yes
VIA Rail, TSB report# R18W0168	No	Heavy rainfall prior to and on the day of occurrence	Culverts of nearby highway accumulated forest debris and failed	No
Canadian National, TSB report# R92T0183	Yes	Extremely heavy rainfall on the day of occurrence	No	No
Canadian Pacific, TSB report # R97T0097	Yes	Rapid melting of snow during the two weeks preceding the derailment	No	No
Canadian Pacific, TSB report # R13C0069	No	150 mm of rainfall in 48 hour period prior to derailment	No	No

Even though the other three companies did not perform any formal risk assessment, the discussions suggested that this kind of analysis is continuously performed by the inspectors on-site. The inspectors are familiar with their territories and from the experience, they know where the problematic sites are located and therefore, they plan for the inspection and mitigation measures accordingly. One of these companies also indicated that they use UAVs to inspect water in areas that are difficult for inspectors to reach. They indicated that UAVs are a useful tool that result in significant efficiency gains; however their use is

limited as they can only be piloted by certified operators. All interviewed companies indicated that the current water inspection procedures are sufficient for safe railway operation but there is certainly room for improvement, especially using the new technologies.

4 METHODOLOGY

This section briefly discusses the basic principles of water detection from satellite and UAV images and introduces two test sections on Canada's rail network

that were used to evaluate the potential of these technologies.

4.1 Water detection using satellite data

One of the significant applications of remote sensing is to identify changes in water bodies on Earth. Surface water bodies are dynamic in nature, because they shrink or expand with time, owing to a number of natural and human-induced factors (Huang, et al, 2018). The process of identifying changes by observing different satellite images taken at different times is called change detection. Change detection quantitatively analyzes temporal effects by using the satellite images taken at different times (Bhavani, et al, 2018).

Satellite remote sensors can be divided into two categories: passive and active. Passive remote sensors respond to external stimuli. They record natural energy which is transmitted or reflected from Earth. Reflected sunlight is the main source of energy for passive remote sensors. On the other hand, active remote sensors use some internal stimuli to obtain information about the objects or phenomenon occurring on Earth (Rana, et al, 2017). Satellite-based optical sensors (also referred to as visible band sensors) are considered as passive, and microwave sensors (also referred to as Synthetic Aperture Radar (SAR) are considered as active types. The multispectral nature of optical sensors provides some advantages for water detection, however their application in detecting surface water is constrained by several environmental factors, such as cloudy sky conditions, cloud shadows that may seriously affect water detection due to similar spectral characteristics between shadow and flood/standing water. Optical sensors also fail to image the water surface beneath flooded vegetation canopies (Alsdorf, et al, 2007).

SAR's use of long wavelength radiation has the ability to penetrate cloud coverage and certain vegetation coverage. It is independent of solar radiation and therefore they can work day and night under any weather conditions. Radar can measure amplitude (the strength of the reflected echo) and phase (the position of a point in time on a waveform cycle). The dielectric and geometric properties of various target surfaces affect the intensity of the backscatter signal and is the basis by which SAR differentiates features on the ground (Irwin, et al, 2017). Fresh water has a high dielectric constant and smooth water surfaces usually provide a specular reflection of microwave radiation; both of these attributes result in very little energy backscatter (appearing dark in the image). In contrast, land surfaces scatter much more energy back to the radar source due to surface roughness and volume scattering (White, et al, 2015). However, SAR-based estimates of the extent of surface water are hindered by wind roughening the water surface for the wavelengths used by most sensors (Alsdorf, et al, 2007). Also due to the side-looking nature of SAR, some areas on the ground surface may be misclassified when terrain or other features create regions of radar shadow (Mason, et al, 2010).

4.2 Water detection using UAV based RGB images

An RGB image is the true-colour image that one can see in everyday life. The abbreviation "RGB" stands for Red, Green and Blue, and it refers to the three colours of light (known as channels) used in image processing science. The three colours can be mixed together to form different colours and to produce colour images on screens.

Water detection methods can be categorized as: traditional methods (hand-crafted features such as colour, texture, etc.) or deep learning-based methods (to learn to extract informative and discriminative features). These also vary depending on whether the camera is fixed or moving, placed on a ground carrier or mounted on a UAV, and whether the methods work on a per-image basis (i.e., only spatial context taken into account) or on videos (i.e., temporal context also considered).

Water is detected based on cues such as colour, texture and reflections detected in stereo-range data, for example by Rankin (2004). That is, a pair of cameras used to generate range images from which reflections are detected. A range image contains the distances of given points in the scene to another point. This method is appropriate for ground vehicles and autonomous navigation. Only colour-based cues are insufficient as they lead to many false detections (e.g., in cases of snow, white rocks, sky regions, and overexposed imagery).

The texture-based water cue proposed in Rankin (2004) is focused on low texture water regions. Although their method was able to detect regions not detected by the colour-based cues only, it required proper thresholds to be set up and yet yielded false detections. In addition to colour- and texture-based cues, Rankin (2004) proposed stereo range reflection-based water cues as well. In Mettes, et al (2017), the dynamic texture of water surface was exploited in differentiating water regions from non-water regions. Their method is not applicable where the camera is moving, e.g., carried by a UAV. Mehra, et al, (2016) proposed a technique to detect stagnant water bodies based on classical hand-crafted features (e.g., Scale-invariant feature transform - SIFT). Besides 2D imaging, some researchers, e.g., (U.S. Patent No. US 9,460,353 B2, 2016), have proposed the use of 3D imaging systems to detect water bodies based on certain characteristics which serve as cues for water's presence, similar to and expanding on the techniques described in Rankin (2004). However, the techniques suffer from many false detections.

Some researchers have addressed the problem of water detection in scenarios such as floods on roadways. An example of such a work is that by Sazara (2019) in which a suite of algorithms were proposed and evaluated. The paper states that remote sensing-based water detection methods lack local details which are needed to gain deeper information regarding the extent and severity of water on the inspected regions.

Kawaguchi (2016) proposed the use of infrared cameras to detect water and dike regions to monitor the water levels in urban water channels. Instead of using

RGB image sequences alone, Ghahremani, et al (2017) employ thermal images as well. However, the dataset and the proposed method seems to consider a fixed camera and as such are based on the assumption that non-water regions have no dynamic behavior compared to water region pixels.

Some researchers have investigated the use of UAVs to detect and monitor water levels. For example, Ridolfi, et al (2018) use a UAV to capture images in dam sites and exploit filter-based techniques to identify boundaries between water and non-water surfaces. Camera-based methods benefit from wide aerial coverage while fetching local details, in contrast to Internet-of-Things or sensor-based methods (Arshad, et al, 2019).

Another methodology for detecting water in single images is that of Han, et al, (2018). In their method, a deep learning method known as fully convolutional network (FCN) was modified to learn the reflection of objects and sky on the water.

Adhering to and in spite of the work required to meet strict safety and security regulations, many railways are investigating and deploying UAV technology for maintenance, surveillance, etc. (Federal Railroad Administration, 2018). The use of UAVs in the railway industry ranges from a passive role, such as rail inspection and surveillance, to an active role, including maintenance and on-site reparation. Even though UAVs are used in many areas such as bridge maintenance and inspection, railway infrastructure inspection, and object detection (e.g., vegetation, water, animals), the exploration of UAVs in preventing the accumulation of water around railway tracks due to flooding or snow-melting has not yet been extensively investigated by researchers. This is due to the fact that the detection of water bodies in RGB images, as those detected by UAV cameras, is a challenge. Water's shape and colour change frequently and continuously throughout a day. Further research work is needed in this area. The recent advancements in deep learning-based computer vision, especially with regards to convolutional neural networks (CNNs), motivates this study to investigate and evaluate CNNs for the task of water detection in RGB images captured by UAVs.

4.3 Test sites

Two sites were selected along Canada's rail network to evaluate the potential of Satellite and UAV images for mapping surface water. Site-I, located on Churchill subdivision (Manitoba), used for assessing Satellite images (optical and RADAR) and Site-II, located on a railway line in eastern Canada, used to investigate the potential of UAV images. The UAV is used by the owner of this line to inspect water bodies located around railway tracks and also in areas that are difficult for their inspectors to reach.

5 RESULTS AND DISCUSSION

5.1 Satellite data

The freely available optical images from Landsat 8 and radar images from Sentinel-1 between May and October of 2019 (one image per month) were obtained for a section of track along Churchill subdivision.

The images obtained from the Landsat 8 are subject to many quality issues; in particular, problems with instrument saturation, topographic shading (where topography does not allow direct solar radiation), cloud shadows and instrument failure. The presence of cloud and cloud shadows decrease the accuracy of the results of remote sensing applications as they obscure the land surface, and the brightening effect of clouds and the darkening effect of cloud-shadows influence the reflectance of each band (Zhu & Woodcock, 2014). The band quality assessment (BQA) data for each image was first reviewed to evaluate its suitability for water mapping. The analysis of this information showed that out of the 8 images, 7 were affected by the presence of clouds and shadows and were therefore not appropriate candidates for water detection (an example is shown in Figure 2). This is a significant limitation that was noted by other researchers as well. This limitation prevented the mapping of the temporal variations of surface water using optical images.

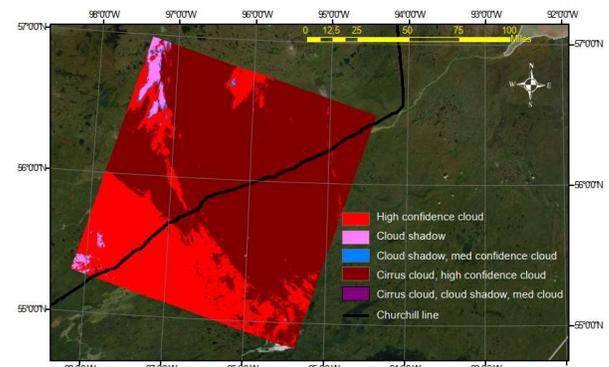


Figure 2. Quality assessment band for the optical image collected from Churchill subdivision on September 07, 2019.

Radar images from Sentinel-1 (Level 1 GRD images IW mode with pixel spacing of 10 m) were collected across the study region using the Alaska Satellite Facility website (<https://search.asf.alaska.edu/#/>). The analysis of radar images conducted based on a technique known as segmentation. This technique is the most commonly used approach to map surface water from radar imagery. In this method, all pixels with a backscatter coefficient lower than a specified threshold in an intensity image are mapped as water. This technique is useful for producing results quickly and inexpensively and is only suitable for calm open water with a specular backscatter response (Martinis, Twele and Voigt 2009) (White, Brisco and Dabboor, et al. 2015). Thresholding could be fixed, in which the threshold value is held constant throughout the image (which is the case in this project) or local (dynamic thresholding) which varies with the position of the pixels in the image. A suitable threshold has a direct impact on the classification of features in an image. It should be noted that before segmentation stage, some pre-

processing techniques are required to calibrate and filter the RADAR images. A more detailed discussion on pre-processing steps can be found in (Roghani, Mammeri, and Siddiqui, 2019).

Figure 3a presents the output intensity images for the image taken on Churchill test site on May 30, 2019 and Figure 3b shows its equivalent histogram and the threshold value used to extract water from other land features. Similar plots were created for each of the images taken between May and October 2019 and used to map variation of surface water over the test section. The results are presented in Figure 4.

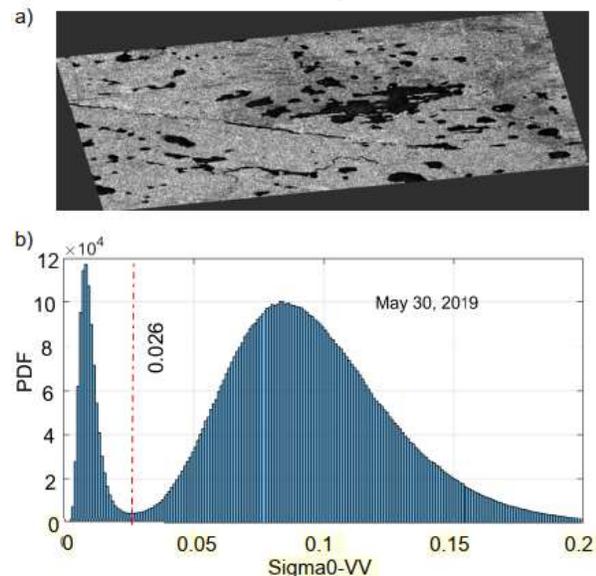


Figure 3. a) The VV polarization intensity image and b) the histogram of the intensity image histogram for the image taken Churchill test site on May 30, 2019 (the red dashed line shows the threshold used to extract water bodies).

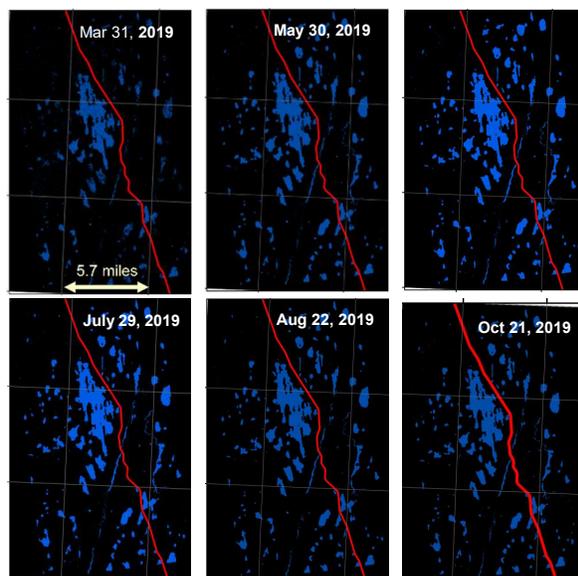


Figure 4. Surface water mapped using satellite data from Sentinel-1 near a section of track along Churchill

line (the red line show the railway line).

The output images (mapped water shown in Figure 4) can be imported into a GIS-based software application (e.g. ArcGIS®) to detect temporal changes in water extent. This information can be aligned with the other datasets such as culverts and bridges locations, geology information, topography, and historical problematic areas allowing the track engineers to determine the high-risk areas.

5.2 UAV images

In this work, a deep learning model based on CNNs was developed for the segmentation of rail-tracks and water regions in UAV-based images, labelled as *RailWater-UNet*. This model is a multi-scale Fully Convolutional Neural network (FCN) designed in an encoder-decoder architecture inspired by U-Net (Olaf Ronneberger, 2015). For a given image, the model learns to identify and differentiate between different pixels (or groups of pixels) belonging to the different classes (water, railway track or other regions). The model is basically used to indicate whether a particular pixel (or group of pixels) belong to water, railway track or other regions in a given image. The dataset is split into training and testing sets for respectively training and testing the model with.

The effectiveness of the model can be assessed using Mean Intersection over Union (mIoU). The mIoU metric gives a measure of how close the model's predicted regions match with the actual ground truth regions for various classes (e.g., water and railway track).

Different versions of the model were trained with the training set, varying the number of training rounds (i.e., epochs) used. Here, we present the results of three such versions: model trained for (i) 100 epochs, (ii) 300 epochs, and (iii) 500 epochs. The mIoU values for each model are summarized in Table 2. The closer an mIoU score is to 1.0, the better it is.

Table 2. Summary of testing set evaluation

Model (#epochs trained)	mIoU		
	Background	Rail-track	Water
100	0.940	0.512	0.660
300	0.944	0.545	0.673
500	0.944	0.544	0.673

The sample outputs of the Model500 are shown in Figure 5. The left column shows the predicted segmentation mask overlaid on the input image while the right column shows the original input image.

The performance of the RailWater-UNet model could be further improved by incorporating ideas inspired by some well-known AI models such as ResNet-like skip connections (He, et al, 2016). Moreover, newer semantic

and unsupervised segmentation models could be studied in the context of rail and water segmentation.

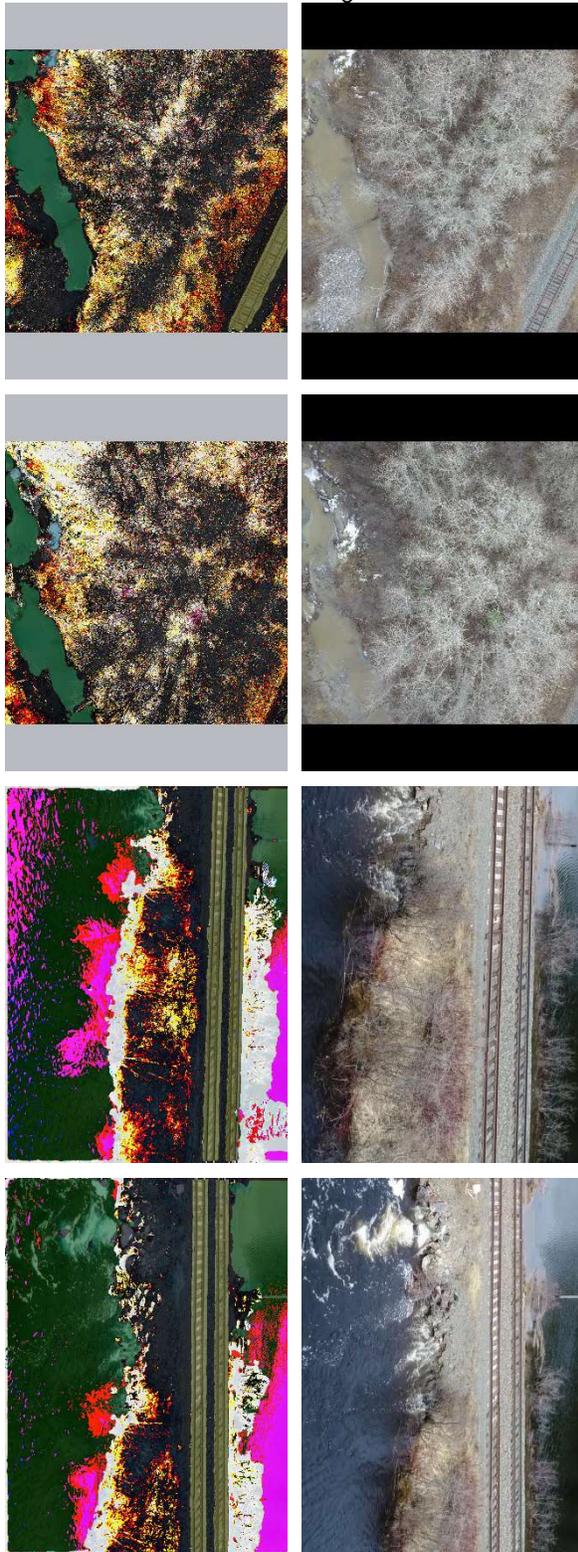


Figure 5. [Left Column] Sample prediction outputs (Greenish regions: Water, Yellowish regions: Railway track, others: background or ignored classes). [Right

Column] Original test images.

5.3 Comparing Satellite and UAV results with hi-rail inspection

Table 3 compares the advantages and limitations of using UAV and satellite data for water inspection with hi-rail/visual inspection by considering five different indicators: spatial and temporal resolution, visibility range, water detectability, and cost of data acquisition. It should be noted that the information in this table reflects authors' opinion based on the preliminary results and may change as more data is available. In addition, for a true comparison, the advantages and limitations of these methods should be assessed by testing over same sections of track.

As it is described in this table, each method has its own advantages and limitations. Given the information obtained regarding the water-related derailments during the interviews and the review of TSB's investigation report, the current state of UAV and satellite technology does not allow for their use as the sole method for water monitoring near railway tracks. However, the limited observations from this project suggested that UAV and satellite imagery can provide supplementary information to hi-rail inspection.

Table 3. Comparing the advantages and limitations of various water monitoring methods.

	Satellite	UAV	Hi-rail
Spatial resolution	10 m x10 m for sentinel-1, filtering would lower it further	Depending on resolution of camera and flight height	Human eye
Temporal resolution	6 days for sentinel-1	Flexible	1-2 times a week, could be more frequent**
Visibility Range	No limit	Depending on flight height, approximately ~10-50 m around the track	Right-of-way
Water detectability	Easy water detection due to water response to Radar signals	Challenging as colour of water differs	Human eye
Cost*	\$	\$\$\$	\$\$
Major limitation	Inadequate spatial and temporal resolution for water issues with ROW	A large dataset of images required for automating water detection process	Not automated, limited in range

* Cost of data acquisition

** Railways also conduct helicopter and/or drone inspections 1-2 times a year to supplement their hi-rail/visual inspections

6 CONCLUSIONS

National Research Council Canada and Transport Canada undertook a collaborative research project to evaluate the potential of satellite and UAV-based

imagery to detect water bodies near railway tracks. Four Canadian railway companies interviewed by NRC indicated that even though the current water inspection procedures are sufficient for safe railway operation, there is certainly room for improvement and complement the visual inspection done by the qualified Rail Inspector, especially using the new technologies. The interviewees also indicated that, in their opinion, water issues have become more frequent in recent years, an observation that is consistent with the results of recently published Canada's Changing Climate Report that suggests that extreme precipitation is expected to increase under both high and low emission climate scenarios. Therefore, to assist in the identification of future impacts from expected climate changes, it is imperative to develop new ways for water inspection in the vicinity of railway corridors, particularly through the use of new and emerging technologies.

The potential of satellite and UAV images were evaluated by testing them over two study sites. The results suggested that even though these technologies cannot replace the current methods of water inspection, they offer an inexpensive and effective method that provides information about the water issues that may not currently be available to track inspectors. The implementation of these technologies into a railway's water inspection methods could provide a better view of the problems occurring far from the track. This is especially useful during the springtime and heavy rainfalls to reduce the risk of water-related derailments. Currently, NRC and Transport Canada are working on the next phase of this project whose main goal is to combine several different technologies into a single framework and evaluate their potential in water detection through limited field trials. These trials will include two types of sensors: LIDAR and high-resolution cameras (providing images within visible and infrared spectrum) from three different platforms: satellites (space), UAV (airborne), and hi-rail truck (at track level). The combination of various sensors and platforms provides a wide range of complementary information with different spatial and temporal resolutions as well as varying fields of view. These results are expected to provide the most comprehensive picture of the water situation near railway tracks and mitigate the risk of water-related derailments on Canada's rail network.

7 ACKNOWLEDGEMENT

The authors would like to thank Transport Canada Innovation Center and NRC Resilient Ground Transportation Program for providing funding for this project. Special thanks to Paul Charbachi from VIA Rail Canada, Brooke Jones and Stephanie Roller from Transport Canada, Steve Henderson from Transportation Safety Board and Robert Caldwell, Sylvie Chénier, Gordon Poole, and Albert Wahba from NRC.

8 REFERENCES

- A. L. Rankin, L. H. (2004, December). DAYTIME WATER DETECTION BY FUSING MULTIPLE CUES FOR AUTONOMOUS OFF-ROAD NAVIGATION. Retrieved February 2020, from <https://apps.dtic.mil/dtic/tr/fulltext/u2/a432750.pdf>
- Arshad B, O. R. (2019). Computer Vision and IoT-Based Sensors in Flood Monitoring and Mapping: A Systematic Review. *Sensors*, 19(22), 5012.
- Billah, A. M., Abdenbi, A., & Agliz, D. (2018). Detection of water-risk zones based on color and texture in home video surveillance. *9th International Symposium on Signal, Image, Video and Communications (ISIVC)*, (pp. 168-173).
- Bush, E., & Flato, G. (2019). *Canada's Changing Climate Report*. Environment and Climate Change Canada.
- C. Sazara, M. C. (2019). Detecting floodwater on roadways from image data with handcrafted features and deep transfer learning. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, (pp. 804-809). Auckland, New Zealand.
- Chen, L.-C. &. (2016). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Federal Railroad Administration. (2018). *Unmanned Aircraft System Applications in International Railroads*. Washington DC: U.S. Department of Transportation Federal Railroad Administration.
- Ghahremani, A., Bondarev, E., & de With, P. H. (2017). Water Region Extraction in Thermal and RGB Sequences Using Spatiotemporally-Oriented Energy Features. *Electronic Imaging, Image Processing: Algorithms and Systems XV, 2017*(13), 78-86.
- Han, X., Nguyen, C., You, S., & Lu, J. (2018). Single Image Water Hazard Detection Using {FCN} with Reflection Attention. *15th European Conference on Computer Vision (ECCV)*. 11210, pp. 105-121. Munich, Germany: Springer.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Computer Vision and Pattern Recognition (CVPR)*. IEEE.
- Irwin, K., Beaulne, D., Braun, A., & Fotopoulos, G. (2017). Fusion of SAR, Optical Imagery and Airborne LiDAR for Surface Water Detection. *remote sensing*.
- Kawaguchi, K. H. (2016). FloodEye: Real-time flash flood prediction system for urban complex water flow. *IEEE SENSORS*, 1-3.
- Khaliq, N. M., & Attar, A. (2017). *Assessment of Canadian Floodplain Mapping and Supporting Datasets*

for Codes and Standards. Ottawa: National Research Council.

Mehra, M., Bagri, A., Jiang, X., & Ortiz, J. (2016). Image Analysis for Identifying Mosquito Breeding Grounds. *2016 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops)*, (pp. 1-6). London.

Mettes, P., Tan, R. T., & Veltkamp, R. C. (2017). Water detection through spatio-temporal invariant descriptors. *Computer Vision and Image Understanding*, 182-191.

Olaf Ronneberger, P. F. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer LNCS, 9351, 234-241. doi:https://link.springer.com/chapter/10.1007%2F978-3-319-24574-4_28

PercepTek Robotics. (2004, December). Detection of small water-bodies. Retrieved February 2020, from <https://apps.dtic.mil/dtic/tr/fulltext/u2/a433004.pdf>

Rankin, A. L., Matthies, L. H., & Bellutta, P. (2016). *U.S. Patent No. US 9,460,353 B2*.

Ridolfi, E., & Manciola, P. (2018). Water Level Measurements from Drones: A Pilot Case Study at a Dam Site. *Water*, 10(3), 297.

Schapire, R. E. (2013). Explaining AdaBoost. In F. i. Vapnik, *Empirical Inference* (pp. 37-52). Berlin, Heidelberg: Springer Berlin Heidelberg.

White, L., Brisco, B., Dabboor, M., Schmitt, A., & Pratt, A. (2015). A Collection of SAR Methodologies for Monitoring Wetlands. *remote sensing*(7), 7615-7645.

Yang, K., Wang, K., Cheng, R., Hu, W., Huang, X., & Bai, J. (2017). Detecting Traversable Area and Water Hazards for the Visually Impaired with a pRGB-D Sensor. *Sensors*, 17(8), 1890.

Zhao, Y., Deng, Y., Pan, C., & Guo, L. (2013). Research of Water Hazard Detection Based on Color and Texture Features. *Sensors & Transducers Journal*, 157(10), 428-433.

Zhu, Z., & Woodcock, C. E. (2014). Automated cloud, cloud shadow and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sens. Environ*, 152, 217-234.

Proposed dynamic load factors developed from instrumented wheelset measurements

Danial Behnia, Parisa Haji Abdulrazagh, and Michael T. Hendry

Department of Civil and Environmental Engineering – University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

Dynamic load factors (DLF) relate the magnitude of vertical wheel forces for rail loads in operation (dynamic loads) to static loads resulting from the weight of the rail car and its contents, as a function of train speed. DLF equations are often used in the selection of rail steel and cross-sections (weight). Equations for DLFs have been put forth by the American Railway Engineering and Maintenance-of-Way Association (AREMA) and others. A limitation of existing DLF equations is that they have been derived from loads measured at instrumented sections of track and observations of many wheel loads but with constant track conditions. For this study, measurements from an Instrumented Wheel Set (IWS) over more than 340 km of freight railway track through the Canadian Prairies provide dynamic loads from one loaded freight car over various track structures at differing train speeds. The IWS system is designed to measure forces originating at the wheel/rail contact point. This system is also instrumented with 16 full-bridge Wheatstone strain gauge circuits placed inside the wheel web. This paper presents a comparison of the magnitude of measured dynamic loads to existing DLF equations and develops new DLF equations specific to types of track structures encountered during the IWS measurements (tangent track, bridges, grade crossings, curves, and switches). An important finding to emerge in this study is DLF values in non-tangent track is around 10% more than tangent track. In addition, this difference increases with increasing speed.

1 INTRODUCTION

Rail breaks and failures in track components are among the most common causes of derailment in Canada. The incidence of derailment at higher train speeds also leads to more significant hazards for industry and the environment (Leishman *et al.* 2017). Moreover, revisiting the variations of the magnitude of dynamic loads along the track is worthwhile given the sustained increases seen in wheel loads and train speeds. Experts consider the dynamic load factor (ϕ : ratio of the vertical dynamic load (P_d) to static load (P_s)) as a significant factor in the design of track structures and selection of rail steel and cross-sections (weight) (Peters 2010; Sadeghi 2012; Van Dyke *et al.* 2017; AREMA 2018). This factor (Eq. 1) is often developed as a function of train speed. Table 1 gives examples of equations for the upper envelope of ϕ in freight lines and the variables for these equations, which show a linear correlation between train speed and ϕ in freight lines.

$$P_d = \phi P_s \quad [1]$$

The main limitation of existing ϕ equations is that they are based on measured dynamic loads from trains passing over an instrumented section of track (Dybala & Radkowski 2013; Van Dyk *et al.* 2017; Yu & Hendry 2019). However, a limitation of this approach is that data are representative of the loads generated over that particular segment of track, which is typically a well-supported and maintained tangent track. Thus, these dynamic load factors are unlikely to be representative of either the average or worst-case conditions for generating dynamic loads (Van Dyk *et al.* 2017).

The Instrumented Wheel Set (IWS) provides direct, continuous, and real-time measurements of the dynamic loads at the wheel-rail interface and thus provides an opportunity to quantify the dynamic loads along longer segments of track and thus develop ϕ envelopes that are representative of the full variety of track conditions.

This paper utilized the IWS system to measure the vertical dynamic loads along a 340 km section track of a North American Class 1 freight railway. These measurements are limited to dynamic loads from a

single car and suspension type, wheel diameter and static load, but do provide dynamic loads generated from the range of track conditions, track assets (bridges, grade crossings, curves, and switches), and operational train speeds encountered on a Class 1 freight railway main line. To date, these factors have not been accounted for when assessing the magnitude of dynamic loads rail structures will be subjected, and specifically in the development of ϕ equations. Therefore, this paper aims to evaluate the variation of dynamic loads (through ϕ) versus train speed along various track assets (i.e., tangent, curves, bridges, grade crossings, and switches). The data are also tabulated to provide ϕ values representative of the loading conditions generated within the range of speeds permissible on North American classes of the track (1 through 4).

Table 1. Summary of the dynamic load factor (ϕ) equations presented in the literature for freight railways.

Expression for ϕ	Eq.	Reference
$1 + \alpha + \beta + \gamma$	[2]	ORE ⁷ / Birmann (1965)
$1 + 5.21 \frac{V}{D}$	[3]	AREMA (2018)
$1.099 + 0.00621V$	[4]	Van Dyk et al. (2017)
$1 + 0.00466V$	[5]	Used by track operator (Peters 2010)

Variable Definition
V: Train speed (km/h)
D: Wheel diameter (mm)
 α, β, γ : Empirical coefficients derived from train speed, vehicle, and other track parameters

* Office of Research and Experiments

2 MATERIALS AND METHODS

The IWS data collection system was mounted on a 15.8 m (52 ft) gondola car (Roghani and Hendry 2016, 2017; Roghani et al. 2015, 2017; Fallah Nafari et al. 2018a,b) to investigate the track performance. The car was loaded with gravel to a total weight of 1175 kN (264 kips). Two IWS were mounted on one end of the car over two-class F wheel plates with a diameter of 915 mm (36"). This system is made up of 16 full-bridge Wheatstone strain gauge circuits, which interpret forces applied to the wheels from the rail and resolve these forces into vertical, lateral, and traction forces for each of the four wheels (Woelfle 2016); only vertical forces (dynamic loads) are considered within this paper. Figure 1 shows an installed IWS on the Gondola railcar. The IWS measurements were taken at a frequency of 200 Hz and filtered with a 20 Hz low-pass filter applied to the data during acquisition (Higgins et al. 1992; Bracciali et al. 2014; Cakdi et al. 2015; Barbosa 2016; Ren & Chen 2019). In addition, a Garmin GPS18X global positioning system (GPS) was set up to determine the location (latitude and longitude), time, and speed. A load measurement accuracy of 0.3% was determined based on the difference between measured static load by a scale after the installation of the IWS system and the sum of the static loads from the four IWS wheels. The wheels had very little wear and were free of defects that would increase dynamic loads.

The section of track included in this study is operated as a Class 1 railway and is part of a high traffic subdivision (>50 GMT/year) through the Canadian Prairies. Table 2 provides more detailed information in this regard. The main limitation was performing data collection without control over the type or weight of adjacent cars or the speed of travel as the car was in revenue service.



Figure 1. Photograph of an IWS installed on the Gondola railcar. Wheatstone bridge strain gauges and connectors covered in yellow epoxy coating are visible in the photograph (photo credit Hendry 2014).

Table 2. Detailed information regarding the section of track on which data were collected.

Information Type	Description
Run (#)	4 (two in each direction)
Bridge and overpass (#)	30
Switch (#)	50
Grade level crossing (#)	100
Curve (km)	83
Maximum speed (km/h)	95
Rail type	Continuously welded rail
Tie type	Concrete (supported primarily by)
Recording date	July-August (2015)

3 RESULTS

To evaluate the repeatability of the data, an 800 m section of track was selected, including a switch, grade crossing, and relatively short bridge. Figure 2a presents an overview of the selected section. Data collection involved four passes of the instrumented car (Table 2), two at slower speeds (24 and 50 km/h, respectively) in the westbound direction (Figure 2b) and two at higher speeds (77 and 88 km/h, respectively) in the eastbound direction (Figure 2c). Figures 2a and b show track assets (i.e., switch, crossing, and bridge) initiate oscillations in dynamic loads. Figures 2b and c also show the magnitude and pattern of dynamic loads are similar and repeatable among all passes.

Random and systematic uncertainties can affect the accuracy of measurements and, consequently, the analyses (Willink 2013; Grassie et al. 1999). The IWS system's accuracy is examined by investigating the

repeatability of measurements between different sensors and measurement runs. Since the train has travelled at various speeds between different runs, the data are preprocessed and adjusted for distance lags and noises. The four steps for preprocessing are including (1) cross-correlation, (2) Empirical mode decomposition (EMD) and Hilbert spectrum (HS), (3) filtering, and (4) data smoothing.

$$x(t) = \sum_{i=1}^n imf_i(t) + r(t) \quad [6]$$

The hypothesis is that the pattern of measurements in different runs is independent of the travel speed and direction. The datasets are smoothed and filtered for low-frequency noises smaller than 0.06 Hz before the preprocessed data are examined for the similarity of

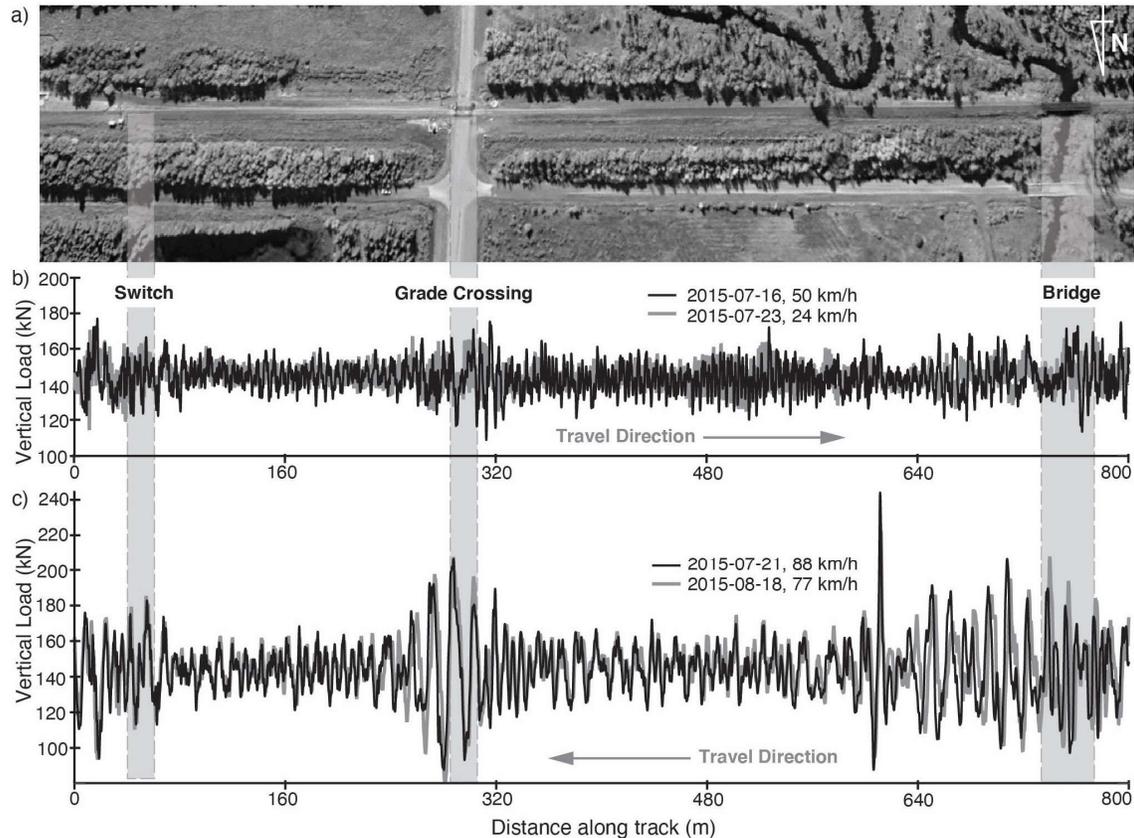


Figure 2. Example vertical load data from the IWS system including tangent track, a switch, a grade crossing, and a short bridge (steel span). These data are presented as (a) a satellite image of the section of track, (b) measurements from two slower moving west-bound trains, and (c) measurements from two faster moving east-bound trains (after Behnia et al. 2021).

Cross-correlation is used first to find the distance lag between measurements in different runs and determine the benchmark. In the second step, EMD is implemented to break down signals to design an appropriate filter to preprocess the data. EMD is a highly-adaptive time-space analysis method appropriate for preprocessing series that are non-linear and non-stationary. The main feature of EMD is to decompose locally high-frequency components of a signal into a separate intrinsic mode function (IMF), and the Hilbert spectral analysis of IMFs provides frequency information. In this method, the signal is sequentially decomposed into the highest frequency component (imf_1) to the lowest frequency component (imf_n) for some finite, n , and a residue signal, r . Finally, there are n IMF's and a residue signal as below (Eq. 6) (Kim & Oh 2009).

variation of patterns between the four runs.

Figure 3 presents the statistical distributions of the magnitude of measured vertical dynamic loads for different track characteristics (i.e., tangent, curve, switch, crossing, and bridge). Table 3 provides the mean (μ), standard deviation (σ), 99.9th percentile, and contribution percentage for these values. The contribution percentages show a significant disparity in the amount of data collected for each type of track asset. Because the mean values are very close to the static load (144 kN), these distributions are naturally normal.

The tangent track has the narrowest distribution (due to μ and σ values) and fewest extreme values. Curved sections and bridges have similar distributions (Figure 3a), as do grade crossings and switches (Figure 3b).

The maximum measured dynamic loads are very high and are representative of outlier conditions that do not appear to provide a reasonable condition on which to base designs. Thus, dynamic loads envelopes based on the removal of outliers (e.g., 99.9th percentile values) are more useful. The highest 99.9th percentile dynamic vertical load occurs for switches.

opportunity to prevent the loss of effective data points. Because the available datasets are approximately normal, Grubbs' test is one of the best procedures to remove outliers (ASTM 2002; Pearson 2005). The Grubbs test is a simple statistical test for outliers, and involves a binary decision based on ratios of the sum of squares of deviations for a reduced sample (Tietjen & Moore 1972). In this study, the authors considered two

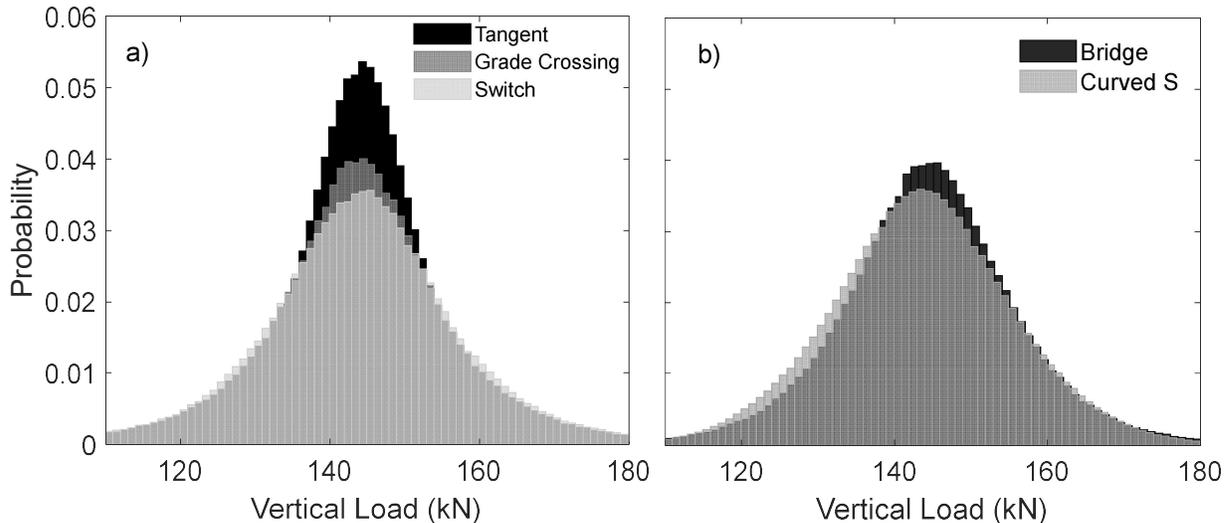


Figure 3. Probability distributions of vertical load measurements from the IWS on (a) tangent track, grade crossings, and switches; and on (b) curves and bridges (based on all four instrumented wheels, and all four passes) (after Behnia et al. 2021).

Table 3. Statistical measures from probability distributions of dynamic vertical load measurements from the IWS (based on all four instrumented wheels, and all four passes) (after Behnia et al. 2021).

Track Type	μ (kN)	σ (kN)	Maximum (kN)	99.9 th (kN)	Contribution (%)
All	143.9	11.3	384.4	177.9	100
Tangent	143.9	10.9	384.4	176.5	69.9
Curved	143.7	12.2	292.8	180.2	24.7
Bridge	144.6	12.4	288.9	181.9	1.6
Grade Crossing	144.2	13.1	264.2	183.4	2.2
Switch	144.1	14.4	302.5	187.2	1.6

4 DISCUSSION

The primary purpose of ϕ equations is in determining design loads. These equations are mainly based on the upper envelope of measured dynamic loads. As mentioned earlier, the measured vertical dynamic loads are naturally normal. Because the available datasets are nearly normal, they confirm outliers in the raw data. Therefore, a first step should be outlier detection. Many existing ϕ equations were generated using 3σ , or the 99.9th percentile, as an upper value exclusive of outliers (Srinivasan 1969; Hay 1982; Esveld 2001; Sadeghi & Barati 2010). However, various alternatives can be considered; the critical part is to discern the outlying observations for particular interests. In many cases, outlying observations come from various operational and device conditions and thus should be removed to prevent adverse effects on the desired results. In addition, outlier detection methods provide an

main steps (data division) to apply Grubbs' test:

- Based on track characteristics; and
- Based on speed increment (every 5 km/h).

Grubbs' test determines a Grubbs value (G), where μ is the mean value of the sample, σ is the standard deviation, and x_i is the value of the i^{th} element of the dataset (Eq. 7).

$$G = \max_{i=1, \dots, n} |x_i - \mu| / \sigma \quad [7]$$

As an example, Figure 4 presents the ϕ derived from the IWS measurements over tangent track, evaluated for 5 km/h increments for the maximum measured dynamic loads (inclusive of outliers), the 99.9th percentile value, and G from the Grubbs' test. In addition, Figure 4 provides a comparison with the AREMA (2018) equation (Eq. 3), as a commonly used measure of ϕ representative of North American and primarily transport freight railways such as the study site. Figure 4 shows the ϕ evaluated from the Grubbs test for removing outliers consistently increases with increasing speed, whereas the increasing trend of ϕ evaluated from the 99.9th percentile is variable. The figure also shows the AREMA (2018) equation (Eq. 3) provides a close representation of the ϕ derived from the Grubbs test applied to IWS data from tangent track and train speeds above 60 km/h. In other words, Eq. 3 increasingly underestimates ϕ with decreasing train speed below 60 km/h. Another significant aspect of estimating the dynamic load factor is the difference between actual and

estimated ϕ values for track assets (non-tangent track), which are more severe than for tangent track.

A quadratic equation fits the ϕ values for tangent track derived using Grubbs' test as a function of train speed (Eq. 8). This equation provides a close fit with an R^2 value of 0.97. Due to the similar statistical distributions of the measurements collected for grade crossings, switches, curves, and bridges in Figure 3, these non-

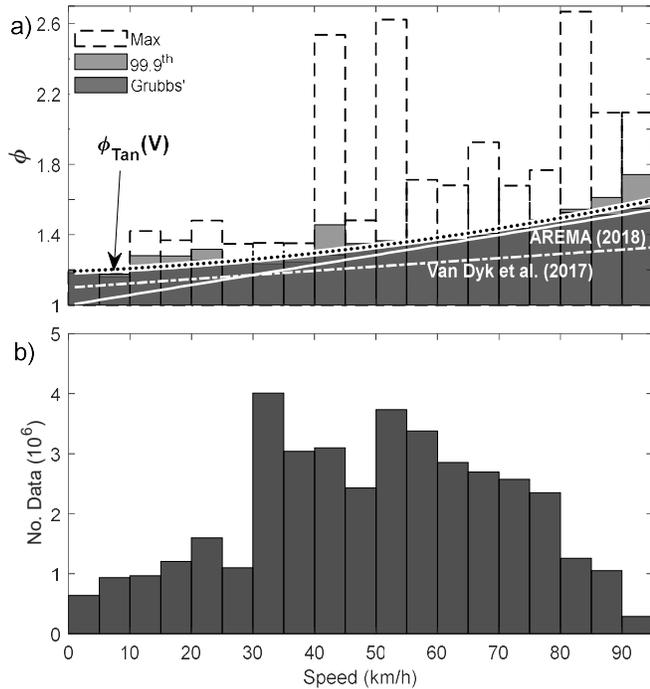


Figure 4. IWS measurements from tangent track as (a) dynamic load factor (ϕ) for 5 km/h increments of train speed, and (b) the number of measurements used to develop the ϕ for each increment of train speed (after Behnia et al. 2021).

$$\phi_{Tan} = 3 \times 10^{-5}V^2 + 1.4 \times 10^{-3}V + 1.19 \quad [8]$$

$$\phi_{NTan} = 3 \times 10^{-5}V^2 + 2.5 \times 10^{-3}V + 1.27 \quad [9]$$

Finally, the presented ϕ values (i.e., AREMA, tangent, and non-tangent) were re-evaluated over the range of permissible train speeds for various track classes (Table 4). Because the IWS wheelsets have very low wear and are free of defects, the measured dynamic loads are a result of speed, variable track conditions, and dynamic characteristics of the car. Thus, the results in this paper go beyond previous studies (such as those presented in Van Dyk et al. (2017) from WILD sites) by showing the effect of track conditions is substantial and exceeds previous estimates (as provided by AREMA 2018). The authors recommend the consideration of the combined effect of various parameters (e.g., different locomotives and car types) in future studies.

tangent track datasets were grouped and plotted in Figure 5, and a quadratic equation fit to ϕ values derived using Grubbs' test was generated (Eq. 9). This equation provides a close fit with an R^2 value of 0.91.

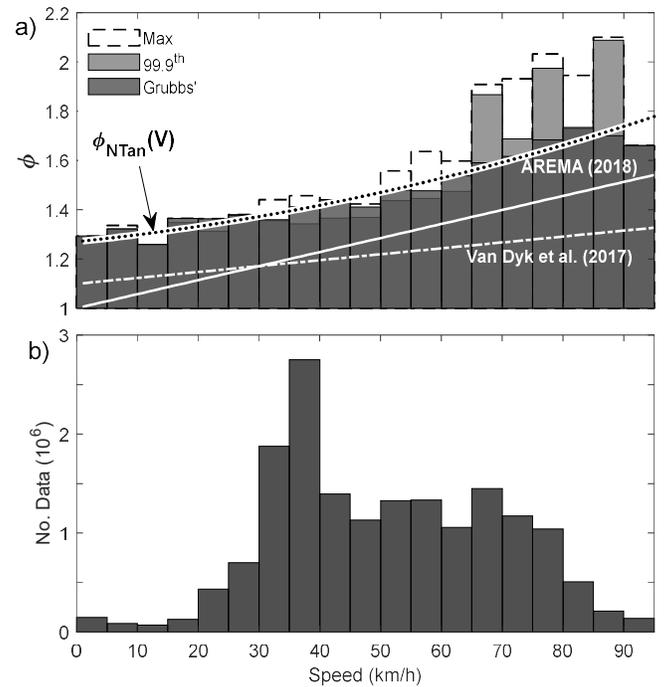


Figure 5. IWS measurements from non-tangent track as (a) dynamic load factor (ϕ) for 5 km/h increments of train speed, and (b) the number of measurements used to develop the ϕ for each increment of train speed (after Behnia et al. 2021).

Table 4. Dynamic load factors (ϕ) derived from Grubbs' test values from the IWS data up to the maximum allowable freight train speed within each North American class of track (after Behnia et al. 2021).

Class	Max. Speed (km/h)	AREMA (Eq. 3)	Tangent (Eq. 8)	Non-Tangent (Eq. 9)
1	16	1.09	1.22	1.36
2	40	1.23	1.28	1.41
3	64	1.36	1.37	1.54
4	97	1.55	1.55	1.73

5 CONCLUSIONS

This study analyzed the variation in magnitude of vertical loads along a section of Canadian freight railway, derived from data collected from IWS. This investigation differs from past studies because it considers various track characteristics. The IWS provided measurements of dynamics loads under

constant car type and conditions, but for the variety of track characteristics found over 340 km of in-service track. In addition, this study mainly focused on ϕ in consideration of the dynamic load factor as the main parameter in design. The data show the impact of track conditions is significant, resulting in ranges of dynamic loads and values of ϕ that track may experience well over those provided by typical means of estimating; this was especially notable for non-tangent track (inclusive of curves, switches, crossings, and bridges). Equations considering ϕ as a function of speed (Eqs. 8 & 9) were developed to quantify this range. The authors propose further work is yet required to develop ranges of dynamic loads and values of ϕ that incorporate both track conditions and variations in car type and condition.

6 ACKNOWLEDGEMENT

The authors thank the Canadian National Railway (CN) and National Research Council Canada (NRC) for their support and facilitation of this project, specifically Tom Edwards and Albert Wahba for facilitating the collection of these datasets. This research was made possible through the Canadian Rail Research Laboratory (www.carrl.ca).

7 REFERENCES

AREMA (American Railway Engineering & Maintenance of Way Association). (2018). Manual for railway engineering, Lanham, MD.

ASTM, E. 178-02, Standard Practice for Dealing with Outlying Observations. (2002). American Society for Testing and Materials.

Barbosa, R. S. (2016). Evaluation of railway track safety with a new method for track quality identification. *Journal of Transportation Engineering: Part A- Systems*, 1; 142(11):04016053.

Behnia D., Hendry M.T., Abdulrazagh P.H., & Wahba A. (2021). Railway dynamic load factors developed from instrumented wheelset measurements. *Journal of Transportation Engineering: Part A- Systems*, (*in Press*).

Birmann F., (1965). Paper 5: Track Parameters, Static and Dynamic. In: Proceedings of the Institution of Mechanical Engineers, Conference Proceedings June (Vol. 180, No. 6, pp. 73-85). Sage UK: London, England: SAGE Publications.

Bracciali A., Cavaliere F., & Macherelli M. (2014). Review of instrumented wheelset technology and applications. In: The second international conference on railway technology: research, development and maintenance (Vol. 167, pp. 1-16).

Cakdi S., Cummings S., & Punwani J. (2015). Heavy haul coal car wheel load environment: rolling contact fatigue investigation. In: Joint Rail Conference. March.

American Society of Mechanical Engineers Digital Collection.

Dybała, J., & Radkowski, S. (2013). Reduction of Doppler effect for the needs of wayside condition monitoring system of railway vehicles. *Mechanical Systems and Signal Processing*, 38(1), 125-136.

Esveld C. (2001). Modern railway track. Zaltbommel, Netherlands: MRT-productions.

Fallah Nafari S., Gül M., Hendry M.T., Otter D., & Cheng R.J. (2018a). Operational vertical bending stresses in rail: Real-life case study. *Journal of Transportation Engineering: Part A- Systems*. March. 1; 144(3): 05017012.

Fallah Nafari S., Gül M., Hendry M.T., & Cheng J.R. (2018b). Estimation of vertical bending stress in rails using train-mounted vertical track deflection measurement systems. Proceedings of the Institution of Mechanical Engineering, Part F: Journal of Rail and Rapid Transit. May. 232(5):1528-38.

Grassie S.L., Saxon M.J., & Smith J.D. (1999). Measurement of longitudinal rail irregularities and criteria for acceptable grinding. *Journal of Sound and Vibration*, 227(5), 949-964.

Hay WW. (1982). Railroad engineering (Vol. 1). John Wiley & Sons; June. 16.

Higgins R.L., Otter D.E., & Martin R.W. (1992). High accuracy load measuring wheelset. In: 10th International Wheelset Congress: Sharing the Latest Wheelset Technology in Order to Reduce Costs and Improve Railway Productivity. Preprints of Papers (p. 181). Institution of Engineers, Australia.

Kim D., Oh H.S. (2009). EMD: a package for empirical mode decomposition and Hilbert spectrum. *The R Journal*, 1(1), 40-46.

Pearson R.K. (2005). Mining imperfect data: Dealing with contamination and incomplete records. Society for Industrial and Applied Mathematics. January 1.

Leishman E.M., Hendry M.T., & Martin C.D. (2017). Canadian main track derailment trends, 2001 to 2014. *Canadian Journal of Civil Engineering*. 44 (11):927-34.

Peters N. (2010). CN Railway Engineering Course. Retrieved from McGill University Civil Engineering Graduate Studies class.

Ren Y., & Chen J. (2019). A new method for wheel-rail contact force continuous measurement using instrumented wheelset. *Vehicle System Dynamics*. February. 1; 57(2): 269-85.

Roghani A., & Hendry M.T. (2016). Continuous vertical track deflection measurements to map subgrade

condition along a railway line: Methodology and case studies. *Journal of Transport Engineering*. December. 1; 142 (12): 04016059.

Roghani A., & Hendry M.T. (2017). Quantifying the impact of subgrade stiffness on track quality and the development of geometry defects. *Journal of Transportation Engineering: Part A- Systems*. July. 1; 143 (7): 04017029.

Roghani A., Macciotta R., & Hendry M.T. (2015). Combining track quality and performance measures to assess track maintenance requirements. In: ASME/IEEE Joint Rail Conference. March 23rd (Vol. 56451, p. V001T01A009). American Society of Mechanical Engineers.

Roghani A., Macciotta R., & Hendry M.T. (2017). Quantifying the effectiveness of methods used to improve railway track performance over soft subgrades: methodology and case study. *Journal of Transportation Engineering: Part A- Systems*. September. 1; 143(9): 04017043.

Sadeghi J. (2012). New advances in analysis and design of railway track system. *International Journal of Reliability and Safety in Railway*. March. 30:75-100.

Sadeghi J., & Barati P. (2010). Evaluation of conventional methods in Analysis and Design of Railway Track System. *International Journal of Civil Engineering*. March. 10; 8 (1): 44-56.

Srinivasan M. (1969). *Modern permanent way*. 555 p., 1969.

Tietjen G.L., & Moore R.H. (1972). Some Grubbs-type statistics for the detection of several outliers. *Technometrics*. August. 1; 14 (3): 583-97.

Van Dyk B.J., Edwards J.R., Dersch M.S., Ruppert Jr C.J., & Barkan C.P. (2017). Evaluation of dynamic and impact wheel load factors and their application in design processes. *Proceedings of the Institution of Mechanical Engineering, Part F: Journal of Rail and Rapid Transit*. January. 231 (1): 33-43.

Willink R. (2013). *Measurement uncertainty and probability*. Cambridge University Press.

Woelfle, A. (2016). Analysis of Wheel-Rail Forces during (MRail) Rolling Deflection Tests (technical Report), Report for the National Research Council, Canada, February, 19th.

Yu F., Hendry M.T. (2019). A new strain gauge configuration on the rail web to decouple the wheel-rail lateral contact force from wayside measurement. *Proceedings of the Institution of Mechanical Engineering, Part F: Journal of Rail and Rapid Transit*. October. 233 (9): 951-960.

Evaluating railway track stiffness variations using instrumented wheelset (IWS) and accelerometer measurements

Reza Mousapour, Parisa Haji Abdulrazagh, Michael T. Hendry and Mustafa Gül

Department of Civil and Environmental Engineering - University of Alberta, Edmonton, Alberta, Canada

ABSTRACT

This paper evaluates the relative stiffness variation along the track using the wheel/rail force and acceleration measured by instrumented wheelset (IWS) and accelerometer. A finite element simulation approach is developed to imitate the actual IWS experiment conducted on a 250-mile railway track in Quebec, Canada. In the site study, the truck and carbody accelerations of the railcar are measured concurrently with the wheel/rail force measurements, and therefore simulation data is collected in the same way. The ratio of the magnitude of cyclic loading and the cyclic displacement is used to estimate the relative stiffness along the track. Cyclic displacements are interpreted by integrating the accelerometer data at the most contributing frequency being common within both loads and accelerations. A method is proposed to calculate the load/displacement ratio by using the detected frequency which is the most relevant between the acceleration and load signals. The method is developed using the frequency domain decomposition (FDD) technique with implementing the singular value decomposition (SVD) as its mathematical tool. To evaluate the proposed methodology a numerical simulation is developed, and the load and acceleration signals are acquired in accordance with the current IWS experiment. The railway track is simulated using a beam on an elastic foundation and the railcar is modeled with a sprung mass. The results obtained from the simulation indicate that proposed methodology could estimate the relative stiffness variation along the beam. In conclusion, the cyclic load/displacement ratio analysis using the predetermined frequencies could be considered as the representative of the variations along the track support.

1 INTRODUCTION

There is a growing interest in evaluating the railway track using instrumented car. In comparison with conventional methods, the instrumented cars could provide continuous information from the track without making any disruption in railway current services (Yang, Y. B., et al, 2020). The measurements are used to identify problematic sections of tracks and help the rail organizations to schedule maintenance programs accordingly (Malekjafarian, A, et al, 2021). Wheel-rail dynamic forces are of the major importance of rail track responses that could be measured using the instrumentation of the track or of the vehicle. It could be used to evaluate the effects of various tracks characteristics on dynamic loads, such as the effect of different track features, stiffness transitions, track geometry roughness and defects in the track infrastructures. Using the dynamic response of the train, the stiffness variation along the track could be estimated (Quirke, P., et al., 2017). Large dynamic loads are generally attributed either to track defects associated with large deflections and soft subgrades (Selig and Li, 1994; Selig et al, 2001; Hendry et al., 2008;), or to variation of track deflections as a result of differential track stiffness (Tzanakakis, 2013; Li and Berggren, 2009; Fröhling,1997). For both cases, track stiffness/

track modulus plays a major role. Quantifying the track modulus and its variations along the track could be used to evaluate the track structure (Do, N. et al, 2017). Track stiffness is the ratio of vertical wheel load to the produced vertical rail deflection. Track stiffness represents track substructure and superstructure, whereas track modulus is not inclusive of rail bending stiffness (Selig et al, 1994). It is used as a key parameter for assessing the condition of track-supporting components (Zarembski, A.M., et al, 1979).

Several methods are developed for determining track modulus and track stiffness. Those methods that involve dynamic excitation and loading are as follows: (i) Rolling stiffness measurement vehicles (RMSVs) are used to calculate the track stiffness using force and acceleration measurements. (Berggren et al.,2002 &2005); (ii) A Portancemeter is used to measure track stiffness using a vibrating wheel suspended by a spring and damper. (Hosseingholian,2006); (iii) A continuous rail deflection measurement system, so-called MRail, is developed at the University of Nebraska, Lincoln. The MRail can be mounted on railcars in service to measure relative rail deflections between the bogie and rail using line-lasers. (Norman et al., 2004; McVey et al., 2005; Arnold et al., 2006). Roghani et al. (2016, 2017a and 2017b) and Fallah Nafari et al. (2018) have studied the MRail measurements to estimate track modulus and to

map subgrade condition along a railway track; and (IV) Using a rail Falling Weight Deflectometer (FWD), measure the modulus of track substructure layers excluding the rails (Haji Abdulrazagh et al., 2016). Haji Abdulrazagh et al. (2019) then developed the Rail FWD device in conjunction with a back calculation program, RAILBACK that estimates the substructure layer moduli by direct impact on the rail.

Existing research mainly focuses on the development of a framework to evaluate relative track stiffness variation using the data obtained from Instrumented Wheel System (IWS) system. The objective is to assess the railway substructure condition by calculating stiffness and evaluating its variation along the track using IWS measurements. The paper describes an evaluation process to examine the developed method using a finite element model that imitates the IWS experiment. The results show it is possible to determine the stiffness variation relatively using the proposed algorithm.

2 BACKGROUND

The instrumented wheelset system (IWS) was developed in 1975 to measure the contact forces at wheel-rail interface to assess the critical running conditions (Prause et al., 1975; Cazzulani et al., 2016). IWS is the most accurate and direct method for measuring wheel-rail contact forces. The IWS system continuously measures the component forces on the wheel using strain gauges attached to the wheel (Ham et al., 2009). The National Research Council Canada (NRC) Automotive and Surface Transportation (AST) Research Centre in collaboration with Iron Ore Company instrumented a freight railcar. The in-service freight railcar is equipped with an IWS and accelerometers to measure the wheel/rail force and acceleration at the wheel and carbody respectively. The railcar instrumented with 12 strain gages on two

wheelsets (IWS-A and IWS-B) to measure wheel-rail forces in three directions. IWS data and acceleration are collected on Iron Ore Company's (IOC) railway track along a round-trip travel from Sept-Îles to north on Ross Bay JCT (milepost 9.00-224.0) in the Wacoua subdivision, and from Ross Bay JCT to Labrador city (milepost 0.00-36.00) in the Northern Land subdivision in the east of the province of Quebec. This data consists of wheel-rail forces and acceleration measurements from the trucks and carbody of the railcar, as well as speed and global positioning system (GPS) coordinates. The track geometry data is also collected over the similar rail route. The lateral wheel contact position, as well as the global positioning systems (GPS), speed, and time are also recorded. Carbody and side frame vertical accelerations, draft sill lateral and longitudinal accelerations, bolster car displacement, brake pressure and coupler force are other signals that are measured by the same railcar. Figure 1 shows the details of the instrumented railcar and configuration and the position of the sensors.

Among the collected signals, wheel-rail vertical forces and side frame vertical accelerations will be used for estimating stiffness parameter along the chosen track sections. The data collected from the site will be used to apply the proposed method to evaluate the track substructure. Track geometry data will be investigated in the future for understanding the range of wavelengths for track geometry defects. These defects might need to be filtered out to obtain clean signals related to the stiffness variation of track substructure. The aim of this paper is to present and evaluate a proposed method to determine the track stiffness variation using contact force and acceleration data. In the future, the validated method will be used to evaluate track stiffness using IWS and accelerometer data from the site measurements.

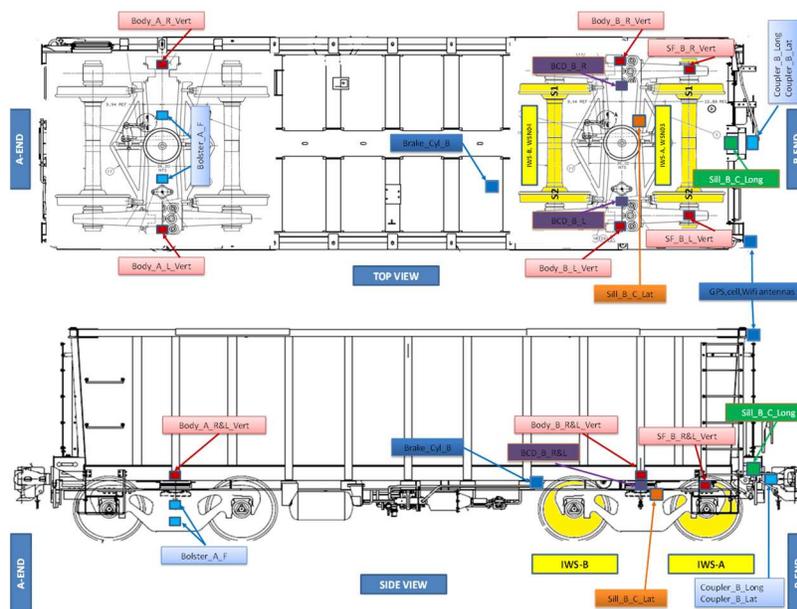


Figure 1. Configuration of sensors mounted on the railcar (a) plan view, (b) side view (NRC, 2020).

3 MATERIALS AND METHODS

To evaluate the implemented methodology, a finite element model is simulated. The details of FEM model and the results of the evaluation of the proposed algorithm is discussed in this section. The simulation is based on modelling a beam on elastic foundation and a sprung mass mounted on a wheel which is moved along the beam. According to the experiments, loads and acceleration are collected from the wheel and the sprung mass.

3.1 MODEL DESCRIPTION

Numerous studies have demonstrated that the beam on elastic foundation simulation works well in evaluating vehicle-track interaction. To imitate the IWS experiment, a sprung mass is modelled and moved along the beam on an elastic foundation. The dynamic analysis of beam on elastic foundation subjected to moving load is discussed in Thambiratnam et al. and it is used in simulation assumptions (Thambiratnam, D., et al., 1996). The acceleration on sprung mass and the vertical load at the contact point are measured. These measurements are used in the proposed methodology to calculate the stiffness and showing its variation. In order to simulate the rail over the substructure, a beam on an elastic foundation is created. The dynamic analysis of the beam is performed when it subjected to a moving sprung mass. The assumptions for the modeling of the beam (rail) and vehicle are presented accordingly.

- Track Model

The track is simulated using a continuous beam with a uniform cross-section over a set of springs. (Figure 2) The rail is assumed as the beam and it is modelled using a rectangular section. The total length of the beam is 110m in this simulation.

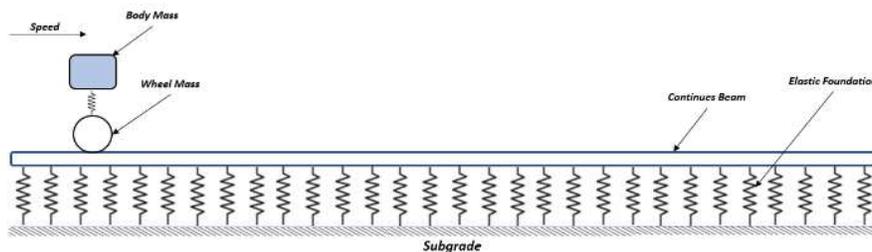


Figure 2. Schematic of moving sprung mass over the Beam on Elastic Foundation.

The Young's modulus, Poisson ratio, and the density of the steel are assumed. The boundary condition of the beam is assumed to be pinned at both ends. The springs are located under the beam and connected to the ground. The beam properties are given in Table 1.

Table 1. Properties of the Beam.

Property	Unit	Symbol	Value
Elastic Modulus	N/m^2	E_b	209×10^9
Beam cross section	m^2	A_b	7.7×10^{-9}
Beam second moment of area	m^4	I_b	8.1×10^{-7}
Springs spacing	m	L_s	0.853

- Vehicle Model

To perform the dynamic analysis, the railcar is modelled with a quarter car vehicle. The single degree of freedom (SDOF) vehicle model is simplified with a sprung mass. The mass is mounted on the spring and will move along the beam on an elastic foundation. The sprung mass moved with the constant speed ($v=1m/s$). The mass of the wheel, the primary suspension stiffness, and carbody mass is presented in Table 2. Figure 3 shows the load and accelerations are measured at the contact point and carbody respectively. The sprung mass properties are given in Table 2.

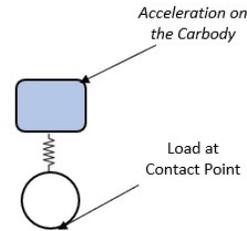


Figure 3. Sprung mass and the location of Load and acceleration measurement.

Table 2. Properties of the sprung mass

Property	Unit	Symbol	Value
Carbody Mass	Kg	m_c	10000
Wheel Mass	Kg	m_w	800
Spring stiffness	N/m	k_s	3×10^7

3.2 METHODOLOGY

The method is proposed to estimate the stiffness variation of track substructure using the acceleration and the IWS data set. Variations in track stiffness cause low frequency changes in the dynamic vehicle's response (Quirke, P. et al., 2017).

Based on the frequency analysis of the vehicle's acceleration and load signals, track stiffness variations could be evaluated. This method is based on calculating the load/displacement ratio and infer the track stiffness accordingly. Both IWS and acceleration data are used to calculate the ratio of the magnitude of cyclic loading to the magnitude of cyclic displacement interpreted from the double integration of the accelerometer data at a common frequency/wavelength by (Selig and Li, 1994):

$$K = \frac{P_{\max_amp}}{d_{\max_amp}} \quad (1)$$

Where K is the stiffness, P_{max_amp} is the magnitude of vertical wheel load, and d_{max_amp} is the maximum rail deflection beneath the wheel. Once the common frequency is determined the rail displacements, d , are calculated from the acceleration data (Hendry et al.,2011):

$$d(t)_{cyclic} = D \cos(2\pi ft - \varphi) \quad (2)$$

$$a(t)_{cyclic} = -(2\pi f)^2 D \cos(2\pi ft - \varphi) \quad (3)$$

$$d(t)_{cyclic} = \frac{a(t)_{cyclic}}{-(2\pi)^2} \quad (4)$$

Where D is the amplitude of the cyclic displacement in mm, a is the acceleration mm/s^2 , f is the frequency of motion in Hz, and φ is the phase shift in rad. In this method, selecting the dominant frequency is a challenging step in the process of calculating the displacement as well as the load/displacement ratio. Figure 4 shows the flowchart to select the common frequencies and calculate the load/displacement ratio accordingly.

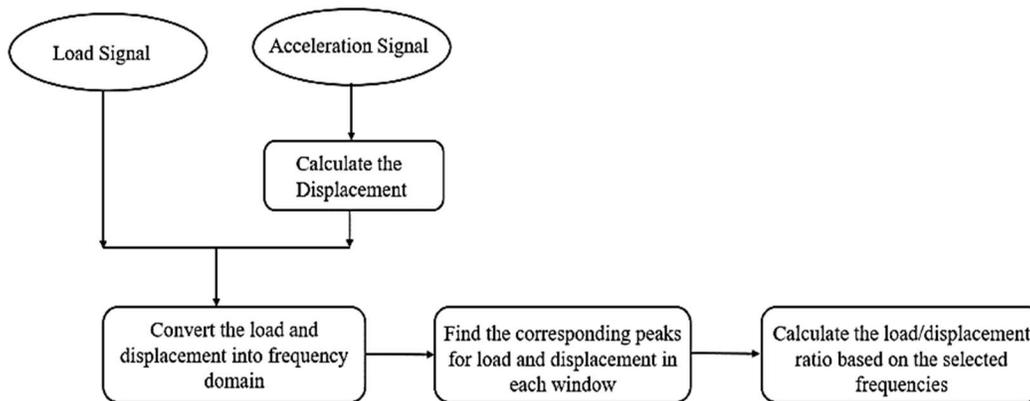


Figure 4. Flowchart to select the corresponding frequencies of vertical load and displacement.

To find the dominant frequency, a mathematical tool is used to decompose the signals' features. In order to decrease the dimension of the problem and classification the data, the Singular Value Decomposition (SVD) has been used. This mathematical approach is used to detect the dominant frequency in different studies by Brincker et al and Malekjafarian et al. (Malekjafarian, A., et al.,2014). In those studies, the power spectral density matrix of the response in each frequency is decomposed by taking Singular Value Decomposition (SVD) of the matrix. The SVD provides a systematic way to determine a low-dimensional approximation in terms of dominant patterns.

In proposed algorithm, the appropriate frequencies to calculate load and displacement are selected using the singular value decomposition method. For the selected frequencies, the stiffness is calculated using the proposed methodology to evaluate the stiffness variation. The results show that by selecting major frequencies using the SVD method, the stiffness variations are consistent. The procedure described in the load-displacement method to calculate the stiffness parameter is repeated for a moving window along the load and acceleration signals obtained along the track. The track stiffness is speed-dependent (Priest et al.,2009). Along the studied track the recorded speed is relatively constant which is a necessary term of applicability of the load-displacement method.

4 MODEL VERIFICATION

For the purpose of verifying the simulation, the beam is assumed to be supported by elastic foundations set with

different spring stiffness. It is essential to confirm that the model results match the analytical calculations. The deflection and stiffness of the beam should be compared to the analytical values. Figure 5 shows the modeling assumptions for stiffness magnitude and distribution of springs as the foundation stiffness. The stiffness of the springs is assumed to be $[4.7 - 9.4 - 2.35 - 4.7] \times 10^7 N/m$ with the presented distribution under the beam (see figure 5). The beam is subjected to the constant and harmonic moving load along the beam. The deflections of the beam are obtained accordingly. Using the output of the model, the stiffness of the beam is calculated and compared with the main assumptions

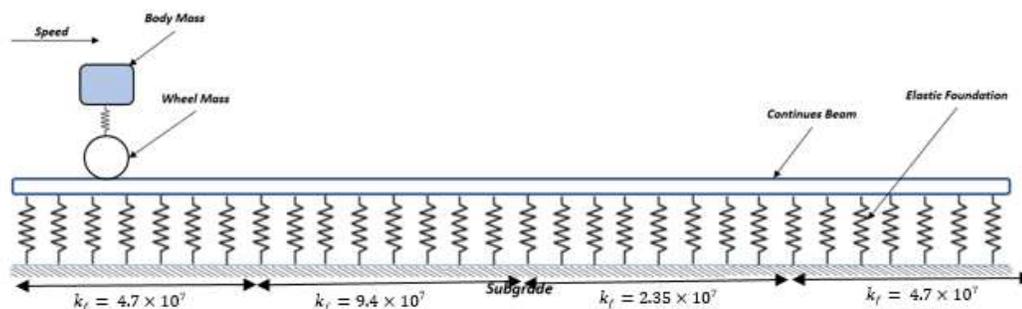


Figure 5. Schematic of Beam on variable Elastic Foundation.

of the simulation to determine whether it follows the magnitude and distribution of the springs' stiffness.

- Constant Moving load

Figure 6 represents the deflection and the stiffness of the beam under constant moving load. The deflections are calculated exactly under the load while it moves over the beam. The variation in the deflection and stiffness plots are associated with the spring stiffness variations in the model. For instance, where the beam is on the stiffer section of foundation, its deflection is small, and its stiffness plot shows greater values.

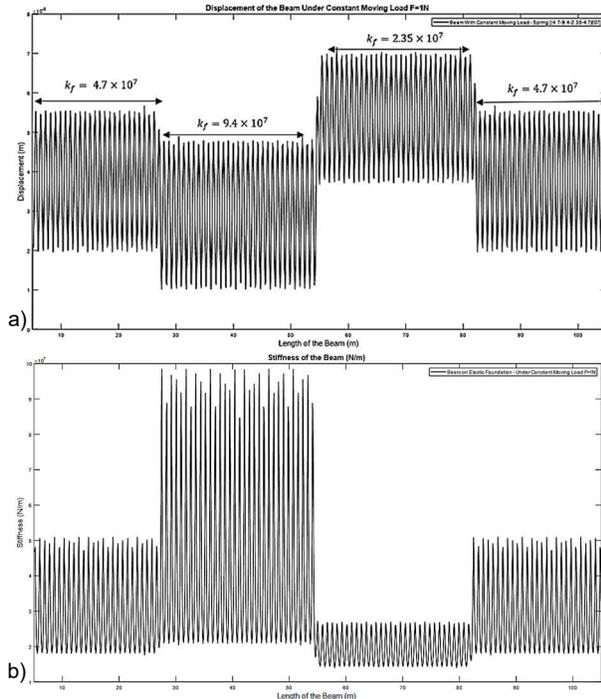


Figure 6. (a)Deflection of the BEF under Constant moving load (b)Calculated Stiffness of the BEF under Constant moving load

- Harmonic Moving load

In the next step, the beam is subjected to a harmonic moving load with a constant speed. The deflections and stiffness of the beam are plotted under this load condition and the results are presented in figure 7. Consistency between the results is evident and associated with spring stiffness variation. The beam is analyzed in different boundary conditions and load cases to see if the developed simulation can obtain comparable results with the analytical calculation as well as if it can capture the spring's stiffness when it is on an elastic foundation. As demonstrated by the results, the model works with different assumptions and conditions, and it could be used for further analysis during the study.

5 RESULTS AND DISCUSSIONS

The recommended methodology for calculating the variation of the stiffness along the track is examined using the signals obtained from the FEM model. The dynamic response and wheel/rail interaction could be

acquired from the created FEM models. This section presents the deflection and stiffness calculations using acceleration and load signals obtained from a quarter-bogie model.

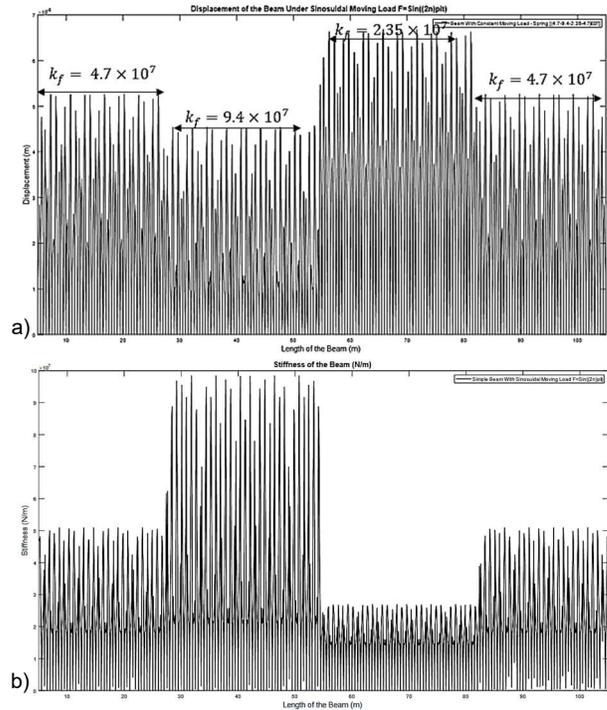


Figure 7. (a)Deflection of the BEF under Harmonic moving load $F = (\sin(2n \cdot \pi \cdot t))$, and (b). Stiffness of the BEF under Harmonic moving load $F = (\sin(2n \cdot \pi \cdot t))$.

The simulated beam on elastic foundation is subjected to the sprung mass that moves along the beam. Regarding the IWS experiment, the acceleration is measured at the sprung mass, and the load is measured at the contact point. Using the frequency domain decomposition technique in the proposed methodology, the deflection of the beam and the variation of stiffness are evaluated. A foundation with diverse stiffness sections is assumed and four different types of springs are used as the foundation stiffness by the following values: $[4.7 - 9.4 - 2.35 - 4.7] \times 10^7$ N/m. The model output signals are used in a proposed methodology to estimate stiffness and determine its variation along a beam. The following plots show the acceleration of the contact point and the body of the sprung mass. It can be seen from these graphs that acceleration varies in consist with the model condition. Regarding the acceleration plot of the contact point, it is evident the stiffer part of the foundation has lower values while the softer part has higher values (see Figure 8(a)). To compare the acceleration variations, Figure 8(b) shows the normalized acceleration plots of the contact point and the body of sprung mass. The plots indicate the primary suspension of the railcar have significant impacts on the measured acceleration. While the acceleration transmitted from the wheel to the sprung mass, the magnitude of the measured acceleration decreased due to the primary suspension. Furthermore,

comparing the variation of the measure acceleration at the contact point with the sprung mass indicates some of the foundation information is missed. The body acceleration and load obtained from the FEM model are considered as input signals to use in the proposed methodology. According to the presented algorithm, the most contributing frequencies being common within both loads and accelerations are detected using the frequency domain decomposition (FDD) technique.

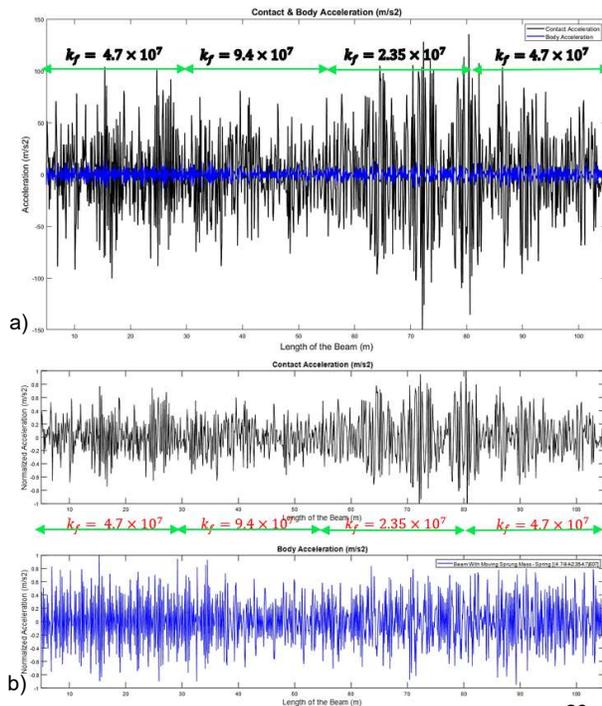


Figure 8. (a) Contact and Body Acceleration, and (b) Normalized Contact and Body Acceleration.

Following plots represent the calculated deflection and stiffness variations of the beam. The deflection plot (Figure 9(a)) shows the variation that results from different stiffnesses of the beam. Accordingly, the deflection variation reflects the change in spring stiffness, so the stiffer part of the beam has lower values than the softer part. Figure 9(b) shows the stiffness variation over the beam with the diverse stiffness section. It is evident from these graphs that calculated stiffness variation is associated with the spring stiffness that was set in the model.

6 CONCLUSION

Stiffness variation in railway track could be detected in vehicle's response signals. Therefore, the stiffness variation can be evaluated by measuring the load and acceleration signals. Using a passing train, the response signals of vehicles are collected by the IWS system and accelerometers located on wheels and rail car respectively. Stiffness variations of the track is a key factor that is used to evaluate the substructure conditions. To evaluate the track substructure and

determine the problematic section of the track, it is essential to quantify the track stiffness and evaluate its variations. A method is presented in this paper to calculate the deflection and stiffness of the beam using the load/deflection method. The deflections are calculated by integrating the accelerometer data at the most contributing frequency being common within both loads and accelerations.

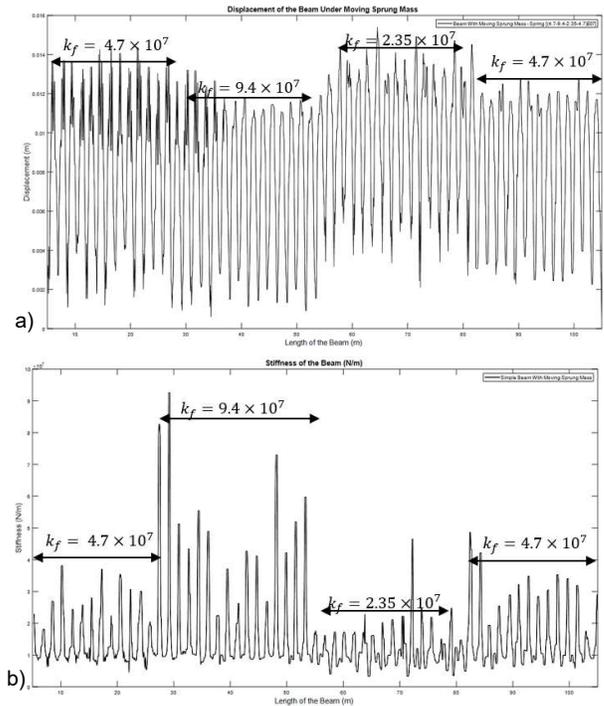


Figure 9. (a) Deflection of the BEF under Moving Sprung Mass – Using Moving median calculation (b) Stiffness of the BEF under Moving Sprung Mass – Using Moving median calculation.

To overcome the challenges in finding the most common frequencies the proposed method is using the frequency domain decomposition (FDD) technique with implementing the singular value decomposition (SVD) as the mathematical tool. For the purpose of validating the proposed methodology, a beam on elastic foundation is created using FEM software. The simulated beam has been verified using both static and dynamic analysis. To imitate the experiment, the dynamic analysis performed by the sprung mass moving along the beam. The deflection and stiffness of the beam were calculated, and the results are consistent with the assigned stiffness under the beam. Regarding the results presented in this paper, it can be concluded that the proposed method could potentially be used to evaluate the stiffness variation of the beam's foundation using the load and acceleration measured on the wheel and carbody. This method could be used to evaluate the track substructure when the load and acceleration are collected along the section. The proposed method will be applied by the data collected from the site to evaluate the track substructure.

7 ACKNOWLEDGEMENT

The authors would like to acknowledge National Research Council Canada (NRC) for providing constructive feedback on this study. This research was made possible through the Canadian Rail Research Laboratory (CaRRL) (www.carrl.ca). Research funding for CaRRL is provided by the Natural Sciences and Engineering Research Council of Canada (NSERC-IRC 523369-18), Canadian National Railway, the National Research Council of Canada, and Transport Canada.

8 REFERENCES

Berggren E, Jahlénus Å and Bengtsson B-E., 2002. Continuous track stiffness measurement, an effective method to investigate the structural conditions of the track. International Conference Railway Engineering 2002, Engineering Technics Press, Edinburg, UK.

Cazzulani G. Gialleonardo ED. Bionda S. Bassetti M. Crosio P. and Braghin F. 2016. A new approach for the evaluation and improvement of the metrological characteristics of an instrumented wheelset for the measure of wheel-rail contact forces. Proc. Inst. Mech. Eng. F J. Rail Rapid Transit, 231 (4), 381-393.

Do, Ngoan T., Saeideh Fallah Nafari, and Mustafa Gül. "Estimation of track modulus over long distances using artificial neural networks." *Procedia engineering* 199 (2017): 2102-2107.

Fallah Nafari S. Gül M. Roghani A. Hendry M and Cheng JJR. 2018. Evaluating the potential of a rolling deflection measurement system to estimate track modulus. Proc. Inst. Mech. Eng. F J. Rail Rapid Transit, 232 (1), 14–24.
Fröhling, RD. 1997. Deterioration of railway track due to dynamic vehicle loading and spatially varying track stiffness, PhD-Thesis, University of Pretoria.

Haji Abdulrazagh P and Hendry MT. 2016. A case study of the use of a falling weight deflectometer to investigate railway infrastructure constructed upon soft subgrades. *Can Geotech J*, 53 (12), 1991–2000.

Haji Abdulrazagh P. Farzaneh O and Behnia C. 2019. Evaluation of railway trackbed moduli using the rail falling weight test method and its backcalculation model. Proc. Inst. Mech. Eng. F J. Rail Rapid Transit, 233(4), 431–447.

Ham, Young-Sam, et al. "Continuous measurement of interaction forces between wheel and rail." *International Journal of Precision Engineering and Manufacturing* 10.1 (2009): 35-39.

Hendry MT., Barbour SL, and Martin CD. 2011. An evaluation of real-time deformation monitoring using motion capture instrumentation and its application in monitoring railway foundations. *Geotechnical Testing Journal* 34.6 (2011): 602-612.

Hosseingholian, M. Froumentin, M. Robinet, A. 2006. Feasibility of a continuous method to measure track stiffness. In Proceedings from Railway Foundations conference. Birmingham.

Li M. and Berggren E. 2010. A study of the effect of global track stiffness and its variations on track performance: Simulation and measurement. Proc. Inst. Mech. Eng. F J. Rail Rapid Transit 224(5), 375-382.

Malekjafarian, A.; OBrien, E.J.; Quirke, P.; Cantero, D.; Golpayegani, F. Railway Track Loss-of-Stiffness Detection Using Bogie Filtered Displacement Data Measured on a Passing Train. *Infrastructures* 2021, 6, 93.

Malekjafarian, Abdollah, and Eugene J. O'Brien. 2014. "Application of output-only modal method in monitoring of bridges using an instrumented vehicle." *Civil Engineering Research in Ireland*, Belfast, UK, 28-29 August 2014.

McVey, B. Norman, C. Wood, N. Farritor, S. Arnold, R. Fateh, M. El-Sibaie, M. 2005. Track modulus measurement from A moving railcar, In Proceedings of the AREMA Annual Conference, Chicago, IL, September 25-27, 2005.

Norman, C. Farritor, S. Arnold, R. Elias, SEG. Fateh, M. Sibaie, ME. 2004. Design of a System to Measure Track Modulus from a Moving Railcar, In Proceedings from Railway Engineering, London.

Prause RH. and Harrison DH. 1975. Data analysis and instrumentation requirement for evaluating rail joints and rail fasteners in urban track, Final Report, Prepared for US Department of Transportation.

Priest, JA., and Powrie. W. 2009. Determination of dynamic track modulus from measurement of track velocity during train passage. *Journal of Geotechnical and Geoenvironmental Engineering*, 135(11), 1732-1740.

Quirke, Paraic, et al. "Drive-by detection of railway track stiffness variation using in-service vehicles." Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 231.4 (2017): 498-514.

Roghani A and Hendry MT. 2016. The use of continuous vertical rail deflection measurements to map subgrade condition along a railway line: Methodology and case studies. *ASCE J Transp Eng*, 142 (12).

Roghani A and Hendry MT. 2017. Quantifying the impact of subgrade stiffness on track quality and the development of geometry defects. *ASCE J Transp Eng*, 143 (7).

Roghani A. Macciotta R. and Hendry MT. Quantifying the effectiveness of methods used to improve railway

track performance over soft subgrades: Methodology and case study. ASCE J Transp Eng 2017, 143 (9).

Selig, ET. and Li, D. 1994. Track modulus: Its meaning and factors influencing it. Transp. Res. Rec.,1470, 47–54.

Selig, ET. and Waters, JM. 1994. Track geotechnology and substructure management, Thomas Telford Service Ltd., London.

Sussmann, TR., Ebersöhn,W., and Selig, ET. 2001. Fundamental nonlinear track load-deflection behavior for condition evaluation. Transp. Res. Rec.,1742,61–67.
Thambiratnam, David, and Yan Zhuge. "Dynamic analysis of beams on an elastic foundation subjected to moving loads." Journal of sound and vibration 198.2 (1996): 149-169.

Tzanakakis K. 2013. The effect of track stiffness on track performance in the railway track and its long-term behaviour. Springer Tracts on Transportation and Traffic, 2013(2), Springer, Berlin, Heidelberg.

Yang, Y. B., et al. "Track modulus detection by vehicle scanning method." Acta Mechanica 231(2020):2955-2978.

Zarembski, A.M., Choros, J.: On the measurement and calculation of vertical track modulus. AREA Bull. 81, 156–173 (1979)

Evaluation of ride quality and rail surface roughness from vibration analysis of in-service passenger rail cars

Parisa Haji Abdulrazagh, and Michael T. Hendry
University of Alberta, Edmonton, Alberta, Canada

Alireza Roghani, and Elton Toma
National Research Council Canada Automotive and Surface Transportation, Ottawa, Ontario, Canada

ABSTRACT

Regulatory and special visual inspection of railway tracks is important to ensure the safe and secure operation of trains. Due to increased traffic volumes, operational speeds and axle loads, new technologies are emerging to collect and monitor data more frequently and in a shorter schedule-window. This paper presents a feasibility study to evaluate ride quality and rail surface roughness from car body and axle box accelerations mounted on service passenger rail cars. Accelerometers are mounted under car body of a rail car and on two axle boxes to evaluate ride quality and rail surface roughness, respectively. The repeatability of measurements, and impact of track features such as bridges, grade crossing and switches on the ride quality is studied by applying a weighted filtering method as per ISO 2631-1997 standard. The ability of axle-box acceleration signals to quantify the rail surface roughness are evaluated by comparing the calculated rail displacements with rail surface profile measurements recorded by a track geometry inspection car. Rail surface relative displacements are calculated from double integration of axle-box accelerations. The results of vibration analysis show that there is a meaningful correlation between increased magnitude of ride quality index and the location of some track features. The comparison between the calculated rail displacements and measured rail surface profile along the entire length of the studied track confirmed that the axle-box acceleration and the applied analytical technique are appropriate for evaluation of rail surface roughness.

1 INTRODUCTION

Railway track testing and evaluation is dating back to the early twentieth century which was conducted commonly with the purpose of studying the behaviour of railway structure and its components (Talbot, 1918; Talbot, 1920; Selig & Waters, 1994). These tests were usually stationary and performed at local points. With the evolutionary enhancement of computing powers and sensors, continuous and real-time measurement of track responses became possible through onboard instrumentation of rail cars since the early 21st century. Some examples of these onboard technologies are Portancemetre (Hosseingholian et al., 2009), rolling stiffness measurement (Berggren, 2009), falling weight deflectometer (FWD) (Sharpe, 2000; Burrow et al. 2007), and vertical track deflection (VTD) measurement (Norman et al., 2004) which have been developed to evaluate the track stiffness/modulus, and become as part of common evaluation techniques later (Haji Abdulrazagh and Hendry, 2016; Haji Abdulrazagh et al., 2019, Roghani and Hendry, 2016).

Studying the vibrations is extensively used over the past years as a non-destructive technique to evaluate the performance of mechanical and structural systems (Ketchum et al., 2012; Casas and Moughty, 2017). National Research Council Canada (NRC) has instrumented one of VIA Rail passenger rail cars with a set of accelerometers installed on electric room floor, car body, trucks and axle boxes. The Canadian Rail Research Laboratory (CaRRL) research team in collaboration with NRC conducted an extensive analysis with the aim of: (1) assessing the ride quality of the studied track and identifying the locations where ride quality could be improved such as bridge approaches, and crossings; and (2) determining the relationship between the axle-box accelerations and the rail surface roughness.

In this study using the weighted filtering method according to ISO 2631-1997 standard, the effect of different track features and change of track ride quality condition over five-month period is investigated. The relationship between the axle-box accelerations and track roughness is studied by performing a novel

computational metrology approach for evaluating the rail track surface.

$$a_w = \left[\sum_i (W_i a_i)^2 \right]^{\frac{1}{2}} \quad [1]$$

2 MATERIALS AND METHODS

2.1 Study Area

Acceleration measurements and track geometry data are collected for 80 km section of VIA Rail line in Ontario, Canada. The studied line consists of 52 kg/m (115-lb/yd) continuously welded rail on wooden ties supported by ballast. The traffic passing the studied subdivisions is approximately 3.5 million gross tonnes (MGT) annually (Haji Abdulrazagh et al., 2021a,b).

2.2 Collection of Acceleration and Track Geometry Data

The NRC Automotive and Surface Transportation (AST) Research Centre in collaboration with VIA Rail Canada instrumented a VIA Rail coach car with seven accelerometers on the underbody of the car (Figure 1(a)), and three on each of four axle boxes (Figure 1(b)). All accelerometers were high sensitivity DC MEMS-type instruments manufactured by PCB Piezotronics (Depew, NY). The underbody car accelerometers had a range of ± 10 g (model 3741B1210G) (at the sides) and ± 50 g (model 3741B1250G) (centerline). The axle-box vertical accelerometers (model 3741B12100G) had a range of ± 100 g (Toma et al., 2019; Haji Abdulrazagh et al., 2021b). The instrumented car was operated in a revenue service train with passengers on board, and data were collected between February and August of 2018 under actual operating conditions. The surface track geometry data were recorded by a track geometry inspection car on July 15, 2018.

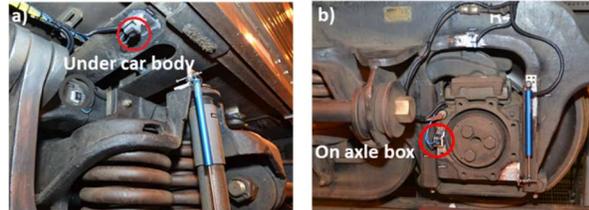


Figure 1. The location of accelerometers mounted on a) the underbody, and b) the axle box. (Images by authors.)

The coordinates of all measurements are recorded by the global positioning system (GPS). Car body and axle box accelerations were collected at the frequencies of 500 and 5000 Hz, respectively. The geometry data were recorded at 0.30-m (1-ft) intervals.

2.3 Calculating Track Ride Quality Index

In order to assess the track condition in terms of ride quality, the vibration contents of the collected car body accelerations related to ride quality were determined by frequency weightings as follows (ISO 2631-1:1997):

In Equation 1, a_w is the frequency-weighted acceleration, W_i is the weighting factor for the i th one-third octave band, a_i and is the root-mean-square (RMS) acceleration for the i th one-third octave band with measuring window length of 5 s as per ISO 2631-4: 2001. W_i is same for vertical and longitudinal directions and is different for the lateral direction.

The vibration total value of weighted RMS, a_v , or the ride quality index (RQI), is calculated from vibration in orthogonal coordinates as follows (ISO 2631-1:1997):

$$a_v = (k_x^2 a_{wx}^2 + k_y^2 a_{wy}^2 + k_z^2 a_{wz}^2)^{\frac{1}{2}} \quad [2]$$

where a_{wx} , a_{wy} , and a_{wz} are the weighted RMS accelerations for the x, y, and z axes; and k_x , k_y , and k_z , are the multiplying factors, which are 1 for the current analysis. In this paper, RQI is calculated from the three directions of the accelerations acquired from acceleration car body.

The ISO standard's classification for RQI is tabulated in Table 1. "A little uncomfortable" and "fairly uncomfortable" in this paper refer to RQI value between 0.315 and 0.5 m/s² and RQI value greater than 0.5 m/s², respectively. These two categories are highlighted with bold italic fonts in Table 1.

Table 1. Vibration levels and the likely reactions (per ISO 2631-1: 1997).

Levels of acceleration (m/s ²)	Likely Passenger Reactions
$a_v < 0.315$	Not Uncomfortable
<i>$0.315 \leq a_v < 0.63$</i>	<i>A little Uncomfortable</i>
<i>$0.5 \leq a_v < 1$</i>	<i>Fairly Uncomfortable</i>
$0.8 \leq a_v < 1.6$	Uncomfortable
$1.25 \leq a_v < 2.5$	Very Uncomfortable
$a_v \geq 2$	Extremely uncomfortable

2.4 Calculating Rail Surface Roughness

A Gaussian filter - a computational metrology method- is applied on the measured geometry datasets to the authors knowledge for the first time to filter the short-wave rail defects such as corrugations from track geometry surface as per International Organization for Standardization (ISO 11562:1996). In this method to quantify the rail surface roughness, the moving RMS, or R_g , of the difference between the elevation of every pair of successively measured points, h_i and h_{i-1} , by a track geometry inspection car is divided by the number of measurements in the chosen length of track, L as follows (Vorburger and Raja, 1990):

$$R_g^{Moving\ RMS}(x) = \sqrt{\frac{1}{L} \int_{x-L}^x (h_i - h_{i-1})^2(x) dx} \quad [3]$$

In this research, the rail surface roughness is evaluated from the vertical axle-box accelerations. The relative vertical displacements are calculated from the double integration of the filtered axle-box accelerations using cumulative trapezoidal numerical integration (Chapra, 2017). The details of noise filtering and removing the unwanted trends as a result of integration are fully described by Haji Abdulrazagh et al. (2021b). The moving RMS is determined using Equation 3, where the calculated relative displacement from double integration, Dh_i , substitutes in $h_i - h_{i-1}$. The rail surface roughness calculated from axle-box accelerations are then compared to those from track geometry measurements.

3 RESULTS AND DISCUSSION

Figure 2(a) shows the original vertical, lateral and longitudinal acceleration data recorded as well as the train speed during the data collection on July 18, 2018. The data show the train predominantly operated at speeds greater than 100 km/h. The stops at two intermediate stations are identified from very slow speed and acceleration values in Figure 2(a) and (b). Figure 2(b) presents the raw vertical axle-box accelerations recorded from left and right sides of the train on July 18, 2018. Figure 2(c) presents the track surface and alignment and corresponding curvature of the track recorded on July 15, 2018 by the track geometry car. The alignment of acceleration and geometry data was carried out by correcting GPS coordinates, accounting for their phase shift and distance lags using a post-processing phase differential method (DiBiase et al., 2013).

3.1 Repeatability and Effect of Track Features on RQI

The RQI is calculated as per ISO 2631-1:1997 and Equation 2. The Calculated RQI values and the train operating speed at two examples of different dates, July 18th and August 1st, 2018 are plotted in Figure 3. The RQI graphs show very similar trends between runs which is as a result of repeatable measurements of car body accelerations in three directions.

Figure 3 also shows that the peak RQI values correspond to major notable features, such as grade crossings, bridges, switches, and culverts along the studied track.

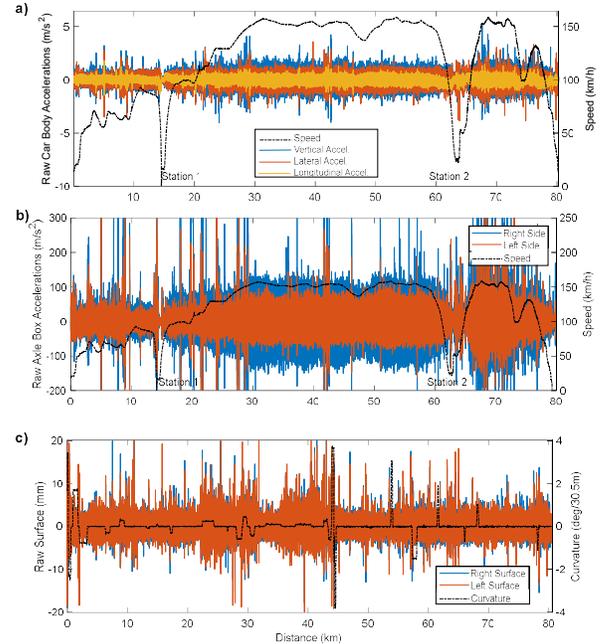


Figure 2. a) Example of raw car body accelerations with corresponding train speed (Haji Abdulrazagh et al., 2021a) b) example of raw vertical axle-box accelerations with corresponding train speed (Haji Abdulrazagh et al., 2021b) and c) the corresponding track surface and degree of curvature recorded along the track (Haji Abdulrazagh et al., 2021a).

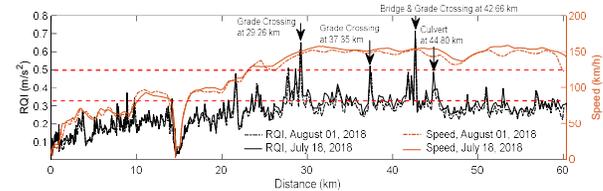


Figure 3. Variation of RQI values from car body accelerations and the train speed versus distance at two examples of different dates at July and August 2018 (Haji Abdulrazagh et al., 2021a), and the corresponding major notable features that resulted in fairly uncomfortable vibrations.

3.2 Variation of RQI Values over Five Months of Data Collection

Figure 4 shows the RQI values being plotted from March to August 2018. No data was recorded on June. The average speed was very similar between the runs with a deviation of only ± 3 km/h. Several track features had RQI values below 0.50 m/s² in March 2018, but from April onwards these values exceed 0.50 m/s²; similar results were noted at locations with RQI values below 0.315 m/s² (not uncomfortable) that then increased to 0.315 - 0.50 m/s² (a little uncomfortable).

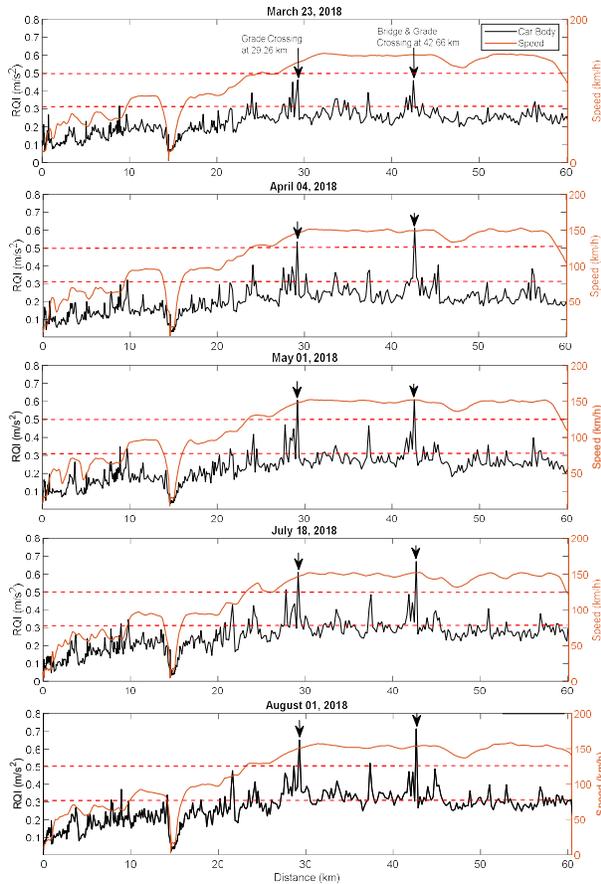


Figure 4. Increase of RQI value from March to August 2018.

The assessment of variation of ride quality values in Figure 4 indicated some locations had recurring spikes of RQI values in “fairly uncomfortable” category ($0.5\text{--}0.8\text{ m/s}^2$). For example, for the grade crossings located at 29.25 and 42.66 km from the origin station, the RQI values on March 23, 2018 was 0.461 and 0.458 m/s^2 , respectively, but for April 17, 2018 and later dates exceeded 0.5 m/s^2 . This recurring observation shows the general degradation of ride quality from March to August 2018 as a result of passing traffic which is 3.5 MGT annually.

3.3 Calculating Relative Rail Surface Displacements from Vertical Accelerations

To calculate the relative rail surface displacements from vertical accelerations, the original acceleration signals are first corrected for noise using a Butterworth filter. The relative vertical displacements are then computed from corrected axle-box accelerations following the procedure described in §2.4. Figure 5(a) shows part of the raw time history of acceleration where the speed of train was relatively constant and the corresponding computed velocity and displacements. Figure 5(b) illustrates the computed velocity after single integration and eliminating non-intrinsic patterns using low-order

polynomial fit. Figure 5(c) plots the displacements calculated after the double integration and baseline correction. The computed relative vertical displacements are the basis for further calculation of the moving RMS and their comparison with the moving RMS rail surface from the track geometry measurements.

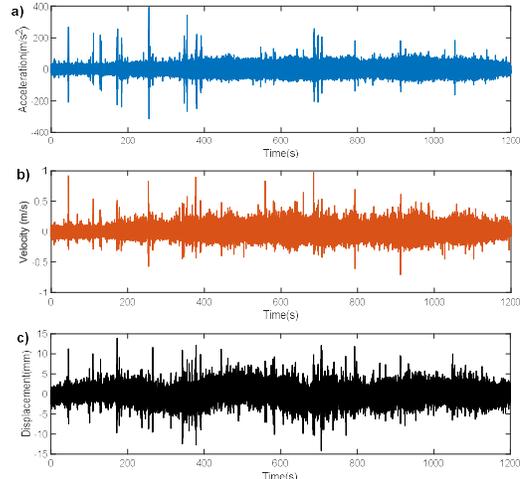


Figure 5. a) raw axle-box vertical acceleration, b) computed velocity, and c) computed vertical displacement between distance 31.8 and 44.0 km (Haji Abdulrazagh et al., 2021b).

3.4 Relationship between the Moving RMS of the Vertical Displacement and Moving RMS of Rail Surface

The moving RMS is calculated for the vertical displacement and rail surface using Equation 3. Hereafter, they are called *calculated rail surface RMS*, and the *measured rail surface RMS*, respectively (Haji Abdulrazagh et al., 2021b).

Figure 6 shows the comparison between the calculated and measured surface RMS values. The major peaks and variations of these two sets of values are very similar.

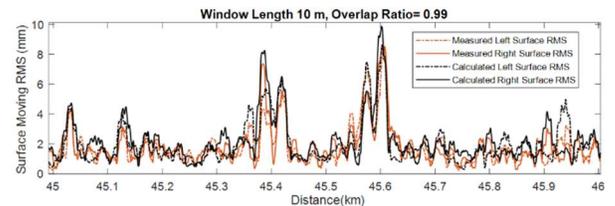


Figure 6. Comparison between the calculated and measured surface moving RMS with a 10-m window length and overlap ratio 0.99.

The moving RMS is calculated with a moving window length 10 m and the overlap ratio 0.99. Different length of moving window and the overlap ratio has been tested to achieve the best relationship and to account for surface metrological considerations and a sufficiently large statistical sample that is representative of a studied behaviour or characteristic. More details of the study can

be found in the research by Haji Abdulrazagh et al. (2021b).

The power spectral density (PSD) is computed for the calculated and measured rail surface moving RMS (with a 10-m window length and overlap ratio 0.99), and is compared in terms of the spatial frequency in Figure 7. PSD allows for comparing the spatial frequency and therefore, the wavelength of the calculated and measured rail surface moving RMS regardless of the speed of measurements by the train and the track geometry car.

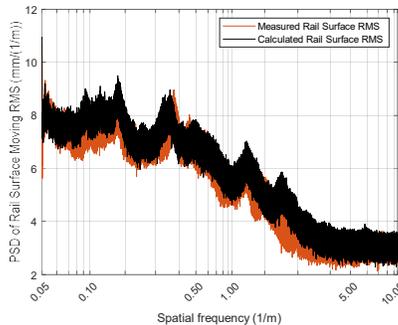


Figure 7. Comparison of the PSD of calculated and measured rail surface moving RMS.

The plot of PSD for both calculated and measured rail surface moving RMS in Figure 7 show that there is a significant peak for at about 0.17 and 0.40 (1/m) of the spatial frequency which are as a result of rail surface deviations with wavelength about 6 and 2.5 m, respectively. The PSD plots present that the Butterworth and Gaussian filters applied on the calculated relative displacements and the measure rail surface are promising filtering techniques for determining the relationship between the axle-box accelerations and the rail surface roughness (calculated as moving RMS in this paper).

4 CONCLUSIONS

Measurements from accelerometers mounted on an in-revenue service passenger car operating on one of VIA rail line in Ontario, Canada were used to quantify the variability of the ride quality of the track over time. RQI values are calculated as per ISO 2631-1: 1997 standard from the measurement of car body accelerations in three directions. The RQI values plotted along the rail route showed recurring peaks at the track features such as bridges, switches and grade crossings which increased from “a little uncomfortable” range (0.315-0.5 m/s^2) to “fairly uncomfortable” range (0.5-0.8 m/s^2) from March to August 2018. This recurring observation might be as a result of general degradation of ride quality under passing traffic. It is suggested that the track features were increased successive RQI values are recorded, potentially need to be improved for the 140 km/h operating speed to enhance the ride quality level. The seasonal effect on the ride quality was not studied in this paper. If a two-year or longer period of acceleration data are recorded, the effect of seasonal cycles such as

frozen subgrade during winter and spring/summer subgrade defrost on ride quality worth further investigation.

The vertical axle-box accelerations which were acquired by the instrumentation in the same passenger rail car were examined to study whether these measurements can be used to evaluate rail surface roughness. A new methodology was developed from double integration of vertical accelerations and applying Butterworth filter to quantify medium-wavelength rail surface roughness by determining the moving RMS of relative vertical displacements. A Gaussian filter was also applied on the surface geometry measurements and the moving RMS was calculated.

Comparison of the moving RMS of calculated vertical displacements and of measured rail surface along the length of the studied track showed they represent identical magnitudes and variations. The PSD was calculated for the moving RMS of the vertical displacements and the measured rail surface to verify whether the new developed methodology is promising in presenting the wavelength of the rail surface variations measured by the track geometry evaluation car. The two major spikes of the PSD values at about 0.17 and 0.40 (1/m) of the spatial frequency shows that the wavelength of the calculated rail surface moving RMS (approximately 6 and 2.5 m) well matches with those measured.

The methodology developed in this research can be applied only when the accelerometer sensitivity complies with the actual range of rail surface wavelengths being recorded by the track geometry evaluation car. This methodology can be improved by testing it for various operating speeds and track geometry measurements.

5 ACKNOWLEDGEMENT

The authors would like to express appreciation for the support of National Research Council Canada (NRC). Special thanks are also extended to NRC's Automotive and Surface Transportation research center testing and evaluation team for installing and maintaining the instrumentation. The authors also gratefully acknowledge the collaboration of VIA Rail Canada Inc., particularly assistance provided by Paul Charbachi. This research was made possible through financial support from NRC and the Canadian Railway Research Laboratory (CaRRL) research program at the University of Alberta [Grant number IRC 523369-18].

6 REFERENCES

- Berggren E. (2009). Railway track stiffness: dynamic measurements and evaluation for efficient maintenance. PhD Thesis, Royal Institute of Technology (KTH), Sweden.
- Burrow, M.P.N., Chan, A.H.C., and Shein, A. (2007). Deflectometer-based analysis of ballasted railway tracks. Proceedings of the Institution of Civil Engineers,

- Geotechnical Engineering, 160(3): 169–177. doi:10.1680/geng.2007.160.3.169.
- Casas, J.R. and Moughty, J.J. (2017). Bridge Damage Detection Based on Vibration Data: Past and New Developments. *Front. Built Environ.* 3:4. doi: 10.3389/fbuil.2017.00004.
- Chapra, S. C. (2017). *Applied Numerical Methods with MATLAB for Engineers and Scientists*, 4th Edition. New York, NY: McGraw-Hill Higher Education.
- DiBiase, D., L. Sloan II, R. Baxter, W. Stroh, B. F. King. (2013). The nature of geographic information. *An Open Geospatial Textbook*. State College, PA: Pennsylvania State University.
- ISO (International Organization for Standardization). (1996). Geometrical product specification (GPS) – Surface texture: Profile method – Metrological characteristics of phase correct filters. ISO 11562:1996. Geneva, Switzerland: ISO.
- ISO 2631-1:1997. (1997). Mechanical vibration and shock—Evaluation of human exposure to whole-body vibration. Part 1: General requirements.
- ISO 2631-4: 2001. (2001). Mechanical vibration and shock: Evaluation of human exposure to whole-body vibration. Part 4: Guidelines for the evaluation of the effects of vibration and rotational motion on passenger and crew comfort in fixed guideway transport systems.
- Haji Abdulrazagh, P., and Hendry. M. T. (2016). A case study of the use of a falling weight deflectometer to investigate railway infrastructure constructed upon soft subgrades, *Canadian Geotechnical Journal*, Vol. 53, No.12, pp.1991-2000. <https://doi.org/10.1139/cgj-2016-0083>.
- Haji Abdulrazagh, P., Farzaneh O., and Behnia C. (2018). Evaluation of railway trackbed moduli using the rail falling weight test method and its backcalculation model *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 233, No.4, pp. 431-447. <https://doi.org/10.1177/0954409718799800>.
- Haji Abdulrazagh, P., Hendry, M.T., Gül, M., Roghani, A., Toma, E. (2021a) Use of measured accelerations from a passenger rail car to evaluate ride quality and track roughness – A case study. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*. <https://doi:10.1177/09544097211041459>.
- Haji Abdulrazagh, P., Hendry, M.T., Roghani, A., Toma, E. (2021b) Evaluating rail surface roughness from axle-box acceleration measurements– A computational metrology approach. *ASCE Journal of Transportation Engineering*, Vol. 147, No. 12, <https://ascelibrary.org/doi/abs/10.1061/JTEPBS.0000601>.
- Hosseingholian M., Froumentin M. and Levacher D. (2009). Continuous method to measure track stiffness: a new toll for inspection of rail infrastructure. *World Applied Sciences Journal*, Vol. 6: 579–589.
- Ketchum C.D. and Wilson. N. (2012). Performance-based track geometry: optimizing transit system maintenance. 2012 Rail Conference, Dallas TX: American Public Transportation Association.
- Norman, C., Farritor, S., Arnold, R., Elias, S.E.G., Fateh, M., Sibaie, M. E. (2004). Design of a System to Measure Track Modulus from a Moving Railcar, *Proceedings from Railway Engineering*, London 2004.
- Roghani, A., and Hendry. M. T. (2016). The use of continuous vertical rail deflection measurements to map subgrade condition along a railway line: Methodology and case studies. *ASCE Journal of Transportation Engineering*, Vol.142 No.12. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000892](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000892).
- Selig E.T. and Waters, J.M. (1994). *Track Geotechnology and Substructure Management*, Thomas Telford ISBN: 0-07277-2013-9.
- Sharpe, P., and Collop, A.C. (1998). Trackbed Investigation: a modern approach. In *Proceedings of the 1st International Conference on Maintenance and Renewal of Permanent Way and Structures*, London. Edited by M.C. Ford. Engineering Technics Press, Edinburgh, Scotland. pp. 37–43.
- Talbot, A.N. (1918). Stresses in railroad track. Reports of the Special Committee on Stresses in Railroad Track, *Proceedings of the AREA, First Progress Report*, Vol. 19, 1918, pp.873-1602.
- Talbot, A.N. (1920). Stresses in railroad track. Reports of the Special Committee on Stresses in Railroad Track, *Proceedings of the AREA, Second Progress Report*, Vol. 21, 1920, pp.645-814.
- Toma, E., Y. Liu, and A. Jahagirdar. (2019). Track monitoring using ride quality measurements. In *Proceedings of the 2019 World Congress of Rail Research 2019*, Tokyo, Japan.
- Vorburger, T. V., and J. Raja. (1990). Surface finish metrology tutorial. NISTIR 89-4088. US Department of Commerce, National Institute of Standards and Technology.

Laboratory investigation of mechanical properties of used rubber tires as Under Sleeper Pad in railway tracks

Morteza Esmaeili, Armin Shamohammadi

School of Railway Engineering - Iran University of Science and Technology, Tehran, Iran

Farshad Astaraki, Majid Movahedi Rad and Jafar Chalabi

Department of Structural and Geotechnical Engineering - Széchenyi István University, Győr, Hungary

Abstract

In railway tracks, using elastic elements is one of the most efficient methods for establishing and changing vertical stiffness, reducing noise and vibration, and consequently decreasing maintenance costs. In this regard, using under-sleeper pads made of Polyurethane, rubber, or Ethylene-vinyl is a practical element. These elements can be used under sleepers both in ballasted and ballastless tracks. Since waste rubber material has not yet been widely used in railway tracks and the availability of a considerable amount of them, examining them as under-sleeper pads is beneficial. Therefore, the current study investigates the applicability of used rubber tires material with different thicknesses of 7, 9, 11 and 13 mm as under-sleeper pads in tramway, metro, high speed, and main ballasted railway tracks. For this purpose, the static and dynamic stiffness's of samples were assessed according to DIN 45673-6. Furthermore, the best performance of the produced pad appeared in a frequency range between 5 to 30 Hz, while its static and dynamic bedding modulus, C_{stat} and C_{dyn} ranges were between 0.042 to 0.14 and 0.1 to 0.18 N/mm^2 respectively.

Keywords: Under-sleeper pad, used rubber, mechanical characters, dynamic bedding modulus, static bedding modulus

1 INTRODUCTION

Although industries are tied to railway networks to carry their heavy resources and products such as oil, crops, minerals, nearby residents, have faced problems because of its induced noise and vibration. Therefore, simultaneously with developing new heavy and fast railway tracks worldwide, the superstructure components must be specially designed to tolerate higher stresses. However, sharper curves under heavy traffic are prone to tense wear and propagate higher noise and vibration levels. Adopting proper solutions can decrease Life Cycle Costs (LCC) and maintenance periods and increase the lifespan of the railway tracks. Until now, different methods have been considered to decrease the level of vibration induced by railway traffic. In general, these techniques can be classified into three categories, reduction of vibration in the source, receiver, and path. However, using specific rubber pads in railway tracks as a solution for vibration problems has received special attention from researchers and industries. This pad's predominant and applicable type involves rubber pads under rails, fastening system plates, and sleepers. Under sleeper rubber pad (USP) has been used in railway track construction, especially in high-speed rail lines, for more than 20 years. This pad can be implemented both in ballasted and slab tracks which in ballasted tracks it places between sleepers and ballast layer.

Zakeri et al. (2017) assessed the effect of USP on reducing vibration by installing several geophones around the sleepers in a test site. The results indicated that the use of rubber pads decreased the amplitude

of vibration. However, the best performance regarding the vibration reduction was reported by using the pads at a frequency magnitude of 50 Hz. USPs have been used in member countries of the international unions of railways (UIC (2009)). The effect of USP with different stiffness on reduction of vibration level investigated by UIC. The result showed that using USPs at frequencies higher than 40Hz decreased vibration between 8 to 15 dB (UIC (2009)). Some other outcomes, as a result of installing USPs, including reduction of imposed stress in the components of the tracks, a decrease of ballast layer settlement, a decrease of corrugation with the long-wavelength in tight curves were also reported by UIC in Germany, Austria, Czech Republic, Spain and France. Gräbe et al. (2016) assessed the effect of USPs on steel and wooden sleeper-ballast interaction using dynamic and static tests carried out in specific ballast boxes. By implementation of USPs, the contact surface between sleepers and pads up to 13% increased, the contact pressure between the sleeper and ballast decreased up to 70%, the ballast settlement up declined to 44%, and the ballast breakage and deterioration went down up to 29% according to the main achievements. Kraśkiewicz et al. (2018) examined the cohesion strength of USPs installed in ballasted tracks using tensile tests and assessing static and dynamic bedding modulus. The defects were also recorded, and then the static and dynamic modulus of USPs were calculated.

A tremendous amount of end-of-life tires (ELT) are daily producing all over the world. The current method, such as burning and storage of this type of wastes, has an improper effect on the environment and human

health. ELT disposal is a chronic issue for many countries, especially developing countries like China and India. Based on the literature and released information from an international tire producer and Thomas & Gupta (2015), 1.5 billion tires are producing every year globally, which will reach 5 billion by 2030 (Thomas et al (2016), Thomas & Gupta (2015)). Currently, The United States, Europe, Japan, and India produce nearly 80% of tires globally. Based on a statistical investigation in 2015 carried out by Zarei et al. (2018), in Iran, 324 thousand tons of ELT, about 4 Kg per person, were yearly produced. An increasing trend was anticipated for the future.

In recent decades, many studies have been performed to assess ELT use as USPs (Deconstructed Tire Under sleeper pads (DT-USP)) or other elastic elements in railway tracks. The results indicated progress in both recycling of ELT and decrease of maintenance costs of railway tracks. Sánchez et al. (2014) examined the use of ELT as USP (DT-USP) which the results were satisfactory. Based on DIN 45673-6 (2010), they analyzed the effect of thickness and materials of the pads on their dynamic and static responses. The thickness of 5 to 11 mm showed the best performance to meet optimum stiffness and omit the transfer forces which eventually, resulted in a decrease in maintenance costs of railway tracks. However, they did not perform the dynamic tests for primary and high-speed railway tracks. The possibility of the use of DT-USP was assessed based on the DIN standard (2010). To do this, dynamic and static Indentation hardness tests on the pads made of ELT with thicknesses of 7, 9, 11 and 13 mm were carried out in the lab environment.

2 LABORATORY INVESTIGATION OF DT-USP BEDDING MODULUS

2.1 Measuring of static bedding module, C_{stat} ,

In this section, a series of tests were carried out on concrete samples under the sleeper pad according to DIN 45673-6 (2010) which simulates the ballast and the underneath layers rigidity and stiffness, to calculate the imposed pressure to rail or track under service loads. C_{stat} can be calculated during product quality control, and it can be used to compare different DT-USPs together. C_{stat} as a reference parameter can describe the quality of the products. This parameter can determine the magnitude of the applied load for mechanical fatigue strength tests and also calculating the dynamic stiffness ratio.

In order to measure C_{stat} , there needs to pay special attention to the loading and unloading rate and ensure that load is applied continuously. In other words, C_{stat} is a quasi-static parameter. However, the magnitude of C_{stat} and its behaviour should be precisely determined to ensure that the comparison of the pads and tests' processes is feasible.

The examined DT-USPs have the dimensions of 300mm × 300mm × t, where t is the pads' thickness, and in this case of study, they are 7, 9, 11 and 13 mm. In the tests, a cubic concrete sample with 300mm ×

300mm × 200mm was used beneath the pads. The concrete block was reinforced to tolerate the applied loads. A layer of glue also was used between the pad and the concrete block. It should be noted that the contact surface was also smooth and capped (surface levels difference did not exceed 0.2 mm). Figure 1 shows a set-up under loading jack. In this study, three concrete samples were built at the temperature of 23±3.



Figure 1. Static bedding module test, a sample of DT-USPs over concrete block under loading jack.

The load was applied through a steel-plated over the pad according to the given magnitude in table 1.

Table 1. Imposed loads to calculate C_{stat} .

The type of track	Evaluated ranges (N/mm^2)	Loading ranges (N/mm^2)
Tramway	$\sigma_1 = 0.01$ to $\sigma_2 = 0.05$	$\sigma_u = 0.01$ to $\sigma_o = 0.06$
Underground tracks	$\sigma_1 = 0.01$ to $\sigma_2 = 0.07$	$\sigma_u = 0.01$ to $\sigma_o = 0.08$
High-speed tracks	$\sigma_1 = 0.01$ to $\sigma_2 = 0.1$	$\sigma_u = 0.01$ to $\sigma_o = 0.11$
Main railway tracks	For ballast with medium compaction: $\sigma_1 = 0.01$ to $\sigma_1 = 0.1$	$\sigma_u = 0.01$ to $\sigma_o = 0.25$
	For ballast with high compaction: $\sigma_1 = 0.01$ to $\sigma_1 = 0.2$	

The load was non-stopped applied with a rate of loading and unloading of 0.01 (N/mm^2)/s for three loading cycles. The force magnitude, F, was calculated according to $F = \sigma \cdot A$ equation.

In order to do the test, the loading plate should place on the center of the samples, and the load was applied through three non-stopped cycles. Each loading cycle start from σ_u and continue to reach σ_o . Furthermore, unloading was wholly performed. In order to calculate C_{stat} , stresses (σ) and settlements (s) were recorded at the third loading cycle. Then based on the values of s_1 and s_2 , C_{stat} can be calculated for stresses rang of σ_1 to σ_2 as follow:

$$C_{stat} (N/mm^3) = \frac{\sigma_2 - \sigma_1}{s_2 - s_1} \quad (1)$$

where $s_2 - s_1$ is the difference of measured displacement at the third loading cycle in the pressure range of $\sigma_2 - \sigma_1$. The average value of three tests is presented as the final result of the test. The C_{stat} should also be calculated for the ballast layer with high compaction for the main railway tracks.

2.2 Measurement of dynamic bedding modulus at low frequencies, $C_{dyn1(f)}$

In this section, a series of tests were carried out on the DT-USPs based on DIN 45673-6 (2010) to calculate $C_{dyn1(f)}$, which can estimate rail deformation under low-frequency train loads. These deformations are mainly because of the elastic bending behaviour of rails and rubber plates under sleepers. Measuring $C_{dyn1(f)}$ at low frequencies can be used in the dynamic and vibrational calculation of the tracks.

The dimensions, sample number, and loading conditions were the same as those used to measure C_{stat} in the previous section. The imposed loads to calculate C_{dyn} are given in Table 2.

Table 1. Imposed loads to calculate C_{dyn} .

The type of track	Loading ranges (N/mm^2)
Tram	$\sigma_1 = 0.01$ to $\sigma_2 = 0.05$
Underground tracks	$\sigma_1 = 0.01$ to $\sigma_2 = 0.07$
High-speed tracks	$\sigma_1 = 0.01$ to $\sigma_2 = 0.1$
Main railway tracks	$\sigma_1 = 0.01$ to $\sigma_2 = 0.1$

A sample of dynamic loading over a concrete block with DT-USPs is depicted in Figure 2.

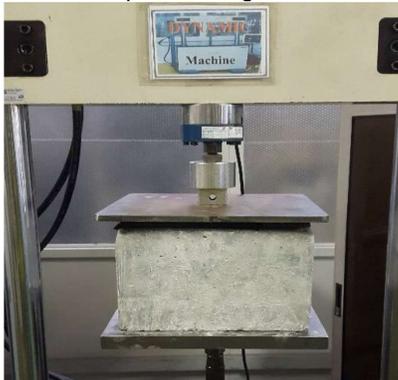


Figure 2. A sample of dynamic bedding modulus test on a specimen of DT-USPs over a concrete block.

For doing this test, the loading plate was placed in the center of the pad. For measuring C_{dyn1} at a specific frequency, the pressure decreased equal to σ_1 and then it remained constant for each tests' frequency for a 3 minute rest time. The data was recorded after 10 seconds of applying dynamic load. Ten complete cycles were performed for each frequency, and the average was calculated (see Figure 3). Finally, for

each frequency, C_{dyn1} was calculated according to the following equation:

$$C_{dyn1} (f) = \frac{\sigma_2 - \sigma_1}{s_2 - s_1} \quad (2)$$

where $C_{dyn1(f)}$ is dynamic bedding modulus (N/mm^3) and $s_2 - s_1$ is the difference of measured displacement (mm) at the third loading cycle in the pressure range of σ_2 and σ_1 (N/mm^2).

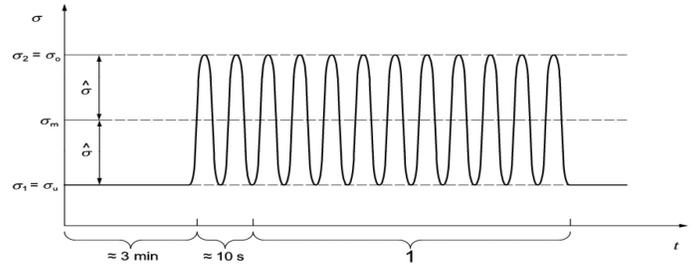


Figure 3. Loading pattern for dynamic bedding modulus tests

σ_m : Average of applied load

δ : Amplitude of applied load

t : time (dimensionless)

1: An evaluation range including ten period

3 TESTS RESULTS AND DISCUSSION

Static bedding modulus (N/mm^3) classifies the mechanical performance of DT-USPs. In this study, the parameter of C_{stat} was calculated for two stress ranges of 0.01-0.1 N/mm^2 (C_{stat1}) and 0.01-0.2 N/mm^2 (C_{stat2}) which was applied to ballast layers. Figure 4 shows the magnitude of C_{stat1} and C_{stat2} obtained for DT-USPs with 7 to 13 mm thicknesses. It should be mentioned that according to the DIN 45673-6 (2010), the DT-USPs were tested three times (a, b and c) to achieve the mean results.

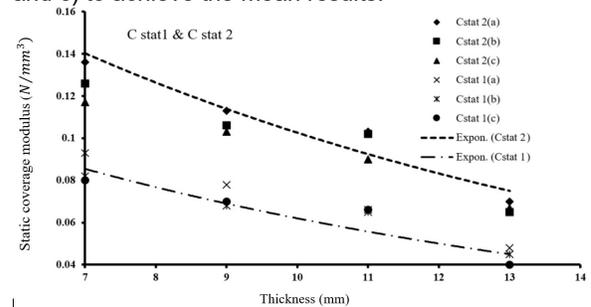


Figure 4. Static bedding modulus versus thickness of DT-USPs.

The results showed that the thinner these pads are, the higher the elasticity modulus, making them harder over time. In addition, by increasing the stress level from 0.1 to 0.2 N/mm^2 the higher values of modulus obtained for the pads, which means that the flexibility of the pads decreases for a more extensive loading range. In continuation, according to data analysis, it was shown that for both stress ranges, the static

modulus follow a specific trend which is an exponential function with R^2 values of 0.8093 and 0.9253 for C_{stat1} and C_{stat2} , respectively. According to obtained C_{stat1} values for average ballast pressure of 0.1 N/mm^2 , all examined DT-USPs are classified as soft pads suitable for decreasing vibration and stresses induced by train. For ballast pressure higher than 0.2 N/mm^2 , the DT-USPs with a thickness of 7 mm are classified as soft pads. However, thicker pads can be used in transition zones to regulate the track stiffness as these DT-USPs decreases the ballast layer depth and longwave corrugation.

Dynamic bedding modulus was examined for predicting the mechanical performance of DT-USPs under different train traffic velocities. The obtained dynamic modulus for frequencies of 5, 10, 20 and 30 Hz are presented in Figure 5.

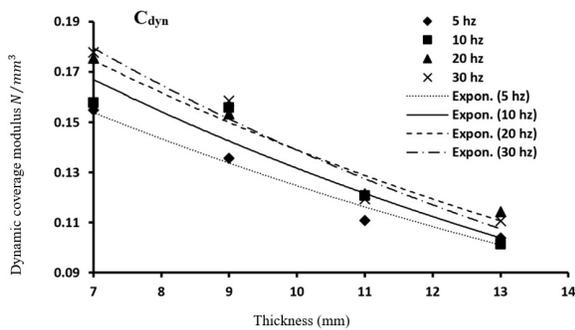


Figure 5. Dynamic bedding modulus versus thickness of DT-USPs.

Furthermore, Figure 6 shows the dynamic to the static ratio for evaluating the hardening of the pads for different frequencies. This parameter was defined and calculated as the ratio of dynamic modulus, $C_{dyn1(f)}$, to static modulus, C_{stat} , for the identical stress range.

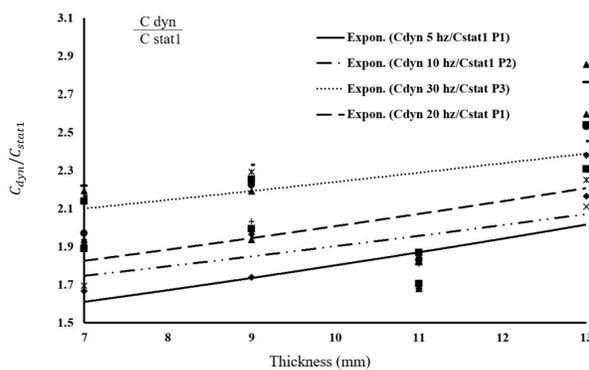


Figure 6. C_{dyn}/C_{stat1} versus thickness (mm).

According to the graph, increasing the DT-USPs' thickness decreased the dynamic modulus for both frequencies, indicating that the thicker pads experienced more vertical deformations than the thinner ones. Besides, it was shown that higher dynamic bedding modulus was obtained as the

frequencies increased, which was in good agreement with the previous studies. Moreover, the results of dynamic modulus were reported as exponential functions with R^2 values of 0.6125, 0.5817, 0.4885 and 0.1029 for frequencies of 5, 10, 20 and 30 Hz, respectively, which showed good agreement with the findings of Sanchez et al. (2014). The results indicated that the stiffness of DT-USPs affects the dynamic performance and the thicker pads are proper for decreasing the stiffness of railway tracks.

4. CONCLUSIONS

This study aims to assess the possibility of using ELT as under sleeper pads (DT-USP) through a series of lab tests. For this purpose, firstly, to determine static behaviour of tracks, DT-USP were prepared in different thicknesses of 7, 9, 11 and 13 mm and then static bedding modulus tests were carried out at different stress levels of $0.01-0.1 \text{ N/mm}^2$ (C_{stat1}) and $0.01-0.2$ (C_{stat2}). To evaluate the dynamic response of the track, loading frequencies of 5, 10, 20 and 30 Hz were also examined for calculating the dynamic bedding modulus, $C_{dyn1(f)}$, at low frequencies. It should be noted that the tests were carried out according to the DIN 45673-6 standard circumstances. The main findings can be summarized as follow:

- 1- Thicker pads are more suitable for railway tracks with high rigidity as they assist in obtaining the optimum stiffness, which eventually decreases maintenance costs.
- 2- The thickness of the pads has a significant impact on their static and dynamic performance, as seen in evaluating stiffness modulus. Thicker pads show higher flexibility than thinner ones which means thin pads have higher elasticity modulus, and their rigidity will increase over time.
- 3- Based on DT-USPs static tests, for ballast pressure of 0.01 N/mm^2 , the pads with the thickness of 7 to 13 mm as soft pads are suitable to decrease induced stresses and vibrations. For ballast pressure of 0.02 N/mm^2 the pads with 7 mm thickness can be used for transition zones and as a solution to decrease corrugation.
- 4- Based on the obtained results from high frequencies loading cycles of 20 and 30, these pads are also suitable for main railway tracks.

5. REFERENCES

DIN 45673 (2010) - Mechanical vibration, Part 6, "Laboratory test procedures for under-sleeper pads of concrete. DIN Deutsches Institut, Germany.

Gräbe, P. J., Mtshotana, B. F., Sebati, M. M., & Thünemann, E. Q. (2016). The effects of under-sleeper pads on sleeper-ballast interaction. *Journal of the South African Institution of Civil Engineering*, 58(2), 35-41.

International Union of Railways (2009). UIC Project, Under Sleeper Pads, 4th Edition. Vienna.

Kraśkiewicz, C., Oleksiewicz, W., Płudowska-Zagrajek, M., & Piotrowski, A. (2018). Testing procedures of the Under Sleeper Pads applied in the ballasted rail track systems. In *MATEC Web of Conferences* (Vol. 196, p. 02046). EDP Sciences.

Sol-Sánchez, M., Moreno-Navarro, F., & Rubio-Gámez, M. C. (2014). Viability of using end-of-life tire pads as under sleeper pads in railway. *Construction and Building Materials*, *64*, 150-156.

Thomas, B. S., & Gupta, R. C. (2015). Long term behaviour of cement concrete containing discarded tire rubber. *Journal of Cleaner Production*, *102*, 78-87.

Thomas, B. S., & Gupta, R. C. (2016). Properties of high strength concrete containing scrap tire rubber. *Journal of Cleaner Production*, *113*, 86-92.

Thomas, B. S., Gupta, R. C., & Panicker, V. J. (2016). Recycling of waste tire rubber as aggregate in concrete: durability-related performance. *Journal of Cleaner Production*, *112*, 504-513.

Zakeri, A.J., Esmaeili, M., & Heydari-Noghabi, H. (2016). A field investigation into the effect of under sleeper pads on the reduction of railway-induced ground-borne vibrations. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *230*(3), 999-1005.

Zarei, M., Taghipour, H., & Hassanzadeh, Y. (2018). Survey of quantity and management condition of end-of-life tires in Iran: a case study in Tabriz. *Journal of Material Cycles and Waste Management*, *20*(2), 1099-1105.

Embankment reconstruction using a wicking geotextile and its effect on the drainage and strength of a fine-grained railway subgrade

Camila Alvarenga
Menard Canada Inc., Vancouver, BC, Canada

Parisa H. Abdulzaragh and Michael T. Hendry
Department of Civil and Environmental Engineering – University of Alberta, Edmonton, AB, Canada

ABSTRACT

As a part of the Canadian Pacific Railway's (CP) Grade Stabilization/Remediation Plan, a 45 m section of a railway embankment located at the Scotford Division in Alberta, Canada, was remediated. The remediation consisted of the replacement of the fouled ballast and sub-ballast soils with clean materials and the installation of both a reinforcing geotextile (Mirafi® RS580i) and a wicking geotextile (Mirafi® H2Ri). The remediation aims to address issues such as cess heave (progressive shear failure) and settlement, which are caused by poor drainage and moisture retention. This site provided the opportunity of observing the enhancement of drainage brought by the performed reconstruction, as it enables testing adjacent remediated and control sections for a direct comparison. Moisture sensors capable of determining the volumetric water content (VWC), bulk electrical conductivity, and temperature from the soil around were installed at both sections aiming to compare their moisture levels directly. A simple method called the antecedent precipitation index (API) was used to convey the precipitation events' influence in the variation of moisture content and strength of the embankment at both sections. Several tests were performed in the site materials, including Soil-Water Characteristic Curves (SWCC) and Direct Shear (DS) tests. These tests allowed the determination of the impact of the reconstruction on the strength levels of the subgrade material according to the field moisture content. In the first year of monitoring, it was observed that the remediated section VWC levels were less sensitive to precipitation events than the ones from the control section. Also, the remediated section presented a significantly higher minimum strength level than the control section during high precipitation events and for most of the study. In conclusion, the remediation potentially improves the embankment drainage and consequently its bearing capacity during heavy rainfall events. Overall, the improvement brought by the reconstruction may be beneficial for mitigating the issues previously seen at the site.

1 INTRODUCTION

Several methods are applied to improve track drainage, such as ditches, trench drains/coarse-aggregate drains, geosynthetics, etc. (Li et al. 2002; Selig and Waters, 1994). In railway track applications, geotextiles are used for separation and retention between layers, but they can also improve track drainage. However, geotextiles usually provide drainage benefits only in saturated conditions. The development of a geotextile with wicking properties made it possible to remove water by capillary action from both the embankment and the subgrade's surface towards the shoulders even in unsaturated conditions.

As part of the Canadian Pacific Railway's (CP) Grade Remediation Plan, the embankment materials of a section at the Scotford subdivision, which presented a small amount of fines, were replaced by clean materials while a reinforcing (Mirafi® RS580i) and a wicking geotextile (Mirafi® H2Ri) were installed. This site was kindly provided by CP based on its history of issues, such as settlement, frost heave, and clay pumping/ballast fouling. These events were attributed to poor drainage

and retention of moisture within both the embankment and subgrade. The remediation performed was expected to remediate the drainage-related issues, but the mechanism behind such remediation had not been assessed yet.

This research aimed to evaluate the impact of the reconstruction using a reinforcing and a wicking geotextile in terms of the drainage properties and soil strength of the embankment materials.

2 METHODOLOGY

For measuring the temporal changes of VWC within both the embankment and subgrade, our site was divided into two adjacent sections of track, one remediated and one control section. The control section had sensors installed within the embankment with no other changes (Fig. 2), while the remediated section had the ballast and sub-ballast replaced with clean materials, and two geotextiles were inserted within the layer interfaces as shown in Figure 3.

Each section was instrumented with five 5TE sensors, which determined the VWC, bulk electrical

conductivity, and temperature from the soil at a 12-hour interval. Two access pipes were also installed to perform readings with a Diviner 2000 probe, which is capable of reading the soil profile at 10cm intervals. The results from the 5TE sensors were prioritized in the analysis due to the greater repeatability of readings in comparison to the Diviner 2000 probe. Tables 1 and 2 present the summary of the locations of the track used in the data collection.

Table 1. Location of 5TE sensors and Diviner access pipes for the control section (Alvarenga, 2021).

Sensor	Depth (m)	Soil Layer	Position at Track	Sensor
T1/T2	0.45	Sub-ballast	Shoulder	5TE
T3	0.45	Sub-ballast	Centreline	5TE
T4	0.65	Subgrade	Centreline	5TE
T5	0.80	Subgrade	Centreline	5TE
D1	0.60	Subgrade	Shoulder	Diviner

Table 2. Location of 5TE sensors and Diviner access pipes for the remediated section (Alvarenga, 2021).

Sensor	Depth (m)	Soil Layer	Position at Track	Sensor
TG1/TG2	0.45	Sub-ballast	Shoulder	5TE
TG3	0.45	Sub-ballast	Centreline	5TE
TG4 ¹	0.91	Subgrade	Centreline	5TE
TG5	1.07	Subgrade	Centreline	5TE
DG1	0.70	Subgrade	Shoulder	Diviner

¹ Sensor is not functional.

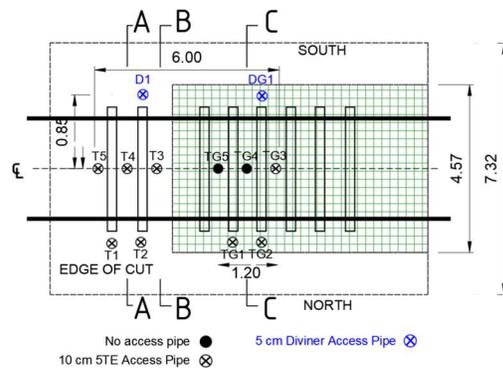


Figure 1. Moisture sensors configuration *in-situ*. (Alvarenga et al. 2020)

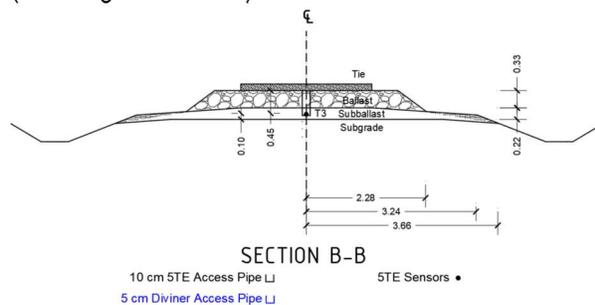


Figure 2. Cross-section of the control section (unchanged, only sensors placed). (Alvarenga, 2021)

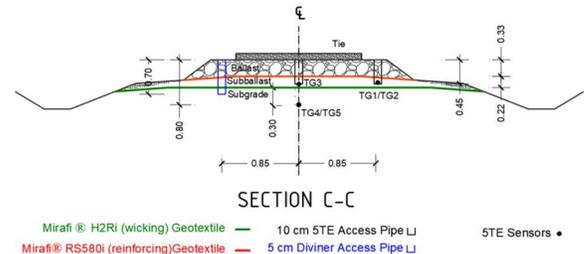


Figure 3. Cross section C-C from Figure 1 (Alvarenga et al. 2020)

In addition, the ballast, sub-ballast, and subgrade of the site were tested to analyze the soil properties and strength behaviour when there was a change in moisture content. Direct shear (DS) and soil-water characteristic curve (SWCC) tests were performed in the sub-ballast and subgrade materials to quantify the effect of moisture on the soil strength.

In order to evaluate the temporal changes in measured VWC concerning precipitation events, daily meteorological data were acquired from three weather stations near the study site along with the VWC data from the 5TE sensors. The antecedent precipitation index (API), a method that relates the VWC and precipitation data from a certain number of days, was used to facilitate the understanding of precipitation events' influence on the moisture content of the embankment at both sections.

The unsaturated strength of the subgrade was estimated by using the results from both DS and SWCC tests based on Vanapalli et al. (1996). Then, the strength levels of both sections were compared to verify which one presented a more significant decrease in strength when precipitation events occurred.

3 RESULTS AND DISCUSSION

3.1 Laboratory tests

Tests such as sieve analysis, moisture content, specific gravity, and Atterberg limits tests were performed to understand the soil behaviour. These tests were performed for the ballast, sub-ballast and subgrade of both control and remediated sections. A summary of the results of the tests is presented in Fig. 4.

Regarding the particle distribution within the field materials, the control section ballast presented some sand from the transition zone between the ballast and sub-ballast layers. In addition, the ballast and sub-ballast of both sections have shown less than 10% of fine content, but the subgrade presented approximately 50% of particles smaller than 0.075 mm.

Soil-water characteristic curves (SWCCs) provide a function used often for the evaluation of unsaturated soil properties by providing a relationship between moisture content and matric suction (Fredlund et al., 2012). The SWCC tests were defined in terms of matric suction vs. volumetric water content for all the embankment soils, except the ballast (Fig. 5). Curve-fitting for VWC vs. matric suction was performed using the Fredlund and Xing (1992) equation.

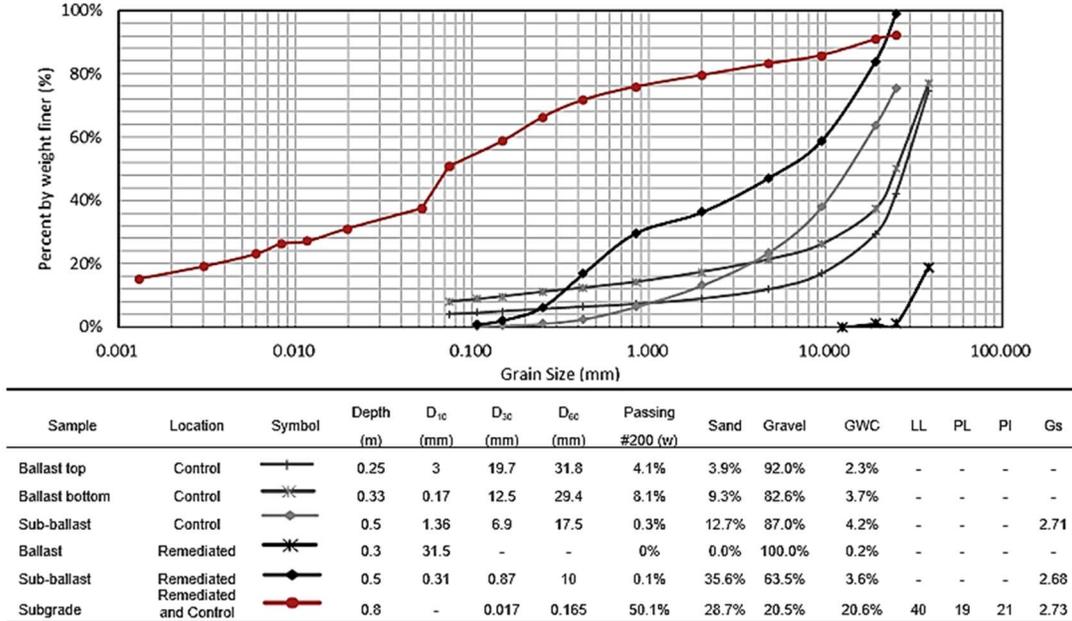


Figure 4. Summary of material characterization tests (Alvarenga et al. 2020)

$$\theta(\psi) = C(\psi) \frac{\theta_s}{\{\ln[e + (\psi/a_f)^{n_f}]\}^{m_f}} \quad [1]$$

$$C(\psi) = 1 - \left\{ \frac{\ln[1 + \psi/\psi_r]}{\ln[1 + 1000000/\psi_r]} \right\} \quad [2]$$

where θ is the VWC corresponding to a selected soil suction, θ_s is the saturated VWC, ψ is the suction, e is equal to 2.72, a_f , m_f and n_f are fitting parameters, $C(\psi)$ is the correction factor defined by Eq. 2, and ψ_r is the suction at the residual value.

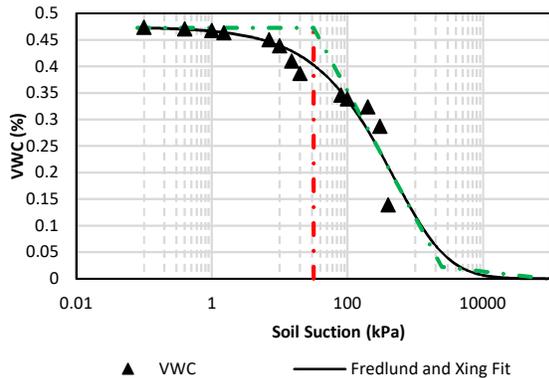


Figure 5. SWCC of the subgrade material (Alvarenga, 2021).

The basic information required for formulating a shear strength function is the saturated shear strength parameters, which can be obtained through direct shear tests and the SWCC (Fredlund et al. 2012). The DS tests

performed on the sub-ballast materials and subgrade made it possible to measure the saturated strength parameters of the embankment materials (Fig. 6) according to Equation 3:

$$\tau = c' + (\sigma_n - u_a) \tan \phi' + (u_a - u_w) \left[\left(\frac{\theta - \theta_r}{\theta_s - \theta_r} \right) (\tan \phi') \right] \quad [3]$$

where θ is the VWC of the soil, and θ_s and θ_r are the saturated and residual VWC values from the SWCC curve, respectively. Figure 4 presents the shear strength parameters for the sub-ballast and subgrade materials.

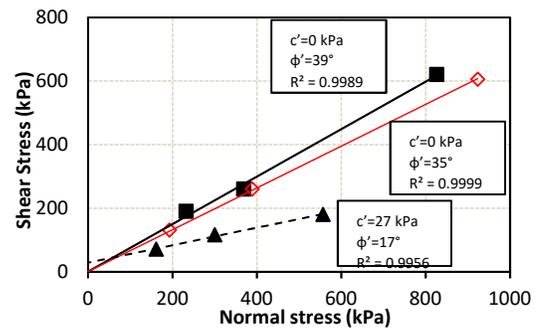


Figure 6. Summary of the Shear Stress results for the materials (Alvarenga, 2021).

3.2 Field Results

3.2.1 VWC Measurements

The VWC fluctuations between 12-hour readings for all sensors were averaged down to a daily value to account

for variations when high precipitation occurred. The winter season was removed from the analysis due to snow accumulation and freezing temperatures occurring during this period. All the information related to the remediated section and the control section in graphs is shown in red and black, respectively, as shown in Fig. 7. The sub-ballast shoulder of the track (T1, T2, TG1, and TG2) has not shown a great difference between the

3.2.2 API Model

The use of an API model allowed us to understand the interrelationship between precipitation and VWC at a daily scale by correlating the precipitation data from the weather stations and the VWC measurements obtained from the 5TE sensors. The first API concept was introduced by Kohler and Linsley (1951) and is defined

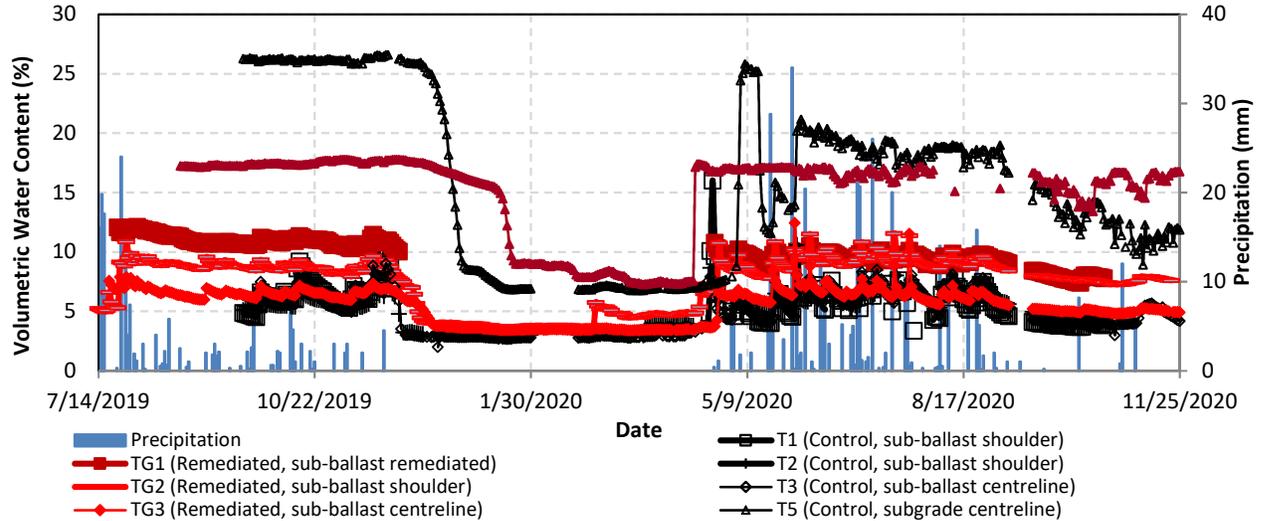


Fig. 7. VWC and precipitation recordings (Alvarenga, 2021).

sensors regarding the influence of precipitation. The remediated section kept higher VWC values throughout the entire study. This behaviour could be attributed to the water release from the geotextile ends to the track shoulder, but this interpretation must be treated with care.

The sub-ballast centreline sensors (T3 and TG3) have shown similar VWC for both sections, but the influence of precipitation was slightly higher in the control section (Fig. 6). During higher precipitation days, the control section has shown higher VWC points than the remediated section, presenting sharper changes in the VWC. In addition, the control section showed a higher VWC in the subgrade at all times throughout the study, except during the freezing period.

All sensors installed at the subgrade presented a decrease in VWC due to a lack of precipitation during a period of 33 days in September/October 2020. Sensor T5 (control section) was the most affected. The presence of a capillary break between the wicking geotextile and the subgrade interface as well as an insufficient hydraulic gradient are believed to be responsible for the lower impact of this period on the remediated subgrade sensor, TG5. The latter is believed to be occurring due to the installation method used for the wicking geotextile. For more effective drainage, the manufacturer (Tencate, 2020) recommends leaving the ends of the geotextiles exposed to promote water evaporation, while the method used kept the ends of the geotextile covered by the sub-ballast soil.

as shown in Eq. 4.

$$API_t = \sum_{t=0}^d P_t k^t \quad [4]$$

where P_t is the precipitation amount occurring on day t , k is the attenuation coefficient ($0 < k < 1$) and t is the time in days. The attenuation coefficient was estimated for each layer according to the best curve fitting value. Determining the value of k requires computations that are not justified as experience has shown that the factor is not critical (Kohler and Linsley 1951). The VWC data analysis went from July 2019- February 2021, which was split into 1-year cycles.

After determining the API for each day of the study, the second step was to separate the VWC data for each year cycle (Summer 2019- Fall 2019, Spring 2020-Fall 2020) to address the yearly precipitation variability. Then, the attenuation coefficient k and the number of days to be used for the API were estimated. These parameters were simultaneously determined by finding the best-fit values (R^2) for all sensors installed at similar locations. Graphical interpretation of the VWC vs. API was performed for each sensor using the linear regression model shown in Equation 5. This process aimed to optimize the quantification of the VWC variation with precipitation.

$$\theta_{v(t)} = \alpha \times (API) + \beta \quad [5]$$

where α and β are fitting parameters, and $\theta_{v(t)}$ is the measured VWC for day t . Figure 8 presents the data for the subgrade centreline sensors T5 and TG5 in 2020. It was observed that the control section sensors presented an outlier cluster of data, which was associated with the wetting period that occurred after the interval without precipitations (October 9 – October 28, 2020).

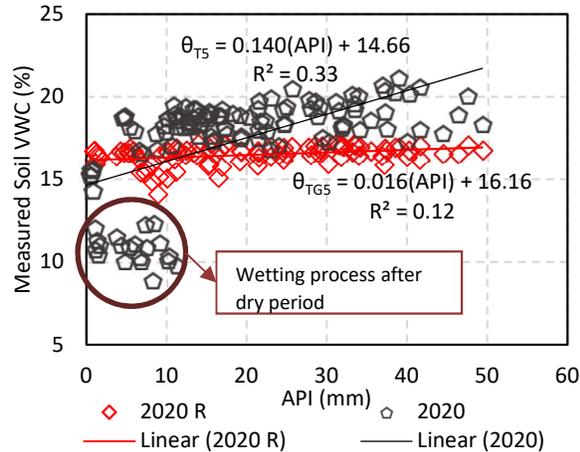


Figure 8. API vs. VWC for the subgrade sensors (Alvarenga, 2021).

Table 3 summarizes the variations of VWC vs. API (represented as the coefficient α) found for each sensor. The remediated section VWC variations were around 0.37-1.02x the variation seen in the control section at the sub-ballast shoulder. For the sub-ballast centreline, the remediated section variation was around 0.8x of the control section.

The period without precipitation greatly affected the coefficient found for T5 (control section, subgrade), as a cluster was formed when rainfall events returned in November/2020 (Fig. 8). For this reason, the effect of the remediation was analyzed after excluding this cluster, which would increase the control section variation considerably. Even after removing the wetting period cluster, it was possible to observe that the remediation favoured the drainage of the track by decreasing the VWC variation coefficient. The rather low R^2 presented by the subgrade is attributed to the low permeability of this layer. Taking this into consideration, the lower R^2 presented by the remediated section may also indicate that this section is less sensitive to changes in the precipitation regimen. Figure 9 presents the linear regression for the sub-ballast shoulder for a comparison between R^2 values for the granular sub-ballast vs. the R^2 found for the fine-grained subgrade.

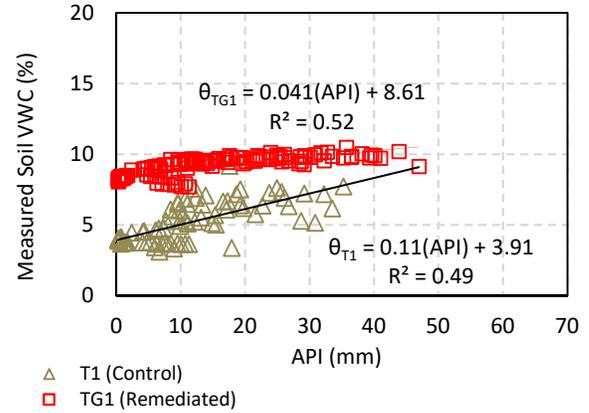


Figure 9. VWC vs. API regression for sub-ballast shoulder sensors (Alvarenga, 2021).

Table 3. VWC variation summary for all sensors (Alvarenga, 2021).

Sensor Location	Sensor #	Coefficient α (VWC vs. API)	
		Control Section	Remediated Section
Sub-ballast Shoulder	1	0.110	0.041
	2	0.041	0.042
Sub-ballast Centreline	3	0.052	0.040
Subgrade Centreline	4	0.076	-
Subgrade Centreline (*excluding wetting period cluster)	5	0.031	0.016

3.2.3 Strength model

The impact of the remediation on the strength of the subgrade is observed in this section. This analysis was made possible using the SWCC and DS results to define the matric suction contribution to the soil shear strength in unsaturated soils according to Eq. 6, developed by Fredlund et al. (1978):

$$\tau = c' + (\sigma - u_w) \tan \varphi' + (u_a - u_w) \tan \varphi_b \quad [6]$$

where τ is the shear strength, c' is the effective cohesion; σ is the total normal stress, $(u_a - u_w)$ is the matric suction, u_w is the pore-water pressure and φ' is the friction angle concerning changes in $(\sigma - u_w)$ when $(u_a - u_w)$ is held constant; and φ_b is the friction angle concerning changes in $(u_a - u_w)$ when $(\sigma - u_w)$ is held constant.

Overall, the suction values found for the sub-ballast of both sections showed that this material is close to dry conditions, and the remediation did not bring an impact on its strength. Therefore, only the subgrade strength was analyzed.

In unsaturated conditions, the presence of lower moisture values leads to an increase in the shear strength and a consequent improvement in the bearing capacity of the material.

Figure 10 presents the strength of the subgrade material through the time of the study. In the figure, it is possible to see that the remediated section strength has maintained levels between 90-97 kPa, due to the VWC

presenting overall constant levels at this section. In contrast, the control section presented an ample range of strength levels, from 71.5 to 101 kPa. The range is caused by the different moisture levels and its direct influence onto the material strength. While the remediated section remained at a constant moisture level, the control section presented highly variable moisture after the thawing of 2020.

Therefore, the remediation improved the drainage of the track embankment after a precipitation event.

In addition, all the remediated section sensors installed at the sub-ballast presented a similar variation of VWC with API, evidencing that homogenization of the variation in moisture may be one of the benefits brought by the reconstruction.

The subgrade may be experiencing some difficulties in

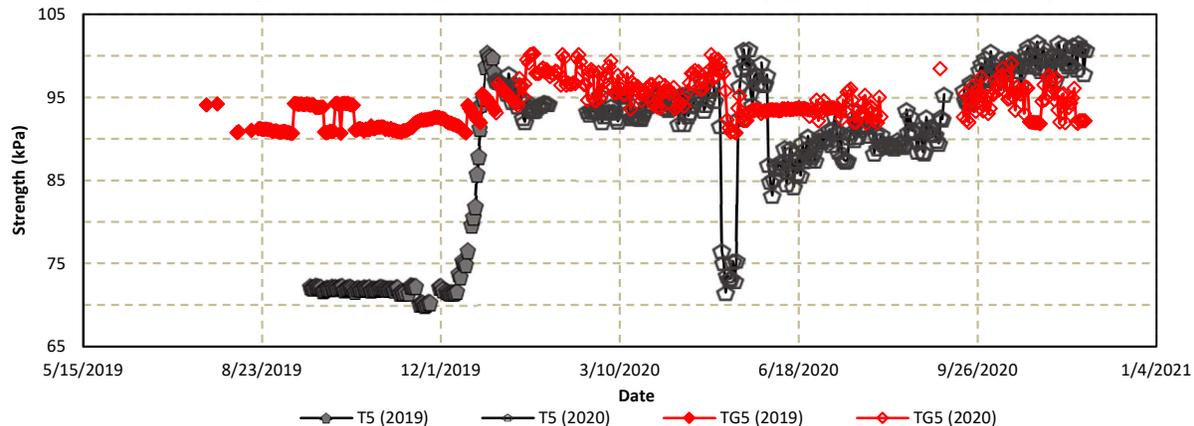


Figure 10. Strength of the subgrade material at the control and remediated sections (Alvarenga, 2021)

The remediated section presented strength levels higher than the control section before the spring-thaw of 2020 and maintained this trend for most of 2020 (Fig. 9), but a dry period in October/2020 promoted a convergence of the estimated soil strength for both sections. In the last months of the analysis (October- November 2020), the control section sensors presented high strength values associated with the VWC response to the dry period. A longer-term analysis from another spring-thaw is desirable to see how these sections behave under a more common precipitation regimen.

Since the remediated section did not present considerable increases in VWC during high precipitation events, this section seems to have its short-term drainage improved by the remediation.

4 CONCLUSIONS

This research had the purpose of understanding the benefits of the reconstruction using a wicking geotextile regarding the drainage performance of a low-permeability subgrade and its possible benefits to the soil strength by soil testing and field monitoring of one control and one remediated section.

During a dry period of 33 days in September/October 2020, all sensors installed at the subgrade presented a decrease in VWC due to the lack of a water source. Sensor T5 (control section) presented a considerable decrease in its VWC during this period. The presence of a capillary break between the wicking geotextile and the subgrade interface is believed to be responsible for the lower impact of this period on the remediated subgrade (sensor TG5).

The remediated section presented lower VWC variations than the control section in all sensors, i.e., this section tends to be less sensitive to precipitation.

releasing moisture at a VWC below 17%; therefore, it is recommended to study the underlying reason through changes in the field setup or the use of laboratory tests. Some of the believed causes is the development of a capillary barrier between the subgrade and the geotextile, clogging of the wicking fibers by the fine particles and the way the wicking geotextile was installed on the field.

The installation method used for the wicking geotextile in this study is recommended by Tencate (2020) only for water redistribution. For drainage improvement, Tencate (2020) recommends leaving the ends of the geotextile exposed to create a hydraulic gradient to facilitate water movement. Further research in a setup where the edges of the wicking geotextile are exposed to air is desirable for better drainage results.

The subgrade strength has maintained consistent levels at the remediated section with values around 94 kPa, while the control section presented subgrade strength values varying between 70-100 kPa. However, the control section only reached high strength levels after an unusually long period without precipitation, making it possible to conclude that the remediation also improved the subgrade strength levels when high amounts of precipitation were present.

This study addressed solely the improvement in drainage and strength brought by the reconstruction and the wicking properties of Mirafi® H2Ri. It is recommended to perform large-scale laboratory tests with both control and remediated soil samples, or the use of finite element model to separate the effects of the remediation from the difference in the soil composition of both sections. This study also did not consider the other possible benefits of Mirafi® RS580i and Mirafi® H2Ri, including reinforcement, stress redistribution, frost heave mitigation, separation, and filtration, neither its effect on

the soil stiffness. Further research is recommended to assess these aspects of the remediation.

5 ACKNOWLEDGEMENT

This research was made possible through the support of the Railway Ground Hazard Research Program, funded in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Pacific and Canadian National Railway.

REFERENCES

Alvarenga, C. B. S. de (2021). *Effect of an Embankment Reconstruction Using a Wicking Geotextile in the Drainage and Strength of Subgrade* [Master's Thesis, University of Alberta]. DOI: <https://doi.org/10.7939/r3-3rv2-9q79>.

Alvarenga, C., Haji Abdulzaragh, P., Hendry, M. T. (2020). Monitoring of the change of moisture beneath a railway embankment and the effectiveness of a wicking geotextile. *Proceedings of the 73rd CGS Conference, GeoVirtual 2020*, Calgary, September 14, 2020.

Selig, E. T., & Waters, J. M. (1994). *Track Geotechnology and Substructure Management*. Thomas Telford.

Environment and Climate Change Canada (ECCC) (2020). *Historical Data: 2008-2018*. Accessed April 09, 2020, from https://climate.weather.gc.ca/historical_data/search_historic_data_e.html.

Fredlund, D.G., Morgenstern, R. N., & Widger, R. A.(1978). *The shear strength of unsaturated soils*. Accessed April 17, 2020, from <https://www.nrcresearchprdoi/10.1139/t78-029>.

Fredlund, D.G., Rahardjo H. & Fredlund M. D. (2012). *Unsaturated Soil Mechanics in Engineering Practice*. New York, NY: John Wiley & Sons Inc. doi: 10.1002/9781118280492.

Fredlund, D. G., Xing, A. (1994). Equations for the soil-water characteristic curve. *Canadian Geotechnical Journal*, **31**(4), 521-532.

Indraratna, B., Salim, W., & Rujikiatkamjorn, C. (2011). *Advanced Rail Geotechnology: Ballasted Track (1st ed.)*. Boca Raton: CRC Press/Taylor & Francis Group. DOI: 10.1201/b10861.

Kohler, M. A. & Linsley, R. K. (1951). *Predicting the Runoff from Storm Rainfall*. Accessed April 18, 2020, <https://www.nrc.gov/docs/ML0819/ML081900279.pdf>.

Li, D., Hyslip, J., Sussmann T. & Chrismer, S. (2002). *Railway Geotechnics (1st ed.)*. London: CRC Press. DOI: 10.1201/b18982.

Tencate (2020). *Installation Guidelines for Mirafi® H2Ri Moisture Management System*. Retrieved March 23,

2021, from https://www.tencategeo.us/media/0dcf4b31-c2bb-4844-9132-61b6a6fcb1a/CTkwZQ/TenCate%20Geosynthetics/Documents%20AMER/Installation%20Guidelines/Roadway%20and%20Railway%20Construction/IG_H2Ri%200520

Vanapalli, S. K., Fredlund, D. G., Pufahl, D. E., & Clifton, A. W. (1996). Model for the prediction of shear strength with respect to soil suction. *Canadian Geotechnical Journal*, **33**(3), 379-392. doi:10.1139/t96-060

Session 4

RISK AND RELIABILITY

From paper notes to digital twins: The evolution of geospatial models along railway corridors and digital convergence

Robin Harrap, Paul-Mark DiFrancesco, David A. Bonneau and D. Jean Hutchinson
Department of Geological Sciences and Geological Engineering, Queen's University, Kingston, Ontario, Canada K7L3N6

Zach Sala, and Matthew Ondercin
BGC Engineering, Vancouver, B.C.

ABSTRACT

Many workers in the rail sector have worked through the arc from spatial data and associated records being entirely paper based to the current heterogeneous mix of CAD, GIS, databases, and focused simulation tools. While there is a digital divide between those records that are and are not digital, there is an increasing digital 'dark age' divide between digital tools that do not inter-operate, and in some cases are no longer supported. Looking to the civil construction sector, the convergence of many forms of plans and records has followed a similar trajectory and there is widespread convergence on just a few tools that interoperate efficiently: GIS (for regional data), BIM (for site data), and visualization tools (for site renderings and public outreach). As graphics processing power increases, the ability to use one or multiple tools on a typical desktop or mobile computer has allowed very sophisticated 3D models that integrate scan data (LiDAR and Photogrammetry), equipment inventory (including maintenance and operational limits records), simulation, and visualization, and these will soon be available on tablet and smart phone platforms. A significant behind-the-scenes agent in this convergence is integration around ideas from game design since games provide an integrated environment for events to happen in complex 3D spaces. In work supported by the RGHRP, we look at the implications for digital convergence towards 'digital twins' of real environments in game engine environments, how different source data can be represented, and the relevance of this convergence for site visualization and analysis, as well as the implications for data integration across teams in an organization. We briefly examine the implications of digital convergence for the utility of existing and forthcoming tools for use on sites, including an example using game technology.

1. INTRODUCTION

Railway corridors comprise infrastructure set in a spatial context, whether that be urban or rural. Many corridors pass through areas with significant geotechnical issues of concern such as rockfalls and landslides. Evaluating, monitoring, remediating, and in some cases avoiding areas with very significant challenges requires a combination of technical knowledge, field data, standards of practice, and strategies for ongoing monitoring.

The information used to prepare for, carry out, and resulting from field data collection is complex. It involves information that will be used to determine where:

- a) A problem is indicated that can be addressed with well-established methods,
- b) A problem is indicated that is beyond well-established methods and will require ongoing surveillance,
- c) A problem is indicated that is an immediate and critical concern,

d) Specific data is required currently or in an on-going basis to evaluate a site,

e) Specialized expertise must be brought in to assess or suggest remediation.

In many if not most cases a combination of these factors will be appropriate.

In addition, in many cases a field visit will require preparatory and follow up work that involves examining historical records, records for similar sites, and related regional datasets such as weather, geological mapping, usage data, and so.

In essence, the problem of understanding a hazard situation along, or near a track, from new and existing data is a problem of constructing a model of the world and using engineering knowledge and best practices to address it. This may include hazards well outside of the scope of geoengineering.

In most cases the decision point is based on a complex mix of knowledge in the minds of the on-site engineers, advice, standards or requirements from experts or

governing agencies, data in a wide variety of forms, and information captured in reports. In many, if not most, cases the primary output besides on-site action is more reports.

A significant problem here is that this world-model, which comprises a number of different components with different levels of preservation (from none at all to repositories of historic data), is not in any ways complete and often cannot be well understood without (at best) consultation with the people active at the site at the time. Much of the knowledge is tacit (Polanyi, 1966, Crandall et al. 2006), and unarticulated even in a longer form report. Furthermore, the knowledge may lie in the space between the mental state of several relevant personnel, with their personal notes, maps and drawings in various formats, instrument readings in various formats, and links to external documents such as reports and guidelines and regulations all being part of a complex mix.

In addition, in at least some cases, and not uncommonly in geotechnical site investigations, software tools will be used to evaluate possible outcomes, an example being rockfall simulation tools. These are, again, at best poorly connected to all of the other data, information, and knowledge that goes into decision making and is captured into an ongoing record.

A further difficulty arises because generally the final repository and common sharing approach for a site model is a collection of reports, some of which may be difficult or impossible to obtain. Historically these ended up in filing cabinets or in storage, accompanied by their source data, and the transfer of the knowledge in these reports – and perhaps even of their existence – is poor with regards to later works.

Furthermore, there is rarely an overarching interpretation and organization to this form of data. The integration was in the mind of the writer and co-workers.

While digital methods introduced over the last several decades may have been meant to address this issue of site information capture and efficient use, in many cases they have instead made the situation more complex. Individual datasets and report components may be explicitly tied to software that is no longer available and are often in vastly more complex forms, and create situations where an old fashioned report would be welcome. This is referred to as the 'digital dark age' problem in information science (Bollacker, 2010). It includes software specific, operating system specific and peripheral specific issues. It also includes the questionable physical preservation of storage formats including tapes, removable disks, CD and DVD media. There is no reason to believe that loss of some if not most of archived digital material will be a problem. For example, if a major cloud storage vendor that had not been in common use for a number of years closed down, what would be lost?

We emphasize that, while these are often very serious issues, the reality is that much of the actual site knowledge remained in the mind of the workers and was never recorded. This situation is a very serious issue that is rarely addressed at all, and results in loss of the experience based interpretation of the significance and

usefulness of the data (Klein, 1992) . Due to the intergenerational lifespan of infrastructure, we are in need of ways to digitize and preserve the knowledge of experts across the various mediums in which they do their work. This includes at least some of the perspectives, opinions, and insights that those that did the work had, but which were not sufficiently technical to be recorded.

This paper examines some larger scale issues in information technology and software best practices and suggests that convergence as seen in the CAD sector and the entertainment industry provides significant future pathways to address these issues. It frames problems and solutions against the reality that we are seeing dramatic increases in network speeds, the use of remote computational and storage resources, and new approaches to model integration driven by media and game production. In the end this points towards the use of digital twins along rail corridors, with some specific potential applications being highlighted. Digital twins are models of environments that include site characteristics, simulations, and site sensor integration (e.g. Hodgkinson & Elmouttie, 2020) and originating in the manufacturing and space sciences sectors. An example using game engines to construct simulation tools illustrates the practicality of the approach discussed.

We will begin by framing some of the overarching problems, and the nature of technology change in tools used, in the context of CAD.

2. FROM PAPER TO CAD TO BIM

Railway infrastructure, and to some extent the geotechnical environment at hazardous sites, is rooted in paper maps, historically produced by drafting technicians. Over the last 40 years there has been an almost complete transition of site mapping, infrastructure plan production, and to an extent change mapping, to digital formats dominated in many cases by computer aided drawing (CAD), with the use of Geographic Information Systems (GIS) also being commonplace.

CAD uses the metaphor of layers of drawn information at some scale, each layer representing related information, with control over which layers are visible. Importantly, the layers seen in a drawing can include layers from other CAD files, including files that are remote. This allows re-use, although at the cost of requiring significant planning, network access, and of creating a more complex working environment.

Chains of tasks to be done to accomplish something with software are referred to as a workflow. In some cases CAD was the *point* of a workflow – producing detailed drawings of infrastructure to be constructed on site would be an example. In other cases, CAD provides spatial context for a larger decision, or for simulation, or for comparing different sites, or for comparing changes to a site over time.

In the construction sector, workflows include planning documents, but also as-built drawings, and drawings used for site maintenance and ongoing monitoring.

A change of focus, here, is revealing. Instead of thinking of CAD (or any other data) as the point of work being done, it can be thought of as an element of a decision-making process. In this case CAD is part of a what is often referred to as a decision support system (Turban & Aronson, 2000) that comprises multiple individuals and several semi-independent data components, much like the description of a site introduced earlier.

CAD files suffer from loss due to version changes, due to the lack of documentation of idiosyncratic styles of use, and to misplacement or loss of their files. In some cases the CAD file will be gone but some printouts may be preserved in some state. As such, they suffer from the problems introduced earlier as much as any other form of data.

In the construction sector several needs 'beyond traditional CAD' have emerged. CAD is fundamentally extensible since it strongly supports user scripting. From many individuals automating (scripting) their own tasks emerged a wide range of common add-ons; we will focus on two types here:

a) Extensions to add stronger support for tabular data about features in a drawing, or assemblies of features, and

b) Extensions to perform simulations in support of decision making.

The first option is similar to the core concept of Geographic Information Systems – linking descriptive attributes to specific spatial features, and represent transitional steps towards using those tools in the first place, as is becoming more common.

The second is an attempt to take computational tasks that would previously have been done in a separate program (perhaps written by a 3rd party) and execute it inside the CAD environment. This might be well integrated, if all the data that the tool uses is found in that file, or may involve user input or linking to other external data. These hybrid approaches are often quite susceptible to software failures if any of the components are upgraded.

These two approaches led to the creation of a subset of CAD over the last two decades: Building Information Management (BIM) tools. BIM differs from GIS in that GIS addresses more regional mapping, while BIM is focused on buildings or groups of buildings. It has been applied to geotechnical problems, e.g. Morin et al. 2014. Often built on and strongly resembling their CAD predecessors, these add specific, decision-support focused abilities to off-the-shelf CAD tools, including maintenance scheduling and probability of failure analysis, links to inspection workflows, simulation capabilities to support decision making during construction and maintenance, and more. An example of simulation capability would be the use of computer graphics and numerical simulation approaches to simulate lighting on buildings to model both illumination and HVAC load.

The key here is that a range of tools, some of which were from vendors and some were user constructed, converged to support stronger workflows and provide stronger decision support capabilities. This idea of

convergence (AT&T, 2017)(and at attempt to look forward to what convergences may be about to affect rail corridor work) is what will occupy the remainder of this paper, with some consideration of the challenges of *where the information and knowledge resides, and how decision makers interoperate in and out of the field*. Our concern here is whether current or upcoming convergences will address the problems introduced.

3. INFORMATION ALONG RAIL CORRIDORS

The emphasis on historical site inspections and reports introduced above does not do justice to how complex the actual data collection, integration, evaluation, and communication process has become over the last few decades.

The records available for a site will include rail and near-rail infrastructure, records on maintenance and repair requirements, conditions at the site, conditions near the site, and more. We will focus here on the site and near-site records and how they have changed.

Historically a site record would comprise a map, some notes in a notebook, some photos, and perhaps some collated data from a sensor.

With the introduction of CAD, drawings and maps might be updated manually in the field for re-working in an office environment.

With the introduction of word processors and spreadsheets, much of the field data would be collated into documents either while in the field, or out of the field, which constitute a collection typically more extensive than what ends up in a report and may or may not be preserved. In some cases databases might also be used. At times significant scripting (often in a spreadsheet) was used, and the scripting itself (and the information about 'how' inherent in the code) was not contained in the following report.

With the introduction of digital cameras and then cell-phone-based cameras, the record from a site might include far more photos, some of which are in resulting reports and communications, some of which are archived, and some of which are discarded (intentionally or accidentally).

With the introduction of Personal Digital Assistant (PDA) hardware, tablets, and smart phones it became possible to take notes, annotate drawings, and integrate photos and notes in the field. Again, these field documents may or may not be available later.

Out of the field, the use of remote sensing data and simulation tools constitutes another collection of data, information, observations, and processes that, while part of the workflow at or about a site, is somewhat disconnected from the other components discussed above, despite often ingesting data from one or more of them.

Three new technologies have become common in the last decade that very significantly add to the data and intermediate outputs that must be managed: LiDAR, Photogrammetry, and UAS data.

In each of these cases a new technology is used to provide direct assessment of a site with remote capability allowing safe observations of, for example, the

upper parts of an unstable slope. We will focus on the LiDAR and Photogrammetry data as both are immediately about the geometry of a site.

Individual slope models built using either of these methods can directly constrain the geometry of a site for direct inspection, integration into numerical models, comparison to other sites, and the like. Models built over time, by repeat data acquisition, not only allow change detection to be carried out, they also can be integrated with numerical modeling approaches to get at the nature of underlying processes, such as trajectories of fragments during a rockfall. This is not distinct from, for example, earlier tools for rockfall simulation, but simply constrains the site model to a greater degree, taking one from a 2D simulation environment to a 3D environment for example. An example of tools that do exactly this is discussed below.

The problem here is that we are dramatically increasing the number of individual tools being used in decision support workflows (as well as dramatically increasing the size of datasets). This increases the number of places where expert judgement is being used (but perhaps not documented), creating far more opportunities for data and context loss, and creating an environment where keeping track of what data exists at a site and quickly inspecting it for relevance is very difficult.

4. TECHNOLOGY CHANGE IN THE NEAR FUTURE

Smart phones are near ubiquitous among technical users. Tablets are increasingly common, and some have (limited) near-field LiDAR capabilities. The line between software that will run on a laptop or desktop computer and a phone or tablet is increasingly blurry. In many cases control of sensors is through an application on a phone. The value of a phone is the connection to a network, and it is increasingly rare to be outside of network coverage (though not impossible along rail corridors).

We want to emphasize three near-future technical capabilities. All of these exist now, but are at the experimental or local availability stage. Taken together, and combined with new supporting technology, they allow a new convergence of digital tools for use at rail sites.

The first is the availability of next-generation '5G' network infrastructure. This infrastructure, accessed with a compatible digital device, is sufficiently fast that remote data resources can be efficiently accessed, that a field worker can capture data directly (either solely or in addition to local storage) into a shared repository, and that this can happen while remote workers are communicating with a field worker with access to camera and audio. Network coverage at this level is limited but will expand very rapidly over the next several years.

The second is the upcoming availability of remote computational resources *over this network*. Numerical models that require very fast computers, very large amounts of storage, and very large amounts of RAM can now be run remotely, integrated with shared

repositories, and will allow sophisticated 'what if' modelling on a site. This would, for example, allow a field worker to determine that the model would be better constrained by a 'bit more data' and save a further field visit. This includes both access to remote CPU resources and the use of GPUs as computation engines (c.f. Eberly, 2015).

The third is the upcoming availability of remote graphics processing *over the network*. Similar to the capability described above, this is the use of a remote high-end computer or set of computers, but instead of driving computation, the point here is to drive graphics output on a phone, tablet, or other portable digital device. Again, this furthers decision making in the field since it allows complex data to be co-visualized in a way that was previously only available on high-end desktop computers. It also opens the door to easy in-field use of augmented reality and virtual reality interfaces.

Since the computational and graphical support offered are programmatic, an application can be built that is transparent to the fieldworker, allows shared use between on-site and off-site workers, and gracefully degrades in the case of network limitations.

Is this possible? Is this reasonable? A first point is that we already use distributed, multi-user, user-friendly applications that have very high computational and graphical support on a daily basis: this is what many websites we access (often with our phones) are. The core technology of, for example, Amazon Web Services delivers many of the capabilities discussed above.

Is this reasonable, though, for complex sites? A second examination of convergence in the software industry supports the idea that workflows and decision support built around high performance visual and numerical worlds is not only possible, it has been available for some time.

5. CONVERGENCE IN MEDIA TOOLS

We will examine two parallel cases from media production. The first is the use of data and software tools to build digital worlds for special effects – from overlaid and inserted digital components overlaid on film through entirely virtual sets – in other words, animation. The second is the creation of virtual environments and assets for games.

Building a spatial feature or environment for use in animation or gaming involves building a model that comprises at least two components:

1. The geometry of the feature or the environment.
2. Data to support simulating the visual appearance of features.

Most animation environments used in film, historically, had to integrate with an existing real environment and were overlaid on film taken on a set or in the world. There are a variety of techniques used, but the relevant issues include templating, individualization, and grouping.

Building assets for animation is complex (Kerlow, 2008) and expensive. Once one is built, it will be re-used in a production, and perhaps made available in a marketplace. Libraries of features are widely available

and include feature from the commonplace (a specific model of car) to the rare (a specific temple in Japan).

Individualization of templates involves situating them at a location and tweaking them to integrate, as well as to 'not look like a copy.' For example, in a movie, scuff marks might be added to sections of (templated) pavement.

Grouping refers to assemblies. Is the model of the temple a single feature? It is actually dozens of individual components that can be placed as a unit, or individually.

There is a strong parallel here with how drawings are made in CAD. Templates are common. Re-use is emphasized. The development of tools and techniques is driven by the expense of human animators.

In a typical workflow in the entertainment industry many tools are used to make these models, and their final destination is often an animation package, which handles motion, animation and some additional special effects. There is a trade-off between converging individual tools into larger animation packages and keeping tools distinct to address specialist workflows, for example the differences between motion-capture animation and digital sculpting. Nevertheless there is convergence.

Game studios add to the requirements for building visually realistic environments. First of all, the user has agency to move around and (to some degree) modify the environment. Second, the interface must be simple enough that it can be learned without significant training (though expert gamers have vastly more sophisticated understanding of what interactions are possible). Third, the system must integrate local data with remote data, (for example, accessing other players' locations, or additional site data from a server, or graphics pre-processing using remote servers to facilitate advanced graphics on mobile phones. The game is designed to be real-time. Simulations, movement, interaction, and peer-to-peer communication all happens simultaneously and if it isn't smooth and responsive, the game will be badly received. The field of virtual world development integrating all of these aspects has evolved over several decades and best practices are well established (Bartle, 2004).

Tools to build game worlds are evolving rapidly, as are tools to integrate with other applications and approaches like LiDAR, photogrammetry, UAS data collection, and to ingest data from GIS.

Games also include significant security measures to control who can see what, what effect players can have on each other, and to recover in the case of a failure. Game companies are also the main drivers of efforts to stream high resolution graphics to mobile devices.

One key capability that has been integrated both into virtual production and game software toolkits is physics simulation. The ability to run physical process simulations (Flynt & Vinson, 2005, Eberley, 2004), for example of falling material, is key to visual realism and reducing manual work by animators. The capabilities of physics engines embedded in animation or game engines have grown dramatically in recent years. The

next section presents a case study of doing exactly this for railway corridor decision support.

Game engines represent the convergence of many ideas that originated in CAD, hand-drawn animation, and animation software tools. The convergence in terms of where these tools are used, in other words between what used to be completely different tools, has been dramatic in the last few years. Here are three examples.

Television shows are now being shot (routinely, not experimentally) inside game engines, with the actors being digitally inserted. This includes many live action shows like TV news.

Concerts are now being held as virtual events, with thousands of participants interacting in (so far, simple) ways with musicians or actors on a stage. This is discussed in online seminars at UnrealBuild, 2021.

Individuals can now use their phones to capture the geometry of features and upload them to shared environments, collaboratively building worlds.

The key aspect of convergence here is that, increasingly, tools are used in building worlds or portions of worlds. Information from sensors, templates, and participants are converging in a shared virtual representation of an (often real world) space.

This is, of course, augmented by different techniques for interacting with digital artifacts such as augmented and virtual reality.

The convergence around building local world models rather than building individual features has now extended beyond the entertainment sector. GIS and CAD vendors now offer direct integration of their data into animation and game engine tools. The field of architectural visualization (seeing a building before it is built, on site) is now merging with game design. The convergence between real and virtual models and many other tasks is increasing.

Not only does this allow new techniques and new approaches to be introduced, not only does it save on the complexity of tools used, but these tools are written by expert teams with an eye to long term form support, security and backup capabilities, and graceful transition to newer software standards rather than vendor lock-in. An example of how game and animation studies can drive open standards that will support interoperability between tools and virtual worlds is the Universal Scene Description movement led by PIXAR, NVIDIA and others (Pixar, 2021). This means that the ease of moving models between different tools is a significant focus with major vendors agreeing on interchange formats so different specialists *can examine the same world model while preserving their own tools and approaches*.

In some cases more powerful vendor-agnostic tools are also making data that had lost software support relevant again, since these tools (in attempting to be as flexible as possible) support many import and export formats.

6. CONVERGENCE WITH RAILWAY SITE INVESTIGATION

If, in the near future, workers on sites have even a few of the capabilities, one can imagine a few scenarios.

These are based on the dramatic increase in network speed, the ability to use computational / graphical resources remotely, and the ability to host large and complex shared models (in other words, worlds) on centralized servers.

First of all, simply being able to use one 3D graphics visualization environment (whether on a handheld device, or using AR or even limited VR) to see all of the historical and recent data in an overlaid spatial context is powerful. It also justifies putting an effort into finding historic datasets and integrating them.

Second, the ability to then run spatial simulations, for example on rockfall trajectories, in that same environment, visualize it, and then plan and carry out further site investigations and mitigation planning is powerful.

Third, the use of the network capability to 'call an expert' while in the field who can see that same model, make suggestions, add supplementary data, integrate outputs from other tools using interchange formats, and perhaps show analogous examples is powerful.

Finally, the direct integration of sensors such as LiDAR and, perhaps more importantly, various sensors deployed by UAS into a single environment may allow new efficiencies in fieldwork. Being able to select an area in the field, and have a UAS use AI-based planning methods to go collect and then deliver up-to-the-minute spatial data into the shared environment would fundamentally change what we mean when we talk about field observations, especially in environments where direct observation is dangerous.

We are stating that the convergence seen in media and gaming tools can and should result in new tools for site investigation along rail corridors. To demonstrate that this claim is reasonable, we now present an example where we have used these tools to do exactly that.

7. AN ILLUSTRATIVE EXAMPLE – ROCKFALLS AS GAMES

In work carried out by two of the authors, our team built a 3D rockfall simulation tool using the Unity game engine. An overview can be found in Harrap et al. (2019) and the detailed documentation is available in Ondercin (2016) and Sala (2018). An example of a detailed study is published in Sala et al. (2019).

The core idea reflects the convergence argument given above. We used LiDAR and photogrammetry data from the White Canyon rail corridor within the southern Thompson River Valley in Interior British Columbia to build a world model that included slopes, along-track infrastructure, and the rails themselves. Although we did not choose to represent vehicles, it would not have added significant complexity.

The world model in the game was supported by the game development environment. Unity supports data import, scaling, characterization, and features strong support for visual realism. It is an example of a game engine, which can be thought of as an integrated authoring environment where many if not all of the tools needed to make a game converge. Although in some

cases other tools are very useful – for example, data to preprocess LiDAR – almost all of the spatial tools are present in Unity to build the 'built' environment. In addition, Unity supports a physics engine that can simulate real-world processes given models and parameters.

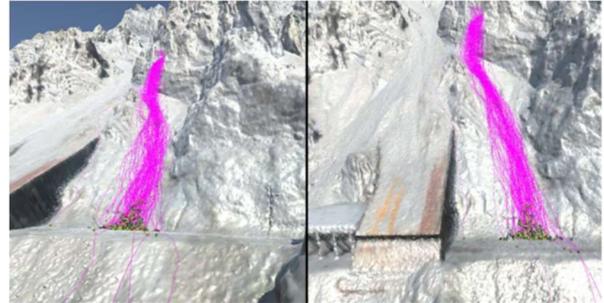


Fig. 1: Two views of the same rockfall simulation run in the Unity rockfall simulation tool. The slope is captured by terrestrial LiDAR scanning. The model has 287 blocks of cubic shape and .064 m³ size. Modified after Ondercin (2016).

Developing in this environment includes importing and scaling spatial assets, writing user interface tools to allow the tool-user to control the 'game,' and attaching scripts to individual features that are active as the game launches, runs, and completes.

In the case of the Unity rockfall simulation tool, the game loads an environment with pre-set rocks that can then be released. Individual falling objects are tracked, as are all of their interactions, and on completion (i.e. when everything stops moving) the results are exported in tabular format for inspection in a data analysis package. An example for homogeneous rockfalls showing the ability to view the same simulation from different points of view in the game engine is shown in Fig. 1. There is no reason that the decision support analysis tools could not be integrated into the game itself, but this was beyond the scope of the student work.

The question, of course, is how the performance of this system compares to real-world situations and to other software tools. By building game environments that mimicked situations where rockfalls were extensively documented, Sala (2018) demonstrated that the Unity rockfall simulation tool gives realistic results useful for decision making. This involved not only comparing rockfall simulations to before- and after-LiDAR and photogrammetry models for rockfall events in the White Canyon, but also beginning to examine how failure-induced fragmentation of blocks might be represented in a game environment. An example of a simulation with different shapes based on those observed in the field is shown in Fig. 2.



Fig. 2: 1000 fragment rockfall simulation result using discrete and variable fragment shapes. Total volume is 100m³. Modified after Sala et al. (2019).

Work on fragmentation, efficiency of larger simulations, interface issues, and calibration is ongoing for this project.

It represents convergence for several reasons. First, the integrated environment made it practical to build world models in the first place. Second, the physics engine in Unity allows the physical process simulation to be carried out. Third, the game environment provides an interface that simplifies (from the users point of view) interacting with a complex 3D environment. Software developed in a game engine also supports including the specifics of how modeling is done in a distributable form, and supports a wider range of input data than many stand-alone packages. For example, the Unity rockfall simulation can use shapes imported from change detection work whereas some rockfall packages support only predefined shapes.

8. DISCUSSION AND CONCLUSIONS

The idea of using tools based on the integration of ideas from media, GIS, CAD, animation and more is not new. Digital Twins (Atkins, undated, Hodgkinson et al., 2020) represent models build of specific environments, often linked to in-situ sensors, and often coupled with simulation tools that allow what-if scenarios to be run. In some cases, they are presented as web applications for ease of use and dissemination.

The Unity rockfall simulation example is a first step towards a digital twin tool for a near-track environment. It currently lacks in-situ sensor integration, and more work is needed on the geological aspects of the modeled rockfalls, but it is an element of a future environment that could be used on a device in the field as described above, using the upcoming technological capabilities we introduced.

We introduced one significant challenge that has not yet been discussed. How does a digital twin or a digital twin-like tool based on the technologies discussed in any way address the problem of information being spread

between many reports and at least partly residing in the minds of workers (who may leave with it when they depart a role)?

Data about data is often called metadata (Wikipedia, 2021). For example, data about who made a GIS or CAD dataset might include their name, when it was done, and some limited information about context.

Data that provides context and justification, links to supporting information (for example using URLs) and illuminates possibilities and opportunities are included in a category that is termed paradata. It includes information about how data was collected (US Census Bureau, 2021). What is lacking in the Unity rockfall model is paradata. Is this a reasonable need that convergence may address?

Game studios spend a tremendous amount of money – tens of millions of dollars – developing games. In the case of fictional settings, they employ teams of writers and very large teams of artists. In the case of games set in a historical setting, they employ domain experts – historians, musicians with period expertise, scientists with domain expertise – in support of a more realistic game. This might involve many reports, drawings, files, ... in other words, all the elements we introduced as being present for rail sites. In many cases the experts will reference books and reports that are historical, and may need to be acquired and perhaps digitized. The analogy between the game design situation and the geotechnical situation is quite strong: building a historically accurate model of a building and building an accurate model of rail infrastructure involves exactly the same tools.

In *Assassins Creed Ancient Egypt*, Ubisoft has released a version of their game set in Egypt that includes tours, access to some of the reference material used in the construction of the game, framing material to provide context, and support materials. In other words, they have built the metadata and paradata (and contextualizing information) into the 3D environment itself. The game world is self-describing. This was originally used for in-house development, but was then released as an educational product. Work on paradata and annotation of game worlds is only going to deliver more powerful tools to document sites, and the use of game engines to develop educational packages provides a marketplace for that development (Sheldon, 2012).

With regards to field sites, integration of sensor data is becoming easier and data is becoming less expensive. Bula et al. (2020), for example, present a method to capture point clouds using a low-cost LiDAR system without expensive GNSS positioning or an IMU. While work by Guerin et al. (2019), highlight the use of thermal imaging combined with LiDAR scanning to define stability conditions of fractured rock masses.

Sensors mounted on vehicles, for examples on train cars, may represent a cost-effective way to update a world model. If a site engineer could stand in the field and visualize a year's worth of surface models for a slope, insight into the nature and seasonality of surface changes might be apparent in the field. Linked to

simulation tools, effect of possible events could be evaluated.

There is no reason that the infrastructure that enables this for a site would not be the foundation for corridor-scale data handling, processing, storage, analysis, and visualization. These processing methods, whether at the site or corridor scale, could include new approaches based on machine learning, feature recognition based on new ideas in deep learning, and automation to the point where change detection flags features of concern for inspection in a digital twin. This inspection might involve users at different sites in a multi-player mode, use augmented reality or virtual reality interfaces such as the HoloLens, and involve very high performance computing support for simulations and visualization.

Game engine technology is causing widespread convergence in media, civil engineering, and architecture. We believe tools based on this have great potential for geotechnical site investigation along rail corridors. We envisage a world where a geoengineer will stand at an outcrop while a drone updates their local model, chats with a co-worker about how to handle a difficult situation track-side, runs a simulation to constrain that situation, and annotates the shared site model collaboratively to keep a lasting record of what was done and why, all on a small tablet computer.

9. ACKNOWLEDGEMENTS

This work was funded by an SSHRC grant to the principal author addressing virtual environment construction and pedagogy, and by CN, CP, and NSERC Support through the Railway Ground Hazards Programme to D.J. H. Student work was also supported by NSERC and Queen's University Graduate Scholarships.

10. REFERENCES

- Atkins Undated. Digital Twins For Built Environment. Atkins / SNC-Lavalin White Paper, available online at: <https://www.snclavalin.com/~media/Files/S/SNC-Lavalin/download-centre/en/report/digital-twins-for-built-environment-report.pdf>.
- AT&T 2017 The complete and modern guide to technology convergence. blog article online: <https://developer.att.com/blog/technology-convergence>
- Bartle, R.A. 2004 Designing Virtual Worlds. New Riders, 741pp.
- Bollacker, K.D. 2010 Avoiding a Digital Dark Age. *American Scientist*, 98, 2, 106. <https://doi.org/10.1511/2010.83.106>.
- Bula, J., Derron, M.-H., & Mariethoz, G. 2020. Dense point cloud acquisition with a low-cost Velodyne VLP-16. *Geoscientific Instrumentation, Methods and Data Systems*, 9(2): 385–396. doi:10.5194/gi-9-385-2020.
- Crandall, B., Klein, G., & Hoffman, R.R. 2006 Working Minds: A Practitioner's Guide to Cognitive Task Analysis. MIT Press, 442pp.
- Eberly, D.H. 2004 Game Physics. Morgan Kaufmann, 777pp.
- Eberly, D.H. 2015 GPGPU Programming for Games and Science. CRC Press, 441pp.
- Flynt, J.P. & Vinson, B. 2005 Simulation and Event Modeling for Game Developers. Thompson Course Technology, 542pp.
- Guerin, A., Jaboyedoff, M., Collins, B.D., Derron, M.-H., Stock, G.M., Matasci, B., Boesiger, M., Lefevre, C., & Podladchikov, Y.Y. 2019. Detection of rock bridges by infrared thermal imaging and modeling. *Scientific Reports*, 9(1): 13138. doi:10.1038/s41598-019-49336-1.
- Hodgkinson, J.H. & Elmoultie, M. 2020 Cousins, Siblings, and Twins: A Review of the Geological Model's Place in the Digital Mine. *Resources*, 9, 24. <https://doi.org/10.3390/resources9030024>.
- Kerlow, I. 2008 The art of 3d Computer Animation and Effects, 4ed. John Wiley and Sons, 500pp.
- Klein, G.A. 1992 Using Knowledge Engineering to Preserve Corporate Memory. *The Psychology of Expertise*, pp. 170-187, Springer. DOI: 10.1007/978-1-4613-9733-5_10.
- Morin, G., Hassall, S., & Chandler, R. 2014 Case Study - The real life benefits of Geotechnical Building Information Modelling. *Information Technology in Geo-Engineering*, D.G. Toll et al., IOS Press. doi: 10.3233/978-1-61499-417-6-96.
- Ondercin, M. 2016 An Exploration of Rockfall Modelling Through Game Engines. MSc thesis, Queen's University at Kingston, 200pp.
- Pixar 2021. Introduction to USD: Universal Scene Description. Online Resource available online at: <https://graphics.pixar.com/usd/docs/index.html>.
- Polanyi, M. 1966 The Tacit Dimension. University of Chicago Press, 128pp.
- Sala, Z. 2018 Game-Engine Based Rockfall Modelling: Testing and Application of a New Rockfall Simulation Tool. MSc Thesis, Queen's University, 206pp.
- Sala, Z., Hutchinson, D.J., & Harrap, R. 2019 Simulation of fragmental rockfalls using terrestrial laser scans from rock slopes in south-central British Columbia, Canada. *Nat. Hazards Earth Syst. Sci*, 19, 2385-2404, <https://doi.org/10.5194/nhess-19-2385-2019>.
- Sheldon, L. 2012 The Multplayer Classroom: Designing Coursework as a Game. Cengage Learning, 284pp.

Turban, E. & Aronson, J.E. 2000 Decision Support Systems and Intelligent Systems. Prentice Hall, 867pp.

Unreal Build 2021 Unreal Build: Broadcast and Live Events, Wednesday June 16, 2021. Archived event recordings at: <https://www.unrealengine.com/en-US/events/unreal-build-broadcast-live-events>.

US Census Bureau 2021 What Is Paradata. Available online as web document at: <https://www.census.gov/topics/research/paradata/about.html>

Wikipedia 2021 Metadata. Online article, available at: <https://en.wikipedia.org/wiki/Metadata>.

Risk analysis of hazardous materials release in railway transportation

Hadiseh Ebrahimi¹, Fereshteh Sattari, Lianne Lefsrud
*Department of Chemical and Materials Engineering, School of Engineering Safety and Risk Management,
University of Alberta, Edmonton, Alberta T6G 1H9, Canada*

Renato Macciotta
*Department of Civil and Environmental Engineering, School of Engineering Safety and Risk Management,
University of Alberta, Edmonton, Alberta T6G 1H9, Canada*

ABSTRACT

Transportation of hazardous materials can cause severe damage to people in densely populated areas. The potential consequences of hazardous material release on people can be identified by risk assessment techniques. In this study, the meteorological conditions were evaluated to define the most dangerous meteorological condition. The ALOHA software is used to model the toxic threat zone and create a hazard map of the incident's location. Then, the local social vulnerability indicators are identified, and the ArcGIS software is used to create a vulnerability map. The risk map is created using the hazard map and the vulnerability map. This methodology is tested at a location in a small city in Canada and the risk map confirmed the importance of integrating population vulnerability in risk assessment. The results of the risk map can help risk managers to optimize their emergency response decisions in the case of hazardous materials release and to plan for proper actions to improve the quality of life for people living in higher-risk areas.

1 INTRODUCTION

A large volume of hazardous materials (hazmat) shipments is transported by the railway in Canada (Ebrahimi et al., 2021; Vaezi & Verma, 2017). In 2015, crude oil constituted 25% of hazmat shipments by railway in Canada, which was the highest amount of hazmat shipments (Shaun & Stéphane, 2016). Canada is the fourth largest manufacturer and third-largest exporter of crude oil in the world (CAAP, 2019). Transportation of crude oil is hazardous, and crude oil spillage into the environment can have toxic effects on humans, the environment, and animals (Frittelli et al., 2015). Severe consequences as a result of railway incidents transporting hazmat, even with low probability, may happen to people especially when the trains cross highly populated areas (Bersani et al., 2016; Landucci et al., 2017).

To control and reduce the consequences of railway incidents transporting hazmat, performing risk analyses and risk assessments are necessary (Torretta et al., 2017; Lefsrud et al., 2020). Risk is identified as the relationship between hazard and vulnerability, which shows the degree of influence in an economic, social, and environmental level for a particular location (Sanchez et al., 2018). Risk assessments can provide appropriate information to identify the areas that are at higher risk levels, develop and change land-use

planning, and assign adequate emergency planning (De Friend et al., 2008; Ramírez-Camacho et al., 2017).

In this study, as safety risk depends on the intensity and nature of the hazard and the demographic structure of people (Sanchez et al., 2018), a risk map is prepared by superimposing the hazard map and the vulnerability map.

2 MATERIALS AND METHODS

2.1 Modeling of hazard scenarios

Main track derailments and collisions are responsible for the most serious railway incidents in high-speed trains (Kawprasert & Barkan, 2010), and may result in tank puncturing and hazmat release (U. De Haag & Ale, 2005). In this study, the focus was therefore on main track derailments and collisions.

The meteorological conditions and types of hazmat release directly influence the dispersion of hazmat and creating risks (Ovidi, van der Vlies, Kuipers, & Landucci, 2020). The potential scenarios of hazmat release in main track derailments and collisions are different based on the type of the release (e.g. rupture, hole) and the meteorological conditions (wind speed, wind direction, air temperature, etc.) (Ramírez-Camacho et al., 2017). According to the Purple Book (U. De Haag & Ale, 2005) and Railway Association of Canada (2017), Loss of Contaminant (LOC) is classified for tank cars into two groups, including instantaneous release (rupture of the tank car) and continuous release (leakage from a 3" hole

¹ Corresponding author email address: hadiseh@ualberta.ca

in the tank car). Rupture of the tank cars might happen as a result of derailment of tank cars (TransportCanada, 1999) and puncture of tank cars usually occurs as a result of a hit from broken rails and/or equipment, such as car couplers (Abotula et al., 2019). As predicting the time and atmospheric conditions for railway incidents are hardly possible (Landucci et al., 2017; Sanchez et al., 2018), the most dangerous meteorological conditions based on what presented by Federal Emergency Management Agency (2018) and Miñarro (2004) is considered.

The most dangerous meteorological condition for hazmat dispersion happens in stability class F, which is very stable. Stable air means that the weather is tending to be calm, and it does not change quickly (Bubbico & Mazzarotta, 2008; McCormick, 2013). Dispersion of hazmat is worse during stable atmospheric conditions because the amount of vertical mixing is reduced and pollutants released in stable atmospheric conditions tend to spread horizontally, rather than vertically (McCormick, 2013). The atmosphere is stable under light winds and clear skies.

2.2 Human vulnerability mapping

A vulnerability map is developed using different social vulnerability indicators, such as average population age, level of education, etc.; which provide a means to estimate the population's ability to predict, resist, and recover from the hazard, and their inherent capacity of people to react to hazards (Li et al., 2010; Glade, 2003). Population vulnerability is extremely important for developing emergency response plans to prevent severe consequences of an incident (Tahmid et al., 2020). Population vulnerability is associated with the characteristics of society. Although various social vulnerability indicators are available in the literature (Enarson, 2007; Hewitt, 2014; Blaikie et al., 2014; Barro, 1992), there is not any clear definition of which indicators should be used in specific fields to study population vulnerability in natural or human-made hazards (Sanchez et al., 2018). In this study, the significant social vulnerability indicators affecting population vulnerability are defined based on a review of available literature and guidelines by the Government of Canada (Government of Canada, 2016) (Table 1).

A geoprocessing model is designed in ArcGIS software to calculate the vulnerability map. The model runs the calculations using the geospatial data for the social indicators develops a vulnerability map. The vulnerability map shows the location of the places where a large hazmat release could cause severe consequences. The ArcGIS software categorizes the train derailment location into different vulnerability levels, such as very low, low, medium, high, and very highly vulnerable locations.

Table 1. List of social vulnerability indicators acquired from literature (Cutter et al., 2000; Cutter, 1996; Füssel, 2007; Kates & White, 1978; Robert W Kates, 1985; Cutter et al., 2003; de Souza Porto & De Freitas, 1996; M F de Souza Porto & De Freitas, 2003; Lambert & Stock, 1993; Hewitt, 1997; Morrow, 1999; Blaikie et al., 2014; MacEachren et al., 2006) and the Government of Canada (Government of Canada, 2016).

Characteristics	Social indicators	Vulnerable Categories
Socioeconomic vulnerability measures	Income	Median total income for the population aged 15 years and older in private households
	Labour	Population aged 15 years and older who are unemployed
	Education	The population aged 15 years and older with no certificate, diploma, or degree
Population vulnerability measures	Children	Population aged 0-14
	Elderly	Population age 65 years and older
	Language	The population who does not know the official language (English or French)
	Female population	
	Population in 2016	
Housing vulnerability measures	Occupied private dwellings by dwelling condition	Major repairs needed
	Housing	Housing units with more than one person per room
	Private households by housing condition	Suitable or not suitable

2.3 Risk assessment and mapping

Risk can be calculated as a function of hazard and vulnerability (Li et al., 2010; Kron, 2005). In this study, only safety risks are considered (potential consequences of hazmat release on people). The social vulnerability indicators are combined with the hazard levels for each geospatial zone using a double-entry matrix (risk matrix in Table 1). This matrix is color-coded and is the central tool used for risk assessment. The risk matrix helps to

quantify the overall risk. In this work, the direct sum of the variables was selected because the risk matrix is rectangular (5 levels of vulnerability and 3 levels of threat). The risk matrix has been developed based on the work by (Tahmid et al., 2020; Federal Transit Administration, 2019; Leśniak & Janowiec, 2019). This matrix is meant to illustrate the development of the methodology presented in this paper; however, each organization needs to evaluate if adopting the matrix presented here adequately reflects the characteristics of their operations, their risk appetite, and the risk tolerance; or if a different matrix needs to be adopted. Levels of hazard and vulnerability are categorized into verbal and numerical scales. Hazard categories are high (3), medium (2), and low (1). These three levels correspond to the three output categories from the hazard analysis tool adopted in this study (ALPHA software), however, these could be expanded if other tools are adopted, and the matrix redefined to suit specific contexts and organizations. Vulnerability categories are very high (5), high (4), medium (3), low (2), and very low (1). These five levels correspond to the five output categories from the human vulnerability analysis in ArcGIS software.

Table 1. Risk matrix for risk assessment.

		Hazard level		
		High (3)	Medium (2)	Low (1)
Vulnerability level	Very high (5)	(8)	(7)	(6)
	High (4)	(7)	(6)	(5)
	Medium (3)	(6)	(5)	(4)
	Low (2)	(5)	(4)	(3)
	Very low (1)	(4)	(3)	(2)

In the risk matrix, the higher the risk score, the greater the overall risk for the hazmat release location. This method helps visualize relative risks between areas and therefore can focus efforts for emergency response.

3 RESULTS AND DISCUSSION

3.1 Hazard modeling

An example application of the methodology is presented for a small Canadian city. To make the example relevant to Canadian operations, the details are inspired by a rail incident that has been investigated by the Transportation Safety Board of Canada. The scenario considers the potential train derailment of loaded crude oil tank cars.

Two types of crude oil release because of tank car derailment are considered: instantaneous release (rupture of the tank car) and continuous release (leakage from a 3" hole in the tank car). The characteristics of Bakken oil, a light, low-density, very flammable crude oil, are considered to create the threat zones using the ALPHA software. Petroleum crude oil (UN1267) is a Class 3 flammable liquid with a wide range of flammability and volatility characteristics. Different types of crude oil, including heavy crude oil, synthetic crude oil,

light and medium crude oil, and bitumen are produced and transported in Canada. Bakken oil is selected as this was the type of hazmat involved in the tragic events associated with the Lac-Mégantic derailment (Transport Canada, 2013) in Canada. The flashpoint (the lowest temperature at which ignition can occur) is the lowest for Bakken oil compared to other types of crude oil (the flashpoint is less than 23 °C), which means that Bakken oil is highly flammable and more prone to explosion. Severe railway derailments, transporting Bakken oil have occurred in Alabama, North Dakota, and Quebec (Frittelli et al., 2014). Table 2 shows the results of the threat zone for the most dangerous meteorological conditions (stability class F). In the selected case study, the warm seasons last from April to early September, with an average daily temperature of 18°C. The cold season lasts from October to March, with an average daily temperature of -1 °C (Government of Canada, 2021b). Therefore, the hazard maps are modeled in two seasons (winter and summer), and the variation is analyzed in the affected location.

Table 3. The results of the threat zone for each final scenario acquired from EVT. ALPHA shows threat zones on a composite threat zone plot. Threat zones are differentiated with different colors. The red threat zone demonstrates the highest hazard, and the orange and yellow threat zones demonstrate areas of decreasing hazard. Red, orange, and yellow correspond to hazard categories High, medium, and low, respectively, in the risk matrix used in this study.

Stability class	F (winter)	F (winter)	F (summer)	F (summer)
Type of release	3 inches hole	Rupture	3 inches hole	Rupture
Toxic dispersion (Yards) 1Yard=0.9144 Meter	R:57 O:91 Y:163	R:234 O:379 Y:625	R:364 O:629 Y:1044	R:717 O:1289 Y:1.2 miles
Flash fire (Yards)	R:122 Y:366	R:486 Y:1118	R:815 Y:1miles	R:1684 Y:1.9miles
VCE (Yards)	O:92 Y:127	O:359 Y:1467	O:581 Y:714	O:1368 Y:1527
Pool/Jet fire (Yards)	R:18 O:27 Y:144	R:62 O:90 Y:143	R:42 O:63 Y:100	R:129 O:195 Y:313
BLEVE/Fire ball (Yards)	R:651 O:920 Y:143	R:651 O:920 Y:1436	R:593 O:837 Y:1304	R:593 O:837 Y:1304

3.2 Human vulnerability and risk map

The social vulnerability indicators were coded into the ArcGIS software to find a vulnerability map based on the census information for the area. Figure 1 shows the vulnerability map prepared using the human vulnerability indicators.

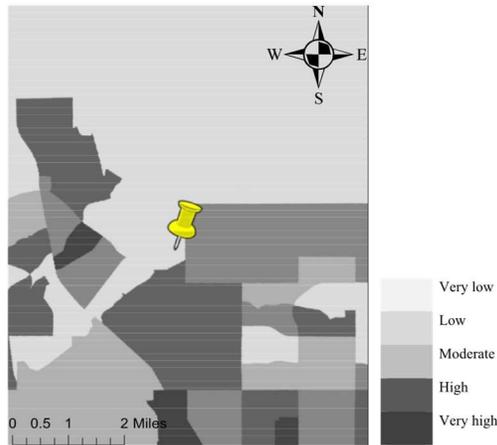


Figure 1. The vulnerability map created by the ArcGIS using social vulnerability indicators (yellow pin shows the source of hazmat release).

The hazard map of the most dangerous meteorological conditions was imported into the ArcGIS software. Figure 2 indicates the hazard map of the most dangerous meteorological condition (stability class F) in eight wind directions.

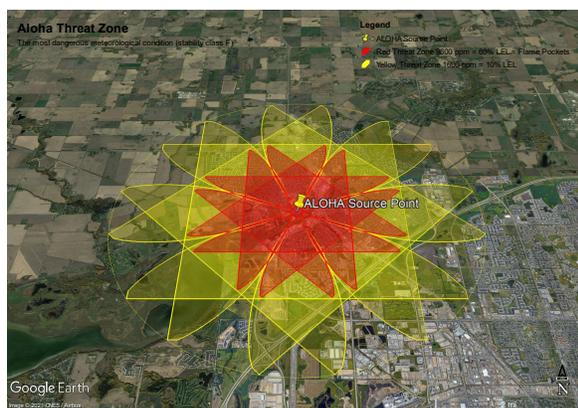


Figure 2. The hazard map of the most dangerous meteorological condition created by implementing the threat zone of the ALOHA in Google Earth (stability class F), (GoogleEarth, 2021).

Figure 3. shows the risk map created by combining the hazard map of the most dangerous meteorological condition and the vulnerability map in the ArcGIS software.

3.3 Discussion

In this study, threat zones for the most dangerous meteorological conditions are overlaid on the vulnerability map to create risk maps and to show a picture of the locations where people are affected by the consequences of hazmat release. The toxic clouds are moved by the wind and disperse in the direction of the wind; the vulnerable people are the ones living in the direction of the wind. The direction of the wind changes

over time, and it is necessary to evaluate the vulnerable population around different directions of the hazmat source. In this study, the risk maps are prepared based on all the wind directions. This risk map can help risk managers to prioritize and optimize their actions to save people's lives in the hazmat release location. For instance, children and older adults are not quickly capable of escaping the hazmat release location. They are vulnerable in case of evacuating the release location, and risk managers should know that they need more help and assistance. Also, the people living in crowded houses might not be able to evacuate their places quickly. Overcrowding can cause problems for them, and special assistance should be considered for these vulnerable people. It would be very challenging to relocate the alignment of the railway network to less populated areas, and therefore risk mitigation strategies should aim at reducing the likelihood of rails incidents and reducing the potential consequences of hazmat release. Emergency response strategies based on the characteristic of the people and meteorological conditions of the location, therefore, play an important role to reduce the risks of hazmat railway transportation in Canada.

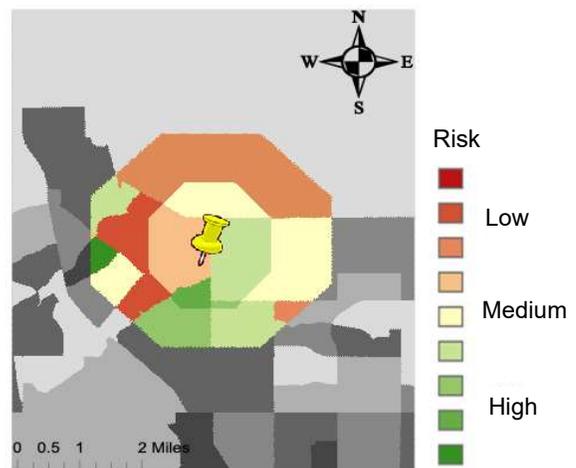


Figure 3. The risk map of the most dangerous meteorological condition created by ArcGIS (stability class F, yellow pin shows the source of hazmat release).

4 CONCLUSION

Incidents involving rail transportation of hazmat can pose a threat to people living in the areas close to the railway track. Risk analysis and risk assessment are necessary to inform decision-making to reduce the consequences of hazmat release on people. In this study, the hazard map for the most dangerous meteorological condition is simulated using the ALOHA software. The Social vulnerability indicators, which have great importance in the risk analysis of hazmat release, are identified to simulate the vulnerability map in the ArcGIS software. By superimposing the hazard map and the vulnerability map

in the ArcGIS software, risk map is prepared for the most dangerous meteorological condition that show the areas from higher risk to lower risk, following a double-entry risk matrix. This risk map can provide accurate information for risk managers to plan for land-use planning in the future, to prioritize their actions based on the vulnerability of people in emergency situations, and to improve the characteristics of people living in higher risk areas, such as education, quality of housing, wellness, etc.

5 REFERENCES

Abotula, S., Konda, B. R., Walsh, W. J., Challa, V. S. A., Ros, T., & Manohar, M. (2019). Analysis of Puncture Resistant Steel for Tank Cars. *Steel Research International*, 90(4), 1800339. <https://doi.org/https://doi.org/10.1002/srin.201800339>

Barro, R. J. (1992). Xavier Sala-i-martin. *Economic Growth*, 2.

Bersani, C., Papa, F., Sacile, R., Sallak, M., & Terribile, S. (2016). Towards dynamic exposure-based schedule for hazardous material trains. *Journal of Rail Transport Planning & Management*, 6(2), 116–127. <https://doi.org/https://doi.org/10.1016/j.jrtpm.2016.05.003>.

Blaikie, P., Cannon, T., Davis, I., & Wisner, B. (2014). *At risk: natural hazards, people's vulnerability and disasters*. Routledge.

Bubbico, R., & Mazzarotta, B. (2008). Accidental release of toxic chemicals: Influence of the main input parameters on consequence calculation. *Journal of Hazardous Materials*, 151(2–3), 394–406.

CAAP. (2019). *Crude oil Forecast, Markets and Transportation*.

Cutter, S. L. (1996). Vulnerability to environmental hazards. *Progress in Human Geography*, 20(4), 529–539.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly*, 84(2), 242–261.

Cutter, S. L., Mitchell, J. T., & Scott, M. S. (2000). Revealing the vulnerability of people and places: A case study of Georgetown County, South Carolina. *Annals of the Association of American Geographers*, 90(4), 713–737.

De Haag, P. A. M. U., & Ale, B. J. M. (2005). *Guidelines for quantitative risk assessment: purple book*. Ministerie van Volkshuisvesting en Ruimtelijke Ordening (VROM).

de Souza Porto, M F, & De Freitas, C. M. (2003). Vulnerability and industrial hazards in industrializing countries: an integrative approach. *Futures*, 35(7), 717–736.

de Souza Porto, Marcelo Firpo, & De Freitas, C. M. (1996). Major chemical accidents in industrializing countries: the socio-political amplification of risk. *Risk Analysis*, 16(1), 19–29.

DeFriend, S., Dejmek, M., Porter, L., Deshotels, B., & Natvig, B. (2008). A risk-based approach to flammable gas detector spacing. *Journal of Hazardous Materials*, 159(1), 142–151. <https://doi.org/https://doi.org/10.1016/j.jhazmat.2007.07.123>.

Ebrahimi, H., Sattari, F., Lefsrud, L., & Macciotta, R. (2021). Analysis of train derailments and collisions to identify leading causes of loss incidents in rail transport of dangerous goods in Canada. *Journal of Loss Prevention in the Process Industries*, 72(April). <https://doi.org/10.1016/j.jlp.2021.104517>.

Enarson, E. (2007). *Identifying and Addressing Social Vulnerabilities*. Emergency Management: Principles and Practice for Local Government. WL Waugh and K. Tierney. ICMA Press.

Federal Emergency Management Agency). (2018). *Threat and Hazard Identification and Risk Assessment (THIRA) and Stakeholder Preparedness Review (SPR) Guide*.

Federal Transit Administration. (2019). *Sample Safety Risk Assessment Matrices for Rail Transit Agencies*.

Frittelli, J., Andrews, A., Parfomak, P. W., Pirog, R., Ramseur, J. L., & Ratner, M. (2014). *U.S. Rail Transportation of Crude Oil: Background and Issues for Congress*. Congressional Research Service. Retrieved from <https://fas.org/sgp/crs/misc/R43390.pdf>.

Frittelli, J., Andrews, A., Parfomak, P. W., Pirog, R., Ramseur, J. L., & Ratner, M. (2015). *U.S. rail transportation of crude oil: Background and issues for Congress*. *Crude Oil Properties and the Hazards of Transport: Background, Data and Literature Summary*, 97–131.

Füssel, H.-M. (2007). Vulnerability: A generally applicable conceptual framework for climate change research. *Global Environmental Change*, 17(2), 155–167.

Glade, T. (2003). Vulnerability assessment in landslide risk analysis. *Erde*, 134(2), 123–146.

GoogleEarth. (2021). Image © 2021 CNES/Airbus. GovernmentofCanada. (2016). <https://open.canada.ca/data/en/dataset/9a91d63b-a41f-4d8d-8e54-1e066700d666>.

GovernmentofCanada. (2021). https://weather.gc.ca/past_conditions/index_e.html?station=psr.

- Hewitt, K. (1997). *Regions at risk: a geographical introduction to hazards*. Harlow, Essex, UK: Addison Wesley Longman Limited.
- Hewitt, Kenneth. (2014). *Regions of risk: A geographical introduction to disasters*. Routledge.
- Kates, Robert W. (1985). The interaction of climate and society. *Climate Impact Assessment*, 3, 3–36.
- Kates, Robert William, & White, G. F. (1978). *The environment as hazard*. New York: Oxford University Press.
- Kawprasert, A., & Barkan, C. P. L. (2010). Effect of Train Speed on Risk Analysis of Transporting Hazardous Materials by Rail. *Transportation Research Record*, 2159(1), 59–68. <https://doi.org/10.3141/2159-08>.
- Kron, W. (2005). Flood risk= hazard• values• vulnerability. *Water International*, 30(1), 58–68.
- Lambert, D. M., & Stock, J. R. (1993). *Strategic logistics management (Vol. 3)*. Irwin Homewood, IL.
- Landucci, G., Antonioni, G., Tugnoli, A., Bonvicini, S., Molag, M., & Cozzani, V. (2017). HazMat transportation risk assessment: A revisit in the perspective of the Viareggio LPG accident. *Journal of Loss Prevention in the Process Industries*, 49, 36–46.
- Lefsrud, L., Macciotta, R., & Nkoro, A. (2020). Performance-based regulations for safety management systems in the canadian railway industry: An analytical discussion. *Canadian Journal of Civil Engineering*, 47(3), 248–256. <https://doi.org/10.1139/cjce-2018-0513>
- Leśniak, A., & Janowiec, F. (2019). Risk assessment of additional works in railway construction investments using the Bayes network. *Sustainability (Switzerland)*, 11(19). <https://doi.org/10.3390/su11195388>
- Li, F., Bi, J., Huang, L., Qu, C., Yang, J., & Bu, Q. (2010). Mapping human vulnerability to chemical accidents in the vicinity of chemical industry parks. *Journal of Hazardous Materials*, 179(1), 500–506. <https://doi.org/https://doi.org/10.1016/j.jhazmat.2010.03.031>
- MacEachren, A. M., Pike, W., Yu, C., Brewer, I., Gahegan, M., Weaver, S. D., & Yarnal, B. (2006). Building a geocollaboratory: Supporting Human–Environment Regional Observatory (HERO) collaborative science activities. *Computers, Environment and Urban Systems*, 30(2), 201–225. <https://doi.org/https://doi.org/10.1016/j.compenvurbsys.2005.10.005>
- McCormick, R. A. (2013). Air pollution climatology. *Air Pollution (Edited by AC Stern)*, 1, 275–321.
- Miñarro, M. D. (2004). *Zonas de planificación para accidentes graves de tipo tóxico: GUÍA TÉCNICA (EN EL ÁMBITO DEL REAL DECRETO 1254/99-SEVESO II)*. EDITUM.
- Morrow, B. H. (1999). Identifying and mapping community vulnerability. *Disasters*, 23(1), 1–18.
- Ovidi, F., van der Vlies, V., Kuipers, S., & Landucci, G. (2020). HazMat transportation safety assessment: Analysis of a “Viareggio-like” incident in the Netherlands. *Journal of Loss Prevention in the Process Industries*, 63, 103985.
- Railway Association of Canada. (2017). *RAILWAY EMERGENCY AWARENESS GUIDE*.
- Ramírez-Camacho, J. G., Carbone, F., Pastor, E., Bubbico, R., & Casal, J. (2017). Assessing the consequences of pipeline accidents to support land-use planning. *Safety Science*, 97, 34–42. <https://doi.org/https://doi.org/10.1016/j.ssci.2016.01.021>
- Sanchez, E. Y., Represa, S., Mellado, D., Balbi, K. B., Acquesta, A. D., Lerner, J. E. C., & Porta, A. A. (2018). Risk analysis of technological hazards: simulation of scenarios and application of a local vulnerability index. *Journal of Hazardous Materials*, 352, 101–110.
- Shaun, S., & Stéphane, G. (2016). *Transportation of Dangerous Goods NEWSLETTER 2016_Vol.36 No 1*.
- Tahmid, M., Dey, S., & Syeda, S. R. (2020). Mapping human vulnerability and risk due to chemical accidents. *Journal of Loss Prevention in the Process Industries*, 68, 104289. <https://doi.org/https://doi.org/10.1016/j.jlp.2020.104289>
- Torretta, V., Rada, E. C., Schiavon, M., & Viotti, P. (2017). Decision support systems for assessing risks involved in transporting hazardous materials: A review. *Safety Science*, 92, 1–9. <https://doi.org/https://doi.org/10.1016/j.ssci.2016.09.008>
- TransportCanada. (1999). *RAILWAY INVESTIGATION REPORT R99T0256*.
- TransportCanada. (2013). *R13D0054 Runaway and Main-Track Derailment Montreal , Maine & Atlantic Railway Freight Train Mma-002 Lac-Mégantic , Quebec*.
- Vaezi, A., & Verma, M. (2017). An analytics approach to dis-aggregate national freight data to estimate hazmat traffic on rail-links and at rail-yards in Canada. *Journal of Rail Transport Planning & Management*, 7(4), 291–307. <https://doi.org/https://doi.org/10.1016/j.jrtpm.2017.12.001>

Prediction of emergency response strategies based on combustion signatures from FTIR spectroscopy using machine learning techniques

Sijie Tian, Chi-Guhn Lee
University of Toronto, Toronto, Ontario, Canada

Dexen D.Z. Xi, Yoon Ko, Nour Elsagan
National Research Council of Canada, Ottawa, Ontario, Canada

ABSTRACT

In the emergency response to freight transportation fire incidents, first responders routinely confront unknown hazards from incidents (e.g., flammable, explosive, toxic, corrosive). The Emergency Response Guidebook (ERG), promoted by Transport Canada, helps first responders to identify hazards based on the materials involved in a transportation incident, and provides guidance for protecting the surrounding during the initial response. Yet, in a scenario where the materials involved are mixed or unknown and when the expert knowledge is absent among the responders, relevant studies suggest that artificial intelligence (AI) tools have the potential to support decision-making. The National Research Council of Canada (NRC) and the University of Toronto are working to develop an AI tool for assisting first responders in identifying hazards and determining the most effective response strategy for freight fires. In this work, we propose a machine learning algorithm using combustion signatures of chemical components, measured from the gaseous effluents of the burning material over time using a Fourier Transform Infrared Spectroscopy (FTIR). We demonstrate the feasibility of this model by training the algorithm on a small set of laboratory samples provided by the NRC to predict the correct emergency response guide in the ERG. The combustion signatures are shown to be uniquely related to the corresponding hazard and are linked to the emergency response strategy.

1 INTRODUCTION

In rail transport, safety is of the utmost importance. If freight catches on fire, it is crucial to identify the type of the fire and scope of the associated hazards in an accurate and timely manner to protect people and the environment from the fire. In the emergency response to freight transportation fire incidents, the first responders are trained to leverage the Emergency Response Guidebook (ERG), which allows them to identify response strategies based on the hazardous materials involved in the transportation incident, and to follow the appropriate responses in the guidebook to mitigate the freight fire incident. However, first responders routinely confront unknown hazards from freight fire incidents such as flammable, explosive, toxic, and corrosive materials. Such cases occur when the goods transported are unlabeled or the materials burning are unknown, and based on the ERG, the first responders are instructed to use the blanket guide: "Guide 111 -

Mixed Load / Unidentified Cargo" which may be too general. As a result, there is a need to aid the first responders in the emergency response to such goods being transported.

The Fourier Transform Infrared Spectroscopy (FTIR) analyzer is a device capable of detecting an array of chemical components from the effluents of the burning material over time. The spectral data obtained from an FTIR produces unique fingerprints of the sample materials (Titus et al., 2019), making the FTIR well-suited to collect data for machine learning (ML) models where patterns and structures are expected to exist within the data. As an early work, Chen et al. (2000) used data from an FTIR analyzer capable of detecting the concentrations of 18 gas species [BD(1)] to train an artificial neural network (ANN) for classifying the corresponding fires as either flaming, smoldering, or nuisance. The training data were composed of repeated tests on eight different materials or chemical compounds, and the validation data came from fire and non-fire cases of materials not observed in the training

set. The trained three-layer ANN correctly predicted 96% of the 248 test cases, demonstrating the feasibility of using FTIR data to detect fires. To the best of our knowledge, there has been few other studies utilizing machine learning algorithms to make classification decisions with the FTIR on combustion processes.

Although ML has yet to be adopted as a candidate method in the analyses of combustion processes, it has been gradually introduced into and increasingly used in the field of fire sciences, especially in fire detection. In a comprehensive review of chemical-based indoor fire detection systems and associated algorithms, Fanollosa et al. (2018) identified that chemical gas sensors can improve fire sensitivity and early detection, but they suffer from a high rate of false alarms. The study suggested pattern recognition algorithms as the only path to improve false alarm immunity and cited numerous works that utilized ML techniques such as ANNs, K-nearest-neighbors (KNN) and decision trees to decrease the chances of false alarms. However, in their discussion of gas sensors for combustion products, the authors only discussed the use of chemical sensor components, namely electrochemical cells, metal oxide sensors (MOX), and non-dispersive infrared cells (NDIR) for fire detection, and explicitly excluded the FTIR analyzer from this discussion without providing rationale for this choice. This suggests a knowledge gap in the understanding of how FTIR analyzers compare to chemical sensor components in fire detection. These chemical gas sensors are similar to FTIR analyzers since they are both used to detect and to quantify target gases, and both take measurements over time. However, the FTIR analyzer has the advantage of being able to detect and collect the signatures of a variety of gases simultaneously, while typically an array of chemical sensor components needs to be manually selected and assembled to measure the target gases. Nevertheless, both tools produce multivariate, time series data. As a result, machine learning models that have been demonstrated to work well for gas sensor data should also work well for FTIR data. In the field of fire detection using images, Park et al. (2019) proposed a fire detection system incorporating a deep neural network for time series sensor data analysis as well as a convolutional neural network (CNN), which is a type of deep neural network used to efficiently analyze temporal or visual data, for fire detection in images among other multi-functional components. Overall, machine learning techniques have become increasingly recognized in the area of fire detection, using either time series data or image data.

In a broader context, ML techniques have been widely adopted to the classification of gases using data collected from gas sensors such as the electronic nose (E-nose). In a review of smart gas sensing technologies, Feng et al. (2019) outlined and compared various smart gas sensor arrays, signal processing methods and gas pattern recognition algorithms including support vector machines (SVM), KNN and ANNs. They highlighted that machine learning now plays a key role not only in making gas classification decisions but also in its ability to adapt for sensor drifts, which alleviates the need to

manually manipulate the signals captured by the gas sensors. In recent years, as neural networks proved to be universal function approximators, gas classification also began to adopt neural networks to replace traditional approaches. For gas classification that uses time series as input data, an intuitive model to use is a CNN, which employs convolutional kernels to reduce the number of trainable parameters in the model and also to detect the same feature at different time points within a time series. Peng et al. (2018) proposed a deep-CNN model, GasNet, with up to 38 layers to classify four types of gases using data from eight MOS sensors, and this model significantly outperformed a trained SVM and ANN. Zhao et al. (2019) proposed a one-dimensional deep CNN, 1D-DCNN, to classify three pure gases and two binary mixture gases. The 1D-DCNN model was able to automatically extract important features from the data through convolutional kernels, and it significantly outperformed a SVM, a KNN, an ANN and a random forest as benchmarks. Another approach is to apply existing deep CNN architectures such as VGG[BD(2)] (a deep CNN architecture named after its creator, the Visual Geometry Group) and Residual Network (ResNet) to gas classification as demonstrated in Han et al. (2019), however such deep architectures would require a large amount of training data, which is often difficult to obtain in gas classification. Since the field of gas classification gradually favored CNN architectures to capture the temporal information effectively, it is natural to approach the combustion signature classification task using combustion data measured by an FTIR via CNNs as well.

Another research area for consideration is the research on the time series classification task itself, which focuses on developing algorithms leveraging the temporal information, and much work has gone into algorithmic development and comparison. A recent paper by Ruiz et al. (2021) compared state-of-the-art multivariate time series classification (MTSC) methods on 26 equal-length time series datasets and highlighted four models: Canonical Interval Forests (CIF), the Hierarchical Vote Collective of Transformation-based Ensembles (HIVE-COTE [BD(3)]), the Random Convolutional Kernel Transform (ROCKET)[BD(4)], and InceptionTime as significantly outperforming the previous time series classification benchmark named dynamic time warping (DTW). In particular, ROCKET, developed by Dempster et al. (2020), was referred as the state-of-the-art method and the recommended starting point for benchmarking in future research on time series classification tasks.

In this work, our objective is to obtain the hazard characteristics of fires using the effluents collected with a FTIR analyzer, which we call combustion signatures, from fires conducted in a bench-scale test set-up, cone calorimeter, by training ML models with the combustion signature data. The combustion signatures used in this study are collected in a laboratory environment; however, our end goal is to augment the dataset, refine the trained model, and eventually apply the model to real freight fire scenarios. For the machine learning models, we use CIF and ROCKET as suggested by Ruiz et al.

(2021), and we will evaluate the classification performance and the time required to train each model. This work a first step taken to understand the fire hazards, which will allow us to later link them to the emergency response strategies in the ERG where possible and provide more specific instructions to the first responders when the burning materials are unknown.

Our main contributions are summarized below:

- We review the applications of artificial intelligence and machine learning in the field of fire sciences and more broadly, classification of gases, to demonstrate that machine learning has been gradually adopted in these fields.
- We apply the classification of six hazard categories on the combustion analysis dataset using existing top-performing multivariate time series classification methods.
- We demonstrate empirically the potential of using machine learning in fire sciences and provide our results as a first step towards allowing first responders to identify fire hazards in unlabeled, burning freights

2 MATERIALS AND METHODS

In this section, we will describe the experiment setup used to collect the combustion signature data, the classification method used, as well as setup for training the machine learning models.

2.1 Experiment Setup

Experiments were conducted using a Cone Calorimeter coupled with a FTIR gas analyser to analyze the emissions. The cone calorimeter is used to detect the fire behaviour of a fuel sample under ambient conditions, where a material of specific size is subjected to a pre-specified heat flux, and heat release and mass loss rates are measured. The heat flux was set to 50 kW/m² in all tests used for this work.

MKS Multigas 2030 continuous FTIR gas analyzer was connected to the exhaust of the cone calorimeter to analyze the gaseous effluents. The gas analyzer was equipped with a 200 mL gas cell heated to 191°C and 5.110 m optical path with a liquid nitrogen cooled detector.

Before sampling, the gas analyzer was flushed with pure dry nitrogen gas (grade 5.0) at a rate of 3L/min for at least 1 hour before a background spectrum of nitrogen was recorded. Sampling was performed by switching the sampling port to a 5' heated sampling line (191°C). The heated sampling line was connected to a port sampling from the ventilation system. The gas ran through a stainless steel filter body, equipped with a bonded microfiber filter element. The sample gas was pulled into the FTIR at 3L/min via a MKS 2380 heated pump (191°C) at atmospheric pressure. The heated pump was fitted with two particulate stainless steel body filters with bonded microfiber/PTFE filter elements to prevent particulate matter from reaching the gas cell. A 5' heated line (191°C) was also used to connect the pump to the

gas analyzer. The gas velocity was calculated to be 20.7 ft/sec. The expected lag between sampling and detection was therefore around 1 second.

Determination of the chemical concentrations present in the effluents was performed using the included MKS software (MG2000, Analysis Validation Utility and Gas Search Utility). Quantitative calibrations were performed by MKS and installed within the MG2000 software.

To confirm the validity of the analysis performed by MG2000, the analysis validation utility (AVU) was used on specific spectra with elevated concentration of targeted compounds. This was to determine an estimate on the detection limits, confidence limits, maximum bias and values specified in ASTM D6348 and EPA 320. The AVU was used strictly to confirm the analysis and check spectral residuals remaining. The values recorded by the AVU were not reported and used as a check. The Gas Search Utility was also used to help predict gases that may have been present and were subsequently added to the recipe.

As seen in Table 1, most materials were only tested once, four materials were tested twice, and three materials were tested three times. Overall, we obtained 30 samples from 20 distinct materials.

Table 1. List of materials tested, and the number of experiments conducted on each material.

	Materials with one experiment	Materials with two experiments	Materials with three experiments
Materials Tested	Blue CAT-6 wiring, carpet flooring, consumer electronics shell, electrical wiring, epoxy, heptane, intumescent caulking, melamine board, PMMA, polyethylene, polyurethanes foam, tar shingle, wall.	Crude oil, diesel, phenolic panel, polystyrene.	ABS, polyisocyanurate, PVC.
Number of materials	13	4	3

2.2 Classification Methods

2.2.1 Canonical Interval Forest

The Canonical Interval Forest is an interval-based forest of trees ensemble method proposed by Middlehurst et al. (2020). For each tree, the model randomly samples an interval from the time series data, computes a set of summary statistics using the selected interval, and concatenates the summary statistics into one vector. A decision tree, which makes predictions by recursively splitting the data on different attributes using a tree structure, is then built using the concatenated feature vector. This process is repeated to obtain r decision trees, resulting in a forest of trees. A classification decision will be made based on the majority vote of the

forest of trees. The detailed process to build the CIF model is shown in Algorithm 1.

Algorithm 1 buildCIF(A list of n cases of length m with d dimensions, $\mathbf{T} = (\mathbf{X}, \mathbf{y})$)

Parameters: the number of trees, r , the number of intervals per tree, k , and the number of attributes subsampled per tree, a (default $r = 500$, $k = \sqrt{d} \cdot \sqrt{m}$, and $a = 8$)

```

1: Let  $\mathbf{F} = (\mathbf{F}_1 \dots \mathbf{F}_r)$  be the trees in the forest
2: for  $i \leftarrow 1$  to  $r$  do
3:   Let  $\mathbf{S}$  be a list of  $n$  cases  $(s_1 \dots s_n)$  with  $a \cdot k$  attributes
4:   Let  $\mathbf{U}$  be a list of  $a$  randomly selected attribute indices  $(u_1 \dots u_a)$ 
5:   for  $j \leftarrow 1$  to  $k$  do
6:      $b = \text{rand}(1, m - 3)$ 
7:      $l = \text{rand}(b + 3, m)$ 
8:      $o = \text{rand}(1, d)$ 
9:     for  $t \leftarrow 1$  to  $n$  do
10:      for  $c \leftarrow 1$  to  $a$  do
11:        if  $u_c \leq 22$  then
12:           $s_{t,a(j-1)+c} = c22\text{Feature}(u_c, \mathbf{X}_{t,o}, b, l)$ 
13:        else
14:           $s_{t,a(j-1)+c} = \text{tsfFeature}(u_c, \mathbf{X}_{t,o}, b, l)$ 
15:       $F_i.\text{buildTimeSeriesTree}([S, y])$ 

```

Algorithm 1. The model build procedure for CIF (Middlehurst et al., 2020).

2.2.2 The Random Convolutional Kernel Transform (ROCKET)

ROCKET, developed by Dempster et al. (2020), leverages a large number of random convolution kernels to learn feature vectors and uses a linear classifier to make the classification decisions. Unlike CNN models, ROCKET is not a deep learning model and does not learn the weights of the convolutional kernels. Instead, it randomly initializes 10,000 convolutional kernels of various parameters such as length and weight of the kernel and apply every kernel to each instance of the data. This random initialization step makes the model extremely fast to run because no training stage is required to identify important patterns from the time series data.

After passing the inputs through the random convolutional kernels, the results, which we call the feature maps, will be used to compute two summary statistics. One statistic is the maximum value and the other is the proportion of positive values within each feature map. As a result, from the 10,000 kernels, we obtain 10,000 feature maps and subsequently 20,000 summary statistics. These summary statistics are then concatenated into a vector of length 20,000 and used to train a linear classifier such as a ridge regression or a logistic regression classifier. In Ruiz et al. (2020), the authors found that ROCKET achieved state-of-the-art accuracy even with small datasets, whether in the number of training examples or the length of the time series, in addition, this model was the fastest classifier among the classifiers tested. As a result, this model is a

very good candidate for our small combustion analysis dataset.

2.3 Training Setup

The classification goal in this study is to answer six binary questions for each sample of burning material, namely: is this sample flammable; toxic; explosive; water extinguishable; corrosive and oxidizing. Each hazard category is independent of each other and as a result, we train one machine learning model per category and present their results separately. Since each material in the combustion signature dataset may take different lengths of time to complete burning, we perform zero-padding on the dataset which means that all experiments have been aligned to take the length of the longest experiment by adding zeros to the end of shorter experiments. Using the combustion signature dataset with 30 data samples, we perform a 70-30 train-test split, meaning 70% of the data (or 21 data samples) will be used for training the models, and 30% of the data (or 9 data samples) will be used to test the generalization properties of the trained models.

For the machine learning models tested, which are CIF and ROCKET, we use the implementations from the Sktime library (Markus et al. 2021), which is a Python library for time series analysis. For each model, we use the default hyperparameter values used in the original papers, for example, we use 500 trees for CIF, and 10,000 random convolutional kernels for ROCKET.

For each experiment, we record the training accuracy, test accuracy and more importantly, the time taken to train and test the model. As concluded by Ruiz et al. (2021), we expect the classification accuracies of the three models to be similar while taking drastically different times to train.

3 RESULTS AND DISCUSSION

Table 2 shows the classification accuracies of ROCKET and CIF, and the time requirements (in seconds) to train each model. The rows show results from each hazard category as well as the average result across all hazard categories. For each row, the superior test accuracy and training time are bolded.

Overall, we observe that all models trained achieved training accuracies of 100% and test accuracies that are less optimal. This indicates overfitting, which means that all models have memorized the training dataset too well which prevented it from generalizing to unseen data points. There exist measures to alleviate the problem of overfitting, including increasing the size of the overall dataset through either data collection or through data augmentation techniques that generate synthetic data through manipulating the existing dataset, or regularizing the model to reduce the complexity of the model. Since the combustion analysis dataset is very small, the ideal solution to alleviate overfitting would be to increase the dataset through both additional data collection and synthetic data generation. We leave the exploration of such tasks as the next step in our future work.

From the test accuracies in individual hazard categories, we observe that ROCKET and CIF tied in three categories: explosivity, water reactivity, and oxidizing; ROCKET outperformed CIF in one category: flammability; and CIF outperformed ROCKET in the remaining two categories: toxicity and corrosiveness. As a result, CIF achieved a slightly higher average test accuracy across all hazard categories. However, looking at the training time requirements, we observe that ROCKET is roughly 1,000 times faster than CIF for all hazard categories, taking 3-4 seconds to complete training, while CIF takes about one hour to train.

Overall, we believe that ROCKET is the clear winner, since its test accuracies are comparable to those of CIF, and it is also 1,000 times faster to run. Considering that the current results are obtained from a preliminary 30-sample dataset, we expect the performance of both models to be improved when a larger dataset is available. In addition, CIF would naturally take longer to train with a larger dataset, making ROCKET even more favorable. These results are analogous to those presented by Ruiz et al. (2021), and we once again demonstrate the superiority and speed of ROCKET due to its random convolutional kernels.

Hazard Category	ROCKET (SOTA)			Canonical Interval Forest		
	Train Accur acy (%)	Test Accur acy (%)	Time (s)	Train Accur acy (%)	Test Accur acy (%)	Time (s)
flammability	100	67	3.99	100	56	4036.81
toxicity	100	56	3.89	100	67	4017.65
explosivity	100	89	3.91	100	89	4016.17
water reactivities	100	67	3.97	100	67	4026.90
Corrosiveness	100	78	3.91	100	89	3114.11
Oxidizing	100	100	4.33	100	100	3047.02
Average Performance	100	76	4.00	100	80	3709.78

Table 2. Results on the combustion analysis dataset from a 70-30 data split.

4 CONCLUSIONS

In this work, we presented a review of applications of machine learning in fire sciences and gas classification. We presented the combustion signature dataset collected using the FTIR and cone calorimeter and used this dataset to demonstrate that ML can be applied to the characterization of freight fire hazards using two existing implementations of multivariate time series classification methods. However, in this study we have been limited by the small scale of the dataset, with only 30 data samples available. All data were obtained in the laboratory, indicating a need to test the models developed using real-world data before adoption. We

have also been limited by a lack of repetition in experiments used. Previously, published classification tasks in related fields such as fire detection or gas classification typically only collected data on a few distinct materials and would run tests on these limited materials for a large number of times, for example, 1000 times per material. However, in our case, we only tested 20 distinct materials and collected a total of 30 data samples, meaning that most materials were only burned once, and no repetitions of tests are available. Since this is a first attempt at using machine learning models on fire hazard detection, using such a complicated and small dataset limited the performance of the machine learning models tested. As a result, the most imminent work is to collect more data in the laboratory and also explore data augmentation methods. As a future step, we will also develop our own multivariate time series classification method for the combustion analysis dataset and to contribute to the research community. These tasks will aid us in eventually applying our designed ML model trained on an augmented, large-scale dataset, to real freight fire scenarios and provide decision support to first responders when they encounter unlabeled, burning goods.

5 ACKNOWLEDGEMENT

This project was supported by collaborative research funding from the National Research Council of Canada's Artificial Intelligence for Logistics Program. This project was supported by the Vector Scholarship in Artificial Intelligence, provided through the Vector Institute.

REFERENCES

- Chen, Y., Serio, M. & Sathyamoorthy, S. (2000), 'Development of a fire detection system using ft-ir spectroscopy and artificial neural networks', *Fire Safety Science* 6, 791–802.
- Dempster, A., Petitjean, F. & Webb, G. I. (2019), 'ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels', *CoRR* abs/1910.13051. URL: <http://arxiv.org/abs/1910.13051>.
- Feng, S., Farha, F., Li, Q., Wan, Y., Xu, Y., Zhang, T. & Ning, H. (2019), 'Review on smart gas sensing technology', *Sensors* 19(17). URL: <https://www.mdpi.com/1424-8220/19/17/3760>.
- Fonollosa, J., Sol'orzano, A. & Marco, S. (2018), 'Chemical sensor systems and associated algorithms for fire detection: A review', *Sensors* 18(2). URL: <https://www.mdpi.com/1424-8220/18/2/553>.
- Han, L., Yu, C., Xiao, K. & Zhao, X. (2019), 'A new method of mixed gas identification based on a convolutional neural network for time series classification', *Sensors* 19(9). URL: <https://www.mdpi.com/1424-8220/19/9/1960>.
- Markus Löning, Tony Bagnall, Matthew Middlehurst, Sajaysurya Ganesh, George Oastler, Jason Lines, ViktorKaz, Franz Király, Martin Walter, Patrick

Rockenschaub, Jesellier, Taiwo Owoseni, Lovkush, AidenRushbrooke, oleskiewicz, Yi-Xuan Xu, Patrick Schäfer, RNKuhns, Hongyi, ... Ayushmaan Seth. (2021). alan-turing-institute/sktime: v0.7.0 (v0.7.0). Zenodo. <https://doi.org/10.5281/zenodo.5093663>.

Middlehurst M, Large J, Bagnall A (2020). The canonical interval forest (CIF) classifier for time series classification. In: Proceedings of the IEEE international conference on big data.

Park, J. H., Lee, S., Yun, S., Kim, H. & Kim, W.-T. (2019), 'Dependable fire detection system with multifunctional artificial intelligence framework', *Sensors* 19(9). URL: <https://www.mdpi.com/1424-8220/19/9/2025>.

Peng, P., Zhao, X., Pan, X. & Ye, W. (2018), 'Gas classification using deep convolutional neural networks', *Sensors* 18(1). URL: <https://www.mdpi.com/1424-8220/18/1/157>.

Ruiz, A. P., Flynn, M., Large, J., Middlehurst, M. & Bagnall, A. (2021), 'The great multivariate time series classification bake off: a review and experimental evaluation of recent algorithmic advances', *Data Mining and Knowledge Discovery* 35, 401–449.

Titus, Deena, E. James Jebaseelan Samuel, Selvaraj Mohana Roopan (2019). Chapter 12 nanoparticle characterization techniques. Ashutosh Kumar Shukla, Siavash Irvani, eds., *Green Synthesis, Characterization and Applications of Nanoparticles. Micro and Nano Technologies*, Elsevier, 303–319. doi:<https://doi.org/10.1016/B978-0-08-102579-6.00012-5>. URL <https://www.sciencedirect.com/science/article/pii/B9780081025796000125>.

Zhao, X., Wen, Z., Pan, X., Ye, W. & Bermak, A. (2019), 'Mixture gases classification based on multi-label one-dimensional deep convolutional neural network', *IEEE Access* 7, 12630–12637.

The use of terrestrial laser scanning for the development of slope process models

David A. Bonneau, D. Jean Hutchinson, Paul-Mark DiFrancesco and Rob Harrap
Queen's University – Department of Geological Sciences and Geological Engineering, Kingston, Ontario, Canada

ABSTRACT

Idealized conceptual models play a key role in the development and advancement of the geosciences. These models are used for a variety of purposes, including prediction and explanation. They are used for scientific research as well as for planning purposes, policy, and hazard mitigation. Datasets are becoming more available at higher resolution as data collection and processing systems are being improved and, increasingly, automated. With an increase in the spatiotemporal frequency of data acquisition, there is now an unprecedented ability to investigate proposed models of geomorphic processes using data from remote sensing systems and to directly compare them to laboratory simulations and more local, detailed natural environment studies. Terrestrial laser scanning (TLS) has been demonstrated to be a key tool in geomorphic monitoring around the world. TLS systems can achieve a high spatial measurement resolution and can be used for long-range analysis of geomorphic processes. Furthermore, these scans record the geometry of the slope at a unique point in time allowing comparison with scans taken at later dates. In this work, we draw on experience from monitoring four study sites in the Thompson-Fraser rail corridor in Interior British Columbia as part of the RGHRP program. We demonstrate how long-term TLS monitoring can be used to develop slope process models of failure mechanisms in postglacial river terraces, and for debris channel accumulation and erosion dynamics. These slope process models can be used to inform how the slope is expected to behave under current climatic scenarios and establish a base-line level of activity. Furthermore, the development of geological slope process models facilitates the investigation of landform development and identification of the active geomorphic processes operating on the slope. Each of these elements contribute to the advancement of the practices of natural slope mitigation planning, risk reduction and ongoing control.

1 INTRODUCTION

Models are central to development of geosciences and the success of geological engineering projects (Fookes 1997). They are used not only for basic scientific research (i.e. theoretical tools for advancing insight and understanding) but also for planning purposes, policy, and hazard mitigation (Bokulich and Oreskes 2017). Therefore, it is imperative to understand the uncertainties present, the risk implications of these uncertainties and ultimately strive for sound understanding of underlying failure mechanisms (Carter and Marinos 2020). Practitioners in the geosciences commonly make interpretations on sparse or incomplete datasets when constructing models.

Geoscientists generally work with sparse and incomplete datasets but work by Bond et al. (2007, 2015) highlights that conceptual uncertainty is inherent to data interpretation and associated model development. Furthermore, they raise significant questions about discussions between experts that are based solely on models. These challenges give rise to difficulties in

model development. Nevertheless, the use of remote sensing and recent advances in state-of-the-art 3D data capture, processing and analytical techniques have expanded the understanding of geomorphic processes operating on natural slopes (Lato 2020). Across the geosciences, researchers are taking advantage of the high degree of spatial sampling to investigate earth surface processes and systems. In some cases the acquisition of remote-sensed data is part of routine fieldwork; for example, the use of drones during geological mapping is fast becoming commonplace.

Remote sensing tools and techniques are increasingly being used for the characterization and management of geohazards. Over the last couple of decades there have been significant advances in the use of terrestrial, airborne and spaceborne remote sensing techniques for Earth surface observations, including rock slope characterization and monitoring (Telling et al. 2017). Terrestrial laser scanning (TLS) has been demonstrated to be a key tool in geomorphic monitoring around the world (Abellán et al. 2014). TLS systems can achieve a high spatial measurement resolution and can

be used for long-range analysis of geomorphic processes when the spatial variability of TLS errors are accounted for (Fey and Wichmann 2017). Advances in data collection systems have demonstrated the ability to capture small scale processes, such as detecting mm-scale rock slope deformation, (e.g. Kromer et al. (2015)), to assessing regional scale rockfall activity (e.g. Benjamin et al. (2020)) in great detail. The ability to integrate such data with in-situ sensors, optical data, video data, seismic sensors offer an unprecedented to better characterize geohazards and propose mitigation measures.

Recent works aim to investigate proposed models of geomorphic processes (e.g. Rapp (1960), van Steijn et al. (1995), Fookes (1997), Krautblatter and Dikau (2007)) with remote sensing systems have benefited from a level of detail unachievable with previous approaches. Hancock et al. (2020) examined snow cornice dynamics with comparison of TLS scans to evaluate existing conceptual models of cornice growth and decay in Svalbard. Through the monitoring period they were able to build on the observational understanding and conceptual model of seasonal cornice dynamics (Vogel et al. 2012). Kenner (2020) used TLS scans to investigate the sediment budgets of geomorphic processes operating on an alpine talus slope in Eastern Switzerland. Each of these studies highlight how remote sensing can supplement the understanding of geomorphic processes, from model development to model verification, and how the dramatic increase in spatial resolution of new sensors enables investigations of processes that previous sensors could not accurately delineate.

In this work, the authors demonstrate how long-term TLS monitoring can be used to develop slope process models of failure mechanisms in postglacial river terraces and debris channels in southern British Columbia (BC), Canada.

1.1 Study Sites

Two study slopes are used within this work. Both sites are located within the southern Thompson River Valley in Interior British Columbia, Canada (Figure 1). Each of these study sites have been monitored with remote sensing technologies as part of the Canadian Railway Ground Hazard Research Program (RGHRP) starting in April 2012.

1.1.1 Sandy Cobble Hill

Sandy Cobble Hill (SCH) (50.287594° , -121.405445°) is located approximately 275 km northeast of Vancouver, British Columbia, Canada. The 20 m tall cliff face consists of a sequence of well stratified sand and gravel layers with marked cobble and boulder horizons (Figure 2). The cliff face is gullied, producing near vertical fin-like structures in the cliff face. The benches and ledges within the cliff face correspond to cobble and boulder horizons. Below the cliff face is a 120 m long talus slope. At the base of the slope is a series of ditches protecting the CN

rail line from debris originating from the slope above. The average angle of the talus slope ranges from approximately 34 degrees at the bottom to 36 degrees at the top. The talus slope has a distinct texture, which results from the sorting of the particle sizes on the slope. In general, sands and gravels are located proximal to the cliff face and coarser grained debris is located further downslope.

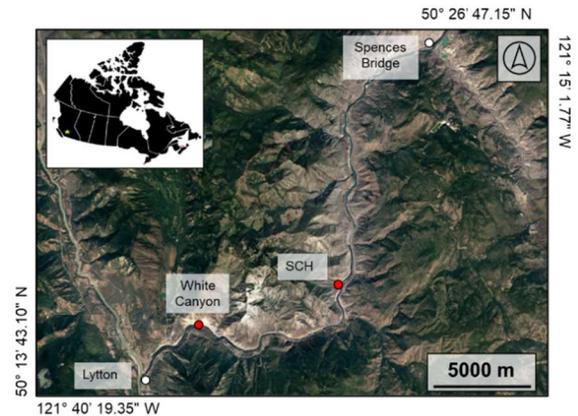


Figure 1. Location of the study slopes in the southern Thompson River Valley in Interior British Columbia, Canada.

The cliff face consists of fluvial gravels referred to as the “Nicoamen gravels”. Both horizontal and crossbedding are displayed within these gravel sequences. At these horizons, the structures indicate that the flow was in the same direction as the modern Thompson River. These gravel units are deposited between upper till and a lower silt unit and were thought to have been deposited during the Olympia interglaciation (36,000 yrs. B.P. to 24,500 yrs. B.P.) (Ryder, 1981). An upper till unit is located above the Nicoamen gravels. This till unit consists of poorly sorted and non-sorted gravels and is interpreted as advance outwash and other drift material from the Fraser Glaciation (Ryder, 1981).

1.1.2 White Canyon

The steep slopes of the White Canyon (WC) (50.266261° , -121.538943°), located 5 km northeast of the community of Lytton, BC, near the confluence of the Thompson and Fraser Rivers, present geohazards to the safe operation of the Canadian National (CN) mainline (Figure 3). Rockfalls and rockslides contribute to the production of debris which accumulates in the steep channels. Dry granular flows and debris flows facilitate the transport of debris downslope, which can result in consequences that range from minimal maintenance and repair of warning systems, to complete closure and rebuilding of the impacted rail lines.

Differential erosion of the WC has formed a morphology that is highly complex and consists of vertical spires and deeply incised channels. The Canyon spans approximately 2.2 km between Mile 093.1 and 094.6 of the CN Ashcroft subdivision. The active portion

of the Canyon reaches up to 500 m in height above the railway track. Two short tunnels mark the entrances to the Canyon. A third portal is located in the middle of the Canyon which separates White Canyon West (WCW) (Figure 3) and White Canyon East (WCE).

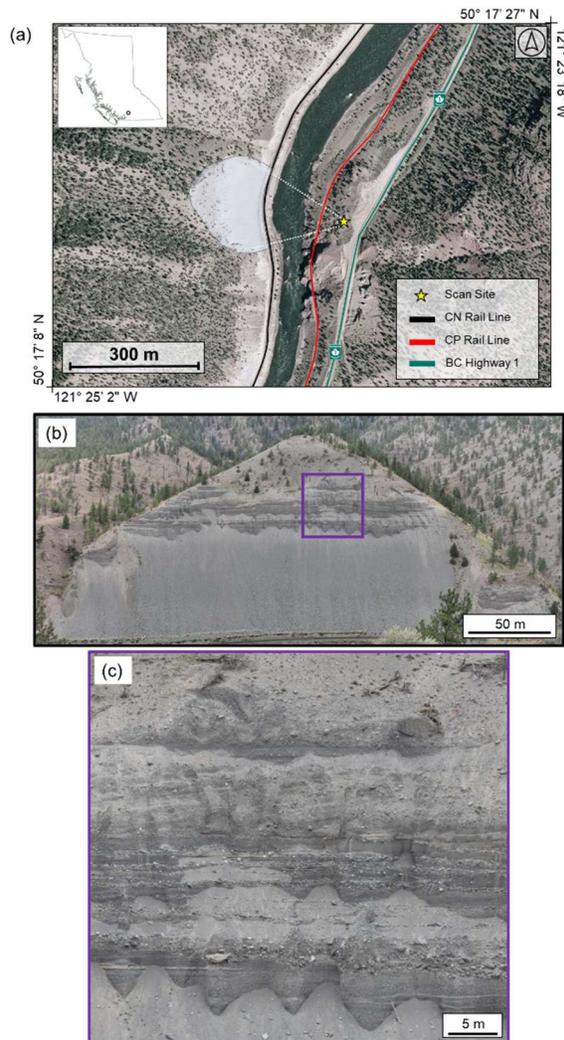


Figure 2. The Sandy Cobble Hill study slope. (a) Location of the study slope and TLS scanning position. The scan extents and area are displayed in the light blue colour. (b) Panoramic photograph of the SCH study slope and surrounding area looking east at the slope from the scan site. Photographs captured on September 25th, 2019. Note scale is not consistent in each of the images. (c) Central portion of the study slope. Note the cross bedding and marked gravel and cobble boulder horizons.

The dominant geological unit in the WC is the Lytton Gneiss. The Lytton Gneiss is composed of a quartzofeldspathic gneiss with amphibolite bands, containing massive quartzite, gabbroic and amphibolite intrusions. Two sets of dykes have intruded the Lytton Gneiss. The first dyke set consists of tonalitic intrusions that are believed to be related to the emplacement of the

Mt. Lytton Plutonic Complex (Brown, 1981). The second dyke set is a series of dioritic intrusions that crosscut the Lytton Gneiss and tonalitic dykes. The dioritic intrusions are believed to be part of the Kingsville Andesites (Brown, 1981). On the western portion of the White Canyon is an outcropping of the Mt. Lytton Plutonic Complex (Brown 1981, Greig 1989). The Mt. Lytton Plutonic Complex in the White Canyon is a distinctly red stained unit which is composed of granodiorite with local diorite and gabbro. The red staining of the rock mass is thought to be a direct result of fluids originating from the weathering of hematite in overlying mid-Cretaceous continental clastic rocks. Sedimentary units primarily composed of sandstone and conglomerate beds unconformably overlie the Mt. Lytton Plutonic Complex in WCW and above the Lytton Gneiss in two areas in WCE. In WCW, the outcrop is gently folded and displays well graded bedding and load structures. The folding is thought to be a result of the local faulting (Brown, 1981).

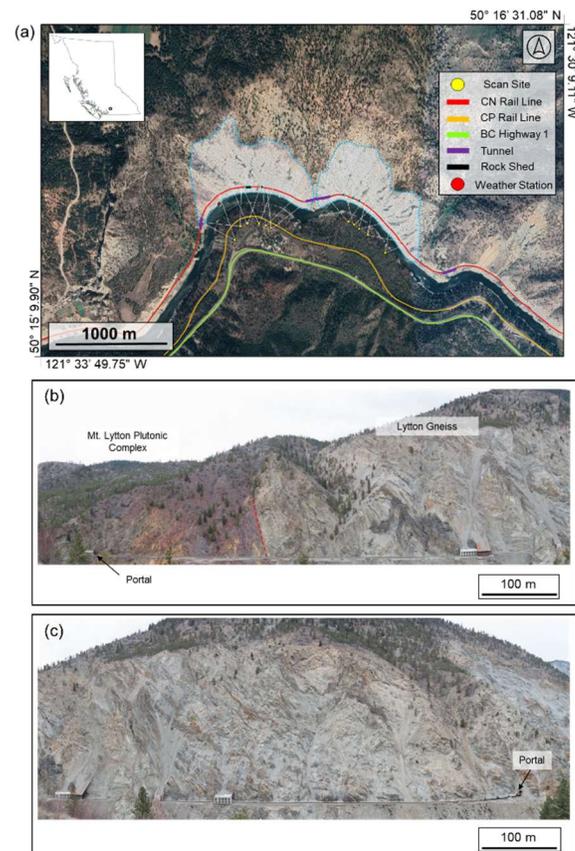


Figure 3. The White Canyon study slope. (a) Location of the study slope and TLS scanning positions. The scan extents and area are displayed in the light blue colour for each site. (b) Panoramic photographs of White Canyon West. Note that the scale is not consistent throughout each of the images. This photo shows the Mt. Lytton Plutonic Complex and the fault contact (red dashed line) with the Lytton Gneiss. (c) This photo shows the central portion of WCW, with the portal to WCE visible near the bottom right corner of the image.

2 MATERIALS AND METHODS

2.1 Remote Sensing Data Capture

Two time-of-flight TLS scanning systems were used to capture high-resolution point cloud datasets of the study slopes. An Optech Illris 3D-ER and a Riegl VZ-400i system were used throughout the study period.

Data capture at each of the study slopes can be broken into three main phases. The initial (low resolution) scans were first captured with the Optech system between April 2013 and August 2014. Between August 2014 to September 2017, high-resolution Optech scans were captured at the study sites. In September 2017, data capture transitioned to using the Riegl system. TLS scans were captured at frequencies that ranged from daily to quarterly between April 2013 and March 2020. In this work, 52 scans captured at WCW and 50 scans captured at SCH are used.

Six scan positions were used to capture a slope model of WCW. The point spacing for data capture was approximately 10 cm for the WCW scans.

At SCH, a single scan position was used to capture TLS point clouds. Point spacing varied between approximately 6 to 10 cm.

High-resolution panoramic images were captured from each TLS scan position. A Gigapan robotic head, mounted on a tripod, was used in conjunction with a DLSR camera to capture the photographs used to generate the panoramic images.

2.2 Point Cloud Processing and Change Detection

After each of the point clouds was captured, the dataset was parsed using the associated parsing software for each TLS system. After the datasets were parsed, vegetation, slope mitigation (i.e., draped mesh) and erroneous points were manually removed.

To align the datasets, all point clouds, both Riegl and Optech, were imported into RiScan Pro for registration. Point clouds were aligned to common baselines based on the epoch in which they captured. The general workflow for the registration process involved a manual point picking of common or shared points between the pairs of time sequential clouds. Following the manual point picking, an iterative closest point (ICP) (Besl and McKay 1992) matching algorithm was applied to refine the alignment.

After the registration process was completed, MultiScale Model to Model Cloud Comparison (M3C2) (Lague et al. 2013) was used for all distance computations between the point cloud TLS data sets. M3C2 is a vector-based distance computation that operates directly on point clouds. The M3C2 implementation in CloudCompare (Girardeau-Montaut 2018) was used for this study with the process outlined by DiFrancesco et al. (2020). In brief, the CloudCompare API was accessed via a custom Python script. The M3C2 distance computations were then performed in batch on all of the TLS datasets.

2.3 Volume Calculations

Volumetric estimates were calculated in two ways in this work. The first method, which was only used in select instances with complex 3D geometry, is the Alpha Solid methodology outlined by Bonneau et al. (2019a) and DiFrancesco et al. (2021). In short, the method reconstructs a manifold, triangulated surface mesh of the input point cloud. The volume is then calculated through the divergence theorem using the triangulated mesh. Readers are referred to the references for further details on the procedure.

All other volume calculations were completed in CloudCompare. To compute the volume in CloudCompare, a raster is generated in a user defined projection direction. After this process is completed, the contributions of each cell are summed together. The contribution is the volume corresponding to the cell footprint multiplied by the difference in heights between sequential 3D models. To calculate the volumes, areas of change were first segmented out using the segmentation tool. The area of change was translated to align with one of the principal axes of the scene orientation. This ensured that the projection direction would correspond to the direction of change, to minimize the potential for over or underestimation of the volume (e.g. Benjamin et al. (2016)).

2.4 Debris Monitoring Methodology

Bonneau et al. (2019b) discuss the debris monitoring methodology used at the White Canyon site. Figure 4 displays a visual representation of the methodology developed. At Time 1 (T1), a preliminary TLS scan of the channel is completed. The volume of debris in the channel, and the depth to the bedrock base of the channel, is at this point unknown. Preliminary estimates of the volume of channel material can be made utilizing approaches developed by Jakob et al. (2005). Locations in the channel where debris is accumulating can also be documented from visual inspection of the panoramic imagery. Subsequently at T2, a debris movement has occurred and scoured the channel to bedrock at select locations along the channel length. With an additional TLS scan and panoramic imagery, the channel bed and geometry at this time can be captured. Locations of exposed bedrock along the channel length are first confirmed with the panoramic imagery and these locations within the TLS scans are stored to generate a bedrock baseline model of the channel. The areas of exposed bedrock within the channel serve as the baseline for subsequent monitoring. It may take several debris flow movements to expose the bedrock along the full length of the channel, in which case the baseline elevation may be the result of a compilation of measurements at different times. As time progresses (T3), the channel bed begins to recharge with debris from rockfall and rockslides. Debris from accumulations on benches moves into the main channel. When a scan taken at T3 is compared to the bedrock baseline model (T2), volume estimates demonstrate the spatial and temporal location of debris accumulating in the channel.

These estimates are all confirmed with visual inspection of the panoramic imagery. Finally, a debris flow occurs at T4. Comparing a TLS scan captured after the debris flow to the bedrock baseline (T2), providing bedrock has been observed in the location previously, permits the calculation of the degree of entrainment and bedrock incision.

It should be noted that not all areas in the channels exposed bedrock over the course of the monitoring period. As a result, to supplement the bedrock baseline models, a lowest elevation cell approach is used to fill in the baseline models. The Open3D Python library (Zhou et al. 2018) is used to grid the point clouds of each of the channel watersheds. Once a grid has been established, the lowest points within each grid are extracted, resulting in a lowest elevation point cloud for the entire channel catchment. A cell size of 10 cm was used for the grid. After these points were extracted, they were filtered by a verticality computation using the eigenvectors from a principal component analysis. This process is used to segment out near vertical outcrop from the remaining channel areas. A statistical outlier filter is also then applied to the point cloud. A merged point cloud of the manually mapped bedrock baselines and the lowest elevation point cloud is used to assess all in-channel stored debris.

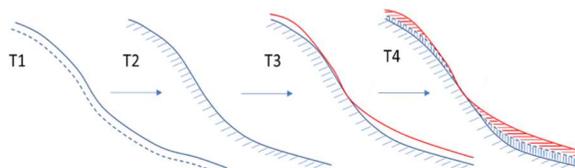


Figure 4. Overview of the proposed recharge monitoring methodology. T1 – baseline scan of channel. T2 – debris flow has occurred and scoured channel to bedrock (idealized). T3 – channel begins to recharge from rockfall and rockslides. T4 – debris flow occurs. Using baseline, the degree of debris and bedrock incision can be calculated. Modified from (Bonneau et al. 2019b)

This process was used to analyze catchments on White Canyon West (WCW) (Figure 5).

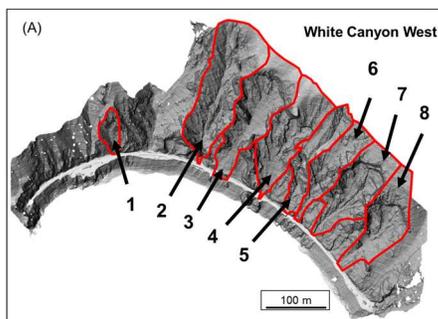


Figure 5. White Canyon West catchment delineations.

A visual example of the outputs from the processing described, in terms of in-channel stored volume, is displayed in Figure 6.

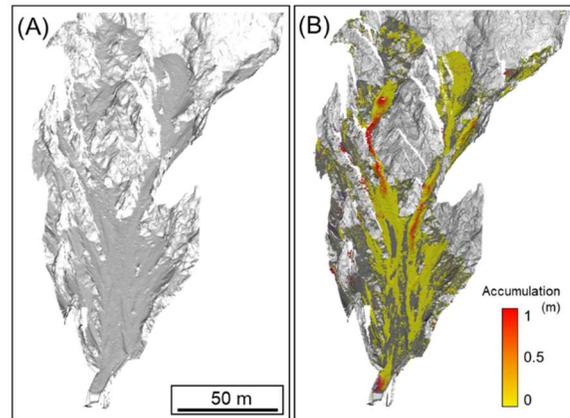


Figure 6. Visual example of the in-channel stored debris volume estimate using the March 2nd, 2020 WCW Channel 4 scan. (a) overview of the channel; (b) distance computation between the bedrock baseline model and the March 2nd, 2020 scan. Colour scale denotes distance from the bedrock baseline model.

The output of this process provides an estimate of the spatial-temporal accumulation of debris. This process can only be conducted on areas within the field of view of the scanner.

2.5 White Canyon Digital Rockfall Database

In this work, the author used a digital rockfall database derived from sequential TLS scans captured at the White Canyon (DiFrancesco et al. (This Conference)). In general, the process involves computing change detection both forward and backward for the two datasets. The change detection is then filtered based on the limit of detection. The limit of detection is defined as two-times the standard deviation of the alignment error. A threshold is applied to the change detection results and the results are filtered and clustered using a 3D clustering algorithm (i.e. DBSCAN – Ester et al. (1996)). Each identified cluster is then meshed to calculate volume (Bonneau et al. (2019a), DiFrancesco et al. (2021)). Each of these changes is assigned a classification (i.e. rockfall or debris movement) using the approach presented in (Bonneau et al. 2019c). This approach uses a majority voting processes where the identified change points search the mask point cloud, and a class is assigned. Further details on the process used in this work can be found in DiFrancesco et al. (2021).

3 RESULTS AND DISCUSSION

3.1 Sandy Cobble Hill

Fifty TLS scans were processed and change detection analyses using M3C2 were conducted between each of the sequential scans. All erosion events captured with the M3C2 change detection algorithm were confirmed by reviewing change that could be observed on the

panoramic photographs. The change detection results were then interpreted to understand and characterize the underlying mechanisms potentially contributing to the cliff face evolution. One example is described below.

Bonneau and Hutchinson (2019) previously reported on the progressive failure in the lower cobble-boulder horizon at SCH. They were able to document the undercutting and erosion of the finer grained material, which destabilized the cobble-boulders. Once these cobble-boulders detach from the horizon, the material being buttressed above also fails and there is retrogression in the cliff face. Retrogression occurs until another cobble-boulder is exposed in the horizon. When this occurs, material can begin to accumulate until the cobble-boulder eventually becomes unstable and the process repeats. **Error! Reference source not found.**7 highlights this interpreted sequence of events and demonstrates that it occurs multiple times at the same horizon and also at different locations on the cliff face. Examples of this failure sequence are displayed in the associated sections. In these sections, over approximately 2 m of slope parallel retreat occurred in these locations over the 7 year period. In addition, the sections highlight that the material being added to the talus slope begins to bury the cliff face. In the example sections presented in **Error! Reference source not found.**, the cliff face is buried by approximately 1.5 m.

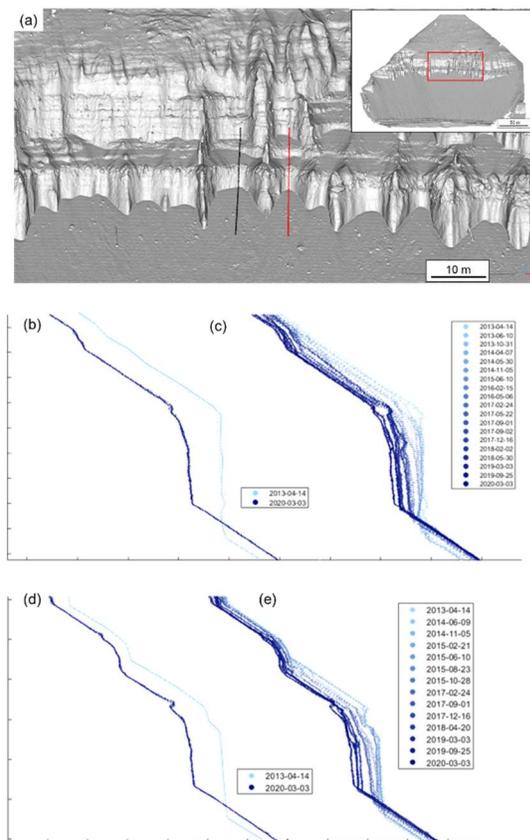


Figure 7: Sections through the cobble-boulder horizon in the central portion of the cliff face at SCH; (a) location of the sections with respect to the fin. The red box displays the location of the fin with respect to

the entire slope. The vertical sections: (b) and (c) correspond to the black line and (d) and (e) correspond to the red line.

This sequence highlights the need to monitor hillslopes for extended periods of time. In many situations, monitoring periods only last the length of graduate programs. Therefore, estimates of erosion rates from these shorter monitoring periods are unlikely to reflect long-term retreat, and certainly not over geological time scales (Krautblatter et al. 2012). Furthermore, observations made during periods of heightened geomorphic activity will result in overestimated erosion rates (Ballantyne and Benn 1994). Hartmeyer et al., (2020) notes that interpretation of erosion rates requires careful consideration of the geological setting and data collection rate.

3.2 White Canyon West

The in-channel stored sediment for each catchment in the WC was estimated for each TLS scan date. The estimated in-channel stored debris (black) and rockfall volume detected (blue) in the catchment is displayed for three catchments in WCW in Figure 8.

Comparison of the estimates of in-channel stored debris and rockfall volume at the White Canyon reveals that many of the channels are being depleted during the time frame monitored (WCW – Channel 2). This could be the result of limited larger magnitude rockfall/rockslide events occurring within the observation period. The largest event that was observed over the course of the monitoring period was an approximately 2600 m³ rock fall event that occurred in June 2013, very near the elevation of the track. This event occurred near the start of the monitoring period and an event of similar magnitude has not been observed since.

Cycles of recharge and evacuation of debris were observed during the monitoring period. Channels on White Canyon West experienced large fluxes of debris into the channel which were then evacuated at a later time, ranging from weeks to years. Understanding the timing of these cycles is of key importance in the development of models describing debris transfer processes. Temporal censoring of debris movements can occur if multiple debris movements occur between sequential scans: the movement of debris can overwrite evidence of previous activity.

4 CONCLUSIONS

The use of remote sensing can assist in the development of slope process models. More formally, a slope process model relies on understanding the mechanisms which actively contribute to the evolution of the slope. The possible failure mechanisms that are operating can therefore be assessed and an appropriate mitigation and remediation strategy can be implemented (Fookes 1997). The combination of long-term monitoring and co-development of a slope process model provides a

greater understanding of how the slope can evolve and behave over time, supporting work within a geotechnical asset management (GAM) framework (Thompson et al. 2016). The cause-and-effect impact of mitigation installation can be better optimized to achieve a maximum unit risk reduction per dollar spent.

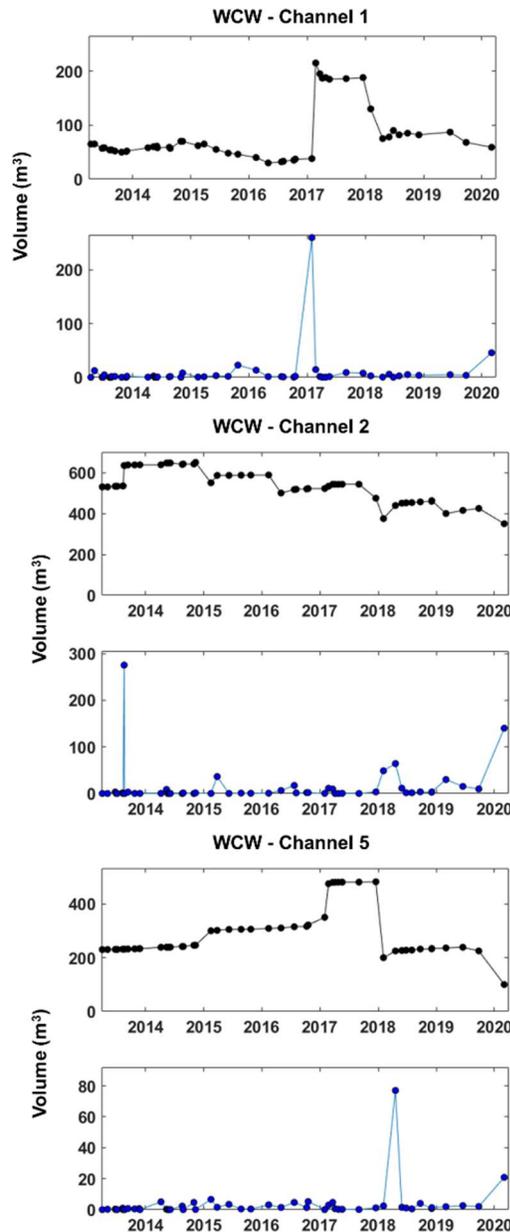


Figure 8: Overview of the estimated in-channel stored sediment for three of the analyzed channels on WCW. The graph pairs correspond to the volume of estimated in-channel stored sediment for each TLS scan (black) and the volume of rockfall volume added to the channel during the period elapsed between scans (blue).

5 ACKNOWLEDGEMENTS

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development Grant held by D. Jean Hutchinson (CRDPJ 470162-2014), in collaboration with the University of Alberta, and supported by CN and CP Rail. The research group was also supported by the NSERC Discovery Grant (RGPIN: 05668-2019) held by D. Jean Hutchinson. Support was also provided to David A. Bonneau and Paul-Mark DiFrancesco through NSERC's Graduate Scholarship Program.

REFERENCES

- Abellán, A., Oppikofer, T., Jaboyedoff, M., Rosser, N.J., Lim, M., and Lato, M.J. 2014. Terrestrial laser scanning of rock slope instabilities. *Earth Surface Processes and Landforms*, **39**(1): 80–97. doi:10.1002/esp.3493.
- Ballantyne, C.K., and Benn, D.I. 1994. Paraglacial Slope Adjustment and Resedimentation Following Recent Glacier Retreat, Fabergstolsdalen, Norway. *Arctic and Alpine Research*, **26**(3): 255. doi:10.2307/1551938.
- Benjamin, J., Rosser, N., and Brain, M. 2016. Rockfall detection and volumetric characterisation using LiDAR. *Landslides and Engineered Slopes. Experience, Theory and Practice*, (July): 389–395. doi:10.1201/b21520-38.
- Benjamin, J., Rosser, N.J., and Brain, M.J. 2020. Emergent characteristics of rockfall inventories captured at a regional scale. *Earth Surface Processes and Landforms*, **45**(12): 2773–2787. doi:10.1002/esp.4929.
- Besl, P.J., and McKay, N.D. 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**(2): 239–256. doi:10.1109/34.121791.
- Bokulich, A., and Oreskes, N. 2017. Models in the Geosciences. *In Springer Handbook of Model-Based Science*, 1st edition. Edited by L. Magnani and T. Bertolotti. Springer. p. 1260.
- Bond, C.E., Gibbs, A.D., Shipton, Z.K., and Jones, S. 2007. What do you think this is? “Conceptual uncertainty” In geoscience interpretation. *GSA Today*, **17**(11): 4–10. doi:10.1130/GSAT01711A.1.
- Bond, C.E., Johnson, G., and Ellis, J.F. 2015. Structural model creation: the impact of data type and creative space on geological reasoning and interpretation. *Geological Society, London, Special Publications*, **421**(1): 83–97. doi:10.1144/SP421.4.
- Bonneau, D., DiFrancesco, P.-M., and Hutchinson, D.J. 2019a. Surface Reconstruction for Three-Dimensional Rockfall Volumetric Analysis. *ISPRS International Journal of Geo-Information*, **8**(12): 548.

doi:10.3390/ijgi8120548.

Bonneau, D.A., and Hutchinson, D.J. 2019. The use of terrestrial laser scanning for the characterization of a cliff-talus system in the Thompson River Valley, British Columbia, Canada. *Geomorphology*, **327**: 598–609. doi:10.1016/j.geomorph.2018.11.022.

Bonneau, D.A., Hutchinson, D.J., and Mcdougall, S. 2019b. Characterizing debris transfer patterns in the White Canyon, British Columbia with terrestrial laser scanning. *In* 7th International Conference on Debris-Flow Hazards Mitigation. Golden, CO. pp. 565–572.

Bonneau, D.A., Jean Hutchinson, D., DiFrancesco, P.M., Coombs, M., and Sala, Z. 2019c. Three-dimensional rockfall shape back analysis: Methods and implications. *Natural Hazards and Earth System Sciences*, **19**(12): 2745–2765. doi:10.5194/nhess-19-2745-2019.

Brown, D.A. 1981. *Geology of the Lytton Area*, British Columbia. Carleton University.

Carter, T.G., and Marinos, V. 2020. Putting Geological Focus Back into Rock Engineering Design. *Rock Mechanics and Rock Engineering*, **53**(10): 4487–4508. Springer Vienna. doi:10.1007/s00603-020-02177-1.

DiFrancesco, P.-M., Bonneau, D., and Hutchinson, D.J. 2020. The Implications of M3C2 Projection Diameter on 3D Semi-Automated Rockfall Extraction from Sequential Terrestrial Laser Scanning Point Clouds. *Remote Sensing*, **12**(11): 1885. doi:10.3390/rs12111885.

DiFrancesco, P.-M., Bonneau, D.A., and Hutchinson, D.J. 2021. Computational Geometry-Based Surface Reconstruction for Volume Estimation: A Case Study on Magnitude-Frequency Relations for a LiDAR-Derived Rockfall Inventory. *ISPRS International Journal of Geo-Information*, **10**(3): 157. doi:10.3390/ijgi10030157.

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *In* *Knowledge Discovery and Data Mining 1996*. Elsevier. pp. 226–231.

Fey, C., and Wichmann, V. 2017. Long-range terrestrial laser scanning for geomorphological change detection in alpine terrain – handling uncertainties. *Earth Surface Processes and Landforms*, **42**(5): 789–802. doi:10.1002/esp.4022.

Fookes, P.G. 1997. Geology for engineers: The geological model, prediction and performance. *Quarterly Journal of Engineering Geology*, **30**(4): 293–424. doi:10.1144/GSL.QJEG.1997.030.P4.02.

Girardeau-Montaut, D. 2018. *CloudCompare (Version 2.10 - Alpha)*.

Greig, C.J. 1989. *Geology and Geochronometry of the Eagle Plutonic Complex, Coquihalla Area, Southwestern*

British Columbia. The University of British Columbia.

Hancock, H., Eckerstorfer, M., Prokop, A., and Hendrikx, J. 2020. Quantifying seasonal cornice dynamics using a terrestrial laser scanner in Svalbard, Norway. *Natural Hazards and Earth System Sciences*, **20**(2): 603–623. doi:10.5194/nhess-20-603-2020.

Hartmeyer, I., Keuschnig, M., Delleske, R., Krautblatter, M., Lang, A., Schrott, L., and Otto, J.-C. 2020. Enhanced rockwall retreat and modified rockfall magnitudes/frequencies in deglaciating cirques from a 6-year LiDAR monitoring. *Earth Surface Dynamics Discussions*,: 1–23. doi:10.5194/esurf-2020-9.

Jakob, M., Bovis, M., and Oden, M. 2005. The significance of channel recharge rates for estimating debris-flow magnitude and frequency. *Earth Surface Processes and Landforms*, **30**(6): 755–766. doi:10.1002/esp.1188.

Kenner, R. 2020. Mass wasting processes affecting the surface of an alpine talus slope: Annual sediment budgets 2009–2018 at Flüelapass, eastern Swiss Alps. *Land Degradation and Development*, **31**(4): 451–462. doi:10.1002/ldr.3462.

Krautblatter, M., and Dikau, R. 2007. Towards a uniform concept for the comparison and extrapolation of rockwall retreat and rockfall supply. *Geografiska Annaler, Series A: Physical Geography*, **89**(1): 21–40. doi:10.1111/j.1468-0459.2007.00305.x.

Krautblatter, M., Moser, M., Schrott, L., Wolf, J., and Morche, D. 2012. Significance of rockfall magnitude and carbonate dissolution for rock slope erosion and geomorphic work on Alpine limestone cliffs (Reintal, German Alps). *Geomorphology*, **167–168**(July 2018): 21–34. Elsevier B.V. doi:10.1016/j.geomorph.2012.04.007.

Kromer, R., Abellán, A., Hutchinson, D., Lato, M., Edwards, T., and Jaboyedoff, M. 2015. A 4D Filtering and Calibration Technique for Small-Scale Point Cloud Change Detection with a Terrestrial Laser Scanner. *Remote Sensing*, **7**(10): 13029–13052. doi:10.3390/rs71013029.

Lague, D., Brodu, N., and Leroux, J. 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, **82**: 10–26. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). doi:10.1016/j.isprsjprs.2013.04.009.

Lato, M.J. 2020. Canadian Geotechnical Colloquium: 3D remote sensing, 4D analysis and visualization in geotechnical engineering: state-of-the-art and outlook. *Canadian Geotechnical Journal*,: cgj-2020-0399. doi:10.1139/cgj-2020-0399.

Rapp, A. 1960. Recent Development of Mountain Slopes in Kärkevagge and Surroundings , Northern Scandinavia. *Geografiska Annaler*, **42**(2): 65–200.

van Steijn, H., Bertran, P., Francou, B., Texier, J. -P, and Hétu, B. 1995. Models for the genetic and environmental interpretation of stratified slope deposits: Review. *Permafrost and Periglacial Processes*, **6**(2): 125–146. doi:10.1002/ppp.3430060210.

Telling, J., Lyda, A., Hartzell, P., and Glennie, C. 2017. Review of Earth science research using terrestrial laser scanning. *Earth-Science Reviews*, **169**(March): 35–68. Elsevier. doi:10.1016/j.earscirev.2017.04.007.

Vogel, S., Eckerstorfer, M., and Christiansen, H.H. 2012. Cornice dynamics and meteorological control at Gruvefjellet, Central Svalbard. *The Cryosphere*, **6**(1): 157–171. doi:10.5194/tc-6-157-2012.

Zhou, Q.-Y., Park, J., and Koltun, V. 2018. Open3D: A Modern Library for 3D Data Processing.

Effects of seasonal track geometry changes on vehicle-track interaction and potential implications for train marshalling

Yi Wang and Kyle R Mulligan
Canadian Pacific, Calgary, AB, Canada

Kevin Oldknow and G. Gary Wang
Simon Fraser University, Surrey, BC, Canada

ABSTRACT

Railway operations in Canada must adapt to large temperature fluctuations between winter and summer. As the rail and track structure cycle through seasonal changes in longitudinal stress, track geometry naturally deviates from the designed nominal, leading to changes in vehicle-track interactions. This work takes a systematic approach in examining the seasonal changes in track geometry for specific territories on a North American Class I railway. The study includes long-wave parameters, such as gage, curvature, and superelevation, and short-wave parameters such as alignments and cross-level deviation. The effects of seasonal changes on vehicle and train dynamics are then explored to identify significant impacts under identical coupler loads and vehicle conditions. Potential implications for train marshaling are considered, including the potential for incorporating near-real-time track geometry measurements into a dynamic, simulation-based train marshaling approach. This would deviate from the traditional approach of static rule-based train marshaling that is condition-independent. The results have shown that on regularly maintained mainlines, there is no apparent correlation between ambient temperature and track geometry. Furthermore, short-wavelength track geometry parameters contribute more significantly to L/V ratios than long-waveform parameters.

Keywords: seasonality, track geometry, train marshaling, derailment safety, L/V, vehicle dynamics

1 INTRODUCTION

Railway operations in Canada must endure significant seasonal temperature fluctuations. In many parts of the country, temperature fluctuations can vary by more than 100 °F (38 °C) between summer and winter. This introduces a potential for large amounts of stress in the track both in tension and compression depending on the season. This in turn may lead to rail breaks in the winter and track buckles in the summer.

This paper takes a different perspective on seasonal temperature fluctuations and focuses on their less-immediate effects on track geometry, and the subsequent influence on vehicle track interaction. As railroads continue to develop more advanced train marshaling methods to further reduce derailment risks due to train make-up, a better understanding of wheel-rail interaction through quasi-real-time track geometry measurements, including the effects from seasonal temperature fluctuations, could lead to better decision making and safer railroad operations.

1.1 Train Marshalling and Track Geometry

Historically, railroads have relied on marshaling rules that specify restrictions on car combinations based on the length and weight differences while also limiting the number of cars having certain mechanical characteristics (Aronian et al., 2013). The marshaling rules are often derived over time, based on a combination of experience, theoretical calculations, and simulation results of longitudinal train dynamics, which assume nominal track and vehicle conditions.

Derailment investigations have shown that the introduction of quasi-real-time vehicle and track measurements could alter the computed risk levels enough to alter the decision on whether a given train might be considered safe under a given set of operating conditions (Wang et al., 2021). The demonstration of such cases gives rise to the need for further research on train marshaling using more advanced simulation techniques that incorporate wheel-rail contact. This

research has also been made possible through the wider adoption of wayside detectors and more frequent track geometry measurements by railroads.

1.2 Rail Expansion due to Temperature

In the absence of longitudinal constraints, rails would contract and expand as temperature changes just as any ordinary steel structure. The amount of linear expansion that would occur can be calculated using the formula:

$$\Delta L = (T_1 - T_0)\alpha L, \quad (1)$$

where L is the initial length of the rail, T_0 and T_1 are the initial and final temperatures, respectively, and α is the linear expansion coefficient. For steel, α is in approximately $6.5 \times 10^{-6} \text{ in}\cdot\text{in}^{-1}\cdot\text{°F}^{-1}$ (Engineering Toolbox, 2021). If no expansion or contraction is allowed, then the internal thermal stress can be calculated using the formula:

$$\sigma = \frac{\alpha E}{L_0}(T_1 - T_0), \quad (2)$$

where E is Young's Modulus. Assuming a rail temperature of 95 °F (35 °C) at the time it is laid, a decrease in rail temperature to 10 °F (-12 °C), which is not considered cold for Canadian winters, would result in approximately 5.5 feet of contraction in a 1000 feet section of track without any constraint.

In practice, rails are not allowed to freely contract and expand. The use of track anchors or elastic fasteners prevent the movement of the rails along their lengths. Especially for track structures using continuously welded rail (CWR), the longitudinal expansion is largely prevented, and the temperature change results primarily in the tensile (or compressive) stresses mentioned above instead. Given that the track structure is not perfectly rigid in the lateral direction, however, there is also a possibility for some expansion to be taken up through "breathing" of curves (i.e., a decrease in curvature under tensile stress, and an increase in curvature under compressive stress).

2 MATERIALS AND METHODS

2.1 Curve Selection and Track Geometry Surveys

Three curves of various curvatures are selected for analysis in this work. All three curves are in mountainous territories of a North American Class I railway, where both extreme summer and winter temperatures are experienced. The curves and their design curvatures are shown in Table 1 below.

Table 1. Selected curves and their curvature and total length.

Curve	1	2	3
Design curvature [degrees]	1.5	6	10
Total length [feet]	1000	1650	1520
Number of surveys	11	15	15

Note that Curves 1 and 2 are simple curves and Curve 3 is a compound curve preceded by a curve of 8.5 degrees in the reverse direction. Approximately 200 feet of tangent track are included in each track geometry survey immediately before the curves to minimize the effects of transient forces on the simulation results (TTCI, 1989).

Track geometry measurements have been collected by a combination of Track Evaluation Cars (TEC) and Autonomous Track Geometry Measurement Systems (ATGMS) over approximately 15 months, from April 2020 to July 2021. The number of surveys collected for each curve is also stated in Table 1. The channels collected from each survey include the following and are measured in one-foot intervals:

- Curvature,
- Gage,
- Superelevation,
- Lateral alignments (both rails, 62-foot mid-cord offsets), and
- Surface alignments (both rails, 62-foot mid-cord offsets).

2.2 Track Geometry Data Processing

The track geometry measurement files are exported in one-mile segments to ensure that the curves of interest are captured in full. The footage of each survey differs between each run and needs to be aligned before the curves can be extracted programmatically and consistently.

The alignment of the surveys for each curve is accomplished using the `finddelay` function in the MATLAB Signal Processing Toolbox with only the footage and curvature measurements. The `finddelay` function aligns the signals by first normalizing each signal against itself, then the function matches the locations of the peaks to calculate the offset (Mathworks, 2021).

Once the surveys are aligned and the curves of interest are extracted, filtering is done using TrackVu (the track preprocessor for NUCARS) to separate the long- and short-wavelength track geometry characteristics as follows (TTCI, 1989):

- Curvature, gage, and superelevation are filtered using a 1/400 feet non-Recursive low pass filter to retain only the long-wavelength features,
- Lateral and surface alignments are filtered using a 1/400 feet non-Recursive high pass filter to retain only the short-wavelength features.

Note that in the pre-processing of track geometry data for NUCARS, the gage can include both long-wavelength measurement data from the gage channel of the track data or a combination of the gage channel and the lateral alignment channel data, which is a short-wavelength measurement.

2.3 Vehicle Dynamics Simulation

The goal of this work is to identify the effect of seasonality on track geometry, and the potential for consequential influence on vehicle-track interaction. A

generic loaded hopper car model and theoretical wheel and rail profiles are used in the simulation. The dimensional information for the car, the wheel, and rail profiles are shown in Table 2.

Table 2. Simulation parameters.

Car type	Covered hopper
Truck center [in.]	1.5
Total weight [lbs.]	263000
Side bearing type	Constant contact
Wheel profile	AAR-1B WF
Rail profile	136RE
Coefficient of friction (tread and flange)	0.4
Speed [mph]	20

The simulation speed is held at a constant value, corresponding to the typical speed trains traverse over the three curves that have been measured.

Two scenarios are simulated. The first scenario uses the full track geometry measurement surveys and all the parameters. This scenario includes the changes in track geometry due to all factors, including the possible environmental changes and deformations under vehicle loads over time. The second scenario only incorporates the long-waveform parameters from the track geometry surveys (i.e., curvature, gage, and superelevation). The reason for including only long-wavelength track geometry parameters is to attempt to isolate only the track design parameters that have changed overtime, possibly due to environmental factors. For both scenarios, the same filtering technique, as described in Section 180.2 has been applied. In the simulation using long-wavelength track geometry characteristics only, the short-wavelength lateral alignment data are not included in the gage data.

Derailment risk is most commonly quantified by L/V ratios (Wilson, 2019). In the freight industry in North America, the thresholds are set by the time/distance-based criteria in Chapter 11 of the AAR Specification M-1001 (AAR, 2020), and the maximum L/V ratio is taken over 50 milliseconds up to a maximum distance of 3 feet. At 20 mph, the distance travelled in 50 milliseconds is approximately 1.5 feet, therefore, the maximum L/V ratios are calculated over 50 milliseconds. For the purpose of the comparisons conducted in this paper, only the axle sum L/V from the leading axle of the vehicle is considered. Without external perturbations, wheels on the leading axles experience the highest L/V in a curve due to their higher angle of attack (Wilson, 2019). The axle sum L/V is also sign-independent, which simplifies the postprocessing of the simulation results.

The results from each simulation run using the full track geometry measurements are further compared against each other for a given curve. The comparisons are presented using the standard deviation as a percentage of the mean for each curve. Comparisons are also made between the simulation results using only the long-waveform geometry characteristics and the full measurements, and the averages of the differences are calculated.

2.4 Temperature Data and Correlation Analysis

The ambient temperature data are collected from the nearest wayside hotbox detectors. At the hotbox detectors, the ambient temperature is recorded with every train pass. The final temperature for the day of the survey is taken as the average ambient temperature from all the train passes at each detector.

The study of temperature effect is done through a set of correlation analyses using the Kendall Rank Correlation (Kendall, 1975). Correlation analysis is a standard method of determining how strongly two variables are related linearly. Results are normalized to a real number between -1 and 1. A value near 1 and -1 is considered strongly correlated, while a value near 0 is considered not correlated. Kendall is one of the standard methods used in correlation analysis and is selected because it does not assume a distribution for the data being analyzed a priori. The method's nonparametric property means that there is no need for statistical analysis on the datasets and is especially useful when the dataset is small.

For this paper, Kendall's correlation analysis is used to compare the following datasets:

- Ambient temperature and track geometry parameters (curvature, superelevation, gage, and cross level),
- Ambient temperature and L/V ratios, and
- Track geometry and L/V ratios.

The correlation analyses on the track geometry parameters are performed based only on the magnitude of each variable, meaning that the absolute values are taken before calculating the maximum or mean values.

The three curves are situated on the busy mainline which is scanned regularly by track geometry cars. The typical survey frequency is one in 14 to 60 days. However, this also means that regular track work occurs which affects the perceived evolution of track geometry. To help remove the potential influence on the correlations between temperature, track geometry characteristics, and simulation results, two sets of correlation analyses are conducted. The first set of correlation analyses include the entire dataset, regardless of the potential influence from track work. The second set of correlation analyses are conducted in time windows where there are no records for track work.

3 RESULTS AND DISCUSSION

3.1 Results

3.1.1 Track Geometry Survey Alignment

The before and after curvature plots for each curve are shown in Figure 1, Figure 2, and Figure 3, respectively.

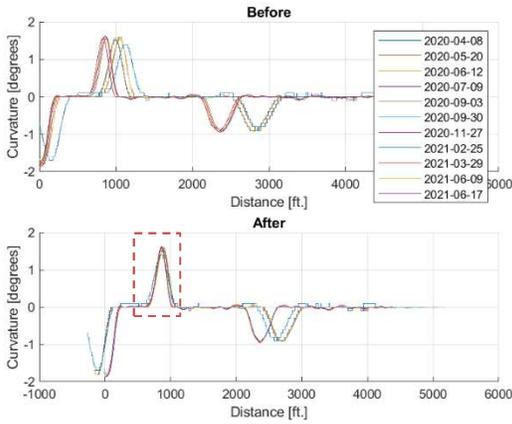


Figure 1. Before and after alignment for Curve 1 (approximately from 500 ft. to 1200 ft. on the bottom plot).

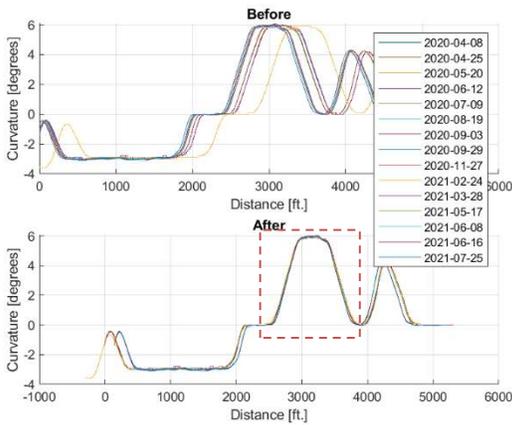


Figure 2. Before and after alignment for Curve 2 (approximately from 2100 ft. to 3800 ft on the bottom plot).

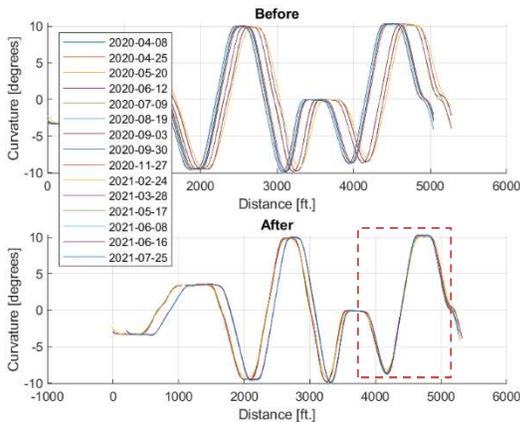


Figure 3. Before and after alignment for Curve 3 (compound curve approximately from 3000 ft. to 5000 ft.).

3.1.2 Temperature Fluctuation and Track Geometry

The visual trends for daily average temperature versus maximum curvature for each curve are presented in Figure 4, Figure 5, and Figure 6. The green boxes in the figures represent the time windows where no track work has been recorded.

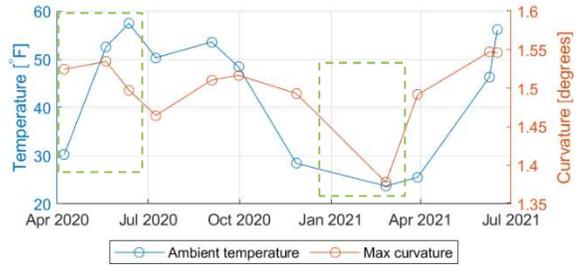


Figure 4. Ambient temperature and maximum curvature trends over time for Curve 1, the green boxes represent the windows of no track work.

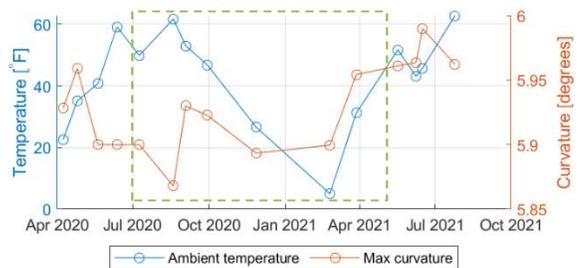


Figure 5. Ambient temperature and maximum curvature trends over time for Curve 2, the green box represents the window of no track work.

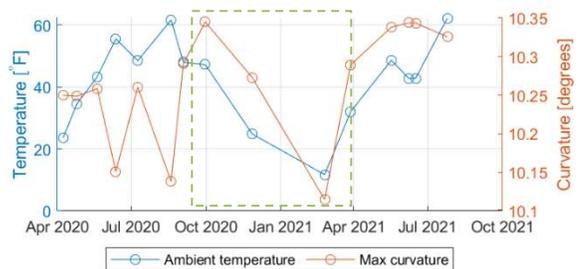


Figure 6. Ambient temperature and maximum curvature trends over time for Curve 3, the green box represents the window of no track work.

Kendall's correlation coefficients between track geometry variables and ambient temperature are plotted in Figure 7 and Figure 8, where Figure 8 has only included surveys taken during the windows of no track work. The track geometry variables are numbered as follows:

1. Max curvature,
2. Mean curvature,
3. Max superelevation,
4. Mean superelevation,
5. Max gage,
6. Mean gage,
7. Max crosslevel, and
8. Mean crosslevel.

Each marker corresponds to the correlation coefficient (y-axis) between the ambient temperature and the variable on the x-axis, e.g., the markers above the number 1 on the x-axis are the correlation coefficients between ambient temperature and the maximum curvature for each curve.

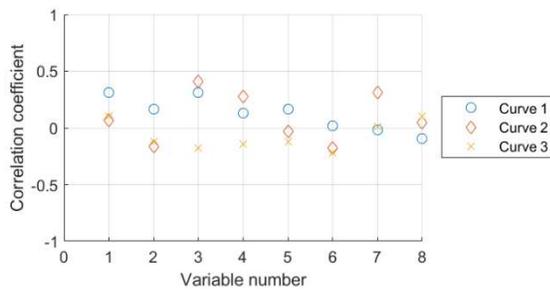


Figure 7. Kendall's correlation coefficients between ambient temperature and track geometry variables from all surveys.

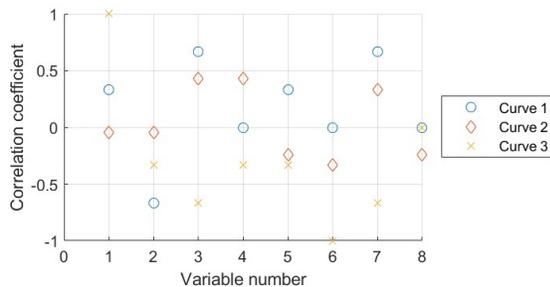


Figure 8. Kendall's correlation coefficients between ambient temperature and track geometry variables from surveys taken during windows of no track work.

3.1.3 Vehicle Dynamics Simulation Results

The maximum L/V ratios from each simulation output are calculated over 50 milliseconds, a sample of which is shown in Figure 9.

The maximum L/V ratios for each curve are presented in Figure 10, Figure 11, and Figure 12, respectively. The ambient temperature data are also plotted for a visual representation of their trend over time. The green boxes in the figures represent the time windows where no track work has been recorded. Note that it is found that some track measurement surveys are missing some or all short-wavelength measurement data. These specific surveys are listed in Table 3.

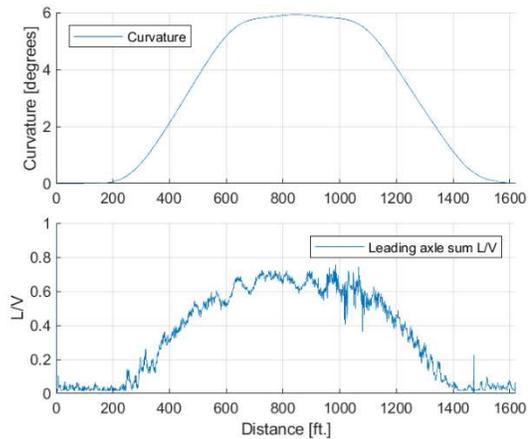


Figure 9. A sample raw vehicle dynamics simulation output for Curve 2 showing curvature and leading axle sum L/V.

Table 3. Surveys with missing shot-wavelength geometry data.

Channel	Curve 1	Curve 2	Curve 3
Left alignment	2020-04-08	2020-04-08	2020-04-08
	2020-07-09	2020-07-09	2020-07-09
	2021-03-29		
Right alignment	2020-04-08	2020-04-08	2020-04-08
	2020-07-09	2020-07-98	2020-07-98
	2021-03-29		
Left surface	2020-07-09		
Right surface	2020-07-09		

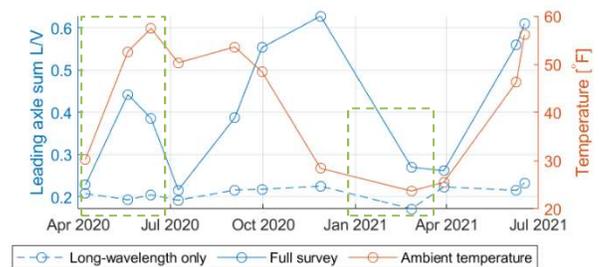


Figure 10. Leading axle sum L/V ratios and ambient temperatures for Curve 1, the green boxes represent the windows of no track work.

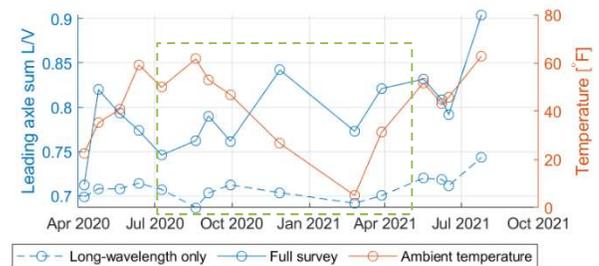


Figure 11. Leading axle sum L/V ratios and ambient temperatures for Curve 2, the green box represents the window of no track work.

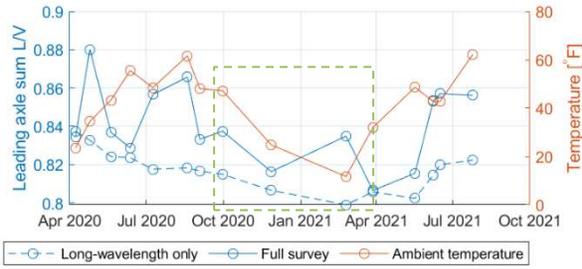


Figure 12. Leading axle sum L/V ratios and ambient temperatures for Curve 3, the green box represents the window of no track work.

The comparisons between simulation results between surveys and the simulation results from using different wave-length geometry characteristics are shown in Figure 13 below.

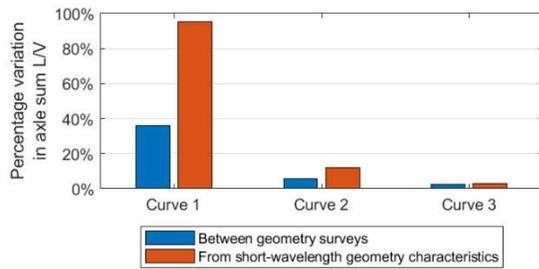


Figure 13. Comparison of simulation results across surveys and using different wave-length geometry characteristics for a given survey.

The Kendall's correlation coefficients for the comparisons between both the ambient temperature and leading axle sum L/V ratio and other variables are presented in Figure 14 and Figure 15, where Figure 15 has only included surveys taken during the windows of no track work. The variables in the figure are numbered as follows:

1. Ambient temperature,
2. Leading axle sum L/V ratio,
3. Max curvature,
4. Mean curvature,
5. Max superelevation,
6. Mean superelevation,
7. Max gage,
8. Mean gage,
9. Max crosslevel, and
10. Mean crosslevel.

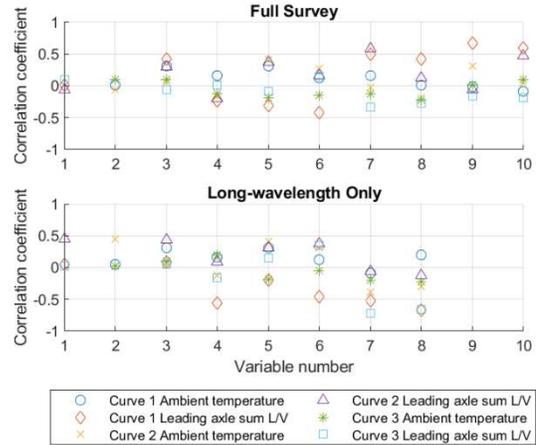


Figure 14. Kendall's correlation coefficients between track geometry variables for all three curves from all surveys. Top: both full-wavelength and short-wavelength geometry variables are included. Bottom: only long-wavelength geometry variables are included.

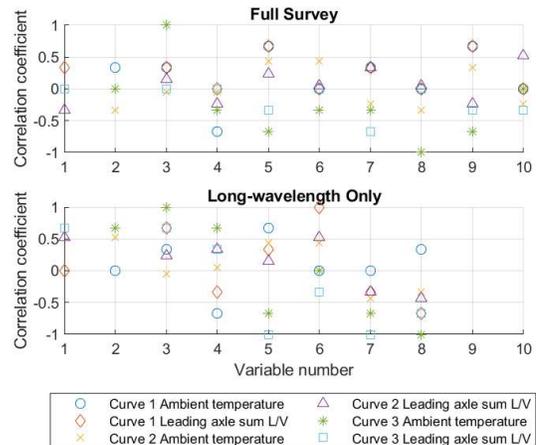


Figure 15. Kendall's correlation coefficients between track geometry variables for all three curves from surveys taken during windows of no track work. Top: both full-wavelength and short-wavelength geometry variables are included. Bottom: only long-wavelength geometry variables are included.

3.2 Discussion

3.2.1 Track Geometry Survey Alignment

Track geometry surveys from different measurement systems have been collected for this work. It is found during the alignment process that the measurements from different systems are not off by a simple scaling factor. This could be due to the difference in how location data are handled by design.

Traditional track geometry vehicles measure distance by using the known wheel circumference on the geometry car and an operator who manually aligns the car with a known location on a subdivision. This method is only feasible on staffed track geometry trainsets. For

the newer autonomous track geometry measurement systems (ATGMS), which exclusively operate as a part of revenue service trains, a method of position has been implemented which relies exclusively on GPS positioning and data in the geographic information system (GIS). The ATGMS estimates its position by comparing its GPS coordinates to the known coordinates of each segment of track in the GIS. This discrepancy may lead to challenges in the wider adoption of the analyses presented in this work when data from multiple track measurement systems exist.

3.2.2 Effects of Ambient Temperature

The visual trends, as seen from Figure 4 to Figure 6 and Figure 10 to Figure 12, are approximately in line with the hypothesis that in theory, especially during the windows of no track work. Rails “attempt” to contract as temperature decreases, which could cause curves to “straighten.” The associated decrease in L/V from the straightening could be observed more clearly for the plots using long-waveform data only, where higher transient forces are not introduced by the short-wavelength track geometry characteristics and also due to some surveys missing short-wavelength data as listed in Table 3.

3.2.3 Ambient Temperature vs. Track Geometry

The visual observation of the correlation between ambient temperature and track geometry and the plotted correlation coefficients in Figure 7 and Figure 8 suggests that the effect of track maintenance activities has a more pronounced effect on influencing track geometry compared to ambient temperatures. Removing the windows of no trackwork appears to strengthen the correlation between temperature and track geometry for Curves 1 and 3 but has also significantly reduced the number of data points. Several reasons could be causing the observations from Figure 7 and Figure 8. The weak correlations could simply be due to the changes caused by trackwork for the entire dataset, and for the reduced dataset, it is caused by the small sample sizes. There could also be differences in measurement magnitudes from different track geometry systems in the data, and further analyses are required by either looking at a much larger dataset or analyzing data collected from one track geometry system.

When the simulation results in Figure 10 to Figure 12 are examined closely, they appear to suggest that the L/V ratios follow the temperature trends more closely for the cases where only the long-waveform track geometry characteristics are considered. This observation is confirmed by the correlation coefficients in Figure 14. The comparisons using long-wavelength data only have led to consistently stronger correlations between the ambient temperature and leading axle sum L/V ratios. This observation is consistent with the fact that given a set of simulation parameters, L/V ratios increase as curvature increases.

The most interesting observations come from the comparisons between simulation results, as shown in Figure 13. For each curve, two sets of track geometry data are simulated. The results are compared both within a specific set of data across all surveys or between the two data sets for a particular survey. The comparisons suggest that at a constant speed, the effects from both ambient temperature and short-wavelength geometry variations diminish as curvature increases. This is due to the fact in shallower curves, short-wavelength geometry deviations could introduce additional angles-of-attack (AOA) that are relatively large compared to the AOA caused by curvature. In sharper curves, where the AOA is already large due to curvature, the increase from track geometry would appear much less. This observation also correlates well to the decreasing slope of the Creep Force Curve as creepage increases. When AOA increases and creepage saturates in higher-degree curves, the rates of change in both longitudinal and lateral creep forces decrease, as they eventually reach a maximum.

Furthermore, track geometry variations have a much greater effect than ambient temperature, especially on Curve 1, which has the lowest curvature of the three curves. When short-wavelength track geometry data are used in addition to long-wavelength data as inputs to simulations, the resulting L/V has increased by as much as 95%, versus the 36% difference ambient temperature has accounted for. This observation implies that having a complete set of track geometry measurement matters far greater than knowing the ambient temperature along the route. The comparisons are consistent with the industry practice of neglecting the track curvature variations that ambient temperature could introduce in the train marshaling design.

The results have shown further motivation to include track geometry data to furthering the research in train marshaling. Track geometry measurements could result in a significant increase in L/V ratios, especially on curves having lower curvatures. The increase highlights the importance of considering complete sets (both long- and short-wavelength) of track geometry measurements in the train marshaling process.

4 CONCLUSIONS

In conclusion, a clearer correlation between ambient temperature and track geometry can be seen in the sub-datasets where windows of track maintenance activities have been removed. Compared to ambient temperature, trackwork has a more pronounced and significant effect on track geometry, especially on curves where creepage is not saturated. The more dominant effect comes from short-wavelength track geometry variations and has highlighted the importance of the role for complete sets of both short- and long-wavelength track geometry in the advancement of train marshaling, where different decisions could be made on the same train makeup based on evolving track conditions.

5 REFERENCES

AAR, Association of American Railroads, M-1001. Service-worthiness tests and analyses for new freight cars. In: *MSRP-C II: Design, fabrication, and construction of freight cars*. Association of American Railroads, 2020.

Aronian A, Wachs K, Jamieson M. Train marshalling process at Canadian Pacific. Vancouver, BC, 2013, p. 15.

Engineering Toolbox, Coefficients of linear thermal expansion, https://www.engineeringtoolbox.com/linear-expansion-coefficients-d_95.html (accessed 11 September 2021).

Estimate delay(s) between signals - MATLAB finddelay, https://www.mathworks.com/help/signal/ref/finddelay.html?s_tid=doc_ta (accessed 12 September 2021).

Kendall MG. *Rank correlation methods*. 4. ed., 2. impr. London: Griffin, 1975.

TTCI, Transportation Technology Center, Inc. NUCARS version 1.0 user manual, 1989.

Wang Y, Nich S, Mulligan K, et al. Multiscale simulation-based mixed train derailment analysis: a case study. American Society of Mechanical Engineers Digital Collection. Epub ahead of print 8 June 2021. DOI: 10/gmgc9q.

Wilson N, Wu H, Klopp A, et al. Railway vehicle derailment and prevention. In: *Handbook of Railway Vehicle Dynamics*. CRC Press, 2019.

Digital rockfall databases: using LiDAR to observe over 7-years of detailed rockfall activity in the White Canyon, British Columbia, Canada

Paul-Mark DiFrancesco, David A. Bonneau and D. Jean Hutchinson

Department of Geological Sciences and Geological Engineering – Queen's University, Kingston ON, Canada

ABSTRACT

Ground hazards related to slope instability are an ongoing challenge for linear infrastructure systems traversing through rugged mountainous terrain. Over the last few decades, risk assessment frameworks have been utilized to minimize losses across transportation corridors from slope instability. A key component of these assessments is a measurement of the magnitude-frequency behaviour of landslide activity, where a database of events is required. Developments in topographic laser scanning now provide us with the capability to observe rockfalls in digitized terrain datasets, thus removing many of the limitations associated with traditional database compilation strategies. This paper demonstrates the application of assembling a digital rockfall database, from approximately 7 years of terrestrial laser scanning (TLS) data, taken at the White Canyon, a rock slope spanning roughly 2 miles, adjacent to Canadian National rail track, in interior British Columbia, Canada. We discuss some of the current challenges in the digital rockfall database generation. We also discuss how our approach to quantitative rockfall risk assessment must change in the future, as we begin to have higher resolution datasets at our disposal. We finish by discussing the exciting possibilities for widespread digital geohazard databases in the future, as data collection platforms continue to evolve.

1 INTRODUCTION

Rock slope instabilities make up a large proportion of the ground hazards that threaten the safe operation of railways that traverse rugged mountainous terrain.

Rock slope instabilities have been shown to follow magnitude-frequency distributions, where smaller magnitude occurrences have higher spatial-temporal distributions. This makes managing rock slope instabilities in transportation corridors a difficult task. Rockfalls (typically, rockfall < rockslide < rock avalanche) occur frequently and are distributed along the corridor. Collision between a train and rock fragments on the track, or fragments falling onto a passing train, may result in injury or fatalities and damage to the vehicles and rail infrastructure.

In recent years, much of the effectiveness of rockfall risk management has come from the expert knowledge and experience of infrastructure managers – who over their career, observe and identify the problematic sections and develop strategies to mitigate the hazards.

Over the past few decades, data-driven quantitative assessment frameworks have been proposed, to optimize the allocation of resources for reducing risk from rockfall hazards (Bunce et al. 1997, Hungr et al.

1999, Guzzetti et al. 2004, Budetta et al. 2016). These frameworks typically aim to quantify the likelihood of a particular risk scenario occurring. For example: the probability of rockfall fragments striking a vehicle; or fragments covering the track and causing a derailment by wedging underneath a passing locomotive.

The most common tool used in a quantitative assessment is a known inventory of rockfall occurrences. For quantitative rockfall hazard or risk assessments, a magnitude-frequency model is typically created to represent the yearly likelihood of occurrence across a spatial domain, from which hazard or risk calculations are made.

In quantitative susceptibility mapping – detailed knowledge of rockfall occurrence is also required, as it provides a means to validate the susceptibility factors utilized. Although susceptibility mapping can be difficult across long spans of linear infrastructure that contain different geostructural, lithological, and triggering factors contributing to rockfall hazards.

While a quantitative risk assessment is a very powerful tool for infrastructure managers, an assessment is only as reliable as its input data – and rockfall data, in some cases, is hard to come by. The challenges with documenting rockfall occurrence involve

issues with visibility, perception, and the chaotic dynamics of the rockfall process. Is that rock fragment sitting in the ditch, new? Are these two fragments separate, or part of a larger event?

With the evolution in topographic spatial data capture systems, mainly LiDAR and photogrammetry, alongside modern computing and graphics, methods have been developed to extract meaningful rockfall information from 3D spatial data. These digital rockfall extraction workflows rely on a series of algorithms and have been in development for the past decade, with various improvements made with the aim to reduce error, increase scale, and increase autonomy (Tonini and Abellán 2014, Williams et al. 2018, Bonneau et al. 2019a, 2019b, Guerin et al. 2020, DiFrancesco et al. 2020, 2021, Benjamin et al. 2020, Carrea et al. 2021). Some efforts have focused on preserving the 3D dimensionality of the data throughout the workflow by avoiding rasterization, to allow for enhanced 3D visualization and analysis.

It should be noted, that there is also an application within this domain which utilizes similar processes to support active monitoring techniques, used to detect deformation signatures prior to rock failures (Abellán et al. 2010, Kromer et al. 2017).

This paper aims to provide a brief explanation of the process used to reduce LiDAR data of a complex geomorphological environment to produce meaningful rockfall information. This process will work for any sort of multi-temporal point cloud data that is co-registered into a common coordinate space (i.e. it can work with photogrammetry as well). The process of digital rockfall database generation is demonstrated for a 2-mile-long rock slope, located adjacent to the Canadian National rail track in interior British Canada. The occurrence of rockfall is visualized, across a near 7-year span of monitoring with terrestrial laser scanning (TLS). The usage of the digital rockfall inventory in the domain of magnitude-frequency relations is demonstrated. We discuss some of the ongoing challenges, and the implications of these challenges in risk assessments when using high resolution spatial data. Lastly, the exciting opportunities that the future holds for large scale monitoring and database development are also discussed.

2 MATERIALS AND METHODS

2.1 Digitizing terrain with LiDAR

Laser ranging is the measurement of distance, using the active emission and subsequent detection of a laser pulse or beam. Using a robotic head or movement in a principle direction, alongside another degree of freedom provided by a mechanical mirror – laser scanner systems are able to digitize a 3D environment into individual distance measurements relative to the scanner (Telling et al. 2017).

The resulting data format is a collection of 3D points in space and is referred to as a point cloud (Figure 1). Advancements in sensors over the years have resulted in increased precision and speed of collection. The

datasets used in this study were collected using two different LiDAR systems: the Optech ILRIS 3D (2013 – 2017), and the Riegl VZ-400i (2017 – present).



Figure 1. A point cloud (right) captured by a Riegl VZ-400i terrestrial laser scanner (left). Intensity values are displayed in greyscale. Light areas correspond to high intensity returns. Black areas correspond to no data, due to occlusion of view and lack of reflectance from the shale unit exposed mid-slope.

In order to observe changes between sequential LiDAR datasets, datasets must be co-registered into a common coordinate system, which typically includes a coarse registration and fine registration stage (Jaboyedoff et al. 2012). Coarse alignment is achieved by (A) matching ground control points, (B) manually picking corresponding points, or (C) through automatic registration by matching features (Holz et al. 2015). Fine registration is typically conducted using some variation of the iterative closest point algorithm (Besl and McKay 1992, Segal et al. 2009). Points corresponding to vegetation are typically removed from the point cloud for geomorphological studies, as we are concerned with the bare earth surface. Approaches exist for segmenting vegetation automatically from point clouds (i.e. Bonneau et al. 2020).

For more information of processes used to reconstruct a survey of photos into a point cloud, using Structure-from-Motion photogrammetry, readers are referred to Westoby et al. (2012) and Schonberger and Frahm (2016). In this work, TLS point clouds are used however, the workflow would be the same with other point cloud data sources.

2.2 Extracting rockfall information from point clouds

Figure 2, located on the following page, highlights the workflow comprised of a succession of 3D operations, used to reduce point cloud datasets to meaningful rockfall information, suitable for input into a magnitude-frequency model. Some of the operations for which we have developed best practices in recent years (Bonneau et al. 2019a, 2019b, DiFrancesco et al. 2020, 2021) are highlighted in Figure 2

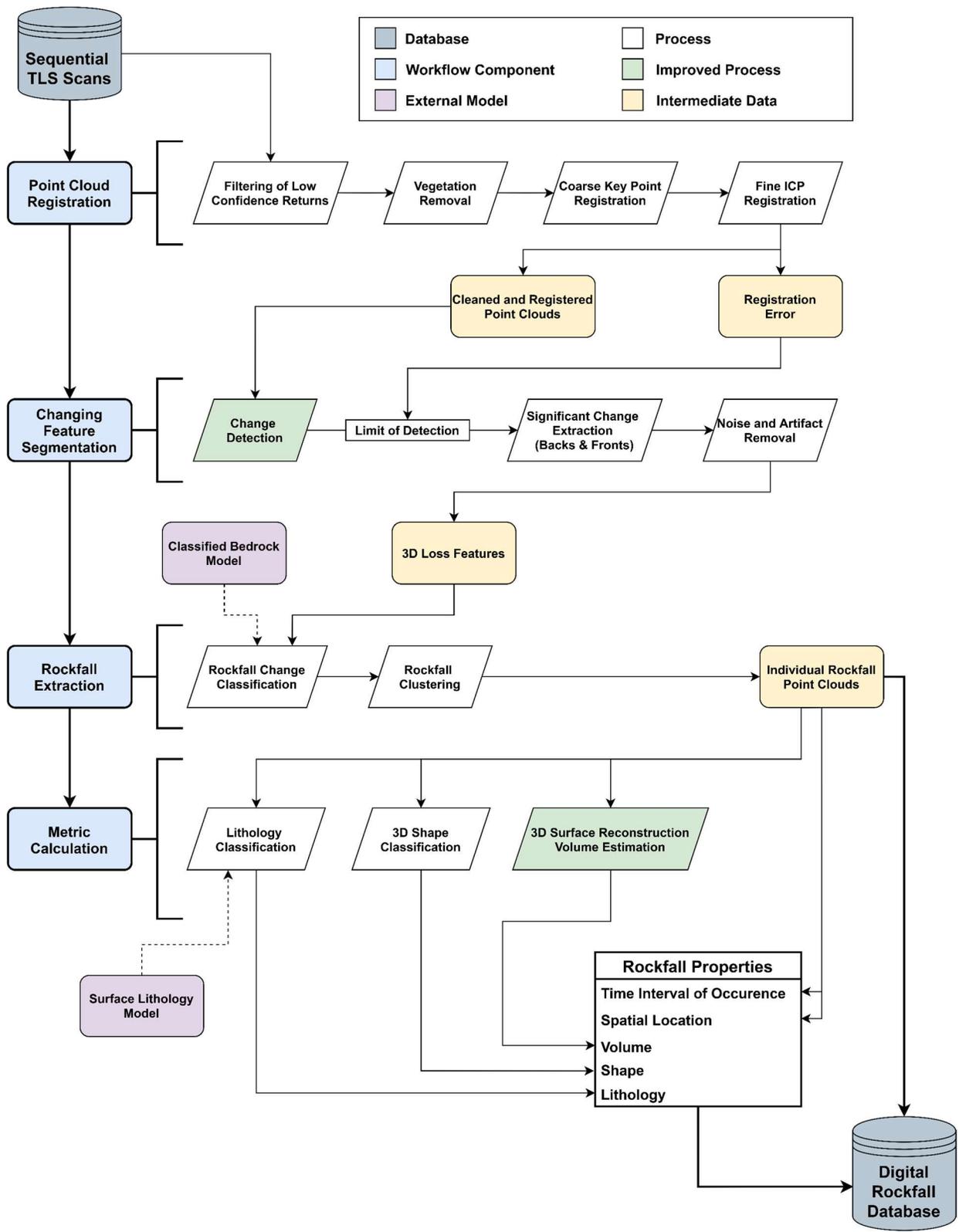


Figure 2. Flow chart of the workflow used to reduce sequential point cloud data, in this case, LiDAR point clouds of a complex geomorphological environment, into meaningful rockfall information. Highlighted in green are improvements to critical processes of the workflow, with efforts to develop best practices for digital rockfall database generation (Difrancesco, 2021).

A brief overview of the workflow depicted in Figure 2 is provided below. First, significant topographical changes can be identified across a set of multi-temporal point clouds using change detection. Change detection aims to compute the distance between the point clouds, after which a threshold can be used to segment significant changes. In our case, these significant changes are the fronts and backs of movements on the terrain, as subsets of the input clouds, or *subclouds*.

The multiscale model-to-model cloud comparison (M3C2) algorithm (Lague et al. 2013) is used to compute the change between the clouds, using appropriate parameters considering a subsampled point cloud resolution of 10 cm (DiFrancesco et al. 2020).

The subcloud of significant change is clustered with a density based algorithm (Ester et al. 1996, Tonini and Abellán 2014), to structure the point cloud into subclouds corresponding to individual changing features.

Because we are interested in measuring the spatial attributes of rockfall events, we do not want the inclusion of other types of movement in our data, such as debris or vegetation changes. Recording other types of movements as rockfall would produce an erroneous magnitude-frequency distribution. The 3D classification approach is conducted by comparing the clustered subclouds with a point cloud geomorphic mask containing a mapping to areas of bedrock or talus (Bonneau et al. 2019b). A hard classification scheme was used, using the 7 nearest-neighbors for each point in the cluster, and then classifying the entire cluster based on the mode class.

Lastly, a 3D volume calculation is done using surface reconstruction methods, which guarantee topologically-correct meshes (DiFrancesco et al. 2021). Because point clouds do not hold any volume, a watertight surface mesh must be reconstructed from point cloud data, in order to estimate a volume that is representative of the shape. A hybrid methodology was used, using the Alpha Solid method for smaller rockfalls, and the Power Crust method for larger complex rockfalls (Edelsbrunner and Mücke 1994, Bernardini et al. 1997, Amenta et al. 2001, Bonneau et al. 2019a, DiFrancesco et al. 2021).

3 STUDY SITE

The White Canyon (50.266261°, -121.538943°) is located near the confluence of the Thompson River and the Fraser River, in Interior British Columbia, Canada. The White Canyon is characterized by steep slopes, deeply incised channels, and a highly weathered host rock of quartzo-feldspathic gneiss. Notable complexity in slope geometry is a result of differential erosion, which has exposed tall spires comprised of tonalite dykes. The Canadian National (CN) main line runs through the canyon on its way to the port of Vancouver. Managing the ground hazards in the White Canyon has historically been a challenge. The site therefore became one of the primary study sites for research in 2012, when the Railway Ground Hazard Research Program (RGHRP) began testing the application of terrestrial laser scanning for rock slope monitoring (Hutchinson et al. 2015,

Kromer et al. 2015, Sala et al. 2019, Bonneau et al. 2019b).

The White Canyon is broken into two subsections, which we denote White Canyon East and White Canyon West. A tunnel connects the eastern and western portions together. Each part of the canyon has 3 rock sheds, designed to carry active channel debris overtop of the rail tracks.

To ease the load of data processing, the western portion was divided into three sections, each comprised of two scan-positions (i.e. 6 scans in total). The eastern portion was divided into two sections, with three and two scan positions (i.e. 5 scans in total). Each section typically has over 100 million points.

To date, the White Canyon is the largest site actively monitored by TLS in the world. This study utilizes 52 change detection intervals in the Western portion of the Canyon, from April 15, 2013, to March 2, 2020 (Figure 3), from 54 TLS datasets (i.e. 6-scans each). An extra scan was taken in September 2017 when transitioning between the Optech and Riegl systems, to avoid comparing data from different scanners, and providing a full suite of change detection data over time.

This study also uses 47 change detection intervals in the Eastern portion of the Canyon, from May 26, 2013, to March 2, 2020 (Figure 3), from 49 TLS datasets (i.e. 5-scans each), with an extra dataset captured when transitioning to the newer equipment.

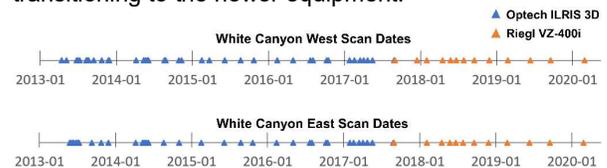


Figure 3. Data collection timelines for the Western and Eastern portions of the White Canyon. In September of 2017, the Optech scanner was replaced by the Riegl.

4 RESULTS AND DISCUSSION

A total of 22,611 rockfalls were extracted from the Western portion of the canyon, for a total volume of erosion due to rockfall of 5,521 m³, and a maximum rockfall magnitude detected of 2,657 m³.

A total of 27,775 rockfalls were extracted from the Eastern portion of the canyon, for a total volume of erosion due to rockfall of 1,933 m³, and a maximum rockfall magnitude detected of 47 m³.

The resulting rockfalls extracted for White Canyon East and White Canyon West are shown on the following two pages in Figures 4 and 5. Closeup images were also rendered, to show the detail of the inventories along the more active areas. The data is layered from most recent, meaning, that newer rockfalls are shown overtop of older rockfalls. This can be seen in Figure 4B, where newer rockfall occurrences are shown overtop of the 2600 m³ failure, which occurred on June 5th, 2013 (Kromer et al. 2015). In some areas, almost a complete resurfacing of the rock slope has occurred over the course of the 7-year period (Figure 4A, 4C, 5B).

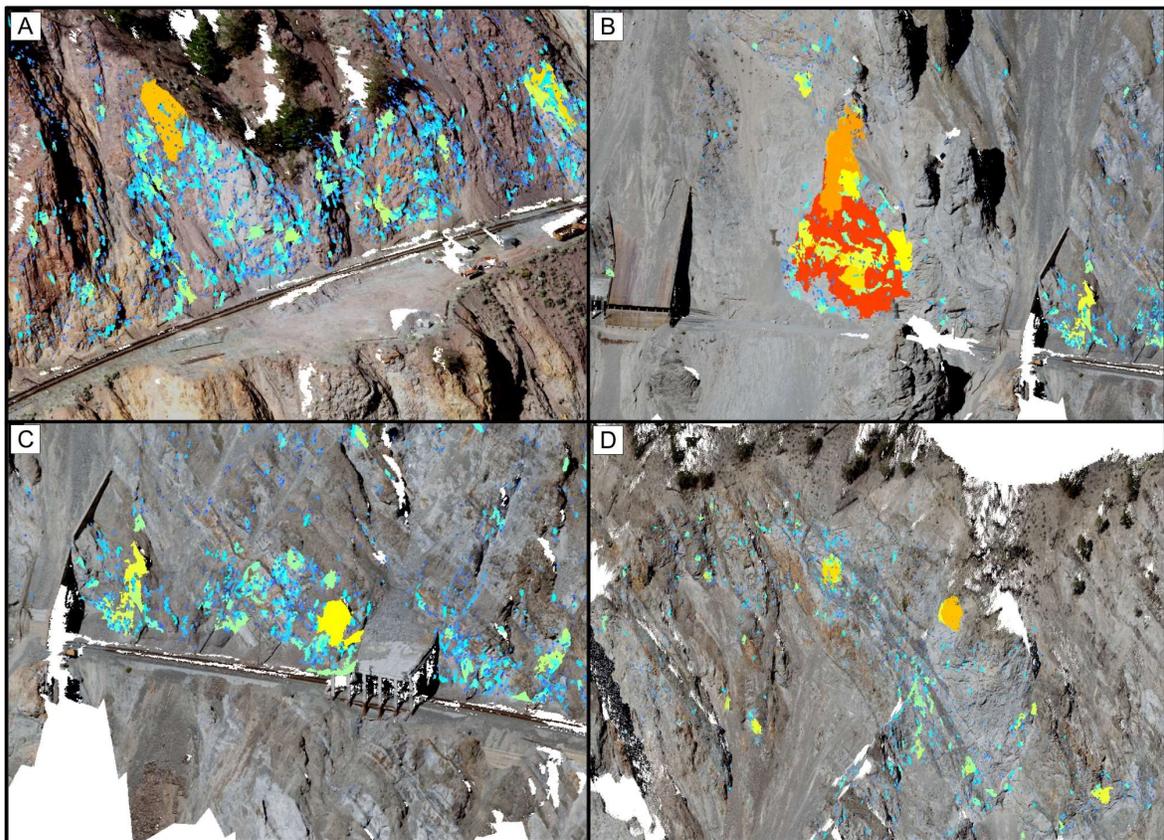
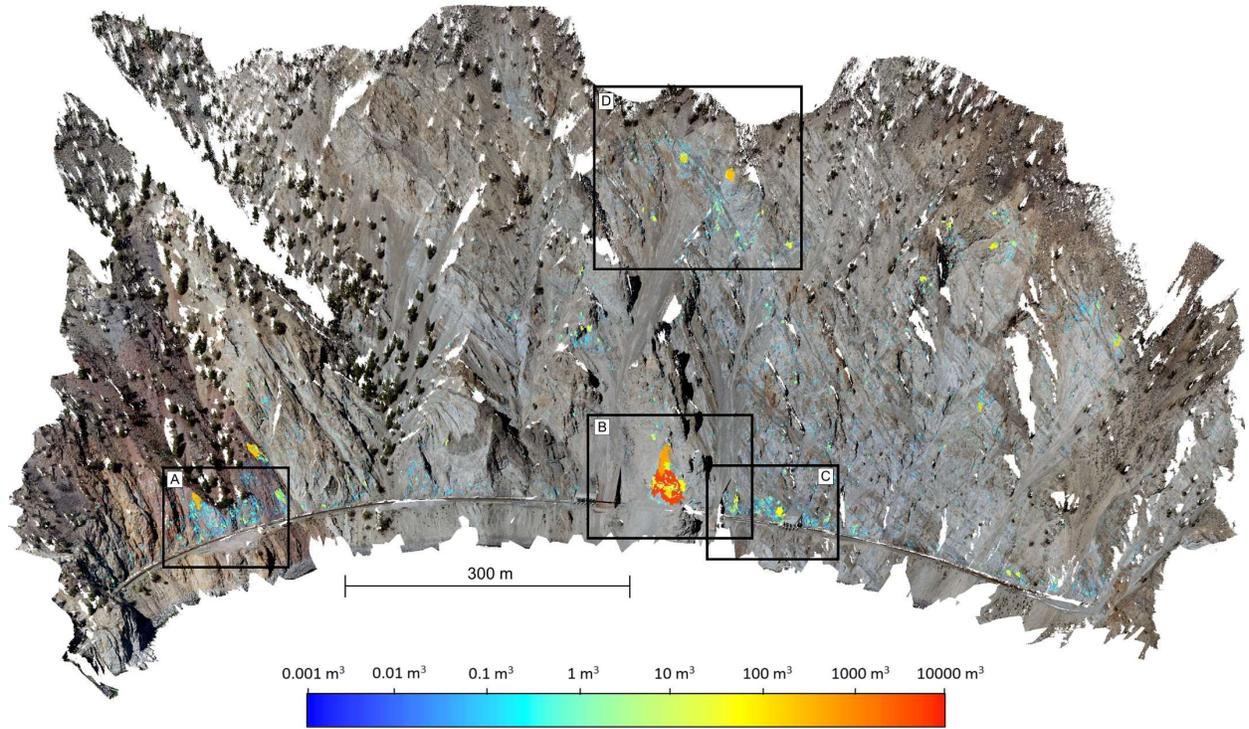


Figure 4. Rockfall data for White Canyon West from 2013-04-15 to 2020-03-02. Individual rockfalls are colorized based on their volume, where cooler colours correspond to lower magnitudes. Areas that are largely covered in cool colours (i.e. insets A and C) have resulted from 10+ low magnitude events that have superimposed through time.

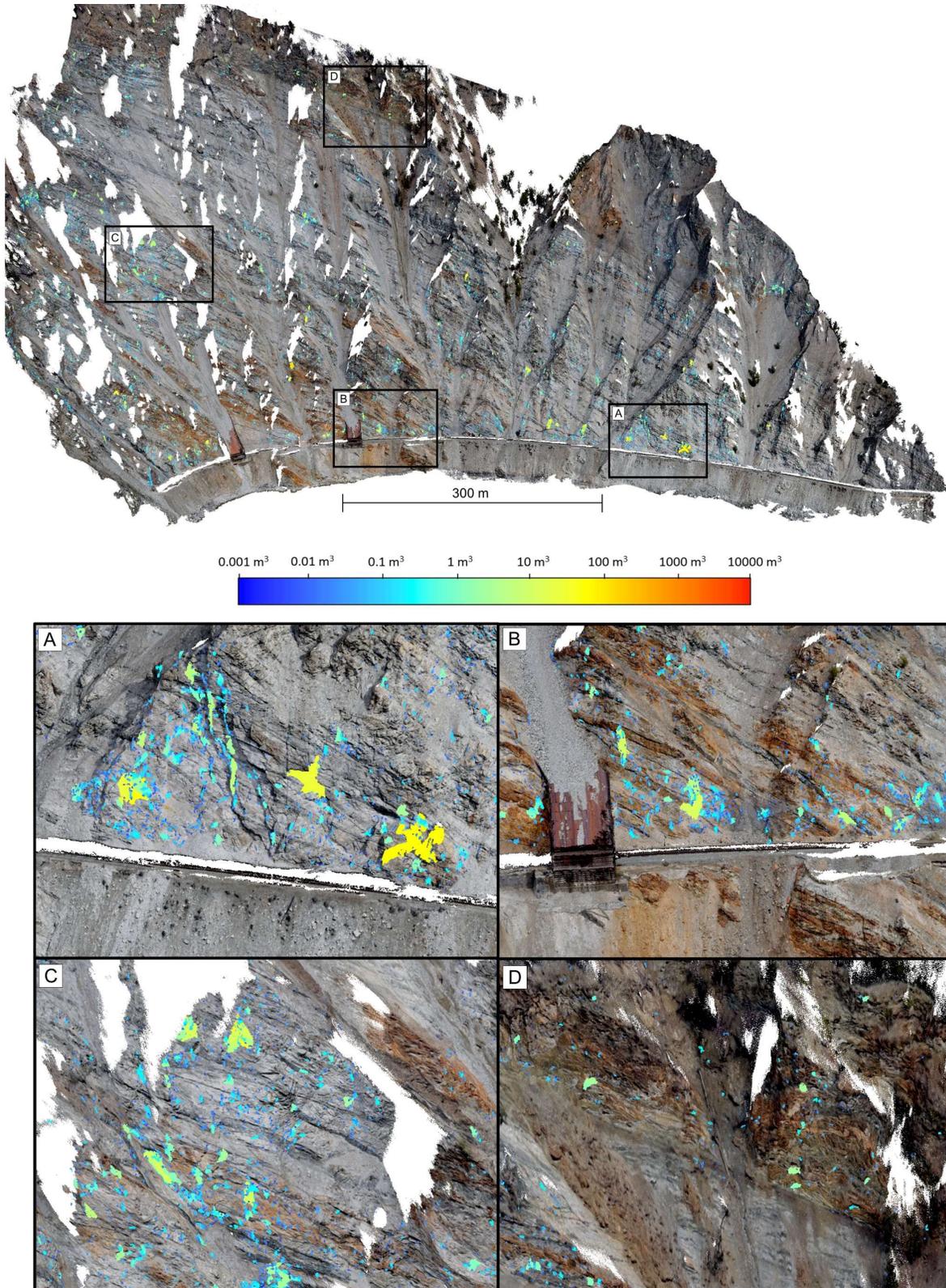


Figure 5. Rockfall data for White Canyon East from 2013-05-26 to 2020-03-02. Individual rockfalls are colorized based on their volume, where cooler colours correspond to lower magnitudes. Areas that are largely covered in cool colours (i.e. insets A-C) have resulted from 10+ low magnitude events that have superimposed through time.

Plotting the rockfall volumes for each of the sites allows for the development of a magnitude-frequency relation. The magnitude-frequency relationship for rockfall is known to follow a power-law distribution (Malamud et al. 2004). At some minimum threshold, the data is no longer represented by the power-law, which is attributed to data censoring (Hungr et al. 1999, Clauset et al. 2009).

In practice, determining the power-law model requires determining the optimal lower-bound cut-off at which the distribution is no longer modelled by the power-law (Clauset et al. 2009). Further, for rockfall – it has been recommended that an upper-bound threshold should be implemented with regards to the geostructural and geomechanical characteristics of the given environment (Corominas et al. 2018). These considerations are still under investigation for our dataset; therefore, we simply present the data with some rank-frequency plots below in Figure 6.

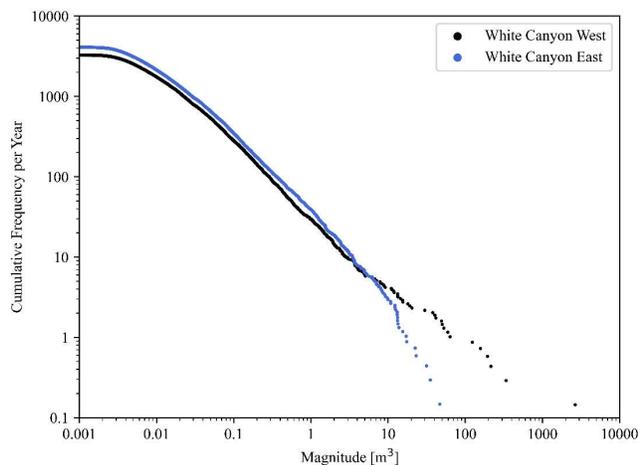


Figure 6. Rank-frequency plot for the rockfalls detected at the Eastern and Western portions of the White Canyon.

Returning to the workflow presented in Figure 2, it can be noted that we have not yet incorporated any significant improvements to the classification of movements detected. While utilizing a 3D mask is better than the previous 2.5D approaches, (van Veen et al. 2017), there are still issues with the methodology. Mainly, mapping such a large site by hand presents some major challenges. It requires attention to detail, significant computer resources to handle the 3D manipulation of data, as well as the usage of a software program that hosts a friendly user experience in its 3D viewport. To date, due to the manual mapping requirements, Agisoft Metashape (Agisoft 2017) has been a time consuming tool to map detailed, and small-scaled, geomorphological classifications on large datasets.

Automated classification methodologies offer a promising alternative to manual mapping. They can work at multiple scales – also, they are parameterized and thus repeatable. Some recent works have demonstrated this, purely geometrically, requiring no training data at all

(Farmakis et al. 2020). As these approaches are validated, they may hold for a much better approach for deducing the types of movements detected at geomorphologically complex sites, such as the White Canyon. We aim to incorporate these methods in the future, to improve the classification between rockfall and talus, before quantifying the level of rockfall activity.

Other considerations with the construction of digital rockfall inventories include the spatiotemporal frequency of data collection. Monthly to seasonal collection of data could allow for the progressive failure of the rockmass to be detected as singular occurrences of rockfall – which can significantly alter the magnitude-frequency statistics of the inventory (Williams et al. 2019). However, while more frequent data can give some intuition as to the true level of rockfall activity, exponentially more data requires a much more automated workflow. Any errors within the workflow could propagate to include false positives, or negatives, into the database. For this study, the ~50 sets of data were intensive, but manageable, for manual quality control.

Another consideration for digital rockfall inventories is an understanding of how the data may change the quantification of risk. There is an inherent bias in the traditional rockfall inventories collected from trackside inspections – that being, that the majority of recorded rockfalls have deposited adjacent to the infrastructure, typically either in a ditch, or alongside the track. Therefore with some of the dynamic spatial hazard components (i.e. the probability that a rockfall reaches the infrastructure), are partially contained within a traditional rockfall inventory. In comparison, we have a much greater understanding of the magnitude frequency relationship of rockfalls in detailed digital inventories, however, we may lack some of the understanding of the spatial component when monitoring is infrequent, or where tracking the deposition of rockfall fragments may be ambiguous (i.e. ditches and track are actively managed and evidence of rockfall deposits may be moved or removed, or their signature is lost with fluxes in talus). However, this can be supplemented with some 3D rockfall modeling software, which aims to provide a probabilistic approach to modelling rockfall trajectory (Dorren 2016, Demmel et al. 2019, Harrap et al. 2019). This means that while our understanding of the spatial-temporal likelihood of rockfall occurrence has improved, we need to move towards more sophisticated modeling tools to further understand the dynamic spatial component of the hazard (i.e. rockfall runoff).

As platforms for data capture advance, we can expect some further innovations and applications for digital geohazard databases. Automated data capture from uncrewed aircrafts could provide corridor-scale coverage, for more regional geohazard assessments. The deployment of fixed data collection provides a source for real-time monitoring for high-priority sites (Kromer et al. 2017). 5G networks and high-performance computing hold potential real-time feedback into infrastructure models, and real-time access of data to users.

5 CONCLUSIONS

Ground hazards related to slope stability remain a significant challenge for linear infrastructure systems traversing through rugged mountainous terrain.

Developments in topographic laser scanning provide us with the capability to observe rockfalls in digitized terrain datasets, and can compress large datasets into meaningful information concerning geohazards. With fast advancements in data capture, high performance computing, and high-speed networks, there is a vast opportunity for large-scale monitoring of geohazards across transportation networks.

6 ACKNOWLEDGEMENTS

The authors would like to acknowledge CN Rail, CP Rail, the Geological Survey of Canada, Transport Canada, and the University of Alberta for supporting the collaborative RGHRP research project. We thank NSERC for their funding via the Alliance and Discovery Grant programs, and Canada Graduate Scholarships.

The current and past members of the Queen's Geomechanics Group are acknowledged for their work in data collection, processing, and development of methodologies.

Lastly, we'd like to thank Daniel Girardeau and developers for the excellent open-source point cloud processing software, CloudCompare, which has had such a positive impact on the geoscience community.

7 REFERENCES

- Abellán, A., Calvet, J., Manuel, J., and Blanchard, J. 2010. Geomorphology Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring. *Geomorphology*, **119**(3–4): 162–171. Elsevier B.V. doi:10.1016/j.geomorph.2010.03.016.
- Agisoft. 2017. PhotoScan Professional. St. Petersburg, Russia.
- Amenta, N., Choi, S., and Kolluri, R.K. 2001. The Power Crust. *In* Proceedings of the sixth ACM symposium on Solid modeling and applications - SMA '01. ACM Press, New York, New York, USA. pp. 249–266.
- Benjamin, J., Rosser, N.J., and Brain, M.J. 2020. Emergent characteristics of rockfall inventories captured at a regional scale. *Earth Surface Processes and Landforms*, **45**(12): 2773–2787. doi:10.1002/esp.4929.
- Bernardini, F., Bajaj, C.L., Chen, J., and Schikore, D.R. 1997. A triangulation-based object reconstruction method. *In* Proceedings of the thirteenth annual symposium on Computational geometry - SCG '97. ACM Press, New York, New York, USA. pp. 481–484.
- Besl, P., and McKay, N. 1992. A Method for Registration of 3-D Shapes. *Bonneau, D.A., DiFrancesco, P.-M., and Hutchinson, D.J. 2019a. Surface Reconstruction for Three-Dimensional Rockfall Volumetric Analysis. ISPRS International Journal of Geo-Information*, **8**(12): 548. doi:10.3390/ijgi/624398.
- Bonneau, D.A., DiFrancesco, P.-M., and Hutchinson, D.J. 2020. A method for vegetation extraction in mountainous terrain for rockfall simulation. *Remote Sensing of Environment*, **251**: 112098. doi:10.1016/j.rse.2020.112098.
- Bonneau, D.A., Hutchinson, D.J., DiFrancesco, P.-M., Coombs, M., and Sala, Z. 2019b. Three-dimensional rockfall shape back analysis: Methods and implications. *Natural Hazards and Earth System Sciences*, **19**(12): 2745–2765. doi:10.5194/nhess-19-2745-2019.
- Budetta, P., De Luca, C., and Nappi, M. 2016. Quantitative rockfall risk assessment for an important road by means of the rockfall risk management (RO.MA.) method. *Bulletin of Engineering Geology and the Environment*, **75**(4): 1377–1397. doi:10.1007/s10064-015-0798-6.
- Bunce, C.M., Cruden, D.M., and Morgenstern, N.R. 1997. Assessment of the hazard from rock fall on a highway. *Canadian Geotechnical Journal*, **34**(3): 344–356.
- Carrea, D., Abellan, A., Derron, M.-H., Gauvin, N., and Jaboyedoff, M. 2021. MATLAB Virtual Toolbox for Retrospective Rockfall Source Detection and Volume Estimation Using 3D Point Clouds: A Case Study of a Subalpine Molasse Cliff. *Geosciences*, **11**(2): 75. doi:10.3390/geosciences11020075.
- Clauset, A., Shalizi, C.R., and Newman, M.E.J. 2009. Power-Law Distributions in Empirical Data. *SIAM Review*, **51**(4): 661–703. doi:10.1137/070710111.
- Corominas, J., Mavrouli, O., and Ruiz-Carulla, R. 2018. Magnitude and frequency relations: are there geological constraints to the rockfall size? *Landslides*, **15**(5): 829–845. *Landslides*. doi:10.1007/s10346-017-0910-z.
- Demmel, S.S.E.S., Ringenbach, A., Christen, M., Caviezel, A., Lu, G., Demmel, S.S.E.S., Ringenbach, A., Bühler, Y., Christen, M., and Bartelt, P. 2019. RAMMS :: ROCKFALL - a modern 3-dimensional simulation tool calibrated on real world data. *In* 53rd U.S. Rock Mechanics/Geomechanics Symposium. New York, NY, USA.
- DiFrancesco, P.-M., Bonneau, D.A., and Hutchinson, D.J. 2020. The Implications of M3C2 Projection Diameter on 3D Semi-Automated Rockfall Extraction from Sequential Terrestrial Laser Scanning Point Clouds. *Remote Sensing*, **12**(11): 1885. doi:10.3390/rs12111885.

- DiFrancesco, P.-M., Bonneau, D.A., and Hutchinson, D.J. 2021. Computational Geometry-Based Surface Reconstruction for Volume Estimation: A Case Study on Magnitude-Frequency Relations for a LiDAR-Derived Rockfall Inventory. *ISPRS International Journal of Geo-Information*, **10**(3): 157. doi:10.3390/ijgi10030157.
- Difrancesco, P.M. 2021. Digital Rockfall Databases: Developing Best Practices for Semi-Automatic Extraction of Rockfall from LiDAR. Queen's University.
- Dorren, L. 2016. Rockyfor3D (v5.2) revealed - Transparent description of the complete 3D rockfall model. ecorisQ paper.
- Edelsbrunner, H., and Mücke, E.P. 1994. Three-Dimensional Alpha Shapes. *ACM Transactions on Graphics*, **13**(1): 43–72. doi:10.1145/174462.156635.
- Ester, M., Kriegel, H., Xu, X., Miinchen, D.-, Sander, J., Xu, X., Miinchen, D.-, Sander, J., and Xu, X. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. Edited by E. Simoudis, U. Fayyad, and J. Han.* AAAI Press. pp. 226–231.
- Farmakis, I., Bonneau, D., Hutchinson, D.J., and Vlachopoulos, N. 2020. SUPERVOXEL-BASED MULTI-SCALE POINT CLOUD SEGMENTATION USING FNEA FOR OBJECT-ORIENTED ROCK SLOPE CLASSIFICATION USING TLS. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XLIII-B2-2**: 1049–1056. doi:10.5194/isprs-archives-XLIII-B2-2020-1049-2020.
- Guerin, A., Stock, G.M., Radue, M.J., Jaboyedoff, M., Collins, B.D., Matasci, B., Avdievitch, N., and Derron, M.-H. 2020. Quantifying 40 years of rockfall activity in Yosemite Valley with historical Structure-from-Motion photogrammetry and terrestrial laser scanning. *Geomorphology*, **356**: 107069. doi:10.1016/j.geomorph.2020.107069.
- Guzzetti, F., Reichenbach, P., and Ghigi, S. 2004. Rockfall Hazard and Risk Assessment Along a Transportation Corridor in the Nera Valley, Central Italy. *Environmental Management*, **34**(2): 191–208. doi:10.1007/s00267-003-0021-6.
- Harrap, R.M., Hutchinson, D.J., Sala, Z., Ondercin, M., and DiFrancesco, P.-M. 2019. Our GIS is a Game Engine: Bringing Unity to Spatial Simulation of Rockfalls. *In GeoComputation.* Queenstown, New Zealand.
- Holz, D., Ichim, A.E., Tombari, F., Rusu, R.B., and Behnke, S. 2015. Registration with the Point Cloud Library: A Modular Framework for Aligning in 3-D. *IEEE Robotics & Automation Magazine*, **22**(4): 110–124. doi:10.1109/MRA.2015.2432331.
- Hungr, O., Evans, S.G., and Hazzard, J. 1999. Magnitude and frequency of rock falls and rock slides along the main transportation corridors of southwestern British Columbia. *Canadian Geotechnical Journal*, **36**(2): 224–238. doi:10.1139/t98-106.
- Hutchinson, D.J., Lato, M.J., Gauthier, D., Kromer, R.A., Ondercin, M., van Veen, M., and Harrap, R. 2015. Applications of remote sensing techniques to managing rock slope instability risk. *In Applications of remote sensing techniques to managing rock slope instability risk.* GeoQuebec. Quebec City.
- Jaboyedoff, M., Oppikofer, T., Abellán, A., Derron, M.-H., Loye, A., Metzger, R., and Pedrazzini, A. 2012. Use of LIDAR in landslide investigations: A review. *Natural Hazards*, **61**(1): 5–28. doi:10.1007/s11069-010-9634-2.
- Kromer, R.A., Abellán, A., Hutchinson, D.J., Lato, M.J., Chanut, M.A., Dubois, L., and Jaboyedoff, M. 2017. Automated terrestrial laser scanning with near-real-time change detection - Monitoring of the Séchillienne landslide. *Earth Surface Dynamics*, **5**(2): 293–310. doi:10.5194/esurf-5-293-2017.
- Kromer, R.A., Hutchinson, D.J., Lato, M.J., Gauthier, D., and Edwards, T. 2015. Identifying rock slope failure precursors using LiDAR for transportation corridor hazard management. *Engineering Geology*, **195**: 93–103. Elsevier B.V. doi:10.1016/j.enggeo.2015.05.012.
- Lague, D., Brodu, N., and Leroux, J. 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). *ISPRS Journal of Photogrammetry and Remote Sensing*, **82**: 10–26. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). doi:10.1016/j.isprsjprs.2013.04.009.
- Malamud, B.D., Turcotte, D.L., Guzzetti, F., and Reichenbach, P. 2004. Landslide inventories and their statistical properties. **711**: 687–711. doi:10.1002/esp.1064.
- Sala, Z., Hutchinson, D.J., and Harrap, R. 2019. Simulation of fragmental rockfalls detected using terrestrial laser scans from rock slopes in south-central British Columbia, Canada. *Natural Hazards and Earth System Sciences*, **19**(11): 2385–2404. doi:10.5194/nhess-19-2385-2019.
- Schonberger, J.L., and Frahm, J.-M. 2016. Structure-from-Motion Revisited. *In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* IEEE. pp. 4104–4113.
- Segal, A., Haehnel, D., and Thrun, S. 2009. Generalized-ICP. *In Robotics: Science and Systems V.* Robotics: Science and Systems Foundation.
- Telling, J., Lyda, A., Hartzell, P., and Glennie, C.L. 2017. Review of Earth science research using terrestrial laser scanning. *Earth-Science Reviews*, **169**(March): 35–68. Elsevier. doi:10.1016/j.earscirev.2017.04.007.

Tonini, M., and Abellán, A. 2014. Rockfall detection from terrestrial LiDAR point clouds: A clustering approach using R. *Journal of Spatial Information Science*, (8). doi:10.5311/JOSIS.2014.8.123.

van Veen, M., Hutchinson, D.J., Kromer, R., Lato, M., and Edwards, T. 2017. Effects of sampling interval on the frequency - magnitude relationship of rockfalls detected from terrestrial laser scanning using semi-automated methods. *Landslides*, **14**(5): 1579–1592. doi:10.1007/s10346-017-0801-3.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., and Reynolds, J.M. 2012. “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, **179**(June 2018): 300–314. Elsevier B.V. doi:10.1016/j.geomorph.2012.08.021.

Williams, J.G., Rosser, N.J., Hardy, R.J., and Brain, M.J. 2019. The Importance of Monitoring Interval for Rockfall Magnitude-Frequency Estimation. *Journal of Geophysical Research: Earth Surface*, **124**(12): 2841–2853. doi:10.1029/2019JF005225.

Williams, J.G., Rosser, N.J., Hardy, R.J., Brain, M.J., and Afana, A.A. 2018. Optimising 4-D surface change detection: An approach for capturing rockfall magnitude-frequency. *Earth Surface Dynamics*, **6**(1): 101–119. doi:10.5194/esurf-6-101-2018.

Development of a daily updated train derailment impact map and analytics system for the national railway network in Canada

Yan Liu, Chengbi Dai, Zach Schenk and Luke Steinginga
National Research Council Canada, Ottawa, Ontario, Canada

ABSTRACT

A systematic analysis of the past train derailments and their impact can help industry and regulators to identify potential safety gaps so as to develop an effective method to mitigate the future risk. There exists a large body of data in the public domain to support such analysis, including the historical derailment records, geography, environment, demography and weather conditions along the railway network. The challenge is how to integrate and make use of the data for improving rail safety in Canada. This paper describes a daily updated Train Derailment Impact Map and Analytics (TDIMA) tool recently developed. A multidimensional database was first developed to integrate data from multiple data sources. Besides the historical number of derailments, the potential consequence was integrated as an additional element in the tool. This is the key difference between the developed tool and the existing ones that use number of accidents as the main variable to trend the risk. Multiple smart interfaces were designed to facilitate visualization-style analytics. With a few clicks, the tool can be used to perform analytics effectively. Case studies show that the trend of rail safety obtained only by number of past derailments can be different from that obtained by using the impact defined as the product of derailments and consequence. Ranking the cause of derailments by the defined impact resulted in different top causes than that ranked by only using number of derailed cars. This result provides the industry and regulator a different perspective to review the maintenance priority. Case studies on crossing, fire, and broken rail / wheel accidents are also discussed. It is also demonstrated how the tool can be used to review and monitor the latest rail accidents on an interactive map.

1 INTRODUCTION

There are typically more than 50 main track train derailments in Canada every year. This has many detrimental effects on public safety, environmental and financial damage. It is therefore important to identify and mitigate future risks caused by derailments and other railway accidents.

The purpose of this study was to develop a data integration and analytics tool that can be used to conduct a systematic analysis of past train derailments and their impact in order to help industry and regulators to identify potential safety gaps and mitigate future risk. There is a substantial quantity of data concerning derailments and the rail network in general that is publicly available. This data includes information such as historical derailment records, geography, environment, demography, and weather conditions along the railway network. The challenge is how to integrate and make use of the data for improving rail safety in Canada.

Transport Canada (TC) has published a map of Rail Networks and Crossings that interactively shows the

locations of past accidents (Transport Canada, 2021). This provides the public an excellent tool to browse the historical derailments. However, the tool doesn't integrate information from environment, demography and weather data. It has a limited capability for user to filter, sort and trend the historical data. Transportation Safety Board of Canada (TSB) has a publicly available set of Rail Transportation Safety Investigations and Reports that contains detailed information and analysis for individual accidents¹. These reports only cover a small portion of past derailments or other occurrences and are difficult to be applied to conduct trending analysis. A complete raw database called Rail Occurrence Database System (RODS) is published by TSB monthly (Transportation Safety Board of Canada, 2021). This database has been used as one of the main data sources in the present study.

Conducting analytics using the historical data of derailments can help to trend safety performance and identify safety gaps. However, such analytics could be incomplete if the consequence of derailments is not considered. For example, a derailment occurring in a high population area poses higher risk than the same derailment occurring in a low population area. Needing

¹ TSB occurrences include two different categories, accident and incident. In the present paper, we use accident to represent both.

a tool for Class 1 railways to select low-risk routes for transportation of dangerous goods on the USA railway network, American Association of Railroad (AAR) supports the development and operation of the Rail Corridor Risk Management System (RCRMS) (Visual Risk Technologies, 2013). The system is only available for Class 1 railways. Human health and environment factors were included as key consequence metrics among other factors in RCRMS.

This paper describes the development of a tool for the national railway network in Canada that is based on both historical derailment records and potential consequences. The tool also combines other data such as geodatabase and weather conditions. The developed tool can be used to effectively conduct analyses and look for trends in past data while taking different factors into account. Case studies are provided to demonstrate how the developed platform can be used to trend historical data from impact point of view.

2 METHODOLOGY AND TOOL DEVELOPMENT

2.1 Impact Definition

Several frameworks (Kaplan & Garrick, 1981; Apostolakis, 2004; Möller, Hansson, & Holmberg, 2017) can be used to assess the risk of derailment. A well-established concept in risk analysis is that the risk equals to the product of failure probability and the consequence caused by the failure event (such as derailments). This concept has been applied as the foundation of the present study and the developed tool, although the historical data has been used to replace the predicted / estimated failure probability as in the conventional risk assessment. Extension of the present tool to include predictive models of the derailment probability based on train, track and operation conditions will be covered in a separate future paper.

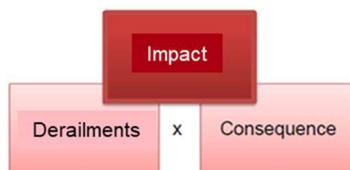


Figure 1. Derailment, consequence and Impact.

Besides the historical number of derailments and other occurrences, the potential consequence was integrated as an additional element in the tool. This is the key difference between the developed tool and the existing ones that use number of accidents as the main variable to trend the risk. As shown in Figure 1, a risk measure called “impact” is defined. It is a numerical value equal to the product of derailment and consequence. Here the derailment was defined as the number of derailments or derailed / released rail cars, obtained from RODS. The consequence was given by numerical variables that quantitatively describe the

extent and scope of damage potentially caused by derailments or other rail accidents. Two consequence variables have been determined by the geographic analysis described in the following subsection. Built on this impact concept and definition, the developed tool is called the Train Derailment Impact Map and Analytics (TDIMA) system.

2.2 Derailment Measures

To measure the scale and scope of a derailment, different features recorded in RODS can be used. This paper focuses on two key features, i.e., number of the derailed cars and that of released dangerous goods cars. Many other derailment features such as total number of cars in derailed trains, the involved railcars and loading conditions have been built in TDIMA for a user to do their own analytics.

The total number of cars derailed in an accident is a numeric representation that can properly measure the overall severity of the derailment. On the other hand, the scope of dangerous goods (DG) release during a derailment can be reasonably measured by the number of released DG cars. A more accurate estimate to the DG release requires more detailed information about the released material and its volume. The developed tool can be easily updated to include these details when more complete data becomes available.

2.3 National Railway Network (NRN)

In order to place railway occurrences on an interactive map, and to perform the geographic analysis for estimating consequence, the custom version of national railway network was built using the National Railway Network (NRWN) geodatabase (Natural Resources Canada, 2016) collected from Natural Resources Canada (NRCan). Both Esri's ArcGIS and Microsoft SQL Server were utilized to construct the NRN geodatabase which has been used as the geographic foundation of the developed tool. Special algorithms were developed to generate the main track segments that can be accurately correlated with railway, subdivisions and milepost based on a well-defined linear reference system.

For many historical derailments, the track data such as curvature is incomplete in the RODS database. Using the geographic features available in the NRN geodatabase, an algorithm was developed to estimate curvature for most of historical derailments. Combining NRN data and the elevation of a location from the Elevation API provided by NRCan (Natural Resources Canada, 2018), track grades for most occurrences were also estimated and integrated in the TDIMA tool.

2.4 Consequence Analysis

To estimate the consequence of a derailment on human health, a measure was required to evaluate the potential for impact on human life. This was done by calculating the potentially influenced population along railway tracks in various regions across Canada. The calculation was conducted using the 2016 CENSUS data collected from

Statistics Canada (Statistics Canada, 2016). This data came in the form of Dissemination Areas. Dissemination Areas do not have a completely consistent population density across their area but are intended to respect the boundaries of census subdivisions and census tracts. This allows their population density to be relatively consistent, and ensures they remain stable over time. Once the population density for each Dissemination Area was calculated, a geographic analysis was performed by placing a half mile by one mile buffer over the NRN geodatabase. The analysis resulted in the total population within the buffer zone for each mile of the national railway network. A half mile was selected as it is the distance most frequently recommended by Transport Canada in the 2016 Emergency Response Guidebook (Transport Canada, 2016) in response to a fire, including if the dangerous good is crude oil, the most frequently transported dangerous good.

A similar methodology was used to determine environmental consequence. The key difference was that instead of population density, the 'water surface area' within the half mile radius was combined with 'distance to water' to reflect the likelihood of surface water being contaminated by a leaking dangerous good. While environmental damage can occur in any geographic location, the complexity and cost of resolving environmental damages rise exponentially when contaminants enter a water source, where they can quickly disperse from the original spill location and create a large area of concern. In addition, contamination of water can cause major problems for humans who may rely on that water as their primary source of drinking water. For these reasons, surface water is the currently the only variable involved in the analysis of environmental damages.

2.5 Impact Metrics

The results of the consequence analysis were combined with derailment measures (section 2.2) to create several impact variables as described in section 2.1 and Figure 1. These impact variables include

- Number of derailed cars X People
- Number of released cars X People
- Number of derailed cars X Water
- Number of released cars X Water

where "X" represents product, "People" means number of people living in the 0.5 mile by mile buffer zone and "Water" means the water area in the buffer zone (square meters).

2.6 Weather Data for Railway Network

Weather data for the entire railway network in Canada was collected from a commercial weather supplier. The data includes daily and hourly weather conditions for NRN in every 20 miles along the railway track and all the accident locations reported in RODS. The historical daily weather data in the last 5 years for the occurrence location with frequently fire accidents was also collected

to examine the effect of climate change on the railway operation.

2.7 Derailment Cause Data

Another important dataset integrated in the TDIMA tool is the derailment cause category. The cause categories on a higher level include track, equipment, environment, human factor and train operation. There are also more detailed causes such as broken rail, track geometry, brake, weather and others. The data was originally prepared and provided by TSB for a study conducted by Canadian Rail Research Laboratory (Leishman, 2017). Upon request, the dataset was provided by TSB for the present study. The received data covers the period between January 2006 and December 2019.

2.8 Text Mining of Occurrence Summary

For each rail accident, railways submit a brief occurrence summary which is available in the RODS database. In a free and plain text format, the summary could include some important key words about weather, emergency brake, presence of a fire from the accident, and many others. Several text mining procedures were developed to extract these key words and integrate them into the TDIMA database.

2.9 Data Pipeline and Automatic Update

A data process pipeline equipped by multiple ETL (extract, transform and load) procedures was built on the framework of SQL Server Integration Services (SSIS). The pipeline was used to integrate and automatically update a variety of datasets in a data warehouse used by the developed tool.

While RODS accident records have been updated monthly from TSB portal, a special procedure was developed to make daily occurrence update from the daily notification report (in PDF format as email attachment) provided by TSB every workday. A Python code parses out the important information from the email attachment and pushes it to an SQL database through the pipeline. This is a key feature of the TDIMA tool which ensures up-to-date analytics to support timely monitoring the safety of the national railway network.

The commercial weather data is also extracted, transformed and loaded into the data warehouse to assist the analytics of weather effects on rail accidents.

Multiple ETL procedures were readily available for updating other two important datasets, NRWN and CENSUS, when new versions of these data become available.

2.10 Multidimensional Data and Analytics Interface

A data structure with multiple dimensions is designed to integrate all rail accident and the previously discussed conditions (or dimensions). Two main categories of data are contained in the data structure. The first group are various accident and risk measures such as number of derailments, derailed cars and newly introduced impact

variables. Another category contains many conditions or dimensions related to rail accidents. Some of these dimensions include date, speed, train length, train weight, track curvature, grade, weather, key words in summary and causes. This integrated dataset can be used to conduct various data analytics directly. In the developed TDIMA tool, multiple intuitive Tableau (TABLEAU SOFTWARE, LLC, 2021) interfaces were also designed and developed for a user to conduct visualization-style analytics on the interfaces using the database. With a few mouse clicks, the interfaces can be used to perform analytics tasks such as

- Find how many derailments with more than 10 cars derailed occurred in the last 5 years and check if these derailments occurred in high population areas.
- Identify top 10 track segments where the highest number of cars derailed or the highest impact to people occurred in the last 3 years.
- Examine if derailments that happened at low temperatures lead to more derailed cars.
- Search derailments under high wind condition.

Several case studies, other than the above mentioned examples, using the integrated database and / or the tool interfaces will be discussed in the next section.

3 CASE STUDIES AND RESULTS

The following case studies and results were all based on the developed TDIMA database and interfaces. Only accidents that occurred on the main track were included.

3.1 Trending Derailments by Impact

It is expected that values of the impact metric defined in Figure 1 and section 2.5 vary depending on geographic locations even if the same number of cars derailed or released in a derailment. The higher the potential consequence, the higher the impact for the same derailment event. Therefore, the trend of rail safety indicated by the past accidents can be different if the impact variables are used in the trending analysis.

Figure 2 (a) shows the total derailed cars on the main track of whole network between 2013 and 2020. It can be observed that the total number of derailed cars was trending down in three years starting from 2014, and the number has been trending up mildly since 2017. Using the impact variable of derailed cars X people, Figure 2 (b) gives the trend similar as Figure 2 (a) between 2013 and 2017. However, in recent three years (from 2017 to 2020), Figure 2 (b) shows that the impact of derailments has been trending down significantly. This suggests that the rail safety in the high population areas has been improved in the last a few years. Such improvement would be the results of better risk management and oversight placed in key railway subdivisions located in high population areas. It should be noticed that the improvement came after a “jump” deterioration in 2017.

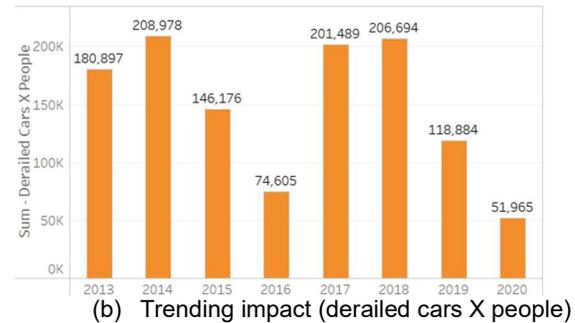
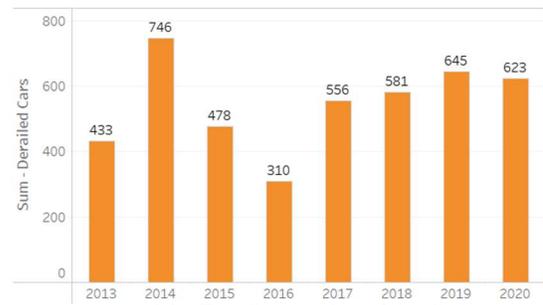


Figure 2. Main Track Train Derailments since 2013.

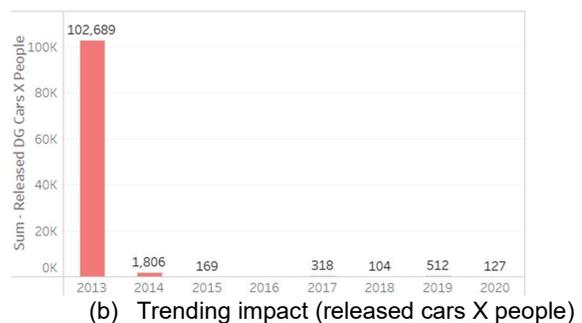
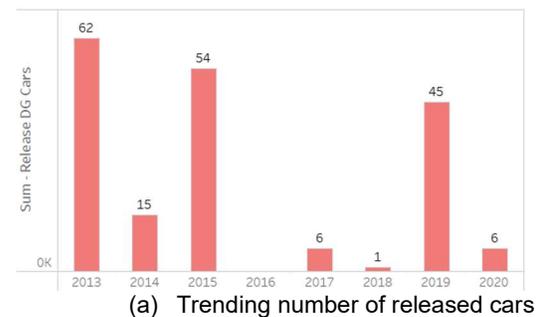


Figure 3. DG Release since 2013.

A more noticeable difference between using raw accident number and impact variable can be seen in Figure 3. The number of released DG cars as reported in RODS is trending in Figure 3 (a), in which the 2015 and 2019 derailment scopes were not too much lower than 2013. However, when the impact (product of released cars and the affected people) is used for trending, a very different assessment can be seen. The

2013 derailment has a much higher impact than all other years after 2013. This assessment agrees with how painfully the public was affected in 2013 Lac Megantic derailment and how little the public health was impacted by the large DG railcar release occurred in remote areas where no people live near railway track. The above comparison doesn't suggest that the large number of DG car release pose no risk. It identifies a strong alert for all stakeholders to prevent such DG release and prevent it from occurring in high population areas.

3.2 Top Causes of Derailment

In previous statistics studies (Leishman, 2017; Liu et al., 2012), it was found that the broken rail and track geometry were two main causes for track-related derailments. On the equipment side, bearing failure and brake issues have been identified as important causes. In these studies, number of derailments and the average number of derailed cars were used to rank the cause.



(a) Ranking by number of derailed cars



(b) Ranking by impact – derailed cars X people

Figure 4. Causes of Derailments.

After integrating the recent data provided by TSB, the present tool can be applied to examine if there is any change in derailment causes. The total number of derailed cars and the impact variable derailed cars X people were applied to rank the causes. Figure 4 shows the ranking results in a tree map format in which each

colored rectangle represents a cause and its area corresponds to numerical values of the used ranking variables. When the total number of derailed cars is used for ranking, as shown in Figure 4 (a), broken rail and track geometry are the top causes, corresponding to 575 and 527 derailments, respectively. After that, failed railbed, broken wheel, high in-train force, weather and brake issue can be found to be important causes.

It can be seen in in Figure 4 (b) that ranking by impact (derailed cars X people) resulted in different top causes. The result suggests that brake and switch issues should be placed as high priorities in order to reduce derailment impact. Broken rail and track geometry, together with broken wheel, high in-train force and weather are still important causes, but not top ones. While more studies might be required, the present results provide the industry and regulator a different perspective to rethink the maintenance priorities.

3.3 Crossing Accidents and Cold Weather

It is well known in the rail industry that crossings are one of the most frequently occurring accident types. In fact, a quick analysis using TDIMA tool shows that, of all the accidents currently included ² in TDIMA, crossing accidents are by far the most frequent one, making up approximately one third of all accidents in the last 20 years. Using the tool that was developed, several interesting observations can be made. One is that, while number of non-crossing accidents have a relatively uniform distribution month-to-month, crossing accidents occurred at a significantly higher rate during winter months, specifically December, January and February, as seen in Figure 5. Another observation is shown in Figure 6, demonstrating that while the total number of incidents has remained relatively steady over the past 20 years, the number of crossing incidents has had a significant and consistent rate of decline. One final observation is shown in Figure 8, where crossing accidents that occurred in last 10 years are grouped by the temperature when accidents occurred. As discussed in section 2.6, the temperature was estimated and integrated by using data from a weather supplier. The results shown in Figure 7 indicates that more crossing accidents occurred under cold weather conditions with the temperature lower than 10°C and the average temperature equal to -4°C. By filtering “ice”, “snow” and “fog” key words in summary data of crossing accidents, it was found these weather-related key words were recorded only for the low temperature accident group. This case study suggests that cold weather and other unfavorable weather conditions left grade crossing prone to collision accidents.

3.4 Fire Accident and Climate Change

This section demonstrates the capabilities of the TDIMA tool in identifying and analyzing new trends of rail accidents.

² Trespassing and a few other accident / incident types were not included in TDIMA database at this time.

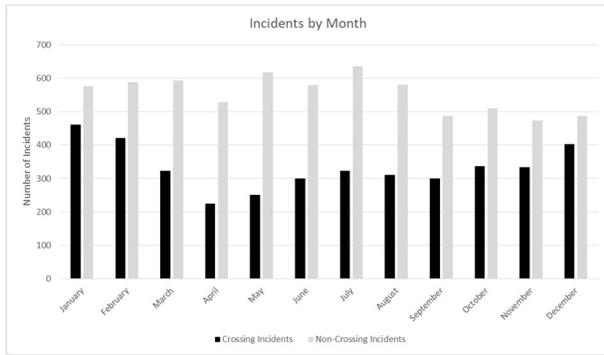


Figure 5. Crossing and other Accidents in Last 20 Years

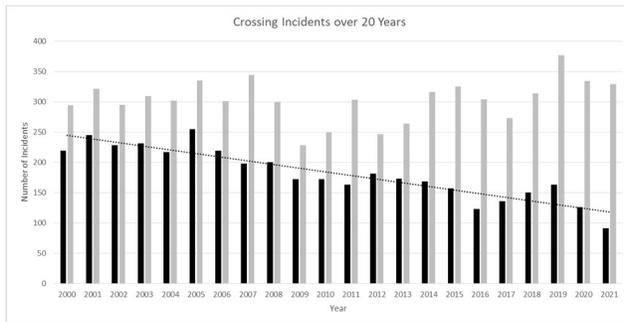


Figure 6. Yearly Crossing Incidents over 20 Years

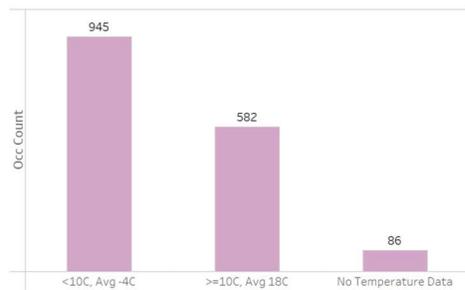


Figure 7. Railway Crossing Accidents and Temperature.

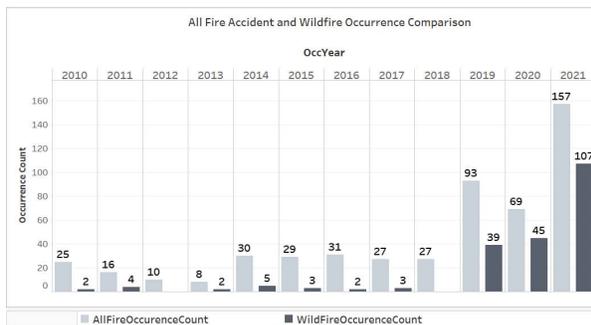


Figure 8. Fire Accident Trend Based on RODS Data.

By trending the overall fire accidents in last 10 years using TDIMA, Figure 8 shows that the main track fire accidents increased exponentially. In the first half of 2021, more than 140 fire accidents were reported which is higher than the sum of 2019 and 2020. There are two

types of fire accidents, equipment related and right of way fire (or wildfire). Using text mining techniques, number of wildfire accidents was estimated and included in the TDIMA analytics interface (see Figure 8). It shows that this increase was mostly contributed by the wildfires along the right of way on the railway tracks. As described in some accident summaries, the equipment fires from the train and rail maintenance vehicles spawned the wildfires along the right of way.

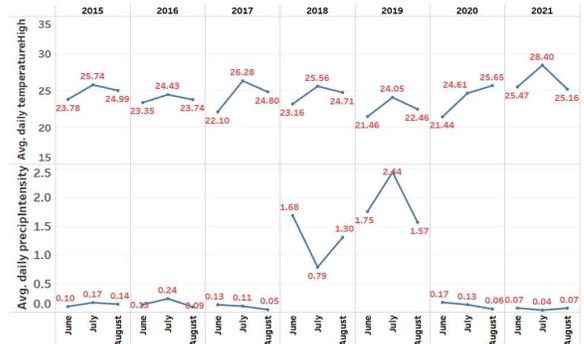


Figure 9. Temperature and Rainfall Trend for High Fire Occurrence Locations.

This case study indicates that recent climate change poses new challenges to the rail sector in Canada especially in the summer season. The daily weather data was collected from January 2015 to September 2021 for the locations with a high number of fire accident occurrences in the recent years. The analysis result of Figure 9 from the historical weather data indicates that the highest monthly average of “daily high temperature” occurred in July 2021, and the lowest monthly average of “daily precipitation intensity” also happened in July 2021.

3.5 Broken Rail and Wheel Analysis Scenario

This example involves studying potential correlation between broken rails and wheels and temperature. This analysis can be easily done using the TDIMA tool. Within the tool, the broken rail and broken wheel cause categories can be first selected. Next, the temperature needs to be selected from the list of available variables. The graph shown in the top plot Figure 10 is what is displayed by the tool. The points on this plot are located according to the date of the occurrence (x axis) and the ambient temperature at the time of the occurrence (y axis). The bottom plot of Figure 10 shows a statistic for each year. A list of statistics such as average, standard deviation, median, 95th percentile, and others are all options, but in this case, the average is used. This results in the lower plot seen in Figure 10, which shows the average temperature for all broken rails and wheels that occurred in each year from 2011 to 2019. The reason it stops at 2019 in this case, is because data about the cause of an accident could only be obtained from 2006 to 2019 as stated in Section 2.7. With access to more data, the analysis could be expanded to more recent years. This analysis indicates that the broken rails and broken wheels tend to occur at low temperatures.

This result should not come as much of a surprise and agrees with the industry observation that the broken rails and wheels occur more frequently under cold weather conditions. However, this case study demonstrates how, with just a few clicks, the tool allows researchers to discover trends and correlations efficiently.

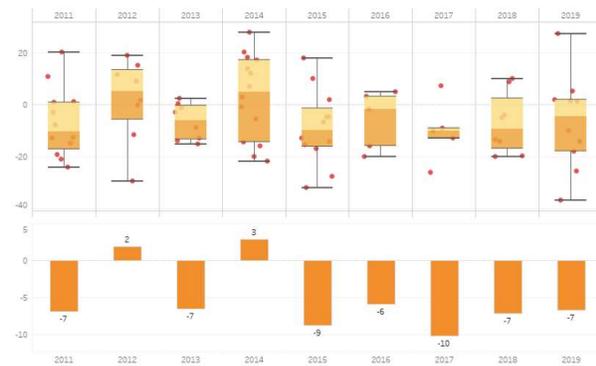


Figure 10. TDIMA Tool Broken Rail and Wheel Example.

3.6 Monitoring Latest Occurrences

In addition to performing analyses on past accidents, the TDIMA tool allows for quickly searching and monitoring the latest occurrences. An interface within the tool was designed to select filters for searching through RODS data to find all the recent occurrences that meet those conditions. Once this is done, individual occurrences can be selected from the resulting list and all the information about the accident is shown on the interface (Figure 11). As the tool contains the most recent data from the Daily Notification Reports, it can be useful for users to monitor recent occurrences and quickly review the accident summary and its geographic location on the map. The impact information is also included, where “P” denotes how many people living at the location and “W” means the water area. The impact variables shown are

- P-DG: product of people and involved DG cars
- P-RDG: product of people and released DG cars
- W-DG: product of water area and involved DG cars
- W-RDG: product of water area and released DG cars

4 CONCLUSIONS

Based on both historical derailment data and potential consequences, a tool called Train Derailment Impact Map and Analytics (TDIMA) for Canada’s national railway network is developed. The database of the tool is the integration of multiple datasets available in public domain, including RODS, NRW, CESUS, weather and others. Besides the historical number of derailments, the potential consequence was integrated as an additional element in the tool. This is the key difference between the developed tool and the existing ones that use number of accidents as the main variable to trend the risk.

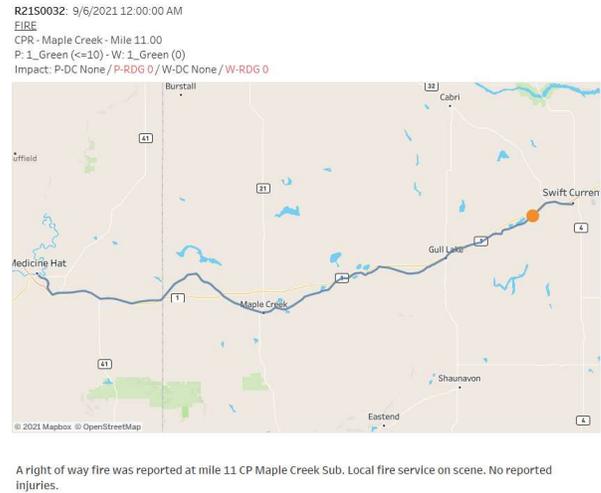


Figure 11. Recent Occurrence Browser of TDIMA Tool.

Multiple smart interfaces were designed to facilitate visualization-style analytics based on the integrated database. With a few clicks, the tool can be used to perform analytics effectively. The following are some key findings from case studies conducted using the developed tool.

The trend of rail safety obtained only by number of past derailments can be different from that obtained by trending the impact variables defined in the present study. The assessment using impact variable of DG release agrees with the observed impacts to public health by DG release derailments since 2013. However, the identified difference doesn’t suggest that the large number of released DG cars in remote area poses no risk. It has to be considered as a strong alert for all stakeholders to prevent such DG release and prevent it from happening in high population areas.

Ranking the cause of derailments by impact (derailed cars X people) resulted in different top causes than that ranked by only using number of derailed cars. The result suggests that brake and switch issues should be placed as high priorities in order to reduce derailment impact. While more studies are required, the present results provide the industry and regulator a different perspective to review the priority of equipment and track maintenance,

After including weather information in the analytics, it was found that the cold weather and other unfavorable weather conditions could make grade crossing prone to collision accidents. A new trend for fire accidents identified by the TDIMA tool indicates that climate change poses new challenges to the rail sector in Canada especially in the summer season.

With just a few clicks, the tool allows researchers to discover correlations between broken rails and wheels and low temperatures efficiently. This result agrees with the industry observations. As the tool contains the most recent data from the Daily Notification Reports, it can be used to review and monitor the latest occurrences on an interactive map.

5 ACKNOWLEDGEMENT

Authors would like to thank the support of Rail Vehicle and Track Optimization (RVTO) and Resilient Ground Transportation (RGT) programs of NRC. The project is also funded by NRC's minor capital project A1-017839. Valuable feedbacks from TSB and TC are greatly appreciated.

6 REFERENCES

Apostolakis, G. E. (2004). How Useful Is Quantitative Risk Assessment? *Risk Analysis*, 515-520.

Kaplan, S., & Garrick, B. (1981). On The Quantitative Definition of Risk. *Risk Analysis*, 11-27.

Leishman, E. M. (2017). *Analysis of Canadian Train Derailments from 2001 to 2014*. a thesis for Master of Science in Geotechnical Engineering, Department of Civil and Environmental Engineering, University of Alberta.

Möller, N., Hansson, S. O., & Holmberg, J. E. (2017). *Handbook of safety principles*. Hoboken, NJ: Wiley.

Natural Resources Canada. (2016, 10 1). *National Railway Network - NRW - GeoBase Series*. (Open government) Retrieved 9 3, 2021, from <https://open.canada.ca/data/en/dataset/ac26807e-a1e8-49fa-87bf-451175a859b8>

Natural Resources Canada. (2018, 06 05). *Elevation API*. Retrieved from <https://www.nrcan.gc.ca/science-and-data/science-and-research/earth-sciences/geography/topographic-information/web-services/elevation-api/17328>

Statistics Canada. (2016). *Aggregate Dissemination Areas, Cartographic Boundary File - 2016 Census*. Retrieved from <https://open.canada.ca/data/en/dataset/1e15ed74-7c58-4f6e-86b3-2e4283e0bbf6>.

TABLEAU SOFTWARE, LLC. (2021). Tableau Desktop Version 2021.1.

Transport Canada. (2016, 02 13). *Edition of the Emergency Response Guidebook*. Retrieved from <https://tc.canada.ca/en/changes-2016-edition-emergency-response-guidebook-erg>

Transport Canada. (2021, 07 13). *Rail Network and Crossings*. (Transport Canada) Retrieved July 29, 2021, from Rail Network and Crossings: <https://tc-tdg.maps.arcgis.com/apps/webappviewer/index.html?id=a540c1fa8c6146e0a6ed4fb39dde34a5>

Transportation Safety Board of Canada. (2021). *Rail transportation safety investigations and reports*. (Transportation Safety Board of Canada) Retrieved July 29, 2021, from <https://www.bst-tsb.gc.ca/eng/rapports-reports/rail/index.html>

Visual Risk Technologies, I. (2013). *Hazardous Materials Transportation Risk Assessment: State of the Practice*. WASHINGTON, D.C.: TRANSPORTATION RESEARCH BOARD.

Session 5

ENHANCED TRAIN CONTROL AND OPERATIONS

Assessment of the LRAIL system to detect and analyze the condition of angle bars on a 260-mile railroad

Luc Faucher, Tony Ducheman, Magali Jarry and Mathieu Cody
Centre d'expertise ferroviaire RAIL - Cégep de Sept-Îles, Sept-Îles, Québec, Canada

David Hébert
ArcelorMittal Infrastructure Canada, Port-Cartier, Québec, Canada

ABSTRACT

Operating trains in cold region draws many challenges regarding the maintenance-of-way. This is particularly true for heavy haul operators, who manage long convoys composed of short wagons filled with iron ore. Changes in temperature, from -60°C during winter to over 30°C in summer add stress to the railroad, which is already impacted by the heavy loads being carried. Angle bars represent a weak link in the railroad infrastructure and are greatly affected by these stresses. Part of a partnership between ArcelorMittal Infrastructure and the *Centre d'expertise ferroviaire RAIL* of the *Cégep de Sept-Îles*, the LRAIL system from Pavemetrics was tested to assess the conditions of all the angle bars on a 260-mile railroad. The LRAIL uses a Laser-Crack Measurement System and captures 3D images of the track, with a 1 mm X and Y, and 0.1 mm Z resolution. With artificial intelligence algorithms, the system detects and locates angle bars, the number of bolts and the measurement of the gap at the joint. To better visualize the data, a software using business intelligence principles was designed to confirm and display the data taken in the field. The interactive dashboards also help rail operators prioritize their angle bar maintenance interventions regarding where to weld first, and how to avoid slow orders that could occur from angle bar fractures. Various warning signs are observed as an indication that angle bars are in the process of breaking or failing. Preliminary results show that the technology tested helps to adequately detect these warning signs and help railway operators act upon them to avoid unwanted failures.

1 INTRODUCTION

With the continuous increase in traffic, length and heaviness of convoys, rail operators must find ways to inspect and maintain railways quickly and efficiently. To meet this objective and increase rail safety, many technologies were developed to assess the health of track components. Some use machine vision to focus on crossties, ballast levels, joint bar cracks, fasteners, rail corrosion, rail wear, rail geometry, etc (Aurora Xiv, 2021; Orrell, 2009; Berry et al., 2008; RailCheck, 2021; MERMEC, 2021). Others use mechanics to evaluate rail geometry (Andian, 2021). For more than twenty years in Canada, the two main causes of declared derailments are either due to the geometry of the track or the breakage of the rail and its welds (TSBC, 2021). The development of improved technologies is essential to reduce the level of derailments and injuries.

An arising Canadian technology, the LRAIL system (Pavemetrics, 2021), has been developed to evaluate a wide range of track components in a single survey. The system uses machine vision according to the principles of a Laser-Crack Measurement System (LCMS) that was previously developed to inspect roads and landing

strips. Previous studies showed its efficiency to also detect and evaluate rail geometry, rail wear, rail defects, spikes, fasteners, rail anchors, ballast, crossties, and joints (Fox-Ivey et al., 2020a; Fox-Ivey et al., 2020b). For this study, the system was operated at hi-rail inspection speeds and captured a scan with 1 mm surface resolution (X and Y) and 0.1 mm depth (Z) resolution.

This study aims at evaluating the technology for angle bar detection on a 260-mile railroad in Northern Québec, at an ArcelorMittal facility (ArcelorMittal, 2021). During operations, angle bars face various strains coming from the heavy load convoys carried and the climate differences on rails. The impacts of a rail seating incorrectly on the plates or on bad ties, just like the stress caused to the rail due to temperature dilatation at spots where anchorage is inappropriate, will lead to a premature wear or break. The data presented in this work shows that the LRAIL system is apt at detecting angle bars, their location (GPS coordinates) and the number of bolts in place. Joint gap and broken angle bars can also be seen on images, but the efficiency of the system in correctly identifying these parameters was not the purpose of this study.

2 METHODOLOGY

The study was conducted according to this methodology:

2.1 The track

The track studied is the one of ArcelorMittal in northern Québec, Canada. It is a Class 3 track mainly used to carry over 26 million tons of concentrated iron ore every year from the mine of Mont-Wright to the port of Port-Cartier. The northern part of the track, a 30-mile section, has a high level of traffic. It is also used to carry 16 million tons of ROM from Fire Lake to Mont-Wright annually, implying the transport of 80 MGT every year on this section. The track is made of continuous welded rail (CWR) and rail grading is 136 lbs. The length of angle bars is 36 inches and they are usually fixed with 6 bolts for permanent joints, and with 4 bolts for temporary joints that will be welded before winter. The survey was conducted on the main line during the month of July 2021, a time of year when both temporary and permanent joints are present on the track.

2.2 The LRAIL system

Three-dimensional scans were captured with LCMS sensors mounted at the back of a hi-rail pickup truck. The LCMS sensors are capable of scanning at speeds up to 28 kHz. The 3D lines were merged into 1 m x 3.6 m images. For a track of 260 miles, this represents over 412 000 images. An inertial navigation system integrates the picture's latitude, longitude, and elevation into each 3D profile. After the survey, the images were all analyzed by a trained Deep Convolutional Neural Network (DCNN). The results can be directly displayed on images, in xml files or in shapefiles. For this study, the information concerning angle bars was extracted from xml files and put into a database. The data was then represented on a dashboard where 2D and 3D images were made available for interpretation.

This study will compare the results obtained with two different versions of a library. The first version was developed before this survey (Library 4.78.4.2) in June 2021, while the second one was updated during this study (Library 4.78.9.7) in July 2021, but with data coming from other railways. For results obtained with both libraries, a human expert has confirmed the initial results obtained for every angle bar detected by the DCNN (Figure 1) in all categories, except for the one of 6 bolts. A sample of 30% of these were analyzed.

2.3 Evaluating the technology

The performance of the technology to detect angle bars and the number of bolts was tracked in terms of precision, sensitivity (recall), specificity and accuracy. These four indicators are calculated as follows:

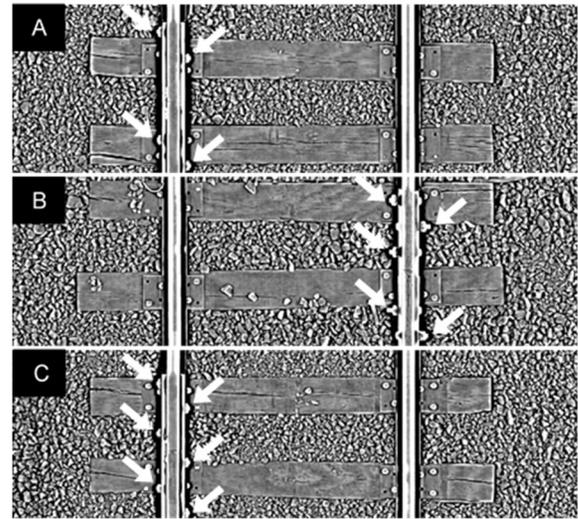


Figure 1. Angle bars detected by the system and having (A) 4, (B), 5 and (C) 6 bolts.

$$Precision = \frac{TP}{FP + TP} \quad [1]$$

$$Sensitivity(Recall) = \frac{TP}{FN + TP} \quad [2]$$

$$Specificity = \frac{TN}{TN + FP} \quad [3]$$

$$Accuracy = \frac{TP + TN}{All\ subjects} \quad [4]$$

TP stands for true positive, TN for true negative, FP for false positive and FN for false negative. The precision expresses the proportion of units the DCNN model says are positives and that are in fact positives. It expresses how much the model can be trusted when it predicts an individual as positive. The sensitivity measures the model's predictive accuracy for the positive class. It means the ability of the DCNN to find all positive units in the database. The specificity is the correctly negative labeled by the model to all those who are negative. Finally, the accuracy is the ratio of the correctly labeled subjects to the whole pool of subjects (Grandini et al., 2020).

To determine the ability of the DCNN to detect angle bars, a section of 3.5 km was analyzed. This section represents a sample of 3 500 images. It turns out that 30 angle bars were present. The expected results are then 30 TP for 3 470 TN. If the DCNN misses the identification of an angle bar, the image will be categorized as a FN. In counterpart, if the DCNN establishes an angle bar where there is none, it will be categorized as a FP.

To establish if the DCNN succeeds at counting the

number of bolts present, the same process was done, but with a random sample representing 30% of all angle bars detected. A confusion matrix for a multi-class classification was produced. The indicators were then calculated for each category, and a weighted average for each indicator was calculated to get an overall view.

Table 1. Precision, sensitivity, specificity, and accuracy of both libraries to detect angle bars.

Indicator	Library 1	Library 2
Precision	100%	96.7%
Sensitivity	96.7%	96.7%
Specificity	100%	100%
Accuracy	100%	99.9%

3 RESULTS AND DISCUSSION

3.1 The detection of angle bars

The indicators for both libraries are displayed in Table 1. For each version, 29 of the 30 angle bars were detected. Library 2 results are a little bit lower, because a false positive was detected. The FP was an angle bar at the junction of a siding in a switch. It was an angle bar, but it was not located on the main line. Despite this error, the DCNN has indicators over 96.7% in each category, which makes it reliable to detect angle bars.

3.2 The count of bolts

Once an angle bar is detected, the DCNN evaluates the number of bolts tightening it. Table 2 shows the initial count of bolts for each angle bar and with both libraries. It also shows the quantity of angle bars that were removed (false positive), and the ones added after a confirmation by an expert (false negative). The last column represents the final count for each library. According to the DCNN, the number of bolts calculated with Library 1 varied from 0 to 9, while it changed from 2 to 6 with Library 2. After the expert reviewed the data, it turned out that all angle bars confirmed had only 4, 5 or 6 bolts. These three categories were the ones expected. The category of 4 bolts is the one of temporary joints that will be welded before winter. The category of 6 bolts represents the permanent angle bars on the track, while the ones with 5 bolts are permanent too but are the ones that previously had 6 bolts. It means that one bolt broke over time due to operations. Finally, there were no angle bars with 3 bolts. If there would have been an angle bar with 3 bolts, an immediate slow order would have been called to respect the speed of a Class 1 track (TC, 2011).

Table 3. Precision, sensitivity, specificity, and accuracy of both libraries to detect the number of bolts.

Indicator	Library 1 Weighted average	Library 2 Weighted average
Precision	82.9%	97.6%
Sensitivity	82.9%	97.6%
Specificity	80.6%	91.7%
Accuracy	86.4%	98.4%

The angle bars that were detected with 0 or 1 bolt with Library 1 were whether switch components or rail lubricators. They were all discarded. Both libraries had results with 2 or 3 bolts. All samples were angle bars, but there was a miscount in the number of bolts. Library 1 also shows results with 7, 8 or 9 bolts. They were all angle bars with miscounted bolts, except for 4 that were not angle bars, and then rejected (not added in another category).

In its entirety, most of the results were in the 4, 5 and 6 bolt categories. However, all angle bars should have been solely sorted in these three categories, without exception. The percentage of angle bars initially sorted in these categories went from 87.4% in the first library to 99.0% in the second one, which is a big step up. However, to really compare the liability of libraries, the precision, the sensitivity (recall), the specificity and the accuracy of both libraries were calculated. The indicators were determined by selecting a random sample representing 30% of all angle bars. The indicators were then calculated for each category, and a weighted average for each indicator was drawn.

As seen in Table 3, all indicators stepped up between libraries. The increase in precision mainly came from the disappearance of false positives in the categories from 7 to 9 angle bars. The increase in sensitivity was driven by a better counts of bolts in the categories of 4 and 6 bolts. The many false negatives became rightly counted (TP) due to the better interpretation coming from Library 2. As TP increases and FN decreases, the sensitivity went from 82.9% to 97.6%.

The gain in value for specificity and accuracy came from better results in the category of 6 bolts. For specificity, less false positives were present (angle bars that really have 6 bolts but were interpreted in another category by the DCNN). The gain was of 11% overall. Finally, the accuracy rise was led by the increase of true positives and true negatives. A gain of 12% was measured.

Table 2. Number of bolts detected with Library 1 and 2

Number of bolts detected	Library 1				Library 2			
	Initial count	Removed	Added	Final count	Initial count	Removed	Added	Final count
0	2	2	-	-	-	-	-	-
1	2	2	-	-	-	-	-	-
2	1	1	-	-	1	1	-	-
3	2	2	-	-	3	3	-	-
4	248	4	25	269	288	3	5	290
5	24	18	5	11	24	15	2	11
6	1771	22	291	2040	2006	5	11	2012
7	2	2	-	-	-	-	-	-
8	57	57	-	-	-	-	-	-
9	228	228	-	-	-	-	-	-
TOTAL	2337	338	321	2320	2322	27	18	2313

3.3 Joint gap and broken angle bars

The next step in the project will be to determine the indicators regarding the analysis of joint gap. The four indicators will be looked at in the same way. Also, a next step in the project is to teach the DCNN to recognize broken angle bars. During a previous survey, an angle bar was broken on the track. As seen on Figure 2, this angle bar can easily be seen on 2D and 3D images. A further step will be to have the DCNN recognize this type of defect.

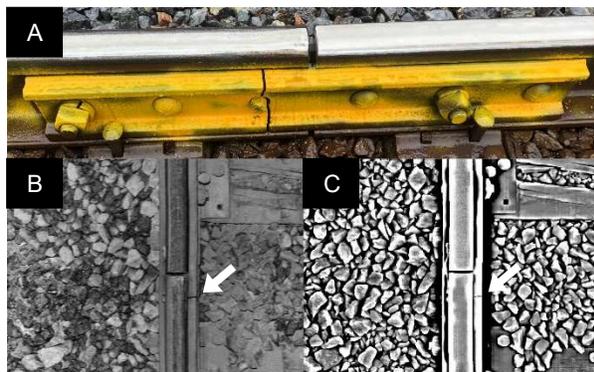


Figure 2. (A) A broken angle bar can be seen on (B) 3D and (C) 2D image representation of a scan by the LRAIL

4 CONCLUSIONS

In conclusion, the performance of the DCNN developed with the LRAIL system to identify angle bars and count the number of bolts meets expectations. Precision, sensitivity, and accuracy to detect angle bars were all over 96.7%. With Library 2, precision, sensitivity and accuracy were over 97.6%. The false positive has diminished the specificity to count bolts at 91.7%, which is still a good result. As many false positives came from real angle bars detected on the wrong path in switches, a better recognition of the main line should be the next step to improve specificity.

5 ACKNOWLEDGEMENT

This study was financed by NSERC, Transport Canada, ArcelorMittal, the Government of Québec, the *Centre d'entrepreneuriat et de valorisation des innovations* and the *Cégep de Sept-Îles*. The authors would like to thank Christopher Braun, Josué St-Amant, Yoland Gallant and Valérie Boucher for their contribution to the surveys, Simon-Olivier Vaillancourt for his work as an intern, as well as Benoit Petitclerc, John Laurent, Thanh Nguyen, Mario Talbot, Sylvain Garant, Jean-François Hébert and Richard Fox-Ivey at Pavemetrics.

6 REFERENCES

- Andian. Accessed September 14, 2021, from <https://www.andian.com/?page=products>.
- ArcelorMittal. Accessed September 14, 2021, from <https://www.transformerlavenir.com/installations/installations-damic/le-chemin-de-fer/>.
- Aurora Xiv® by LORAM. Accessed September 14, 2021, from <https://loram.com/inspection-and-optimization/inspection-services/tie-inspection-services/aurora-xi/>.
- Berry A. & al. (2008) High speed video inspection of joint bars using advanced image collection and processing techniques. Conference: World Congress on Railway Research, Seoul, South Korea.
- Fox-Ivey, R., Nguyen T., Laurent, J. (2020a) Laser Triangulation for Track Change and Defect Detection. Report to the Federal Railroad Administration. Washington, DC. USA.
- Fox-Ivey, R., Nguyen T., Laurent, J. (2020b) Extended Field Trials for LRAIL for Automated Track Change Detection. Report to the Federal Railroad Administration. Washington, DC. USA.
- Grandini, M., Bagli, E., Visani, G. (2020) Metrics for Multi-Class Classification: An Overview. A White Paper. 17 pages.

Orrell, S.C., Nagle, J.A. & Villar, C. (2009) System and Method for Inspecting Railroad Track. US Patent 7.616,329 B2, Austin TX, USA.

MERMEC group, Rail Corrugation. Accessed September 14, 2021, from <https://www.mermeccgroup.com/inspect/track-measurement/61/rail-corrugation.php>.

Pavemetrics. LRAIL. Accessed September 14, 2021, from <https://www.pavemetrics.com/applications/rail-inspection/laser-rail-inspection-system/>.

RailCheck. Accessed September 14, 2021, from <http://bvsys.de/58.html>.

TSBC, Transportation Safety Board of Canada. Rail Transportation occurrences in 2020. Accessed September 14, 2021, from <http://www.bst-tsb.gc.ca/fra/stats/rail/2020/sser-ssro-2020.html>.

TC, Transport Canada. (2011) Rules Respecting Track Safety. Government of Canada.

Development of railway terminals within the context of precision scheduled railroading principles

Richard S. Lanyi, P.Eng., FCSCE

Lanyi Rail Solutions Ltd., Edmonton, Alberta, Canada

ABSTRACT

Rail transportation is a fundamental component of Canada's Economic Corridors, supporting our economy and driving growth potential. Recent disruptions to our national railway service have highlighted the importance of this service to the health of our economy and wellbeing of our nation.

Systematically, our railway networks are comprised of nodes (terminals) and links (rail lines). Efficiency of the network involves efficiency of both Industry terminals and rail lines. Functionally, Railways have adopted the principles of Precision Scheduled Railroading (PSR) in their business and operating practices. PSR's objectives are to create a safe, reliable, cost-effective, efficient, and timely transportation service. It involves controlling costs, maximizing asset utilization, highly precise operations planning, and synergy across all processes.

PSR focuses on the carload (customer shipment). It optimizes the rail network from an endpoint-to-endpoint perspective. Trains must run on time to meet customer commitments. When focusing on carload, Railroaders examine every process effecting on-time delivery, constantly fine tuning.

Industry terminal operations are a key component of the overall railway transportation service. PSR has changed the way Railways operate. It demands that Industry also change how they plan, design, and operate their terminals. This paper will discuss:

- Interfaces between Industry terminal and Railway operations
- Industry terminal operating practices through the lens of PSR
- Improvements in the planning, design, construction, maintenance, and operations of Industry terminals
- Use of new technologies to further improve safety, efficiency, and cost effectiveness of terminal operations

The goal of PSR is excellence in rail transportation service through safe, efficient, predictable, and reliable operations. For the customer, this translates into faster transportation service, lower transportation costs, overall supply chain stability, and reduced demand for transportation assets. For the Canadian economy, this translates to a competitive advantage in global and domestic marketplaces, lower costs for our goods and services, and opportunities for sustainable growth.

1 INTRODUCTION

A supply chain is a cycle commencing with the sourcing of raw materials to the *delivery* of finished products to end-users. The cycle begins with sourcing raw materials which are *transported* by a logistics provider to a supplier, which acts as the wholesaler. The materials are *transported* to a manufacturer that refine and process them into a finished product. These products are *transported* to a distributor that wholesales the finished product, which is then *delivered* to a retailer. The retailer sells the product to consumers. Consumption completes the cycle, but creates demand that drives the production of more raw materials, and the cycle continues.



Figure 1. Supply Chain Cycle (FCI Web site)

The supply chain is a fragile system dependent upon numerous parties to function and communicate efficiently and cooperatively. A disruption in any link will impact the entire chain resulting in systemic delays which drive up costs. This system relies on safe, efficient, cost effective, reliable and dependable transportation services to keep things moving.

Rail transportation is a fundamental component of Canada's Economic Corridors, supporting both domestic and international supply chains, driving our economic growth potential. The health of our national economy is often gauged by the volume of goods being transported.

Due to the expanded reach of railway networks, local disruptions to rail traffic will translate into extensive delays over thousands of kilometres resulting in congestion of rail lines, yards, and terminals. These delays will also translate into a higher demand for resources such as railcars, locomotives, and train crews. Recovery from short term delays locally, will require extended timeframes to re-establish fluidity across the network.

Over the past 15 to 20 years, North American Railways have developed and adopted the principles of Precision Scheduled Railroading (PSR) to guide their business and operating practices. The reliability of efficient and fluid railway networks, however, relies on the seamless interface between rail yard and terminal operations, and rail line operations. Industrial processes and railway operations are an integrated fluid process designed to keep things moving safely, efficiently, and cost effectively. For the benefits of PSR to be truly realized, it demands that Industry also adopt these principles related to how they plan, design, and operate their plants and terminals. For these reasons, Terminal operating practices, as well as, engineering planning and design, construction, and maintenance practices, all need to be evaluated through the lens of PSR.

Why does this matter? The goal of PSR is excellence in rail transportation service leading to a more valued and profitable transportation industry through safe, efficient, predictable, and reliable operations. For Industry, this translates into faster transportation service, lower transportation costs, overall supply chain stability, and reduced demand for transportation assets.

Why is this important? Recent disruptions to our national railway service have highlighted the importance of rail service to the supply chain cycle, the health of our economy, and wellbeing of our nation.

For the Canadian economy, this translates to a competitive advantage in global and domestic marketplaces, lower costs for our goods and services, and opportunities for sustainable growth. Collectively, everyone benefits.

2 PSR – HOW DOES IT WORK? A RAILWAY PERSPECTIVE

PSR's objectives are to create a safe, cost-effective, efficient, and timely transportation service. It involves controlling costs, maximizing asset utilization, highly precise operations planning, and synergy across all processes. PSR is based on optimizing the rail network from a comprehensive end to end perspective of operations and management. This practice deviates from a more regional, and internally competitive, perspective of railway operating practices.

PSR focuses on the carload (customer shipment) rather than the train carrying the carload. Velocity and train length are still important, however, the focus on moving cars takes precedence. Trains must run on time to meet customer commitments. When focusing on carload, every process effecting on-time delivery must be optimized, and constantly fine tuned (plan – execute – monitor – analyze – revise the plan).

Train delays result in network congestion and increased terminal dwell time and congestion resulting in an increased demand for assets (railcars, locomotives, and train crews), higher costs, and less dependable service. PSR promotes higher efficiency, network fluidity, and a well-balanced operation leading to lower costs and more reliable service. A balanced operation ensures the right resources are in the right locations at the right times. This results in shipments moving with discipline, precision, and synchronization.

2.1. PSR's Guiding Principles

North American Railroads have established a fairly common set of guiding principles as the foundation of operating and business efforts. These principles are constant in planning and execution, globally applicable, and form the business and cultural context of the company. These principles are all related and co-dependent. The belief is that strict adherence to these guiding principles will lead to a high level of customer satisfaction and exceptional financial performance. The following are typical examples of the guiding principles adopted by Railroads practicing Precision Scheduled Railroading.

Safety – is essential in protecting the public, the environment, employees, and the company, which translates into good business; the more predictable an operation or process is, the fewer exceptions will occur, thus ensuring a high level of safety

Timely Service – encompasses the belief that a customer who pays for timely service should receive timely service; from a Railroad perspective this is known as “doing what you say you are going to do”.

Cost Control – focuses on executing continually refined, fine-tuned, processes in synch with other processes, removing redundancies; it does not promote indiscriminate cost cutting

Asset Utilization – the more efficient the asset, the fewer assets an operation requires; assets only provide a return when serving the intended purpose, otherwise they are a liability; assets need to earn their keep (idle cars cost money and do not make money).

People – PSR recognizes that people are the foundation in which all other principles are built.

2.2. PSR's Service Design Principles

Service Design Principles are subordinate and complementary to PSR's Guiding Principles, and most directly support Asset Utilization. Due to the complex inter-connectedness of railroad operations, optimizing one principle does not necessarily optimize the others. Too great a focus on any one principle can have a negative impact on the others. Balance is the key.

2.2.1 Optimize Car Asset Utilization

Optimizing car assets include efforts to minimize car dwell in yards and minimize car handling. Minimizing car dwell reduces the cost impact in consumed yard capacity, and the overall number of cars needed. Minimizing car handling improves the car cycle, increases car velocity, allowing more productive car use, ultimately reducing the car fleet while optimizing costs.

2.2.2 Optimize Fuel Efficiency, Power Requirements, and Train Builds

A locomotive is the most expensive piece of rolling stock on the railroad, and fuel is one of the highest operating costs. Overpowering a train typically generates more waste than productive speed. The objective to fuel efficiency and optimized power requirements are to right-size the pulling horsepower with the trailing load, taking into consideration the terrain and relative average speeds.

Optimizing train builds focuses on maximizing train length by territory, thus minimizing train starts which reduces demand for locomotive power, fuel, and train crews while reducing demand on network capacity.

From an operating perspective, the Service Plan governs power utilization by minimizing unused slots to maximize train efficiency, and by ensuring downstream yards are receiving the power they require to maintain their operations. Further, locomotives have prescribed maintenance requirements and schedules, as such, they need to be in the right place at the right time.

2.2.3 Manifest and Unit Train Service

There are two basic types of train service – general-purpose or manifest, and special-purpose or unit trains. Unit trains are often considered an efficient, low-cost alternative to manifest service. This is based on the belief that unit trains provide faster transit times by avoiding intermediate handling thus saving operating expenses.

Often, a total cost approach, involving a detailed operations analysis, can reveal that the opposite is true, if special conditions for unit trains are not met. Unit trains are most efficient and economical when the following conditions apply:

- Loading and unloading operations are completed on a daily basis (within 24 hours)
- Daily carloads are in the range of 100 or more
- Trains cycle in a closed loop between a single origin / destination pair
- Trains operate 7 days a week
- Locomotives remain with the train consist

If any condition is not met, the economics and efficiency of the service diminishes. Inefficiencies can arise due to increased in-transit inventory, poor car velocity, and locomotive and crew imbalances. PSR utilizes mixed traffic, which minimizes the number of train starts and frees up capacity.

2.2.4 Optimizing the Service Plan

The Service Plan is focused on car deliveries, not trainloads. It can be optimized through the use of multiple traffic outlets and by balancing traffic by direction. Multiple traffic outlets involve having more than one way of moving cars to destination. Using multiple outlets increases flexibility and lowers operating risk. Scheduled trains do not necessarily have consistent carloads day to day. Some days they may run short, and overflow other days. Running short wastes train productivity, overflow risks service failure.

In optimizing the Service Plan, PSR strives for train balance (equal number in both directions) to further minimize costs and provide a higher level of service reliability. A balance accomplishes this by reducing empty car moves or asset repositioning, which do not produce revenue. It also avoids holding accumulated cars in yards for strategic departure risking congestion, and increased dwell time. It is focused on overall network operations performance over a regional operations success. Optimizing the Service Plan involves balancing train movements and fluidity to the fullest extent possible. It minimizes costs as well as excess train capacity.

2.3 Supply Chain Considerations

With Precision Scheduled Railroading, there are recognized constraints, not under the direct control of the Railroad, to achieving an optimized Service Plan and network operation. These may include operating efficiencies in the supply chain involving a customer's rail yard or terminal, port terminals, and interchanges with other Class 1 or Shortline Railroads. PSR focuses on establishing shared goals across the supply chain to optimize throughput for all, thus, benefitting all parties involved.

3 PSR - A RAILWAY TERMINAL PERSPECTIVE

3.1 Rail Terminal Components and Configurations

The purpose of the rail terminal is to provide an interface between Industry and the Railroad. This interface is comprised of the following basic components:

- Product loading/unloading tracks
- Railcar storage tracks
- Railroad arrival and departure tracks
- Spur track connecting the arrival and departure tracks to the Railroad's mainline

Rail terminals can have a variety of configurations based on physical site constraints, type of train service (manifest or unit trains), and terminal operations. The two most common configurations are ladder and loop tracks.

Figure 2 provides an illustration of a typical ladder track configuration, which includes supporting runaround and pull-back tracks, used to facilitate staging on inbound and outbound trains, and the movements of cuts of railcars between the loading/unloading, storage, and arrival/departure tracks. Ladder tracks are typically used with both manifest or unit train service.

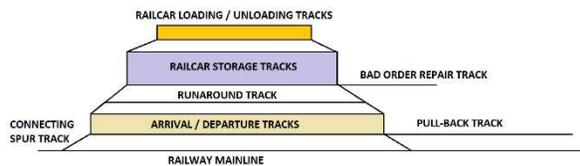


Figure 2. Typical Ladder Track Terminal Configuration

Consistent with their attention to PSR, Railroad's service is primarily focused on, what is commonly called, a "hook and haul" operation. This involves a scheduled arrival to deliver loads/empties to an empty arrivals' tracks, then pick up empties/loads from an adjacent departure track, and depart. Railroad's discourage services involving storage of customer railcars, or switching railcars within the terminal, by imposing significant surcharges for these "additional" services.

Figure 3 provides an illustration of a typical loop track configuration. In this configuration, inbound and outbound trains are accommodated on opposite sides of the loop, with loading/unloading at the middle of the loop. Storage and additional capacity can be accommodated with multiple loops. Loop tracks are typically used with unit train service. They typically require a larger terminal footprint than ladder tracks, though they require fewer turnouts, which are costly trackwork components requiring additional capital and higher maintenance costs. The operational benefits of loop tracks include fewer resources due to less switching in not having to disassemble or assemble arriving/departing trains, as well as, the opportunity of keeping the locomotive power with the train.

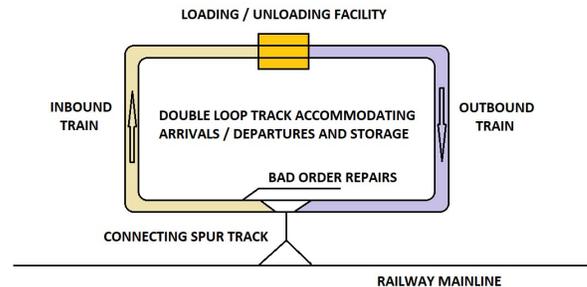


Figure 3. Typical Loop Track Terminal Configuration

As a means of establishing a consistent PSR approach between Railroad and Industry terminal operations, loop track configurations are best suited for the following conditions:

- Unit trains consisting of at least 100 cars
- Daily plant production of unit train capacity
- Seven day per week terminal/plant operations

3.2 Terminal Production

Terminal Production is the basis from which the details of fleet sizing, terminal design, and operations are derived. There are three (3) basic components to Terminal Production:

1. Product to be transported and product properties (i.e., density, solid, liquid, gas)
2. Production rate (annual, weekly, daily)
3. Frequency of rail service

3.2.1 Product Being Shipped

Product type and material properties will define

- Railcar type and capacity
- Terminal material handling equipment
- Material storage requirements
- Spill containment/response
- Terminal operating safety requirements

3.2.2 Railcar Fleet

From a PSR perspective, the objective of railcar selection is to minimize the number of railcars (assets) required to satisfy production and maintain operational reliability. This involves selecting railcars which have:

- The largest gross weight capacity allowed on the rail line
- The largest load limit capacity within the gross weight capacity class
- A consistent railcar length to suit fixed loading/unloading infrastructure configurations
- The shortest railcar length to minimize rail terminal track lengths and maximize railcar useable space

The railway network is made up of series of links and nodes. The links are comprised of mainlines and branchlines, where mainlines service national and regional economic corridors, and the branchlines feed the mainlines, as tributaries to a major river. The nodes

are comprised of terminals, whether they be mainline classification yards or Industrial terminals.

Mainlines and branchlines are broken down into discrete segments called Subdivisions. Each Subdivision is comprised of unique physical components of rail line, sidings, and double tracking, with unique horizontal and vertical (grade) alignment characteristics. They also have defined operating constraints based on these characteristics and the condition of railroad infrastructure (grade, drainage, track, structures, crossings, signals) which may limit allowable Subdivision railcar weight and train speeds. Allowable Subdivision railcar gross weights are classified by industry standard classes of gross railcar weights.

It is not uncommon for branchlines, which typically have less traffic, to be maintained to a lesser standard than mainlines, with reduced allowable railcar weights and slower train speeds. In North America, the standard maximum allowable mainline gross railcar weight is 286,000 lbs. This has set the standard regarding the manufacture of railcars and the supply of railcars within industry. Branchlines, however, can be limited to lower allowable gross railcar weights of 268,000 lbs. (7% reduction) and 263,000 lbs. (9% reduction). Some Subdivisions in the US (Union Pacific, Norfolk Southern) have increased their allowable railcar weights to 315,000 lbs. (10% increase) to handle bulk commodity unit trains. These services require a special fleet of heavier capacity railcars, and these Subdivisions require infrastructure (track and structures) which can handle these additional loads, as well as, a higher level of maintenance.

Reduced railcar capacity will result in the need for more railcars (more assets), which in turn will lead to longer tracks (more assets), and more time for switching and loading/unloading (less efficiency). All of these will drive up capital and operating costs. Reduced train speeds will result in longer origin/destination cycle times, which will also drive up the need for more railcars (more assets), also driving up capital and operating costs. As such, plant/terminal site selection, through the lens of PSR, can be key in managing railcar assets, rail terminal infrastructure and operations, as well as, material handling systems (loading/unloading).

In addition to determining fleet size based on production demand, an allowance for additional assets will be required to accommodate scheduled maintenance and unscheduled bad order repairs. This allowance can be in the order of +/- 5% of the production fleet. Additional assets also need to accommodate downstream terminal and Railroad service delays and disruptions (washouts, derailments, yard and mainline congestion, locomotive and crew shortages, etc.). These are typically assessed in terms of a number of days of reserve production capacity. These additional assets require additional infrastructure (storage) and resources to manage them, incurring additional costs.

Fleet management is essential in assessing utilization of railcar assets (service time, distance traveled, % loaded/unloaded), as well as, asset health (repair history, maintenance schedules). Today, sensors can be installed on railcars to record and monitor this data. Sensors can track location and measure load status, brake status, wheel and bearing performance, and hatch or door securement. As an example, Amsted Digital Solutions have created a tele-metrics platform designed to collect accurate, real-time data which can generate actionable and predictive intelligence which can be used to optimize fleet performance and number of assets required, improve safety, and reduce costs, while delivering the required level of service.

3.3 Rail Terminal Operations

Planning rail terminals begins with initial and forecasted annual production estimates (i.e., tonnes, barrels, gallons, etc.). From these, estimates for daily production can be developed such that daily terminal operations can be defined, along with the associated infrastructure, fleet size, and resource pool. Where future expansions are anticipated, it is essential that these terminal concepts be developed early in order to ensure that adequate space is available, initial configurations can be readily expanded without significant interruptions to existing operations, and that future operations can deliver the anticipated production levels.

Plant production, rail terminal operations, and Railroad services are integrated processes. Disruptions to any one potentially impacts the other two. Figure 4 shows a flow chart illustrating the interfaces between plant production and material storage, rail terminal operations, and Railroad product delivery and empty railcar returns.

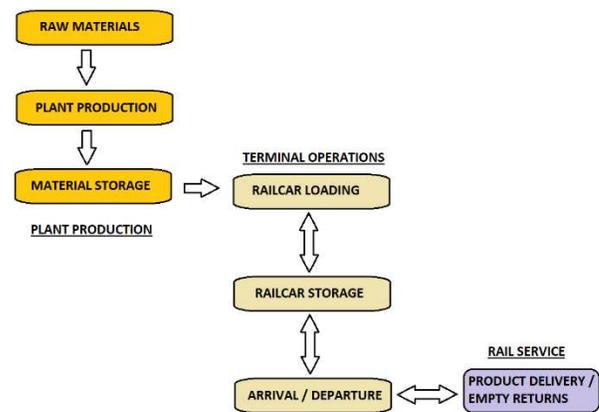


Figure 4. Rail Terminal Interfaces

3.3.1 Storage

To compensate for day to day variations in production and unplanned disruptions, storage is required. As illustrated in Figure 4, there are two types of storage required – material storage and railcar storage. Material storage looks after variability in Plant Production, while railcar storage looks after variability in rail service. From

a PSR perspective, both require assets and resources and attract costs to the operation. Efficient utilization of these assets is the key to controlling these costs while maintaining the desired level of service.

Storage capacity is typically measured in days of production. It is defined by the range between minimum and maximum volumes required to maintain a sustainable work flow and operation. Evaluation of storage capacity comes down to a risk assessment involving the frequency and magnitude of shortages related to plant, rail terminal, and rail service operations, and their associated costs.

Common forms of material storage include tanks, for liquids; bins and stockpiles, for bulk commodities; and stacks for containerized products. Excess storage is not efficient utilization of assets, though it does not produce a risk to plant, rail terminal, or rail service operations. Insufficient storage will shut down the plant and create disruptions to normal plant operations, but not rail terminal or rail service operations.

Railcar storage requires tracks and real estate, along with equipment and resources to switch out railcars. Excess storage is not efficient utilization of assets which does not produce a risk to plant, rail terminal, or rail service operations. Insufficient storage imposes considerably more risk impacting plant, rail terminal, and rail service operations. As previously noted, allowances need to accommodate scheduled maintenance, unscheduled bad order repairs, downstream terminals, and Railroad service disruptions. Keep in mind that downstream terminal and Railroad operations are outside the control of the product producer.

Insufficient access to railcars will result in a reduction in daily railcar loading capacity, which will translate into an increase in material storage, which when exceeded, will result in a slow down, and eventual shut down, of plant production.

Consider the following production with manifest rail service:

- A = 50 railcars loading per day
- B = 15 day transit cycle time
- C = 1 day additional terminal storage
- D = 1 set of railcars on the loading track
- E = 5% additional railcar allowance

Estimated Fleet Size (F) = $A*(B+C+D)*(1+E)$ [1]

Using this simplified approach, the estimated fleet size would be 893 railcars. Of these, 700 (78%) are in transit (outside the terminal) leaving 150 (17%) in the terminal and 43 (5%) being repaired or maintained. This is assuming the terminal is holding one set of 50 railcars on the loading track, one set of railcars on the storage tracks, and one set of railcars has been delivered to the arrivals track.

In this example, as long as the terminal operates 7 days per week, and there are no service disruptions, the flow of railcars will cycle in such a fashion as not to generate additional railcar storage in the terminal. Should this operation work only 5 days per week (Monday to Friday), the following imbalances in railcar flow will occur at varying times:

- Some cuts of 50 railcars will arrive on weekends, thus requiring storage and additional resources to receive these railcars
- Some days of production will not have returning cuts of 50 railcars, thus requiring cars from storage
- Some days will see two (2) cuts of 50 railcars arriving, thus requiring storage for at least one cut

This example highlights the benefits prescribed by PSR service principles involving a steady 7 day per week work flow in the rail terminal. With only one day storage, this terminal could run out of capacity should any other service delays occur.

Also from this example, unit train service of 100 railcars could be considered every second day, assuming a 7 day per week operation. This would have no impact on the railcar loading operations, but would require additional tracks to store the extra days production. Further, it could influence the consideration of a loop track configuration as a means of simplifying the rail terminal operations. As previously highlighted, other than daily unit train service is not considered consistent with PSR service principles. Alternating unit train service will produce inconsistencies in train balance creating the potential for congestion along the mainlines and in classification yards, as well as, imbalances in demand for locomotive power and train crews.

3.3.2 Terminal Production

With daily plant output production defined, the following terminal planning issues can be assessed and decided:

- The type of rail service – manifest or unit trains
- The number of cars to be loaded and shipped
- The number and frequency of Railroad switches
- The number of loading stations, speed of loading, and equipment/resources required for loading
- Number and length of rail terminal tracks, including the use of ladder and/or loop track configurations
- A daily rail terminal operating plan

The Rail Terminal Operating plan will include a detailed, minute by minute, outline of sequential tasks and durations. Some of these tasks could include the following:

- Shift start up – hand-off from the previous shift, reviewing daily switching plans, safety briefing
- Movement of railcars from storage to loading tracks
- Loading of railcars
- Movement of loaded railcars to the departure track
- Breaks

- Inspection of arriving train with empties
- Movement of Bad Orders to repair tracks
- Movement of arriving empties to storage tracks
- Inspection of departing train
- Set up railcars and railcar moving equipment for next shift
- Shift wind down - hand-off to the next shift

Depending on production rates, rail terminal operations can be completed by single or multiple daily shifts. Should daily rail terminal operations have the ability to be completed within a 12 hour shift, daily output can be doubled without the need for additional infrastructure by adding on a second 12 hour shift. Additional resources to man the second shift would, however, be required.

3.3.3 Switching Rail Terminals

Railcar moving equipment are assets which impact the efficiency of rail terminal operations. Examples of such equipment include:

- Railcar indexers for small cuts of railcars along the loading/unloading tracks
- Track mounted, self-propelled, railcar movers (trucks, loaders, car-movers) for general terminal service (loading/unloading and switching)
- Locomotive switching units (smaller locomotives) for general terminal service (loading/unloading and switching)
- Larger locomotives with higher horsepower and tractive effort for moving large cuts of railcars for general terminal service (loading/unloading and switching)

Switching in a rail yard or terminal is hazardous work. Turnouts need to be aligned properly to ensure railcars are lined into the correct tracks. Railcars can be moving on multiple tracks, lines of sight are often restricted, and you will not necessarily hear a cut of railcars rolling in your direction. The outcomes of accidents can be deadly.

Safety is one of PSR's guiding principles and fundamental to rail terminal operations. The best way to minimize safety risk is to minimize the opportunity for an accident to occur. Technologies currently employed to improve safety and operating efficiency may include some of the following:

Remote Control of Railcar Moving Equipment

"Belt Packs" are hand held control panels that allow one to operate switching locomotives from the ground. They can allow switching crews a better visual perspective aimed at avoiding collisions, though the downside is this requires the operator to be on the ground and not on the equipment.

For larger operations, like unit trains on loop tracks, locomotives can be controlled remotely from a Rail Terminal Control Centre. This eliminates the need for resources on the ground altogether. Loop track operations are inherently safer due to the reduced amount of switching required.

Remote Control Turnouts

The ability to remotely operate turnouts significantly improve the efficiency of rail terminal operations and reduces the risks to accidents on the ground. Dual controlled turnouts can be operated either through radio communications (DTMF) or using software from a Rail Terminal Control Centre.

New Rail Yard Hazard Detection Systems

Miller Ingenuity has developed an electronic hazard protection system (ZoneGuard) used to protect workers in active rail operation zones. A sensor can be position on a track tie between the rails to detect track mounted work equipment or railcars and send an alert to workers using wearable devices.

Loram Technologies has developed an area monitoring system which automatically detects objects in the path of track mounted equipment. Objects which come within a defined work zone will activate an alert for the equipment operator.

These technologies are not intended to substitute defined terminal worker safety protocols but to be used as aids in enhancing equipment operator and terminal worker awareness of their surroundings, thus providing an additional layer of vigilance and protection.

3.3.4 Rail Terminal Operating Technologies

3.3.4.1 AEI Railcar Tracking

AEI (Automatic Equipment Identification) is an electronic railcar recognition system used by the North American railroads industry. They consist of passive tags mounted on each side of rolling stock and active trackside readers using RF technology to identify railroad equipment. Historically, AEI railcar readers have been used to capture a consist of trains entering and departing yards. They had been considered not cost effective to implement within railyards; however, the lack of accurate railcar tracking information can result in significant costs.

The reporting of railcar location (Track ID) and position (within a cut of railcars) is important for safety and terminal operating efficiency reasons. Automated railcar tracking eliminates the need for manually inventorying and searching for railcars. It can provide real-time railcar information and location as railcars are moved within the terminal. Improper reporting can result in additional effort, cost, and delays in terminal switching and arrival/departure operations.

3.3.4.2 Custody Transfer Systems

Custody Transfer refers to transactions and accurate measurement of physical substances being transported from one party to another. Track mounted weigh scales are often used to weigh bulk commodity railcars entering and exiting loading and unloading facilities. Two main

types of weigh scales include static scales, which weigh a single railcar stopped on the scale, or the much more efficient weigh-in-motion scales which weigh railcars as they move across the scale without stopping.

Some bulk railcar loading facilities using overhead bins to load open hopper cars have very accurate bin scales which measure the products being loaded into the empty railcars. Liquids commonly use very sensitive flow metering equipment to measure the volume of product being loaded into tank cars.

3.3.5 Arrivals and Departures

Based on their PSR service principles, Railroad's expect to interface with their customer's terminal within a dedicated arrivals/departure yard. Their preferred "hook and haul" operation will involve dropping empties/loads on the arrival track, hooking onto waiting loads/empties on the departure track, and departing.

From a terminal operator's perspective, this means that the arrival tracks need to be clear and ready to receive railcars. They should be prepared to visually inspect the inbound train and identify any bad orders which they will need to setoff on the bad order track for repairs.

This also means that the departure tracks need to be filled and ready for pickup. Departing railcars will need to be blocked (or grouped) by destination for manifest service. This will minimize the amount of switching the Railroad's will need to do in their classification yards when it come time to build their train sets for destination. This practice will reduce yard dwell time, speed up the transit cycle, thus leading to minimizing the railcar assets required to support the service and reduce the demand for railcar storage, all leading to lower capital and operating costs for the required level of service.

Further, terminal workers will need to ensure all departing railcars are properly coupled and the air brake lines are properly connected. It is common to have a compressor and air line installed along the departure tracks in order to pre-charge the air brake lines before Railroad's arrival. This will facilitate a quicker departure, especially in winter months, when it takes longer to confirm end to end air pressure. The quicker terminal operations can deal with the Railroad's, the more time they will have to focus on their other terminal operations.

Custody Transfer information is also required for regulatory purposes associated with transporting products by railcars. Prior to departure from any terminal, a train manifest needs to be generated for the train crew which includes, in proper sequence (front to back), the railcar ID, the gross weight of the railcar, and whether or not the products are classified as dangerous commodities. It is also important to be able to identify, and resolve, any overloaded or imbalanced railcars, prior to departure.

3.4 Rail Service – Origin to Destination to Origin

Every carload gets assigned a trip plan, by the Railroad, between original and destination. This trip plan includes, not just scheduled pickup and delivery from origin to destination, but also intermediate classification destinations where the railcar might get reclassified from one train to another.

Railcars can experience delays anywhere along their route for a wide variety of reasons. Railcars can get setoff due to mechanical repairs; connections can get missed due to network delays (congestion, cold and hot weather slow orders, track condition, etc.); or classification yards maybe congested. Daily tracking of railcars in transit by the terminal operators is important for two main reasons. First, to know where the railcars get hung up, why, and what is being done to get them moving again. Second, to be able to develop a recovery plan at the terminal, with updated delivery information, which will sustain production.

Railroad's provide daily delivery reports to their customers. Since the advent of PSR, and in efforts to become more transparent and better aligned between rail service and terminal operations, access to Railroad real-time data has become a priority for both customer and service provider.

As an example, CN has recently launched an on-line service called Track and Trace API (Application Program Interface) which provides customers with real-time information about shipment location and status. This system monitors shipment location through the use of GPS tracking. It also provides access to the latest shipment status reports, along with current estimated times of arrival at either an interchange point, rail destination, or customer site.

APIs are designed to accommodate an easier and more cost-effective transmission of real-time data from one computer system to another. The railroad industry has undergone a major transformation in the way goods are moved, with complex supply chains becoming the norm. APIs can provide a competitive advantage to customers. They are part of a multi-faceted effort to forge an end-to-end, seamless, digital platform (The Digital Railroad) that can deliver real-time visibility and connectivity through state-of-the-art technologies.

Today, customers have the ability to apply sensors to their fleet and use third party services to, not just track shipments, but also manage their own supply chain. As an example, Cando Rail Services have recently launched a multi-purpose platform called Quasar which utilizes leading edge Internet of Things technology to create a digit twin of the customer's supply chain. Features include asset management, real-time railcar tracking, shipment management, yard/terminal management, and transportation cost management. It provides performance metrics, as well as, predictive and prescriptive analytics.

4 RESULTS

The following is a summary of the results obtained from a review of Industry Rail Terminal operations through the lens of PSR.

Rail Terminal Configurations

Ladder tracks can be utilized effectively for manifest or unit train service, however, loop track configurations are best suited for the following conditions:

- Unit trains consisting of at least 100 cars
- Daily plant production of unit train capacity
- Seven day per week terminal/plant operations

Railcar Fleet Selection Criteria

- Largest gross weight capacity allowed on the rail line
- Largest load limit capacity within the gross weight capacity class
- Consistent railcar length to suit fixed loading/unloading infrastructure configurations
- Shortest railcar length to minimize rail terminal track lengths and maximize railcar useable space

Terminal Location Selection

Preferred terminal locations have the following operating characteristics:

- Railroad mainlines or branchlines which accommodate 286,000 lbs. gross railcar weights
- Railroad mainlines or branchlines with minimal operating restrictions (i.e., speed, steep grades)
- Access to destination through multiple routes and/or Railroads

Storage

Ensure sufficient material storage, insufficient storage will shut down the plant and create disruptions to normal plant operations, but not rail terminal, and rail service operations.

Ensure sufficient railcar storage, insufficient storage imposes higher risks impacting plant, rail terminal, and rail service operations.

Terminal Operating Practices

Utilize a steady 7 day per week work flow to optimize efficient and reliable rail transportation services, reduce the demand for railcars and terminal railcar storage, and to reduce capital and operating costs.

Daily terminal output can be doubled without the need for additional assets and infrastructure when there is the opportunity to load a train within 12 hours.

Arrivals and Departures

Railroads preferred “hook and haul” operations.

Arrival tracks need to be clear and ready to receive railcars. Departing railcars need to be blocked by destination for manifest service.

Rail Service – Origin to Destination to Origin

Daily tracking of in-transit railcars is important to know where they get hung up, why, and what is being done to get them moving again, as well as, to be able to develop a recovery plan at the terminal.

Access to real-time performance and predictive data is essential to manage assets, operations, costs, and service.

Information Systems – The Digital Railroad

Information management systems are fundamental to safe and efficient rail terminal and Railroad operations. Information needs to be accurate, real-time, and easily accessible to all stakeholders within the supply chain. Today, the Railway industry has embraced new data management tools such as Big Data, Machine Learning, Artificial Intelligence, and Digital Twinning in order to streamline their operations with superior predictive tools with the PSR objectives in mind for improving safety and efficiency, controlling costs, improving service, and delivering value to their customers. We are seeing the emergence of a truly integrated and transparent Digital Railroad. Some examples highlighted in this paper include:

- Fleet management systems
- Safety management systems
- Rail Terminal Operating technologies
- Railcar tracking systems
- Railroad service information systems

5 CONCLUSIONS

An idea ceases to be an abstract concept and evolves into something transformational when positive results are established on a continuous basis due to successes and maturing of the concepts. Over the past 15 to 20 years, Precision Scheduled Railroading has taken the railroading industry to a higher standard of performance and a higher level of success. With that said, there remains much more to do as these concepts of railroading continue to evolve and develop focused on growing market share for the transportation of goods across North America and providing value to the shippers this industry services.

Rail transportation is a fundamental component of Canada's Economic Corridors, supporting both domestic and international supply chains, driving our economic growth potential. The reliability of efficient and fluid railway networks, however, relies on the seamless interface between Industry terminal and Railroad operations.

Using a common set of operating practices and principles, set out by PSR, this paper has highlighted what these interfaces are, and how they need to interact to create value through faster and more reliable transportation service, lower transportation costs, and overall supply chain stability. For the Canadian economy, this translates to a competitive advantage in

global and domestic marketplaces, lower costs for our goods and services, and opportunities for sustainable growth.

Precision Scheduled Railroading is founded upon the endless cycle of plan – execute – monitor – analyze – and revise the plan. It is an ongoing and developing concept that recognizes the balance and dependency of its principles. The future and true potential of PSR (PSR 2.0) will be realized as operating practices between competing railways, customers, ports and other supply chain entities, become standardized, generating closer collaboration and alliances, resulting in greater benefits and better service for customers, regardless of the route or the carriers engaged. This will require that all parties work harmoniously for the greater good of Industry, consumers, and the Railroads. Future success will depend upon innovation and the ongoing development and implementation of new technologies through the evolution of the Digital Railroad in support of improving safety, reliability, efficiency, and utilization of assets, while managing costs, and delivering superior service.

6 REFERENCES

Amsted Digital Solutions - fleet management software and services. <https://www.amstedrail.com/amsted-digital-solutions/>

Cando Rail Services - Quasar - multi-purpose platform supply chain management service. <https://www.quasar.tt/>

CN Track and Trace API. <https://www.cn.ca/en/customer-centre/tools/api/>

E. Hunter Harrison (2005) How We Work and Why - Running a Precision Railroad; Canadian National Railway Company

Loram Technologies -area monitoring system. <https://loram.com/>

Miller Ingenuity - ZoneGuard - electronic hazard protection system. <https://www.milleringenuity.com/products/Roadway-Worker-Protection>

Supply Chain Definition – Corporate Finance Institute <https://corporatefinanceinstitute.com/resources/knowledge/strategy/supply-chain/>

Literature review on cognitive impacts of cab warning systems and train control technologies

Mona A. Rad^{1*}, Lianne M. Lefsrud^{1,2} and Michael Hendry¹

¹Canadian Rail Research Laboratory, Faculty of Civil Engineering, University of Alberta

²Chemical and Materials Engineering, Faculty of Engineering, University of Alberta

Daniel Blais

Innovation Centre, Transport Canada

ABSTRACT

Cab warning systems and train control technologies provide alarms to inform train operators of dangerous situations and enforce speed and movement restrictions to avoid or mitigate negative consequences. While these systems can enhance safety as the last resort of safety control, they may have adverse cognitive impacts on train operators. This article reviews publicly published articles and reports that analyzed human factors issues of in-cab technologies. The reviewed technologies include, but are not limited to, cab signaling systems, automatic train control systems, anti-collision devices, train operators' vigilance systems, and train operators' reminder devices. The findings demonstrate that these technologies can cause a variety of human factors issues, including over-load or under-load, over-reliance on the system, complacency, loss of situation awareness, mode confusion, distraction, and automatic responding. To reduce the potential negative impacts related to the design and usability of in-cab technologies on train operators, the authors recommend employing the human reliability analysis and the human-in-the-loop processes to better understand the impacts of in-cab warning systems on train operators, identify the contextual factors influencing the train operators' performance, and develop strategies to mitigate such human-induced risks.

Keywords: In-cab warning systems, Train control technologies, Cognitive impacts, Human factors.

1 INTRODUCTION

Automated control systems are increasingly being adopted in various industries to improve safety and optimize operations. They are changing the nature of work and transitioning the role of human operators from manually operating to passively monitoring. It is assumed that such technologies prevent or reduce known human errors and improve safety, however, past experience in various domains shows that they may cause unexpected impacts on human performance and introduce new sources of human risks. The reported incidents and accidents from high reliability organizations (HRO) such as nuclear power plants and the aviation industry are strong evidence for this claim. Statistics illustrate that human factors account for over 75% of major railway accidents, marine casualties, or aviation accidents (Cullen & Smith 2004, Evans 2014, Tao et al. 2020). This number reaches over 90% for failures in nuclear plants (French et al. 2011).

Due to the significant role of human operators in the safety aspect of the social-technical systems and the importance of employing strategies to reduce human-caused risks, a large number of studies in diverse industries have been conducted to develop and deploy human reliability analysis (HRA) and identify human factors issues associated with automation. The performed HRA studies differ in their scope, the types and levels of decomposition of the tasks addressed, and the factors considered to influence the human error probability (Mkrtchyan et al. 2015). While some contributions focus on the identification, measurement, and reduction of human-induced risks at the design stage, some others aim at evaluating the effects of performance shaping factors (PSFs) on human performance (Patriarca et al. 2020).

The bibliometric analyses conducted by Tao et al. (2020) and Hou et al. (2021) highlighted the main research areas of the HRA publications including human performance in human-machine systems. Papadimitriou

* Corresponding author's email: ahmadira@ualberta.ca

et al. (2020) presented a systematic literature review of the effects of human factors on transport automation safety. They reviewed the role of misaligned trust in automation (i.e. mistrust vs overreliance), situation awareness in automated driving, the transition of control between human and machine, and operators' experience and training in the effectiveness of the automated systems in the road, aviation, maritime and rail domains. The literature review studies disclosed that much less research about the influence of automation on human operators' performance has been carried out in the railway domain as compared to the aviation and road sectors. This was a reason that Papadimitriou et al. (2020) recommended transferring some experiences and lessons learned in relation to human-automation interaction from road and air transport to the railway industry.

One of the limited exploratory reviews in the railway sector about the influence of automation on train operators is Bearman & McCusker (2008), which has been later reorganized and updated by Zimmermann (2015). These review studies were respectively performed under the auspices of the Australian and Canadian governments and were published in non-scientific indexed databases. Bearman & McCusker (2008) categorized the potential human factors issues of new in-cab and train control technologies into "fixation on and distraction by the technology", "disregard or attenuation to warnings", "errors inputting data into the system Increased attention and knowledge demands Lack of insight into how the technology is functioning", "poor communication by the technology about its current functioning", "misplaced trust in the automation", "mismatches between the driver's understanding of a task and the way this task is performed by the technology", "changes to the nature of train driving, particularly relating to change in driving technique required by some new technologies", "difficulties for drivers transitioning in and out of the technology", "shifts in workload", "changes to work-roles and coordination requirements for train personnel other than drivers". Then, Zimmermann (2015) recategorized the factors into design and usability (involving distraction and workload, human-machine communication, ergonomic display design, and assigning final authority), skill retention, the transformation of work, training and mental models, operator expectation and trust, and unanticipated effects and interactions. In 2021, the authors of the current paper prepared a report on the cognitive impacts of in-cab warning systems for Transport Canada (TC) (see Rad et al. (2021)) and this paper is the extension of the report to fill in the research gap.

This paper reviews the relevant publications in both scientific and non-scientific indexed databases and summarizes the main human factors challenges in relation to using diverse train cab warning systems including cab signaling devices (e.g., Automatic Warning System (AWS)), automatic train control technologies (e.g., Positive Train Control (PTC) and European Train Control System (ETCS)), train-to-train anti-collision systems (e.g., Train Collision Early Warning System (TCEWS)), train operator reminder devices (e.g., Driver's

Reminder Appliance (DRA) and In-Cab Signal Reminder Device (ICSRD)), and train operators' vigilance devices (e.g., Driver Vigilance Systems (DVS) and Monitoring Engineer Fatigue (MEFA)). The remainder of our paper is organized as follows: Section 2 summarizes our bibliometric search methodology, and Section 3 provides descriptive analysis. Results and discussion are presented in Section 4. Finally, our conclusions of this review are summarized in Section 5.

2 BIBLIOMETRIC SEARCH METHODOLOGY

We searched academic research databases (i.e., ScienceDirect, Scopus, Web of Science, SpringerLink, and IEEE Xplore), along with non-citation research databases such as FRA, RSSB, and Volpe center libraries for published research in English, up to the date 1 August 2021. We selected several keywords related to human reliability analysis (HRA) in combination with in-cab warning systems and automated train control (ATC) technologies to maximize the number of documents to be analyzed. We also considered references cited by the obtained research as a source for discovering more related scholarly research. This yielded 46 original publications after removing duplicates.

3 DESCRIPTIVE ANALYSIS

This section represents the statistical analysis of the 46 original research publications that have investigated the impacts of railway automation on train operators.

The distribution of the published papers and reports per continent in the years between 2002 and 2020 is illustrated in Figure 1. The year 2017 with the highest share for the European countries, is the most significant period for this research topic.

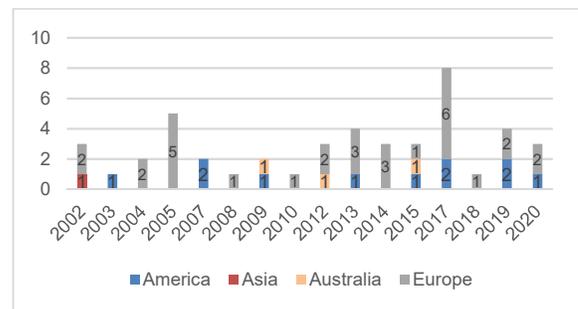


Figure 1. Yearly distribution of the reviewed publications per continent

Detailed analysis discloses that the UK and the US are the leading countries in research on human factors issues of railway automation, with an aggregate share of 65% (see Figure 2).

As shown in Figure 3, the majority of the publicly available studies in the field of research are reports published in non-scientific indexed databases (37%). Among scientific-indexed publications, Journal articles (28%), followed by conference papers (22%) have the main proportions. This result confirms Patriarca et al. (2020)'s claim that some human reliability analysis

(HRA) research and development contributions are publicly available but not recorded in citation databases, and some others are proprietary research.

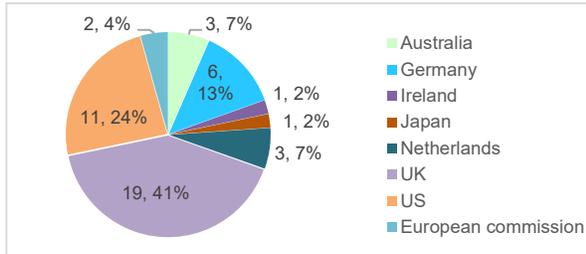


Figure 2. Countries' contributions in the human factors' studies related to railway automation.

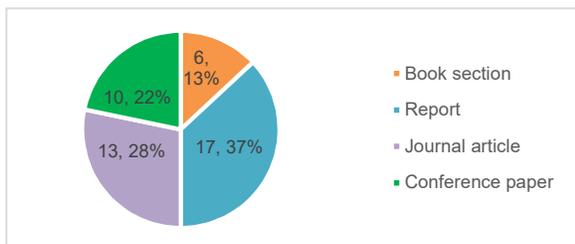


Figure 3. Type of publications

The existing literature is also classified according to the type of data sources (see Figure 4). Like what was reported by Mkrtychyan et al. (2015), three main sources of data, i.e., theoretical data, empirical data, and expert judgment data, were used in the reviewed publications. The theoretical data is excluded from Figure 4 since it was utilized in almost all papers. Moreover, the empirical data is categorized into real-world data (i.e., data obtained through in-cab observations, on-train-data-recorder (OTDR), or occurrence databases) and simulator/simulation tests data.

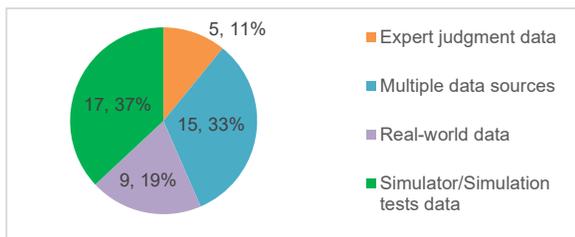


Figure 4. Type of data sources

According to Figure 4, 33% of the researchers combined diverse sources of data to get a better understanding of the ironies of automation. For example, in situations that were a lack of empirical information, they used expert knowledge along with the empirical and/or theoretical information (see e.g., Wreathall et al. (2007a), Rose & Bearman (2012), Nneji et al. (2019)).

4 RESULTS AND DISCUSSION

In this section, we review human factors challenges associated with the introduction and use of a wide range of in-cab warning systems and automated train control technologies.

4.1 Workload

Workload may be characterized as the reaction to demand or stress, with either positive or negative consequences (Oppenheim et al. 2010b). It can be physical (e.g., pressing a button), visual (e.g., scanning the light on the display), and cognitive (e.g., interpreting a signal) (Halliday et al. 2005).

The workload to performance relationship is illustrated in Figure 5. The ideal workload situation happens when "homeostasis" is achieved, which can be described as a balance where coping and adaptation to task demands are optimal. Any deviations from the optimal workload level, either an increase or a decrease, can contribute to lower performance (Oppenheim et al. 2010b, FRA 2014). Under-load can result in loss of situational awareness, boredom, fatigue, frustration, over-confidence, and increased reaction times while over-load causes irrational problem solving, loss of situational awareness, exhaustion, and low self-esteem (FRA 2014, Robinson et al. 2015).

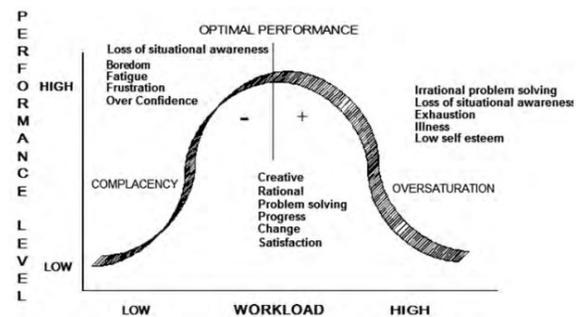


Figure 5. Workload versus performance (FRA 2014)

It is argued in the literature that the integration of any new technology into a system can, on the one hand, cause an overload of mental workload but can, on the other hand, lead to underload because of the increased automation (Robinson et al. 2015). Therefore, several research studies have been conducted to investigate the impacts of a new device on the workload levels of train operators. For example, the RSSB assessed the effectiveness of an In-cab Signal Reminder Device (ICSRD) and concluded it has the potential to increase workload by requiring additional cognitive and physical tasks. This can increase train operator stress and thus response time (Halliday et al. 2005).

The RSSB (2002) compared the imposed workload level of the two types of Driver's Reminder Appliance (DRA) systems on train operators and concluded that the Automatic Warning System (AWS) activated DRA, in which the system is automatically set with regards to the tasks related to AWS, poses a lower workload to the train operator compared to the train operator set DRA. Crick et al. (2004a)'s research demonstrated that the

visual information provided by the automatic train protection (ATP) system is too much for some train operators to handle, causing increased workload and distractions. Verstappen et al. (2017) found that monitoring innovative devices in train cabs during driving requires multiple resources (e.g., visual and cognitive resources), which can conflict with the primary driving task. This can cause an increase in workload and influence driving performance. Van Der Weide et al. (2017) realized that train operators experienced notably lower workload when driving with European Rail Traffic Management System (ERTMS) compared to driving with ATB (i.e., the legacy system in Netherland), and very experienced train operators even reported boredom. Spring et al. (2009) also reported a reduction in the mental workload of train operators, even to a sub-optimal level, because of an increase in the levels of automation (LOAs). Historical data related to head-up displays revealed a substantial decrease in train operator workload (Davies et al. 2012).

Analyses performed by Foulkes (2004) and Buksh et al. (2013) showed the Level 2 ERTMS, in which all signaling indications are shown in the cab and there are no lineside signals, contributes to a lower workload than the current train driving task. However, according to studies conducted, a variety of factors including the level of ERTMS implemented, train operator strategy, type of traction, and transitions into and out of ERTMS could impact workload under ERTMS (Robinson et al. 2015). An increase in workload because of transition in/out of a train protection system, particularly in complex areas such as stations and level crossings was also reported by Foulkes (2004) and Monk et al. (2017). A series of studies about the workload level of the Positive Train Control (PTC) system revealed that frequent, often non-informative audio alarms of the PTC systems and the required data entry during initialization and/or operation are sources of workload (Wreathall et al. 2007a, Roth & Multer 2009, Roth et al. 2013). Wreathall et al. (2003) emphasized that train automation systems usually automate the easy parts of a task, reducing workload during times when the workload is already minimal while requiring extensive human involvement in challenging situations when the workload is high. Therefore, during high-paced high-risk situations where the workload is already very high, there is an increase in workload demands (Wreathall et al. 2003).

During Wreathall et al. (2007a)'s interviews with train operators who have had experience driving with the PTC system, they expressed concerns regarding high numbers of audio warnings that require to be acknowledged. These can create distractions and high workloads for the train operators. The train operators recommended that audible alarms should be restricted to alert them to potential issues (e.g., an upcoming speed restriction that might be missed) and should be avoided for positive circumstances (e.g., when a speed restriction is no longer in effect) (Wreathall et al. 2007a). Brandenburger and their colleagues performed a series of studies and experiments at the German Aerospace Centre about the effects of railway automation on the

train operators (Brandenburger et al. 2017a, Brandenburger & Jipp 2017, Brandenburger et al. 2017b, Brandenburger et al. 2018, Brandenburger & Naumann 2019, Brandenburger et al. 2019). Their experiments illustrated that the higher grades of automation (GOAs) do not always mean a lower workload level. While the transition from GOA-1 to GOA-2 reduced the workload level and kept it at the sub-optimal level (Brandenburger et al. 2018, Brandenburger & Naumann 2019), the transition from GOA-2 to GOA-3 increased workload and made it closer to an intermediate, optimal level of workload (Brandenburger et al. 2019).

In summary, the effect of automation on workload is mixed and automation may increase or decrease the workload level. Nneji et al. (2019), at Duke University, developed the Simulator of Human Operator Workload (SHOW) based on the empirical data from the U.S. railroad industry to quantitatively model freight rail operator workload.

4.2 Distraction

Verstappen et al. (2017) conducted a study about the effects of innovative devices in Dutch train cabs on train operators and highlighted that conflicts between the use of these devices (e.g., communication devices or information devices) and train driving tasks, particularly in critical situations, can be a source of distraction for the train operator. Safar et al. (2020)'s interviews with the US train operators revealed that non-integrated in-cab displays and alarms can be a contributory factor for distraction. The train operators indicated that there are often non-safety/non-critical alarms that may be distracting. These alarms may sound continuously when activated despite being acknowledged, causing annoyance and distraction to the train operator. Furthermore, a frequent false alarm also raises the risk of ignoring safety-critical alarms by the train operator (Safar et al. 2015, Safar et al. 2020). Frequent, often non-informative audio alarms created by PTC systems can also be a source of distraction (Wreathall et al. 2007a, Roth & Multer 2009, Roth et al. 2013). Using devices such as ICSR and DRA, which need to be used on the move, may distract train operators' attention away from signals ahead while using the device (Halliday et al. 2005).

The impact of vigilance devices on distraction is different. Vigilance devices, on the one hand, are argued to reduce distraction through an increase in levels of vigilance, arousal, and attention (Halliday et al. 2005). However, on the other hand, they can divert the train operator's attention away from the primary task of driving (Wilde & Stinson 1983, Rose & Bearman 2012).

4.3 Loss of Situation Awareness

Endsley (1996) defined situation awareness (SA) as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". They categorized SA into three levels: level 1 (perception), level 2 (integration and comprehension),

and level 3 (projection).

SA is reflected in a train operator's actions and response time (Halliday et al. 2005, Park et al. 2020). According to the conscious thinking processes that are needed to attain SA, longer response times are expected in decision making with good SA. When the levels of train operator vigilance or arousal decrease, their attention may deviate from the task of checking signals and thus increase the risks of making skill-based errors. An operator may automatically respond during low arousal despite the loss of SA (Halliday et al. 2005). SA is commonly assumed to be improved with experience; thus, novices have low SA and are more dependent on displays of information (Halliday et al. 2005). Crick et al. (2004a) found that more experienced train operators had considerably longer response times to the AWS, possibly reflecting the greater understanding by these individuals of the need to be aware of the signal before responding. According to information available, expert judgment, and the simulator experiment, Thomas & Davies (2008) and Davies et al. (2012) proposed that head-up display of speed and brake information can help the train operator to maintain situational awareness. In-cab Signal Reminder Device (ICSRD) and a moving map display are other recommendations for enhancing situation awareness of train operators (Halliday et al. 2005, Liu et al. 2017).

4.4 Mode Confusion

Mode confusion (mode error) happens when the user is confused about the system's current mode (i.e., errors in SA) or is unable to recall how the system reacts in the current mode (i.e., slips of action) (Wreathall et al. 2007b). There is a risk that the train operator does not understand or forgets that the mode change has occurred due to distraction and workload, which can result in mode errors (Sebok et al. 2015, Sebok et al. 2017). Safar et al. (2020) clarified that train operators accustomed to driving within cab signal territory may forget they are within no cab signal territory and wait for the speed reduction alarm to adjust the speed.

Two types of mode transitions and their related mode confusion are considered for the PTC system. One type of mode transition is when a train is equipped with a PTC system, but depending on the circumstances, the system could be operating or not. For example, on a PTC-equipped train, the system may not be operational because the train is outside of PTC territory or because the PTC system is malfunctioning. The second type of mode transition is related to the situation in which a train operator works on both PTC territory and non-PTC territory. The potential issue allied with the first type of mode transition is that the train operator may not recognize that the PTC system is not operating or may notice but fail to adequately enhance vigilance to compensate for the lack of PTC protection. When a train is moving between PTC equipped and unequipped areas, i.e., the second type of mode transition, factors including complacency and skill loss could contribute to the train operator errors. If the train operator has become over-reliant on the PTC system and, due to

temporary workload or distraction, fails to notice or forgets protection provided by PTC is not available at the moment, it can result in complacency and therefore human errors. (Wreathall et al. 2007a, Wreathall et al. 2007b, Roth et al. 2013).

Changes in operating conditions can be the main cause of mode confusion. These kinds of problems may be triggered by frequent switching between trains that have the DRA and those that do not, or switching between trains with train operator set DRA and AWS activated DRA (RSSB 2002). This means that the increase in the number of transitions in a route/work shift can raise the probability of error (Monk et al. 2017). Hence, an inkblot strategy for rolling out ERTMS imposes less workload on train operators than a patchwork strategy when there are fewer transitions. Notably, in an inkblot strategy, ERTMS is rolled out from one starting track towards adjacent areas, while in a patchwork strategy, the development of ERTMS is distributed across the network, e.g., based on technical urgency (Van Der Weide 2017).

4.5 Complacency and Over-reliance

Roth & Multer (2009) referred to complacency as a general term that reflects the incapability of a train operator to act as well without a system as they could before the system was installed. Complacency can have various negative consequences, such as train operators not detecting the system failure (or it being off), experiencing delays in identifying and reacting to a system failure, and losing their driving skills and therefore not being able to perform the driving task as well when the system is not available as they previously could have (Roth & Multer 2009, Roth et al. 2013). Complacency tends to criticize the operator for unreasonably depending on a system and is closely connected to the principles of overreliance and excessive trust (Wreathall et al. 2007a, Roth & Multer 2009, Roth et al. 2013).

The more operators trusted the automation, the more they left it in control without supervision. When train operators are passive and observant, they are more prone to perform a task based on system feedback rather than anticipatory, self-identified strategies. This could be crucial because anticipation is a required factor for higher-level situation awareness and a lack thereof could detrimentally impact the error management process when the system fails (Giesemann 2013). Over-reliance on the system can highly increase the risk to accept the displayed information even when it is incorrect (Halliday et al. 2005). Abe et al. (2002) reported that over-trust in warning systems can cause substantial delays in responding to hazards when there is a mismatch between what the train operator expects and the actual state of the system.

McBride et al. (2014) highlighted that the reliability of a system can play a role in excessive trust and complacency and is a double-edged sword. On the one hand, the greater the reliability of an automated system, the better the performance when the system is perfectly operating. On the other hand, the reliable automated

system increases the tendency to get complacent, which makes the operator less vigilant and less capable of reacting to system errors or failures. With a highly reliable PTC system, train operator performance may decline if the information provided becomes unavailable (Wreathall et al. 2007a). A degree of complacency in checking all alarms was reported by Carey (2015) for cases containing an excess of non-critical alarm messages. Furthermore, the results of Brandenburger & Jipp (2017)'s study showed that train operators perform worse in degraded operations for the higher levels of automation, which can partly be compensated by train operators' experience.

4.6 Visual Attention Allocation

Naweed (2014)'s study disclosed that, in spite of the existence of in-cab devices and signaling systems, the outside area still needs to be searched for danger. It takes time for the eyes to refocus from one viewing distance to another one (i.e., visual accommodation)(Halliday et al. 2005). Hence, the transfer of primary information from outside the cab to inside the cab could negatively affect safety due to the shift in attention and visual accommodation increasing the risk of missing out of cab important events (Wreathall et al. 2007a).

An exploratory eye-tracking field study carried out by Naghiyev et al. (2014a) and Naghiyev et al. (2014b) illustrated that some train operators are more dependent on the system and reactively respond to situations, while others rely less on the alerts and alarms and are more proactive. Overall, when train operators used the ERTMS system, they spent considerably more time monitoring the speedometer rather than seeing the out-of-cab environment in comparison to conventional systems. The results of studies conducted by Brandenburger et al. (2017a), Van Der Weide et al. (2017), and Hely et al. (2015) also confirmed that train operators direct considerably less attention to out of cab than in-cab devices when driving trains equipped with ATP systems (e.g., ERTMS) compared to those with conventional systems. The operators of trains equipped with a PTC system also reported a greater need for focusing on in-cab displays, at least initially, thus limiting their ability to check outside the cab (Roth & Multer 2009, Roth et al. 2013). During an examination of PTC systems, the train operators pointed out that they needed to closely track the in-cab display to remain within the braking curve and prevent a penalty brake application. The train operators emphasized that when the train traveled within a time window that allowed no flexibility in schedule variation or approached territories with speed restrictions, attention allocation emerged as an issue (Wreathall et al. 2007a).

Monk et al. (2017) found that when new signaling systems were first introduced, train operators were more focused on the in-cab signaling display and were distracted by it, reducing the amount of attention paid to monitoring outside of the cab. However, after a while, they could better balance their attention between inside and outside of the cab. Some Incremental Train Control

System (ITCS)'s train operators also mentioned that, after 3 weeks to 1 month of working with the system, they spent less time monitoring the in-cab display. However, some train operators indicated no noticeable change in their attention distribution even after they had sufficient experience in running a PTC-equipped train and a remaining inability to have a head-up driving style (Wreathall et al. 2007a).

Operating ICRSD also means a train operator needs to devote more attention to in-cab displays, which can cause head-down driving and constant changes in visual distribution (Halliday et al. 2005). Despite some detrimental effects of in-cab systems on train operators' visual attention distribution, Merat et al. (2002) discovered that AWS can considerably increase the number of looks at signals as great numbers of the first looks at signals of the train operators studied were taken after AWS had sounded.

4.7 Automatic Responding

A train operator may read or hear an alarm without understanding its importance and meaning and show a skill-based, reactive response to it, called automatic responding (Oppenheim et al. 2010a, 2010b, Carey 2015). In some cases, the train operator is mentally fatigued but physically awake enough to press the push button or enter data into the train control system because motor reflex actions generally need a lower level of cognitive endeavor. Therefore, the train operator may trigger automatic responses (Stein et al. 2019). The automatic responding shows the warning system has failed in its primary purpose to alert the train operators and attract their attention to threads (Halliday et al. 2005).

The results of a questionnaire survey of 277 UK train operators illustrated that a considerable number (i.e., 56%) have automatically acknowledged an Extended AWS alarm at least once during their driving experience, although only 2% did it on a daily basis(McLeod et al. 2005). McLeod et al. (2005) also analyzed On-Train Monitoring Recorder (OTMR) data and found that some train operators started pushing the cancellation button before the alarm had started to sound or they responded to the alarm very quickly. These anticipatory and quick responding behaviors may indicate unconscious alarm canceling, with only a physical response and no interpretation (Halliday et al. 2005, Balfe 2020). Because the AWS neither differentiates between caution and stop signal aspects nor has a mechanism to prevent misperceptions, train operators who are confronted with successive cautionary signals (yellow or double yellow) likely respond to the AWS horn without conscious interpretation about the signal aspect (Lawton & Ward 2005).

The existence of warnings related to situations that are less important than restrictive signals (e.g., Temporary Speed Restriction (TSR)), personal factors, the high number of alarms, and signaling issues were reported as the main reasons for automatic responding (McLeod et al. 2005). More analysis revealed that poor alarm management (e.g., excessive, uninformative audible alarms) reinforces a tendency to an automatic

response without completely perceiving the alarm's meaning (Wreathall et al. 2007a). To address this problem, a variety of ideas for effective alarm management were recommended. First and foremost, the warnings must be accurate and informative to reduce the risk of automatic responding. Second, if an audible warning is adopted in a situation that the operator requires to monitor somewhere other than the display screen (in this example, they need to check out the window), it is useful to make distinctions between sounds relating to auditory alarms of different conditions (e.g., a different tone for approach versus stop). Third, using different actions (e.g., a different button push for approach versus stop alerts) for acknowledging different alarms is a good method to mitigate the risk of automatic responding (McLeod et al. 2005, Wreathall et al. 2007a).

4.8 Memory Failures

Short-term memory, also known as “working memory”, is volatile and easily lost or distorted. Not only the passage of time but also interference between the current contents of working memory and newly arriving information can be a reason for information loss (McLeod et al. 2003). Thus, McLeod et al. (2003) recommended that a train operator should never rely on working memory to maintain vital safety data and suggested that external assistance is necessary.

Crick et al. (2004b) reported that 77.1% of their focus group train operators had at least occasionally forgotten the signal aspect after acknowledging the AWS. The evidence showed a possible risk of the train operator being uncertain about what an active alert corresponds to after around 7 seconds (Moray et al. 1983). A memory failure related to the DRA system may include forgetting to set the DRA, pressing a different button (e.g., AWS alarm reset or door release button) instead of the DRA button, resetting the DRA, and starting the trip based on the platform guard's signal without checking whether the signal aspect is clear (RSSB 2002). Davies et al. (2012) proposed that repeating the AWS warning on the head-up display can reduce the risk of train operators forgetting a cautionary signal was shown on the previous signal.

4.9 Skill Loss

Skill loss (skill degradation) is a probable but unpleasant feature of automation (Bainbridge 1983). As supervisory train control technology increases, train operators have a reduced opportunity to carry out tasks themselves, thus contributing to skill loss. The skill loss issue becomes apparent in a situation that requires the operator to take charge of the train (Wreathall et al. 2003). Therefore, maintaining the required skills of train driving is important and can be achieved either through a frequent application or structured training (Balfe 2010). Giesemann (2013)'s study highlighted the effects of the train protection system on task-related competence and control expectations which can result in poor automation failure management. They believed that the system is not a problem itself, but the reasons

for concerns are the lacking feasibility of anticipation and a proactive driving style which might push train operators into passivity, causing loss of situation awareness and thus errors in the event of an automation failure.

4.10 Summary of Automation and Human Factors

A summary of the main human factors issues associated with the use of in-cab warning systems and automated train control technologies is provided in Table 1.

Table 1. Key human factors issues of in-cab warning systems

Author	Workload	Distraction	situation awareness	Mode confusion	Complacency and over-reliance	Visual attention allocation	Automatic responding	Memory/failures	Skill loss
Balfe (2020)							♦		
Safar et al. (2020)		♦							
Oppenheim et al. (2010b)								♦	
Brandenburger and Naumann (2019)	♦		♦						
Brandenburger et al. (2019)	♦								
Nneji et al. (2019)	♦		♦						
Stein et al. (2019)				♦					
Brandenburger et al. (2018)	♦								
Brandenburger et al. (2017a)			♦				♦		
Brandenburger & Jipp (2017)					♦				
Sebok et al. (2017)	♦	♦	♦	♦					
Van der Weide et al. (2017)	♦						♦		
Van der Weide (2017)	♦						♦		
Monk et al. (2017)	♦			♦			♦		
Verstappen et al. (2017)	♦	♦							
Liu et al. (2017)				♦					
Hely et al. (2015)	♦						♦		
Robinson et al. (2015)	♦								
Safar et al. (2015)		♦							
Naghiyev et al. (2014a)								♦	
Naghiyev et al. (2014b)		♦					♦		
Naweed (2014)								♦	
Smith et al. (2013)									
Roth et al. (2013)	♦	♦	♦				♦		

Author	Workload	Distraction	situation awareness	Mode confusion	Complacency and over-reliance	Visual attention allocation	Automatic responding	Memory failures	Skill loss
Gieseemann (2013)			♦		♦				
Buksh et al. (2013)	♦								
Rose and Bearman (2012)		♦							
Davies et al. (2012)	♦		♦					♦	
Scott & Gibson (2012)	♦								
Oppenheim et al. (2010a)	♦		♦						
Oppenheim et al. (2010b)								♦	
Roth & Multer (2009)	♦	♦		♦	♦	♦			
Spring et al. (2009a)	♦								
Thomas & Davies (2008)	♦		♦					♦	
Wreathall et al. (2007a)		♦			♦	♦	♦		
Wreathall et al. (2007b)			♦	♦	♦	♦			♦
Halliday et al. (2005)	♦	♦	♦		♦	♦	♦		
Crick et al. (2004b)					♦			♦	
McLeod et al. (2005a)					♦			♦	
McLeod et al. (2005b)					♦				
Lawton & Ward (2005)		♦	♦					♦	
Foulkes (2004)	♦		♦				♦		
Crick et al. (2004a)	♦	♦		♦					
Wreathall et al. (2003)	♦	♦	♦		♦				♦
RSSB (2002)					♦				♦
Abe et al. (2002)					♦				
Merat et al. (2002)						♦			

5 CONCLUSIONS

This paper summarizes the potential negative impacts of in-cab warning devices and automated train protection systems on train crews. Through this review, we have found that most of the publicly published human-automation interaction research has been performed, or at least sponsored, by governments and regulatory bodies. Furthermore, depending on the technology development stage and available facilities, human factors researchers have collected and utilized real-world data, simulator experiments data, and/or expert judgment data in addition to theoretical data.

This literature review reveals that workload (i.e., under-load or over-load), distraction, loss of situation awareness, mode confusion, complacency, off-balance trust in automation (i.e., over-trust or mistrust), head-

down driving, automatic responding, and skill loss are the key human-automation challenges.

6 REFERENCES

- Abe, G., Itor, M. & Tanaka, K. (2002). Dynamics of drivers' trust in warning systems IFAC Proceedings Volumes, 35(1), 363-368.
- Bainbridge, L. (1983). Ironies of automation. Analysis, Design and Evaluation of Man-Machine Systems. G. Johannsen and J. E. Rijnscorp, Elsevier: 129-135.
- Balfe, N. (2010). Appropriate automation of rail signalling systems: a human factors study. Ph.D. Thesis, University of Nottingham, Nottingham, UK.
- Balfe, N. (2020). Human factors applications of on-train-data-recorder (OTDR) data: an exploratory study. *Cognition, Technology & Work*, 1-15.
- Bearman, C. & McCusker, L. (2008). A review of the human factors literature on new in-cab and train control technologies CRC for Rail Innovation, Australia.
- Brandenburger, N., Hörmann, H.-J., Stelling, D. & Naumann, A. (2017a). Tasks, skills, and competencies of future high-speed train drivers. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 231(10), 1115-1122.
- Brandenburger, N. & Jipp, M. (2017). Effects of expertise for automatic train operations. *Cognition, Technology & Work*, 19(4), 699-709.
- Brandenburger, N., Wittkowski, M. & Naumann, A. (2017b). Countering train driver fatigue in automatic train operation. In Sixth International Human Factors Rail Conference, London, UK.
- Brandenburger, N., Thomas-Friedrich, B., Naumann, A. & Grippenkov, J. (2018). Automation in railway operations: effects on signaller and train driver workload. In Proceedings of the 3rd German Workshop on Rail Human Factors, ITS Mobility.
- Brandenburger, N. & Naumann, A. (2019). On Track: A Series of Research about the Effects of Increasing Railway Automation on the Train Driver. *IFAC-PapersOnLine*, 52(19), 288-293.
- Brandenburger, N., Naumann, A. & Jipp, M. (2019). Task-induced fatigue when implementing high grades of railway automation. *Cognition, Technology & Work*, 1-11.
- Buksh, A., Sharples, S., Wilson, J., Coplestone, A. & Morrisroe, G. (2013). A comparative cognitive task analysis of the different forms of driving in the UK rail system. Rail human factors: supporting reliability, safety, and cost reduction. N. Dadashi, A. Scott, J. R. Wilson and A. Mills, CRC Press: 173-182.
- Carey, M. (2015). The challenges of achieving effective alarm management for railway infrastructure systems. The Fifth International Rail Human Factors Conference. London, UK.
- Crick, J., McCorquodale, B., Chissick, C., Diamond, H., Lear, A., Page, H. & McGuffog, A. (2004a). Driver error

- data collection project: final report. Qinetiq, Rail Safety and Standards Board (RSSB), London, UK.
- Crick, J., McCorquodale, B., Chissick, C., Diamond, H., Lear, A., Page, H. & McGuffog, A. (2004b). Driver error data collection project: report on work package three-data collection. Qinetiq, Rail Safety and Standards Board (RSSB), London, UK.
- Cullen, L. & Smith, S. (2004). Rail-specific human reliability assessment technique for driving tasks. Rail Safety and Standards Board (RSSB), London, UK.
- Davies, K., Thorley, P., Thomas, M. & Gibson, H. (2012). Feasibility of head up displays in driving cabs. Rail human factors around the world: impacts on and of people for successful rail operations. A. M. J.R. Wilson, T. Clarke, J. Rajan, & N. Dadashi, CRC Press. 465: 465.
- Endsley, M. R. (1996). Automation and situation awareness. *Automation and Human Performance: Theory and Applications*, 20, 163-181.
- Evans, A. W. (2014). Fatal train accidents on Europe's railways: 1980-2013.
- Foulkes, J. (2004). Impact of the European rail traffic management system on driver workload. Early human factors analysis for UK-ERTMS driving (Final Version). Human Engineering Ltd, Shore House, Bristol, UK
- FRA. (2014). Collaborative incident analysis and human performance Handbook. Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.
- French, S., Bedford, T., Pollard, S. J. & Soane, E. (2011). Human reliability analysis: A critique and review for managers. *Safety science*, 49(6), 753-763.
- Giesemann, S. (2013). Automation effects in train driving with train protection systems—assessing person-and task-related factors. Rail human factors. supporting reliability, safety and cost reduction. N. Dadashi, A. Scott, J. R. Wilson and A. Mills, CRC Press: 139-149.
- Halliday, M., Gall, B., Gibson, H. & Raggett, L. (2005). Assessing the effectiveness of an in-cab signal reminder device Rail Safety and Standards Board (RSSB), London, UK.
- Hely, M., Shardlow, T., Butt, B., Friswell, R., McIntosh, A. & Williamson, A. (2015). Effects of automatic train protection on human factors and driving behaviour. In Proceedings 19th Triennial Congress of the IEA, Melbourne, Australia.
- Hou, L.-X., Liu, R., Liu, H.-C. & Jiang, S. (2021). Two decades on human reliability analysis: A bibliometric analysis and literature review. *Annals of Nuclear Energy*, 151, 107969.
- Lawton, R. & Ward, N. J. (2005). A systems analysis of the Ladbroke Grove rail crash. *Accident Analysis & Prevention*, 37(2), 235-244.
- Liu, A. M., Oman, C. M. & Voelbel, K. (2017). Development and Evaluation of a Moving Map Display for Rail Applications. Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.
- McBride, S. E., Rogers, W. A. & Fisk, A. D. (2014). Understanding human management of automation errors. *Theoretical Issues in Ergonomics Science*, 15(6), 545-577.
- McLeod, R., Walker, G. & Moray, N. (2003). Extended AWS study: review of the knowledge base. Nickleby HFE Ltd, glasgow, UK.
- McLeod, R. W., Walker, G. H. & Mills, A. (2005). Assessing the human factors risks in extending the use of AWS. Rail human factors: supporting the integrated railway. B. N. J. Wilson, T. Clarke, & A. Mills. London, Routledge: 109-119.
- Merat, N., Mills, A., Bradshaw, M., Everatt, J. & Groeger, J. (2002). Allocation of attention among train drivers. *Contemporary Ergonomics*. P. T. McCabe: 185-190.
- Mkrtychyan, L., Podofillini, L. & Dang, V. N. (2015). Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability engineering & system safety*, 139, 1-16.
- Monk, A., Cross, M.-E., Mcfarland, S., Collis, L., Agutter, R. & Fletcher, D. (2017). ERTMS transitions: What impact do they have on the driver? Sixth International Human Factors Rail Conference. London, UK.
- Moray, N., Neil, G. & Brophy, C. (1983). Selection and behaviour of fighter controllers. Ministry of Defence, London, UK.
- Naghiyev, A., Sharples, S., Carey, M., Coplestone, A. & Ryan, B. (2014a). ERTMS train driving-in cab vs. outside: An explorative eye-tracking field study. *Contemporary Ergonomics and Human Factors S. Sharples and S. Shorrock*, CRC Press.
- Naghiyev, A., Sharples, S., Ryan, B., Coplestone, A. & Carey, M. (2014b). Alerts and Alarms in Conventional and ERTMS Train Driving: An Exploratory Eye-Tracking Field Study. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications
- Naweed, A. (2014). Investigations into the skills of modern and traditional train driving. *Applied Ergonomics*, 45(3), 462-470.
- Nneji, V. C., Cummings, M. L. & Stimpson, A. J. (2019). Predicting locomotive crew performance in rail operations with human and automation assistance. *IEEE Transactions on Human-Machine Systems*, 49(3), 250-258.
- Oppenheim, I., Shinar, D., Carsten, O., Barnard, Y., Lai, F., Vanderhaegen, F., Polet, P., Enjalbert, S., Pichon, M., Hasewinkel, H., Lützhöft, M., Kircher, A. & Kecklund, L. (2010a). Critical review of models and parameters for Driver models in different surface transport systems and in different safety critical situations. IT for error remediation and trapping emergencies. I. Oppenheim. European Commission Directorate-General Research
- Oppenheim, I., Shinar, D., Enjalbert, S., Dahyot, R., Pichon, M., Ouedraogo, A., Lützhöft, M., Carsten, O.,

- Hjälmdahl, M. & Cacciabue, C. (2010b). Description of Unified Model of Driver behaviour (UMD) and definition of key parameters for specific application to different surface transport domains of application. IT for error remediation and trapping emergencies. D. Shinar. European Commission Directorate-General Research
- Papadimitriou, E., Schneider, C., Tello, J. A., Damen, W., Vrouwenraets, M. L. & Ten Broeke, A. (2020). Transport safety and human factors in the era of automation: What can transport modes learn from each other? *Accident Analysis & Prevention*, 144, 105656.
- Park, D., Yoon, W. C. & Lee, U. (2020). Cognitive states matter: Design guidelines for driving situation awareness in smart vehicles. *Sensors*, 20(10), 2978.
- Patriarca, R., Ramos, M., Paltrinieri, N., Massaiu, S., Costantino, F., Di Gravio, G. & Boring, R. L. (2020). Human reliability analysis: Exploring the intellectual structure of a research field. *Reliability Engineering & System Safety*, 107102.
- Rad, M. A., Hendry, M. & Lefsrud, L. (2021). Literature review on cognitive impacts of in cab warning systems. *Transport Canada*
- Robinson, D., Waters, S., Basacik, D., Whitmore, A. & Reed, N. (2015). A pilot study of low workload in train drivers. *Transport Research Laboratory, Crowthorne, UK.*
- Rose, J. A. & Bearman, C. (2012). Making effective use of task analysis to identify human factors issues in new rail technology. *Applied Ergonomics*, 43(3), 614-624.
- Roth, E. & Multer, J. (2009). Technology implications of a cognitive task Analysis for locomotive engineers. *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.*
- Roth, E., Rosenhand, H. & Multer, J. (2013). Using cognitive task analysis to inform issues in human systems integration in railroad operations. *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.*
- RSSB. (2002). Driver reminder appliances: effectiveness study. *Rail Safety and Standards Board (RSSB), London, UK.*
- Safar, H., Multer, J. & Roth, E. (2015). An investigation of passing stop signals at a passenger railroad. *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.*
- Safar, H., France, M., Multer, J. & Roth, E. (2020). Preventing Stop Signal Overruns: Good practices for passenger railroads. *Federal Railroad Administration, Volpe Center*
- Sebok, A., Wickens, C., Laux, L. & Jones, M. (2015). Supporting human-automation interaction in the rail industry by applying lessons from aviation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications.
- Sebok, A., Walters, B. & Wickens, C. (2017). Simulator-based investigations of human error potential with railroad automation in the locomotive cab. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications
- Spring, P., McIntosh, A. & Baysari, M. (2009). Counteracting the negative effects of high levels of train automation on driver vigilance. In *Proceedings of the 45th Annual Human Factors and Ergonomics Society of Australia Conference.*
- Stein, C., Liu, A., Brown, D. A. & Porch, A. (2019). Monitoring engineer fatigue (MEFA). *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC.*
- Tao, J., Qiu, D., Yang, F. & Duan, Z. (2020). A bibliometric analysis of human reliability research. *Journal of Cleaner Production*, 260, 121041.
- Thomas, M. & Davies, K. (2008). Feasibility of head-up displays in driving cabs. *8th World Congress on Railway Research. Seoul, Korea.*
- Van Der Weide, R. (2017). ERTMS roll out from a train driver's perspective. *Sixth International Human Factors Rail Conference. London, UK.*
- Van Der Weide, R., De Bruijn, D. & Zeilstra, M. (2017). ERTMS pilot in the Netherlands—impact on the train driver. *International Human Factors Rail Conference. London, UK.*
- Verstappen, V. J., Wilms, M. S. & Van Der Weide, R. (2017). The impact of innovative devices in the train cab on train driver workload and distraction. *Sixth International Human Factors Rail Conference. London, UK.*
- Wilde, G. J. & Stinson, J. F. (1983). The monitoring of vigilance in locomotive engineers. *Accident Analysis & Prevention*, 15(2), 87-93.
- Wreathall, J., Roth, E. M., Bley, D. & Multer, J. (2003). Human reliability analysis in support of risk assessment for positive train control. *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.*
- Wreathall, J., Roth, E., Bley, D. & Multer, J. (2007a). Human factors considerations in the evaluation of processor-based signal and train control systems: human factors in railroad operations. *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.*
- Wreathall, J., Woods, D. D., Bing, A. J. & Christoffersen, K. (2007b). Relative risk of workload transitions in positive train control. *Federal Railroad Administration (FRA), US Department of Transportation, Washington, DC, USA.*
- Zimmermann, K. (2015). Literature review on the human factors impact of train control technologies or other in-cab automation *Transport Canada Rail Safety, Canada.*

Selection of track forms – an LCC perspective

Wenxu Li, Richard Dwight

School of Mechanical, Materials, Mechatronic and Biomedical Engineering– University of Wollongong, Wollongong, NSW, Australia

ABSTRACT

It has been more than half a century since the first introduction of slab track to the railway industry. Significant understanding of the pros and cons of slab track has been developed. However, apart from situations which clearly favour one track form over the others, people are often perplexed on the right form to choose for their projects. Well applied, life-cycle costing, by assessing the performance of a track form over its whole life cycle, is able to provide a holistic and impartial view. It allows different track forms and variants within basic forms to be compared. Three LCC models of increasing complexities have been developed to cater for different available data situations. All of the models are based on a cost-breakdown. The two simpler models require only cost and track usage information. The third model is an engineering model requiring a substantial amount of data but promising better information and accuracy: being more adaptable to situation and design specifics. The trial results tend to suggest that slab track is more advantageous than the ballasted track based on the data acquired. Constrained by the available data, the work details the two simpler models and demonstrates their implementations with examples. The methodology for the third model has been set out. Directions for further development are extracted from the work identifying required data for more accurate model output.

1 INTRODUCTION

Slab track was first introduced by Japan to their high speed rail links to replace ballasted track. This introduction was to achieve more acceptable maintenance demand in the face of growing requirement for availability in the early 1970s (Anto et al. 2001). Fast growth in applications worldwide has followed since then. Not only used almost as an exclusive option for building high speed lines, slab track is also the preferred option on structures and at places where the ground or subgrade is consolidated. Even at non-conventional sites where ballasted track has always been seen as the only option, slab track has been considered as an alternative. Examples include applications to support heavy haul operations. Promising results have been reported (Li & Bilow 2008, Karsten & Zhang 2015), but it is the structural integrity instead of the life-cycle cost that is of interest in these applications. No solid conclusion as to the overall performance of slab track in comparison with ballasted track can be drawn, which may also be partly attributed to the short durations of these trials: the slab track is far from reaching its service life.

Compared to ballasted track, slab track has many features which make its application attractive. The first and foremost is its apparently low maintenance requirement resulting in high availability. This is particularly important for lines where headway is short and disruption is costly, e.g., urban commuter lines. This is a result of the compact and uniform design of the slab track which renders it resistant to track geometry

deterioration and structural degradation. Slab track also features higher stability and ease of access. However, all these benefits come at an increased cost of design and construction, or, the initial capital investment. When additional measures are taken such as those to mitigate the associated noise and vibration impact, the cost can be pushed up. Frequently, when a decision is made with respect to which track form to choose, despite being the more promising choice in many situations, such as at places where flying ballast is an issue or on elevated structures such as bridges, tunnels and viaducts, slab track is thwarted by no other reasons but its high initial cost in some cases. In others cases, lack of experience on its maintenance and operation may also account for the reluctance in its adoption by the rail industry. But the latter has become less and less of a concern with the advancement of technologies and continual accumulation of experiences around the world.

Understanding the whole of life performance of different track forms therefore becomes extremely relevant under this context. Life cycle cost, LCC, by being able to quantify cost performance, in conjunction with prescribed reliability and availability, provides an effective measure in searching for an optimal track form solution on a project-by-project basis. The importance of LCC evaluation also arises from the recognition that construction cost, a significant contributor to LCC, tends to be only secondary in comparison with operation and maintenance costs when the whole of life of a track is taken into account in most cases (Garry Thuer & Vocks 2012).

Various LCC models have been developed, either as an independent tool (Thokala et al. 2010, Pratico & Giunta 2018) or as part of a decision support system (Zoeteman 2001, Stanislav & Esveld 2001). Pertinent concerns with regards to LCC modelling work include: the cost elements to be included: e.g., agency cost (costs borne by asset owner or manager), user cost (service disruption cost) or externality cost (environmental cost) or all of them. Other considerations may include, what system or components to include: sleepers, rails or the track system overall; how maintenance is carried out, i.e. the maintenance regime and the restoration effect of each maintenance action; and, the interaction between track components on their subsequent degradation development. Degradation modelling and maintenance optimization each by itself may deserve independent and long-term exploration. The uncertainties involved at each step of the modelling process also need to be addressed if models representative of reality are to be developed. The availability of data, however, plays a decisive role on what model can be implemented. Finally, each model developed has to be fit for the purpose: it should be at the right level of complexity for the problem addressed.

This work starts with a brief introduction on a generic process proposed for selecting a track form. Three LCC models developed to support the process are then introduced. These models increase in complexity: level of detail, in order to suit different data situations. LCC here consists of agency cost, i.e. construction cost, maintenance cost, and disposal cost. Disposal cost is considered because the efforts of decommissioning tracks of different track forms may differ and different costs/incomes may arise from disposing the decommissioned track. User and externality costs will be translated into agency costs based on the impact the underlying events have on the agency. Costs of all the track components can be conveniently incorporated into these models. This cost break-down structure is consistent with IEC (2019) specifications.

Trial results of two of the three proposed models are reported by using data supplied by rail managers and extracted from the literature. The principle of the last model is introduced. The theoretical basis and implementation for it can be found in Li & Dwight (2021). Limited by the data accessible, trials of this model are still ongoing. It is considered that these models are applicable to all situations where a decision to select a track form or sub-variant of a track form is to be made.

2 TRACK-FORM SELECTION PROCESS

The selection process targets all decisions where a track form is to be selected. It consists of four stages, as illustrated in Figure 1. In the preparation stage the context of the project is investigated to obtain the attributes of the track of interest (to be constructed). The attributes explored may include: length; usage; traffic type; environmental impacts and construction conditions such as geological and environmental information at the site and the supporting structure requirements. Finally the resources in place need to be understood, which

include the organizational culture; construction resources and maintenance ease and possibility etc.

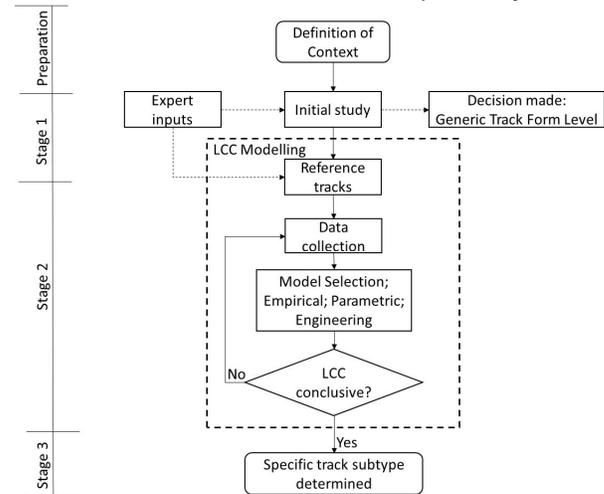


Figure 1. Track form selection process: solid arrows: sequence of flow; dashed arrows: input or output from a step

Stage 1 applies heuristic analysis to identify whether there are attributes of the track of interest which preclude or favor a particular track form, or, if a particular track form implies a lower cost for the project. If no track form can be excluded/determined from this stage analysis, stage 2 proceeds to evaluate LCCs of all the candidate track forms. Otherwise, stage 2 only goes ahead with the track forms or variants of a track form based on the refined scope of options out of Stage 1. This stage may involve: collection of data from reference tracks which have data available and identification of those which can inform the LCC of the track of interest, i.e. with similar attributes; selecting the right LCC model from the three proposed LCC models based on the data situation and the required level of details on the LCC; generation of LCC projections by the use of the selected model; and, lastly, inspecting the results and replacing the model with more accurate ones with more data collected until the results are evident enough to differentiate the candidate options. Other factors affecting the decision such as budget constraints and resources accessible for implementing the decision may also be further considered at this stage. Failure of complying with these constraints also directs the analysis to the data collection step. Stage 3 concludes the process by moving the project forward by preparing and allocating resources necessary to adopt the determined track option.

It is noted that an LCC optimal solution may not be the most suitable when all influencing factors are balanced. Some track options may be missed during the analysis. Nonetheless, even though the final solution may not be optimal in terms of LCC, it is still considered sub-optimal by allowing for all of the interrelated factors and options accessible.

3 LCC MODELS

LCCs that are realised are a function of both the design of a track and the approach taken to its operation and maintenance. In order to generalize the developed LCC models, each cost component is related to the cost drivers such as: usage; the design features; and, the maintenance and disposal activities. In addition, by using cost rates in accordance with the market rates corresponding to the track of interest, localized effects introduced by a reference track: such as labour rate; and, material cost can be minimized.

To enable comparison between different track forms and their components which are likely to have different time horizons of interest, all costs are expressed in terms of Present Value (PV). This value can be easily converted into other cost measures such as annuity and future value (Fabrycky & Blanchard 1991) if required.

Three models are developed to take care of different data situations. Model 1 requires only cost per unit time or use values. Model 2 requires event cost and frequency data. Model 3 relies on degradation modelling to determine the maintenance needs, i.e. use of simulated track component degradation levels and predefined maintenance initiation criteria to determine timings of various maintenance activities for these track components. As a result, the accessible data situations determine the possibility for utilising one of these 3 modelling approaches. When the available data enable more than one models, the more complex model can provide a deeper insight with improved accuracy where this is useful.

Model 1

The following assumptions are made in order to simplify the formulation for the model:

- 1) construction time is so short compared to the life of a track that there is no difference between nominal and present value of construction cost (discount rate can be assumed to be 1.0).
- 2) maintenance cost rate \dot{c}_m , i.e. cost per unit length and per unit time, is a linear function of track usage, e.g., the accumulative load carried or the number of train passages, and headway;
- 3) disposal is dealt with at the whole railway line/section level. It occurs only once when the line or section is decommissioned. The disposal-and-replacement activities or renewals during the life time of a track on some components are considered maintenance actions and the cost is included in the maintenance cost.

It is noted that the model structure allows the assumptions to be relaxed whenever necessary: if construction time is significant (though not likely), the cash flow per annum during construction can be converted to PV by including an inflation rate and a discount rate and \dot{c}_m can also be made a non-linear function of track usage and headway if actual operation data indicate so, e.g., exponential growth over time as a result of accelerated track degradation.

LCC is calculated by

$$LCC = C_c + C_m + C_d \quad [1]$$

In Eq.(1), the construction cost, $C_c = L \bar{c}_c$, where L is the length of the track of interest and \bar{c}_c is the cost per unit length for the reference track; the maintenance cost, $C_m = L \sum_i \dot{c}_m(i, t_i, h_i)$, where $\dot{c}_m(i, t_i, h_i)$ is a function of usage, i.e. tonnage t_i , and headway, h_i at track age i ; the disposal cost, $C_d = L \bar{c}_d$, where \bar{c}_d is cost per unit length for the track disposed.

After taking into account the inflation rate k and discount rate j , the PV of LCC, LCC_{PV} , is given by,

$$LCC_{PV} = \sigma L \left(\bar{c}_c + \sum_{i=1}^n \dot{c}_m(i, t_i, h_i) \frac{(1+k)^i}{(1+j)^i} + \bar{c}_d \frac{(1+k)^n}{(1+j)^n} \right) \quad [2]$$

where σ is an adjustment factor designed to account for the similarity between the reference track and the track of interest. It is a value between 0 and 1. A larger σ indicates greater similarity between the reference track and the track of interest. n is the length of service life. $\dot{c}_{m,i}$ is related to t_i and h_i linearly by,

$$\dot{c}_m(t_i, h_i) = a \frac{t_i}{h_i} + b \quad [3]$$

The model parameters a and b can be estimated by linear regression based on the data obtained from the reference track.

Among all of the parameters used in the model, only σ needs subjective input. It can be directly represented by the similarity level between a reference track and the track of interest or correlated with the similarity level through a function if more appropriate. This similarity level is 'at the track level' which can be assembled from those 'at the attribute level'. The latter typically involves inputs from experts, especially for attributes which are not quantifiable, e.g., ease of maintenance and access. One approach to quantify these attributes is to assign a score to each one of them, e.g., a score within 0 to 5, with a bigger value indicating a higher similarity level. In order to further account for the role an attribute plays in affecting the decision of selecting the track, a weight is further designed for each of the attributes. Ease of maintenance, for instance, may be considered more important than the noise impact as the latter can often be addressed after construction. A bigger weight therefore can be assigned to the ease of maintenance attribute. The similarity 'at track level' can then be assembled from the similarities at the attribute level by summing up the products of these similarities and their corresponding designed weights, further normalized by the sum of the weights.

Though arbitrary as the approach seems to be, by combining results from a number of experts through some controlled process, the subjectivity may be minimized. The level of sophistication of such combination of views ranges from simple averaging of the similarity levels suggested to various consensus processes such as the DELPHI approach (Jillson 1975).

Model 2

Model 2 is constructed based on the same principle as Model 1, i.e. Eq.(1). Model 2 differs from Model 1 on that it assembles LCC from a lower level of cost elements:

those for individual maintenance activities instead of an aggregated cost rate. The benefits are an LCC with a higher accuracy tailored to the actual situation and higher confidence with the analysis among the users. In the case where data are a mixture of aggregated and activity-specific costs, a hybrid model combining Model 1 and Model 2 is probable. This can be easily derived by simplifying the formulation of model 2. When only aggregated costs are available, Model 2 reduces to Model 1.

To implement the model, the same set of assumptions to those of Model 1 are required. As noted for Model 1, these assumptions are imposed to simplify implementation and may be relaxed to suit the situations. In the formulation, the construction and disposal costs are the same to those in Eq.(2). Maintenance cost, however, is assembled from servicing cost, inspection cost, repair cost and accident (unplanned events) related cost, and is given as,

$$C_m = \sum_{y=0}^t \left(\begin{array}{l} \sum_p \dot{c}_{ser,p}(y) + \sum_p \dot{c}_{a,p}(y) \\ \sum_p \sum_q n_{inp,q,p}(y) \dot{c}_{inp,q,p}(y) \\ + \sum_p \sum_q n_{rep,q,p}(y) \dot{c}_{rep,q,p}(y) \\ + \sum_m n_{acc,m}(y) \dot{c}_{acc,m}(y) \end{array} \right) \frac{(1+k)^y}{(1+j)^y} \quad [4]$$

where $\dot{c}_{ser,p}(y)$ is the cost of servicing of type p per unit track length at year y ; $\dot{c}_{a,p}(y)$ is the cost per unit length, due to unplanned/prolonged track unavailability for event p at year y , the determination of which depends on the penalty regime in place or the impact on the service users; $\dot{c}_{inp,q,p}(y)$ is cost per inspection of type q on track component p per unit length at year y ; $\dot{c}_{rep,q,p}(y)$ is the cost per repair of type q on track component p at year y ; $\dot{c}_{acc,m}(y)$ is the cost of an accident of type m per unit track length at year y . The n 's in Eq.(4) represent the respective numbers of occurrences of the various maintenance activities, i.e. frequencies.

The costs can be informed by the reference track data, but the frequencies of accidents or maintenance actions have to be based on the perceived usage and headways of the track of interest. If a linear relationship exists, each frequency of occurrences is given by

$$n(t, h) = a \frac{t}{h} + b \quad [5]$$

The model parameters a and b are determined based on the data collected from the reference tracks. t and h are the traffic load and headway, respectively.

It is noted that, during the implementation of the model: 1) some parameters required by Eq.(4) may not be available individually but as a combination, e.g., the frequency of occurrences of a maintenance activity and the associated cost value may be provided as a cost rate per year for that activity. In this case, the hybrid model of Model 1 and Model 2 is used, i.e. the product of event frequency and cost is replaced by the corresponding annual cost rate;

2) the various cost rates for maintenance actions may also be related to other explanatory variables such as durations of track possession and loading conditions. The new relationship can be plugged into Eq.(4).

Model 3

This model is a simulation-based engineering model (Li & Dwight 2021). It integrates degradation models for track components with maintenance thresholds to determine maintenance needs. By further combining with distributions determined for traffic interference durations, the time required for maintenance and unplanned events can be simulated. LCC is estimated by converting these durations into costs. This model is segment-based from which a larger scale of model can be assembled, e.g., a section or a network. Here a segment refers to the smallest unit length of track where the degradation data are collected.

In the model, only components of the track system are considered. Trackside auxiliary systems such as signalling systems and the overhead wiring system are excluded, but by being modular the model allows extension to take them into account once relevant models are available.

The modular concept is illustrated in Figure 2 where the track of interest is divided into consecutive segments, from 1 to n . Maintenance cost is calculated for each segment separately and then aggregated to obtain the overall maintenance cost for that track.

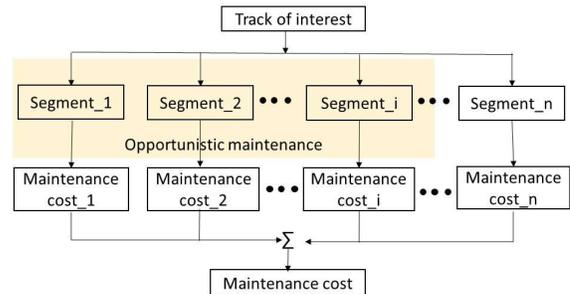


Figure 2. Structure of the proposed model to estimate maintenance cost of the track. The shading region represents an instance of the application scope of opportunistic maintenance: some segments.

The model used to determine the maintenance cost for any given segment is given in Figure 3 which is component-based. LCC is aggregated from the cost for each component.

For a ballasted track, three track components are considered: sleepers, ballast and rails. The failures considered include: sleeper failures; rail rolling contact fatigues, and surface wear including corrugation, and track geometry deterioration. Degradation models of fastening systems are not included as there are still no proper models available for them. Longitudinal level and horizontal alignment are measures used to represent track geometry conditions as they are currently the most widely used indicators to initiate track geometry correction actions. The maintenance actions considered include: sleeper replacement; ballast tamping and stone-blowing, ballast replacement; and, rail grinding;

inspections of various track components. In addition, budget constraints are allowed at each simulation time step if relevant. Spatial correlation is accounted for whenever possible, e.g., between the geometry degradation processes of neighbouring track segments. Maintenance types considered include: preventive, corrective, concurrent and opportunistic maintenance. Opportunistic maintenance is depicted in Figure 2. It refers to the kind of maintenance which takes place on some segments by taking advantage of opportunities arising from work on adjacent segments, e.g., due to availability of repair resources and track possession. Here the scope is restricted to components of same type to those of the relevant segments. A pertinent maintenance strategy is concurrent maintenance which differs with opportunistic maintenance on that different track components from a same segment are maintained at one time. Preventative maintenance and corrective maintenance differ by the timing of application: whether the specified maintenance threshold is reached. The model allows the exploration of various combining strategies of these maintenance types.

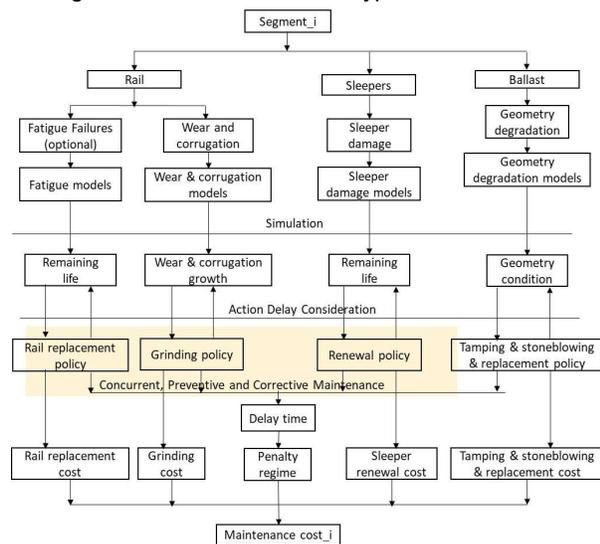


Figure 3. Structure of the proposed model to estimate maintenance cost of a ballasted track form. The shaded region represents the application scope of the noted maintenance types.

To apply the model to a slab track, modifications of what is given in Figure 3 to reflect the specific configuration of the slab track considered are necessary, such as: a degradation model for the supporting track slab should be introduced to replace the ballast model; the sleeper model may be removed if not relevant to the system and new maintenance strategies may be included. It is noted that failure mechanisms for concrete slab are still not well understood.

4 MODEL TRIALS

Model 1 and Model 2 have been implemented based on data collected from both the literature and the industry. For the purpose of these trials, no complete context data

were able to be established. The adjustment factor σ in Eqs.(2) and (4) is thus irrelevant. Comparison of LCC of the ballasted and slab track is reported without the influence of these factors included.

4.1 Model 1 Trial

Cost data (Ove Arup & Partners Ltd 2003) were collected in a feasibility study which was carried out to investigate the performance of a slab track in supporting heavy axle load freight traffic while still maintaining the standards for operating high speed trains.

The attributes of the track of interest are reported in Table 1. It is noted that data on some apparently relevant attributes are not available. Because σ is assumed to be 1 for both the slab and ballasted track, the missing values have no impact on LCC evaluation.

Table 1. Attributes of the candidate track (Li & Bilow 2008)

Aspect of Context	Elements	Measures
Operation	traffic type, e.g., train suspension design, and traffic speed	Mixed traffic: class 9 track (320kmh); freight train: 64kmh
	performance or condition requirements	39-ton axle load
	train schedules (operating patterns)	late night to the next morning
	nature of penalties for lack of performance	N/A
Use profile	traffic volume	170 million gross tons over 3 years
Community impact	local community requirements (or cost as a function of impact on local environment)	N/A
	environmental disturbance: noise and vibration; air pollution	N/A
Construction	Accessibility (such as Open ground or built-up areas; rivers; mountains)	Open ground
	geotechnical and environmental conditions	Silty sand; after compacting, the modulus of the soil is > 6.9Mpa and the strength is > 0.3Mpa.
	existing structures and their condition (such as Bridges; viaducts; streets)	N/A
	availability of construction machinery and the skills of engineers	N/A
	track length	76.2m
	Resources available	organisational context (including management approach)
	location and access to construction and maintenance resources (includes issues of economies of scale)	N/A
	budget allocable for construction	N/A

The accessible cost data from the study are the overall construction cost and annual maintenance cost as replicated in Table 2. This level of information available makes Model 1 the preferable model option.

Table 2. Ballasted track and slab track costs from the reference track: numbers all in US dollars

Track System	Construction cost, \$	Maintenance Cost, \$
Ballasted track	\$1 Million per mile (\$621.37/m)	\$50,000 per mile (\$31.07/m) (per year)
Slab track	\$1.3 Million per mile (\$807.78/m)	\$10,000 per mile (\$6.21/m) (per year)

To implement Model 1, further assumptions are made:

- the discount rate is 4%;
- the inflation rate in US at 2003 is $k = 2.3\%$ (statista 2020);
- the disposal cost is insignificant compared to the construction and maintenance cost, i.e. $\dot{c}_d = 0$ in Eq. (2)
- there was temporal change in maintenance cost rate, i.e., \dot{c}_m is constant and always equal to the value given in Table 2;
- the design life of the ballasted track is 20 years;
- the design life of the slab track is 50 years;
- there is no significant accident within the life time of both track forms;
- the renewal cost of ballasted track at the end of its service life is equal to its construction cost;
- renewal of a track occurs at the end of its service life and includes all of the track components.

It is noted that track length in Eq. (2) cancels out of the equation when the LCC ratio is calculated. The indicated projected cost ratio over 50 years is given in Figure 4. It can be seen that the two track forms have equivalent indicated expenditures approximately 9 years after construction. Subsequently ballasted track is more expensive, driven by higher maintenance costs. A significant cost increase takes place when ballasted track needs to be replaced, as shown by the jump in cost presented in the figure. This figure indicates that slab track is the preferred track form if ownership of the track is more than 10 years. Note that this result is significantly impacted by the limited available data and lack of specifics of the situation under study.

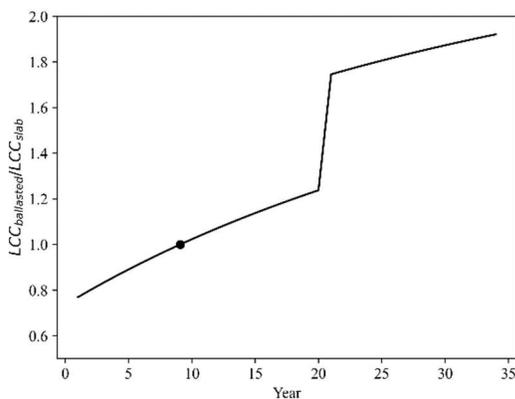


Figure 4. LCC ratio of ballasted track and slab track since construction based on Model 1

A sensitivity analysis has also been conducted where the discount rate, construction cost rate and maintenance

cost rate are independently varied. The resulting ratios of overall costs of ballasted and slab track are shown in Figure 5 a to c, respectively. A review of the results indicates that with the increase of discount rate from 4 to 10%, the LCC of ballasted track remains lower than that of the slab track for additional 3.5 years. An increase in either construction or maintenance cost for slab track puts off the date when slab track LCC equals that of the ballasted track significantly, a signal that slab track becomes less viable. Due to the shorter service life, ballasted track is renewed more frequently. The accumulated construction cost of the ballasted track surpasses that of the slab track after the first renewal. When the discount rate approaches 0.1, the construction cost of the slab track exceeds that of the ballasted track or the maintenance cost of slab track closely approaches that of the ballasted track, the LCC growth rate of the ballasted track slows down to a level below that of the slab track. A negative slope may thus occur as observed in Figure 5 c.

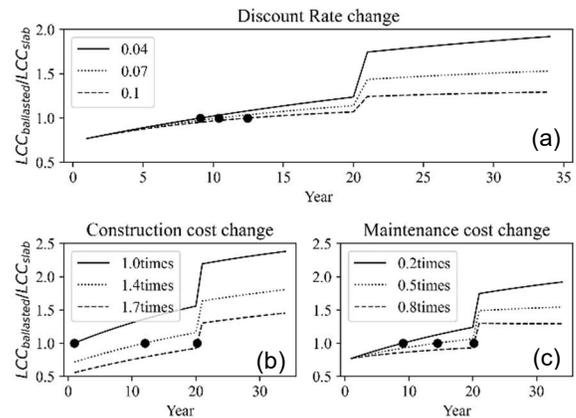


Figure 5. LCC ratios of ballasted track and slab track since construction based on Model 1 by varying discount rate (a); construction cost rate (b) and maintenance cost rate (c)

4.2 Model 2 Trial

The data were from a particular operator who operates two sections of track with equal length but different types: one is a slab track and the other ballasted. Both are constructed on bridges. Various maintenance activities and their costs were recorded. The construction costs for ballasted track and slab track, however, are not available. The values under the US context as shown in Table 2 are taken instead. It is understood that deviation from the actual LCC values is expected because of this data intake but the main focus here is to demonstrate the implementation of Model 2.

The reported activities include various inspections by different groups of personnel and inspection equipment as well as maintenance activities such as routine maintenance, reactive maintenance, rail grinding, surfacing and sleeper pad replacement. Both the frequency and cost rate per meter of the reported activities are available.

Assumptions made with respect to the data in order to implement Model 2 include:

- the purchasing power difference between

- jurisdictions is insignificant;
- the inflation rate from 2003 to 2019 when maintenance cost data were reported is constant and equal to 2%;
- the discount rate is 4%;
- the frequency and the cost rate of each maintenance activity are constant;
- the consumer price inflation rate $k = 1.6\%$ (Reserve Bank of Australia 2020);
- the disposal cost is insignificant compared to the construction and maintenance cost, i.e. $c_d \approx 0$ in Eq. (4);
- the servicing cost is insignificant such that its absence does not impact the overall LCC;
- the occurrence of accidents is so rare that its impact on LCC is negligible;
- unexpected events leading to long track unavailability do not exist;
- the design life of the ballasted track is 30 years;
- the design life of the slab track is 100 years;
- the renewal cost of ballasted track at the end of its service life is equal to its construction cost.

Based on Eq.(4), the LCC ratio of the ballasted track and the slab track is calculated. The result is presented in Figure 6. It shows that LCC of the ballasted track quickly exceeds that of the slab track at around 1.5 years and remains high afterwards. Therefore it is not as viable as the slab track for this particular trial. By varying the construction costs of both track forms, it is expected the outcome will be different. The successive step changes in the figure represent renewals of the ballasted track at the corresponding years, i.e. 30, 60 and 90 years.

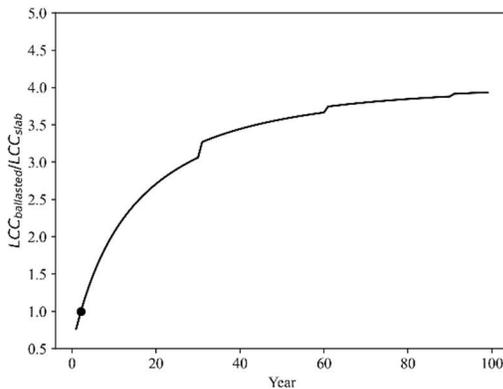


Figure 6. LCC ratio of ballasted track and slab track over years based on Model 2

Sensitivity analysis results are shown in Figure 7. Here it shows that maintenance cost has the biggest impact on the LCC ratio between the two track forms. However, slab track continues to outperform the ballasted track under all of the conditions considered in the sensitivity analysis.

5 RECOMMENDATIONS

For the completeness of LCC estimation, costs associated with the environment, which result in

externality cost, should be included. So are the costs arising from the impact of selecting a track form on customers and other entities. Though considered secondary in this work, these costs will increase in significance partly driven by reductions in construction and maintenance cost due to the continual advancement of technology and accumulation of experience on maintenance and operation. The major contributors to environmental cost may come from vehicle operation, track construction, maintenance as well as disposal (Spielmann & Scholz 2005). Models may be adopted or developed to account for these factors.

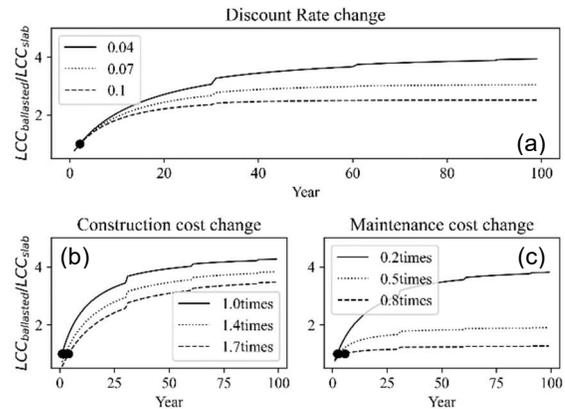


Figure 7. LCC ratios of ballasted track and slab track since construction based on Model 2 by varying discount rate (a); construction cost rate (b) and maintenance cost rate (c)

To facilitate the use of the models developed, an international database pooling experience from a wide range of application scenarios needs to be established. Context, cost and maintenance data collected for various tracks are managed in the database. When used, relevant data can be readily pulled from the database and plugged into the three developed or other models to find the optimal solution swiftly. Joint efforts across the whole rail industry are indispensable for the successful establishment and maintenance of the database.

There is an urgent need to compile data on operating experience for slab track applications to more unconventional areas, such as mixed traffic and heavy haul in terms of project initiatives and life cycle costs. It is reported that mixed traffic applications are common, particularly in Germany (Bastin 2006) and on some French railway lines (Robertson et al. 2015). Based on relatively high vertical stability, slab track is likely able to bear a heavier load than ballasted track, but no systematic study on this aspect is available. Therefore the proposed data compilation work can shed light on the cost-effectiveness of slab track under a series of application contexts and provide better guidance on planning similar future projects.

Ballasted track should not be ruled out too quickly when making a decision in applications where slab track is supposed to perform better and vice versa. Improvements in ballasted track technology have been reported. Those also include new measures to

counteract some of the disadvantages of ballasted track, including flying ballast, e.g., by lowering ballast level (Riessberger 2016). On the other hand, special treatments required on slab track, e.g., mitigation measures to address noise and vibration problems, may impact on its cost.

The support structures though not directly form part of a track physically but are closely related and have to be considered when conducting LCC study of a track. Slab track typically has a lower profile and smaller overall mass when compared to ballasted track, thus capable of reducing the scale and complexity of structure construction, e.g., by resulting in a smaller tunnel which in turn drives down the whole LCC. However, any change of track forms on a railway line may demand a transition zone which is costly to build and maintain, driving up the LCC.

6 CONCLUSION

Three LCC models, of varying complexity and accuracy, allowing LCC of a track form to be evaluated and compared with other track forms have been developed. All of the models are based on cost breakdown. By formulating the relationship between costs and the attributes of a track such as usage and operation pattern, the impact due to localized characteristics of a reference track can be minimized. An adjustment factor is introduced in these models to further account for the similarity in application contexts between a reference track and the track of interest. An empirical approach is proposed to compute this adjustment factor. It relies on expert inputs to determine similarities 'at the attribute level', from which similarity 'at the track level' is aggregated. It is expected that by seeking opinions from a number of experts and averaging the results, the adjustment factor may be reasonably estimated.

Among the three models, model 1 requires only costs per unit time/length; model 2 requires details of maintenance activities, frequencies, and their respective costs; model 3 uses degradation modelling to predict the maintenance needs. Together with the cost of each maintenance action LCC can be estimated. This more complex model, while providing a deeper insight with improved accuracy, requires considerably more input data. Conversely it can be better tailored to a specific situation and provide insights into detailed design decisions.

Sensitivity analysis conducted on the two simpler LCC models support the view that slab track's typically higher relative construction cost is of less relevance to LCC than the ongoing maintenance and renewal costs associated with ballasted track where ownership exceeds 10 years, based on the data used in the trials. Such a conclusion is based on critically limited data and awaits confirmation on a larger data set.

To make further gains in understanding the comparison of different track forms requires the development of an international database collating operating experience for a broad range of track forms applied to a range of situations. This is of particular importance given the significant impact of ongoing

maintenance and operating costs on the LCC of a particular track form.

7 REFERENCES

- Anto et al. (2001). Development of slab tracks for Hokuriku Shinkansen lines. *QR RTRI*, 42.
- Bastin, R. (2006). Development of German non-ballasted track forms. *Proceedings of the Institution of Civil Engineers - Transport*, 159, 25-39.
- Fabrycky, W. J. & Blanchard, B. S. 1991. *Life-cycle cost and economic analysis*, Englewood Cliffs, New Jersey, Prentice Hall.
- Garry Thuer & Vocks, H. 2012. Track slab system technology in Australia-innovation challenge based on European long term experiences. *AusRail 2012*. Canberra, Australia.
- IEC 2019. IEC 60300.3.3: Dependability management Part 3.3: Application guide - life cycle costing.
- Jillson, I. A. (1975). Developing guidelines for the Delphi method. *Technological Forecasting and Social Change*, 7, 221-222.
- Karsten, A. & Zhang, R. (2015). lab track performance under heavy haul revenue condition and its remediation method. Proceedings of the 13th International Railway Engineering Conference, Edinburgh, UK. The Institute for Rail Infrastructure Engineers.
- Li, D. & Bilow, D. N. (2008). Testing of Slab Track under Heavy Axle Loads. *Transportation Research Record*, 2043, 55-64.
- Li, W. & Dwight, R. (2021). A simulation-based LCC model for railway track. 11th International Conference on Modelling in Industrial Maintenance and Reliability, UK.
- Ove Arup & Partners Ltd (2003). Slab track: the commercial case-A scoping study.
- Pratico, F. G. & Giunta, M. (2018). LCC-based appraisal of ballasted track and slab tracks: limits and potential. *The Baltic Journal of Road and Bridge Engineering*, 13.
- Reserve Bank of Australia. (2020). Measures of Consumer Price Inflation. Australia. Accessed 23/07 2020 from <https://www.rba.gov.au/inflation/measures-cpi.html>.
- Riessberger, K. (2016). High performance ballasted track. Fixed track forms for high speed lines, Manchester Conference Center, UK. The Permanent Way Institution and the Union of European Railway Engineer Associations.
- Robertson, I., Masson, C., Sedran, T., Barresi, F., Caillau, J., Keseljevic, C. & Vanzenberg, J. M. (2015). Advantages of a new ballastless trackform. *Construction and Building Materials*, 92, 16-22.
- Spielmann, M. & Scholz, R. (2005). Life Cycle Inventories of Transport Services: Background Data for Freight Transport (10 pp). *The International Journal of Life Cycle Assessment*, 10, 85-94.

Stanislav, J. & Esveld, C. (2001). ECOTRACK: An objective condition-based Decision Support System for long-term track M&R Planning directed towards reduction of Life Cycle Costs. 7th International Heavy Haul Conference, Brisbane, Australia. 199-208.

statista. (2020). United States: Inflation rate from 1990 to 2020. Accessed 23/07 2020 from <https://www.statista.com/statistics/191077/inflation-rate-in-the-usa-since-1990/>.

Thokala, P., Scanlan, J. & Chipperfield, A. (2010). Life cycle cost modelling as an aircraft design support tool. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 224, 477-488.

Zoeteman, A. (2001). Life cycle cost analysis for managing rail infrastructure-concept of a decision support system for railway design and maintenance. *European Journal of Transport and Infrastructure Research*, 1, 391-413.

Session 6

HUMAN FACTORS

Applications of innovative accident analysis methods in railways: A review

Mona A. Rad^{1*}, Lianne M. Lefsrud^{1,2} and Michael Hendry¹

¹Canadian Rail Research Laboratory, Faculty of Civil Engineering, University of Alberta

²Chemical and Materials Engineering, Faculty of Engineering, University of Alberta

ABSTRACT

Accident analysis methods are used to determine the factors and their interrelationships that contributed to an accident. Various methods, including Fault Tree Analysis (FTA), Human Factors Analysis and Classification System (HFACS), and Systems Theoretic Accident Model and Processes (STAMP), were developed to model accident causation. However, the classical methods show weaknesses in accident modeling in sociotechnical systems that have complex dependencies of system components, and uncertainties in system behavior. To address the limitations, newer methods such as Bayesian networks (BNs), Petri nets (PNs), text mining (TM), and machine learning (ML) were used alone or with other techniques to model accidents. This article presents a review of publications in six databases (ScienceDirect, Scopus, Web of Science, SpringerLink, Google scholar, and IEEE Xplore) of these accident analysis methods for the railway industry. The publications are categorized into network-based and artificial intelligence (AI)-based accident analysis methods, and additional categories, such as the type of algorithms and techniques, data sources, and tools applied. The findings show that Bayesian networks and text mining are the most widely used network-based and AI-based methods for analyzing railway accidents.

1 INTRODUCTION

Risk management and accident analysis play a key role in understanding accidents' mechanisms and contributing factors and, thus, the most effective and efficient measures to prevent them. Heinrich (1931) presented the domino theory as the first accident analysis model to better learn from past events. Since then, various accident modeling methods have been developed to better represent our understanding of accidents and the complexity of sociotechnical systems. Traditional accident analysis models are classified into three categories: sequential (simple linear), epidemiological (complex linear), and systemic (complex non-linear) models based on their underlying assumptions (Underwood & Waterson 2013, Klockner 2015). Examples of the sequential accident models that describe an accident as the result of a chain of discrete events occurring in time-ordered sequences are fault tree analysis (FTA), failure modes and effects analysis (FMEA), and event tree analysis (ETA). Human factors analysis and classification system (HFACS) is the most popular epidemiological accident modeling (complex

linear). Systemic (complex non-linear) models such as Accimap, systems theoretic analysis model and processes model (STAMP), and functional resonance analysis method (FRAM) are the best-fit methods to analyze accidents in complex systems.

These classical accident causation models usually make simplifying assumptions and ignore some characteristics of complex systems including dependencies among causes and contributory factors, complex relationships between humans and automation, organizational and human influencing factors, temporal and dynamic system behavior, and uncertainties (Huang et al. 2018, Kabir & Papadopoulos 2019). To address some limitations of these classical approaches, network-based methods such as Bayesian networks (BNs) and Petri nets (PNs) have recently achieved popularity for analyzing accidents and safety risks. Moreover, the advent of artificial intelligence (AI) technology, including machine learning (ML) and text mining (TM), offers more possibilities for automatically exploring occurrence databases and accident reports and discovering implicit, but not immediately obvious, relationships.

* Corresponding author's email: ahmadira@ualberta.ca

Mkrtchyan et al. (2015) reviewed the use of BNs in human risk analysis (a.k.a., human reliability analysis (HRA)) and identified five main groups of BNs' usage involving: the modeling of organizational factors, analysis of the relationships among failure influencing factors, BN-based extensions of existing HRA methods, dependency assessment among human failure events, and assessment of situation awareness. The applications of BNs and PNs in system safety, reliability, and risk assessments were reviewed by Kabir & Papadopoulos (2019). They highlighted the efficacy of the BNs and PNs frameworks in comparison with the classical accident and safety analysis methods and illustrated their strengths and weaknesses as standalone or model-to-model transformation approaches in different practical application scenarios. Weber et al. (2012) also revealed the advantages of BNs over Markov chains (i.e., a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event) and fault trees techniques to model and assess the dependability in risk analysis. The main benefits of BNs were the capability to model complex systems, to make predictions as well as diagnostics, to compute the occurrence probability of an event, to update the calculations according to evidence, to represent multi-modal variables, and to help user-friendly modeling by a graphical and compact approach. Marcot & Penman (2019) showed that how the integration of BNs with other analytical frameworks such as management decision networks, structural equation modeling (SEM), and Bayesian neural networks can enhance Bayesian classifiers and machine learning algorithms, improve model structuring and parameterization, and facilitate the development of time-dynamic models.

Others such as Huang et al. (2018), Ghofrani et al. (2018), and Hegde & Rokseth (2020) put their efforts on reviewing the applications of Artificial Intelligence (AI)-based data analytics methods for incident analysis and their usefulness and gaps. Huang et al. (2018) emphasized a paradigm shift in accident investigation methodology in the era of big data and highlighted the advantages of the modern analysis methods over the classical ones. They concluded that data-driven accident analysis illustrates the circumstances of accidents more objectively, focuses on the relationships between a safety phenomenon and safety data, transforms accident analysis from qualitative to quantitative, recognizes the early warning and early intervention of an accident through real-time data, forecasts potential accidents, and are more congruent with new safety issues. The review conducted by Ghofrani et al. (2018) showed that descriptive analytics such as accident causes and influencing factors, accident frequency and severity, have higher popularity compared to predictive and prescriptive analytics in the data-driven rail safety analysis. They also proposed leveraging big data sources of rail infrastructure and train operations to merge smaller rail accident databases and trace the train accident occurrences based on a series of precursor events. Recently, Hegde

& Rokseth (2020) presented a thorough review of publications using machine learning in engineering risk assessment. They illustrated that risk identification enjoyed the most popularity among three phases of risk assessment (i.e., risk identification, risk analysis, and risk evaluation) in using machine learning algorithms. They also uncovered that the railway industry is third, after automotive and construction industries, for adopting machine learning in risk assessment. The domain-specific applications of ML and TM techniques for accident and risk analysis were also reviewed by Halim et al. (2016) and Gutierrez-Osorio & Pedraza (2020) in road transportation, Ismail et al. (2021) in the mining industry, Yan et al. (2020) in the construction industry, and George & Renjith (2021) in process industries.

To the best of the authors' knowledge, the literature in this field of study suffers from the lack of a thorough review of the applications of these accident analysis methods to the railway industry. Therefore, in this paper, we review studies that use these state-of-the-art methods including Bayesian networks, Petri nets, machine learning, and text mining to delineate factors that contributed to the railway occurrences and how they are correlated. The remainder of our paper is organized as follows: Section 2 describes the fundamentals of Bayesian networks, Petri nets, and machine learning and text mining. Section 3 summarizes our bibliometric search methodology. Descriptive analysis and details review are discussed in Sections 4 and 5, respectively. Finally, our conclusions of this review are summarized in Section 6.

2 BACKGROUND

2.1 Bayesian Networks

Bayesian networks (BNs) are directed acyclic graphs that model a set of variables and their conditional dependencies as nodes and edges (Kabir & Papadopoulos 2019, Qiao et al. 2020). A graph-like representation is a qualitative part of the BNs model and prior and conditional probabilities are the quantitative parts. In the acyclic graph, the nodes, which are shown as circles, represent the random variables and directed arcs illustrate dependencies or cause-effect relations among the nodes (Kabir & Papadopoulos 2019).

Features such as the intuitive graphical representation, modeling uncertainty, developing interdependencies between factors, and the possibility to combine various sources of information (i.e., theoretical data, empirical data, and expert judgment) have made BNs an appropriate tool for accident and risk analysis (Weber et al. 2012, Mkrtchyan et al. 2015, Qiao et al. 2020). They have been deployed as a standalone approach and/or a model-to-model transformation approach (Al-Shanini et al. 2014, Kabir & Papadopoulos 2019). Constructing and using BNs comprise three steps of problem structuring (i.e., identifying variables and network structure and expressing as statistical variables), instantiation (i.e.,

specifying conditional probabilities), and inference (i.e., entering variables, propagating, and interpreting results) (Sigurdsson et al. 2001).

2.2 Petri Nets

Petri Nets (PNs) are mathematical and graphical tools that are appropriate for modeling and analyzing dynamic, distributed, parallel, and concurrent systems with time constraints (Vernez et al. 2003, Wu et al. 2015, Kabir & Papadopoulos 2019). They are described by a set of places, a set of transitions, a valuation function, and an initial marking. In a PN graph as a directed bipartite, a circle represents a place and a thin rectangle stands for a transition, and arrows and tokens respectively illustrate valuation functions and marking.

The required steps to create and analyze a PN model are problem structuring (i.e., identifying places and transitions based on system behavior and draw its PN model), instantiation (i.e., forming the initial marking by putting specified numbers of tokens in the specified places and specifying firing rates for the timed transitions), and analysis (i.e., executing/simulating the model using PN simulator and interpreting results) (Kabir & Papadopoulos 2019).

2.3 Machine Learning and Text Mining

Machine learning (ML) as a subset of artificial intelligence (AI) involves algorithms to create and adapt models which can automatically be improved through experience and by the use of data. The ML is a subset of artificial intelligence while deep learning and text mining are subsets of ML (Hegde & Rokseth 2020). There are three common types of ML algorithms: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning contains classification and regression algorithms. Unsupervised learning mainly deals with unlabeled data and involves clustering and association techniques to discover similarities and differences. And reinforcement learning algorithms learn via feedback from their own actions and experiences and are good for developing an appropriate action model.

ML algorithms can analyze the various format of the input data, from numerical data to textual one, which can be historical, real-time, or a combination of both. Moreover, data mining and text mining are respectively referred to as applying ML itself or in association with other methods such as statistics and natural language processing (NLP) to analyze numerical data and textual data. Text mining contains everything from information retrieval to text classification and clustering, to entity, relation, and event extraction through exploring text corpora.

3 BIBLIOMETRIC SEARCH METHODOLOGY

We searched six databases (ScienceDirect, Scopus, Web of Science, SpringerLink, Google scholar, and IEEE Xplore) for published scholarly research in English,

up to the date 1 August 2021. Our keyword search terms were: “accident modeling”, “accident analysis”, “accident causation”, “accident causal”, “accident models”, “accident prediction”, “safety”, “safety risk management” and “safety risk analysis”. Furthermore, our methods-related keywords are “Bayesian networks”, “Petri nets”, “Network theory”, “Artificial intelligence”, “Machine learning”, “Data mining”, “Text mining”, and “Deep learning”, which we used in conjunction with rail-related keywords “railroad” and “railway” to find published studies in the research subject. Then, we included those studies that have used these methods to analyze at least one railway occurrence. Thus, publications with the aims of safety risk assessment as well as accident prediction were included only if they investigated the railway accidents with such methods. This yielded 54 articles.

3 DESCRIPTIVE ANALYSIS

This section represents the statistical analysis of the 54 original research articles that have used these methods to analyze railway accidents and safety risks.

The distribution of published papers in different years is illustrated in Figure 1. In 2004, the first application of modern accident modeling was observed by the work of Marsh & Bearfield (2004). Years 2018 and 2019 with respectively 13 and 10 publications are the most significant periods for the deployment of this emerging field in rail transportation. The share of the two groups of methods is approximately equal, with 44% for the network-based methods and 56% for the AI-based techniques.

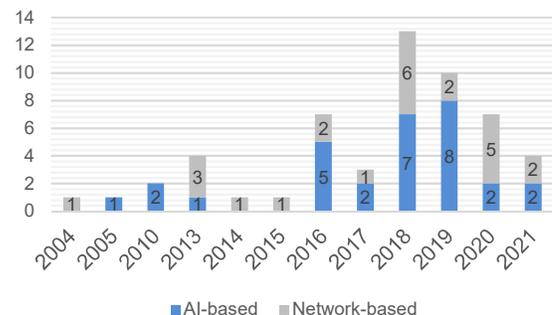


Figure 1. Yearly distribution of the reviewed articles based on the applied methods

As shown in Figure 2, the majority of the reviewed studies are journal articles (65%) with the remainder being conference papers.

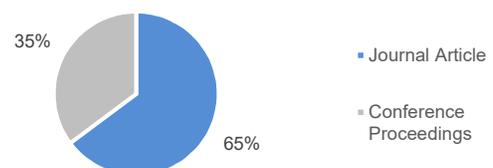


Figure 2. Distribution of the reviewed articles by type of publication

Table 1 reflects the distribution of articles published in journals. As can be seen, the Journal of Safety Science and the Journal of Reliability Engineering & System Safety are the two leading outlets, with a combined share of 37% of published papers.

Table 1. Distribution of the reviewed articles amongst the journal sources

Name of journal	No. of articles	% of articles
Safety science	7	20%
Reliability Engineering & System Safety	6	17%
Accident Analysis & Prevention	3	9%
AIMS Electronics and Electrical Engineering	2	6%
Other journals	17	48%

4 RESULTS AND DISCUSSION

The new accident and risk analysis frameworks can be classified into network-based and artificial intelligence (AI)-based methods, which can have further categorized by the type of algorithms and techniques, data sources, and tools applied.

4.1 Network-based Accident Analysis Methods

4.1.1 Bayesian Networks in Railway Accident and Safety Analysis

Bayesian networks (BNs) may be used alone or in combination with other analytical methods to analyze railway accidents and the associated safety risks. Marsh & Bearfield (2004) used BNs to develop a casual model of signals passed at danger (SPAD) incidents in the UK by incorporating organizational factors. They later applied the same approach to model train boarding and alighting incidents and argued the advantages of BNs over fault and event trees for modeling events (Bearfield & Marsh 2010).

In a series of research, Liang and their colleagues developed BN-based frameworks for railway grade crossing occurrences to identify the important causal factors, their relationships, and their quantitative influence (Liang et al. 2017, Liang & Ghazel 2018, Liang et al. 2018, 2020). In 2017, they produced a causal statistic risk assessment based on hierarchical causal Bayesian networks (called CSRA-CBN) approach to explore the key contributing factors to accidents/incidents at grade crossings as well as their combined impacts on safety. The CBN model was created according to the statistical analysis of the accident and incident databases (Liang et al. 2017). Later, Liang et al. (2018) combined grade crossing accident/incident databases with expert knowledge data to construct a causal network structure and used forward and reverse inferences and Euclidean distance to identify the strength of factor interactions and the

most important ones. Liang & Ghazel (2018) statistically analyzed grade crossing accident/incident database with respect to traffic moment (i.e., the combined traffic at the grade crossing, which is calculated by multiplying road traffic frequency by railway traffic frequency), different kinds of transport mode and different geographical regions and identified causalities between these factors and accident occurrence. Then, they established the BN risk model based on the outcomes of statistical analysis and predicted accident probability, and assessed risk level. To construct a BN structure for accident and safety risk analysis of railway level crossing, Liang et al. (2020) discovered preliminary causality through automatic structure learning including the Bayesian search (BS), naïve Bayes, and greedy thick thinning (GTT) algorithms, and optimized it by causality constraints derived from expert knowledge

Dindar and colleagues also conducted a series of BN studies on derailments. In the first study, Dindar et al. (2018) provided a Fuzzy Bayesian network (FBN) for weather-related derailments at railway switch systems. A causal relationship was built through analyzing 18,000 derailment reports across the U.S., which was quantified using Buckley's probability calculation and confidence intervals. In the second study, Dindar et al. (2019) adopted a new stochastic mathematical modeling technique based on a hierarchical Bayesian model (HBM) to investigate component failure-related derailments at railway switches. They integrated multiple specialized packages, such as MATLAB for image processing, R for statistical analysis, and ArcGIS for displaying and manipulating geospatial data, to better model and display complex solutions. Finally, in 2020, human factors related to derailments on switches and crossings were examined by Dindar et al. (2020) using a Fuzzy Bayesian network (FBN). They combined accident data and expert knowledge data to develop the BN model and employed fuzzy set theory to quantify the linguistic information and compute the human error likelihood of derailment.

A transformation of an event tree (ET) model to a BN model can be seen in Ye & Zheng (2016a). They first developed an event tree (ET) model for the failure of the component of the automatic train protection (ATP) system regarding the human cognitive process in the generic error modeling system (GEMS) (i.e., they considered perception, scenario analysis, decision making, and action-taking as a human cognitive process). They then transferred the ET model into hierarchical BN (HBN) considering rail-performance shaping factors (R-PSFs) and finally established a human-machine bow-tie model. Expert judgments which were aggregated by Dempster-Shafer (D-S) evidence theory were used for selecting appropriate R-PSFs for each cognitive phase, constructing causal relationships, and performing quantification. In another ET to BN model transformation for assessing human risks associated with using ATP, Ye & Zheng (2016b) utilized the fuzzy inference theory to improve conditional probability tables (CPT) building method for BN.

Zhang et al. (2018) defined and classified factors that contributed to high-speed railway accidents in China using the human factors analysis and classification system-railway accidents (HFACS-RAs) method. After that, they built a BN structure according to the HFACS-RAs classification and the results of interaction analysis by the Chi-square test and Odds ratios (OR). Finally, the D-S/AHP evidence fusion method relying on expert knowledge was adopted to infer the conditional probability tables (CPTs) in the BN. A hybrid approach of interpretive structural modeling (ISM) and BN was applied by Huang et al. (2020) to analyze the relationships and interaction strengths among the risk factors and accident causes of railway dangerous goods transportation system (RDGTS). The safety performance of five railways (i.e., Chinese, Japanese, Spanish, French, and South Korean railways) was compared through analyzing accident data sets and developing risk assessment models using Bayesian inference, decision tree, and Petri net techniques (Rungskunroch et al. 2021).

Table 2 provides a summary of the sources of data and software packages that were used to analyze and construct the BN models.

Table 2. Data sources for constructing the Bayesian network models

Author	Country	Software	Data sources	
			Expert judgments data	Occurrence data
Marsh & Bearfield (2004)	UK	-	♦	
Bearfield & Marsh (2010)	UK	-	♦	♦
Liang et al. (2017)	France	GeNIe		♦
Liang et al. (2018)	France	GeNIe	♦	♦
Liang & Ghazel (2018)	France	GeNIe		♦
Liang et al. (2020)	France	GeNIe	♦	♦
Dindar et al. (2018)	UK	MATLAB		♦
Dindar et al. (2019)	UK	MATLAB, R, ArcGIS		♦
Dindar et al. (2020)	UK	-	♦	♦
Ye & Zheng (2016a)	China	GeNIe	♦	
Ye & Zheng (2016b)	China	GeNIe	♦	
Zhang et al. (2018)	China	GeNIe	♦	♦
Huang et al. (2020)	China	GeNIe	♦	♦
Rungskunroch et al. (2021)	UK	Python	♦	♦

As can be seen in Table 2, occurrence data analysis is usually used in combination with expert judgments data to solve the data scarcity problem. In rich data situations such as level crossing incidents, the BN model development relies only on occurrence data. Furthermore, GeNIe is the widely used tool in the BN-based accident and safety analysis.

4.1.2 Petri Nets in Railway Accident and Safety Analysis

Similar to BNs, Petri nets (PNs) have been applied as standalone or as a part of model-to-model transformation approaches to assess railway accidents and risks.

Wu et al. (2015) deployed stochastic Petri nets (SPNs) to present a model of train rear-end collision accidents. The quantitative analysis and uncertainty modeling were respectively undertaken by the isomorphic Markov chain model and Fuzzy random method. Dirk et al. (2013)'s study is one of the proposed model-to-model transformation frameworks for PNs. They proposed formalSTAMP by integrating PNs with the systems theoretic accident model and processes (STAMP) method and employed it to scrutinize the Wenzhou 7.23 accident as the most serious rail accident in China. In 2018, Song & Schnieder (2018) extracted the fault tree (FT) of train head to tail collisions and then mapped it into colored Petri nets (CPNs) to address limitations of the FT method including modeling time-related attributions and non-linear relationships. The accuracy of the framework was verified by using Monte Carlo simulation and state-space analysis. Recently, Zhang et al. (2020) developed a Fuzzy Petri net-fault tree analysis (FPN-FTA) model for the stampede accident of Shijiazhuang high-speed railway station in China and simulated the FTA-FPN model with Stateflow of Matlab software. The accident was first represented in FTA and then converted to FPN through integrating dynamic weighting FPN and FTA. Finally, the optimal risk controls were determined after building a bi-objective risk control model and optimizing with the particle swarm optimization (PSO) algorithm.

4.1.2 Other Network-based Methods

Xin et al. (2013) believed that BNs and PNs methods suffer from having a local view in accident analysis as they focus on point-to-point or part-to-part analysis, which is not enough for complicated railway accidents. Therefore, they employed the complex network theory (CNT) to identify the causation of the Wenzhou 7.23 accident that occurred in China. They found that the inspection of signals and the checking of line conditions before trains run were the main reasons for this accident. The Wenzhou 7.23 accident was also investigated by the complex network theory integrated with the cascading failure theory by (Luo et al. 2014). They concluded that the equipment's failure was the root cause of the accident while the control flaws of the train operation system in preventing or hindering the propagation of cascading failure played an important

role. In another study, Li & Wang (2018) first identified the causes of railway accidents as well as their relationships by analyzing the Federal Railroad Administration (FRA) databases, and then, built the cause-effect network using complex network theory to depict how railway accidents occur.

Aguirre et al. (2013) presented a combined approach based on evidential networks (ENs) and fault tree analysis (FTA) to integrate the human, organizational and technical factors to risk analysis in railway accidents. They used the belief functions theory, also known as the Dempster-Shafer (D-F) or evidence theory, for quantification. Liu et al. (2019) provided a network theory-based accident model merged with topological analysis for understanding rail accidents. They delineated latent patterns of hazards and proposed a practical way to generate an accident causation network from accident reports. This model was later extended in Liu et al. (2021)'s study to better adapt to the heterogeneous characteristics produced by various causes of hazards and accident contributory factors. Recently, Lam & Tai (2020) presented a network analytical approach to clarify incident factors and how they affect each other in railway incident chains. The model was used to survey railway events in Japan from local view analysis, global view analysis, and contextual view analysis perspectives.

4.2 AI-BASED ACCIDENT ANALYSIS METHODS

This literature review illustrates that the applications of machine learning (ML) methods in accident and risk analysis are on the rise. The application of association rule mining to discover patterns among accident data can be seen in Mirabadi & Sharifian (2010)'s study. They utilized generalized rule induction (GRI) algorithm to recognize relationships between the accidents' causes by discovering repetitive patterns within the past accident data of the Iranian Railway (RAI). Ghomi et al. (2016) employed the ordered probit model (OPM), Apriori, and classification and regression tree (CART) algorithms to extract the factors affecting the severity of highway-railway grade crossings accidents. The results of the three algorithms showed that train speed has the highest impact on injury severity. 392 Chinese railway accident reports were collected and processed by Yu et al. (2018) and factors involved in accidents were identified and classified based on the cognitive reliability and error analysis method-railway accidents (CREAM-RAs) taxonomy framework. They called the categorized accident factors multi-attribute railway accidents dataset (MARA-D) which was later clustered and visualized adopting the self-organizing maps (SOM) algorithm. Alawad et al. (2019) employed machine learning, particularly the decision tree (DT) method to analyze accidents occurring at railway stations in the U.K. The significant causes, their interactions, and the traits of passengers influenced by accidents were extracted in order to improve safety at the railway stations. The application of 11 different types of ML algorithms to uncover patterns in the equipment accident database of the FRA and to predict

derailments can be seen in Bridgelall & Tolliver (2021). The extreme gradient boosting (XGB) classifier showed the best prediction performance at predicting derailment accidents among other algorithms.

In another research stream, researchers analyzed semi-structured and/or unstructured textual descriptions of railway incidents and accidents, along with structured data. Williams et al. (2015)'s work focused on applying a topic modeling algorithm, called latent Dirichlet analysis (LDA), to uncover the themes of railway grade crossing accidents embedded in the text body of the FRA investigation reports. They find that additional training of conductors to make the leading freight car more conspicuous can reduce the accidents. In another study, Williams et al. (2016) adopted the LDA topic modeling and k-means clustering algorithms to analyze serious rail accidents in the U.S. and Canada to find key differences. Accidents involving bridges, for example, were more prominent in Canada, while it was not seen in the obtained clusters for the U.S. rail accident reports. Another distinction was the prevalence of accidents containing runaway cars in the Canadian railways.

Williams & Betak (2016) explored the FRA equipment accident reports from January 2010 to February 2015 using LDA as well as text clustering techniques and visually represented the text clusters. Both techniques concluded the main topics in the accident reports are grade crossings and trucks, shoving, and hump yards. They also discovered that major accidents themes are those related to lining switches and accidents involving the actions of railroad personnel. Brown (2016) investigated a role that text mining can play in a better understanding of accident characteristics and accident factors. They examined over 11 years of railway accident reports with and without incorporating text analytics. They concluded that the accuracy of prediction for accident severity can be improved by incorporating factors found by text mining and modern ensemble methods (i.e., random forests and gradient boosting).

Syeda et al. (2017) studied the rail accident investigation branch (RAIB) reports by exerting natural language processing (NLP) techniques. They first defined the entities of interests (EOIs) according to the traditional accident analysis approaches, e.g., human factors and organizational factors, and then determined the frequency, sequence, and co-occurrence of words and the EOIs to help accident investigators for surveying causal relationships.

Two textual analysis techniques, i.e., latent semantic analysis (LSA) and latent Dirichlet allocation (LDA) were used and compared in the study by Williams & Betak (2018). These methods uncovered the most frequent rail accidents (e.g., switching accidents and grade crossing accidents) but also uncovered the less frequent (e.g., accidents involving ballast maintenance equipment). Moreover, Williams & Betak (2018) showed that applying two methods for mining texts can identify more accident topics compared to applying only one text mining technique. Karthi & Priscilla (2018) offered using the ID3 algorithm

to classify the semi-structured part of accident reports and to extract the causes of major rail accidents. Li et al. (2018) applied full-text retrieval and text classification techniques to analyze the accident and fault reports of the Taiyuan railway bureau. They first used the TF-IDF algorithm to identify the most important user input keywords in the given documents. They later made classification applying a special type of artificial recurrent neural network (RNN) algorithm called long short-term memory (LSTM). Kamekar et al. (2018) proposed the utilization of the ID3 algorithm, the naïve Bayesian (NB) classifier, and the agglomerative hierarchical clustering (AHC) method to outline the factors that influenced the accidents of the Indian Railway industry.

Heidarysafa et al. (2018) examined if text mining is helpful to extract accident causes from accident narratives that include terminologies that are not easy to understand by non-expert readers. Moreover, to know whether the reported causes in the structured format are consistent with those explained in the narratives or not. To answer these questions, they adopted three main deep learning approaches, i.e., convolutional neural nets (CNN), recurrent neural nets (RNN), and deep neural nets (DNN), along with word embeddings such as Word2Vec (i.e., word to Vector) and GloVe (i.e., Global Vectors for Word Representation). The results indicated that applying deep learning techniques for exploring railway accident descriptions can accurately classify the causes of an accident and detect important inconsistencies in accident reporting.

In Hua et al. (2019)'s paper, 283 Chinese railway accident reports were classified into accident description and causal analysis classes using the multichannel convolutional neural network (M-CNN) model. After that, the accident factors were derived from the identified causal analysis sentences by using the conditional random field (CRF) model and summarized into the main categories of human factors, mechanical equipment factors, operating environment factors, and management factors. Soleimani et al. (2019a) investigated 48,080 highway-railway crossing incidents in the U.S. to discover the reasons for the incidents from the textual descriptions. The critical reasons for the incidents in every state were identified utilizing the TF and TF-IDF techniques. Furthermore, the incident similarities between all the states were assessed with the pairwise correlation calculation. Finally, machine learning methods (i.e., random forest and logistic regression) were employed to classify incidents into "car struck train" or "train struck car" categories, using both the fixed fields of the FRA reports and the narrative ones. The defined categories help understand whether the incidents were more related to the driver's behavior or warning devices. The model was later developed by incorporating decision tree (DT), random forest (RF), XGboost (XGB), and logistic regression (LR) machine learning algorithms as well as geospatial analysis (Soleimani et al. 2019b, Soleimani et al. 2021).

Table 3 summarizes the AI models and algorithms

that were used to analyze railway safety databases. As can be ascertained, the applied methods are categorized into machine learning (ML) and text mining (TM). The latter involves those articles that utilized TM and/or ML algorithms to explore railway accident and incident narratives while the former only focuses on analyzing the structured occurrence data. Moreover, text mining studies usually contain two steps: first natural language processing (NLP) are applied to transform unstructured text into structured data and then machine learning (ML) algorithms are used to extract further information.

Table 3. Summary of the applied machine learning (ML) and text mining (TM) models and algorithms

Author	Method	Model	Algorithms	Data source
Mirabadi and Sharifian (2010)	ML	Association rule mining	GRI	RAI accident database
Ghomi et al. (2016)	ML	Association rule mining	Apriori, OPM	FRA database
		Classification	CART	
Yu et al., (2018)	ML	Clustering	SOM	Chinese Railway accident reports
Alawad et al. (2019)	ML	Classification	DT	RAIB database
Soleimani et al. (2019b)	ML	Classification	DT, RF, XGB, LR	Rail Inventory Management System (RIMS) database
Bridgelall and Tolliver (2021)	ML	Classification	DT, RF, AB, XGB, GB, K-NN, NB, SVM, ANN, SGD	FRA database
Williams et al. (2015)	TM	Topic modeling	LDA	FRA database
Williams et al. (2016)	TM	Topic modeling	LDA	NTSB and TSB reports
		Clustering	K-means	
Williams and Betak (2016)	TM	Topic modeling	LDA	FRA database
		Clustering	K-means	
Brown (2016)	TM	Topic modeling	LDA	NTSB reports
		Classification	OLS, PLS, RF, GB	
Syeda et al. (2017)	TM	Topic modeling	LDA	RAIB database
		Clustering	Ward's method	

Author	Method	Model	Algorithms	Data source
Williams and Betak (2018)	TM	Topic modeling	LSA, LDA	FRA database
Li et al. (2018)	TM	Word importance	TF-IDF	Taiyuan Railway Bureau (TRB) accident reports
		Classification	LSTM	
Kamerkar et al. (2018)	TM	Classification	ID3, NB	Indian Railway accident database
		Clustering	AHC	
Karthi and Priscilla (2018)	TM	Classification	ID3	FRA database
Heidarysafa et al.(2018)	TM	Word importance	TF-IDF	FRA database
		Classification	CNN, RNN, DNN,	
Hua et al. (2019)	TM	Classification	NB, SVM, RF, M-CNN	Chinese Railway accident reports
		Pattern recognition	HMM, CRF	
Soleimani et al. (2019a)	TM	Word importance	TF, TF-IDF	FRA database
		Classification	RF, LR	
Soleimani et al. (2021)	TM	Word importance	TF, TF-IDF	FRA database
		Classification	XGB	

*Abbreviations: GRI (Generalized Rule Induction), OPM (Ordered Probit Model), CART (Classification and Regression Tree), SOM (Self-Organizing Maps), DT(Decision Tree), RF (Random Forest), XGB (Extreme Gradient Boost), LR (Logistic Regression), AB (Ada Boost), GB (Gradient Boost), k-NN (k-Nearest Neighbors), NB (Naïve Bayes), SVM (Support Vector Machine), ANN (Artificial Neural Network), SGD (Stochastic Gradient Descent), LDA (Latent Dirichlet Analysis), OLS (Ordinary Least Squares), PLS (Partial Least Squares), LSA (Latent Semantic Analysis), TF-IDF (Term Frequency-Inverse Document Frequency), LSTM (Long Short-Term Memory), ID3 (Iterative Dichotomiser 3), AHC (Agglomerative hierarchical clustering), CNN (Convolutional Neural Nets), RNN (Recurrent Neural Nets), DNN (Deep Neural Nets), M-CNN (Multichannel Convolutional Neural Network), HMM (Hidden Markov model), CRF (Conditional Random Field), TF (Term Frequency)

Table 3 also reveals that classification, clustering, and topic modeling were the common techniques to investigate railway accident databases. Furthermore, accident data associated with the US railways was investigated more than other railways using these new methods.

Overall, the AI-based accident analysis methods have some advantages over the classical methods. First, they can analyze high-volume data in a short time. Second, they provide an opportunity to automatically explore the narratives of railway accident reports that usually offer a much richer amount of information regarding accident characteristics and the potential reasons behind them. Third, they help understand similarities and differences between accidents, derive hidden relationships among factors and accidents, discover implicit information, and predict accidents, which are not easy to process manually. Finally, they can leverage rail infrastructure monitoring data, train

operations data, human performance data, etc., for accident and safety analysis and acquire a deep understanding of their relationships.

5 CONCLUSIONS

This paper summarizes new approaches for analyzing railway accidents and safety risks, and categorizes them into network-based and artificial intelligence (AI)-based methods. Through this review, we have found that advances in computer technology have produced a paradigm shift in accident modeling and made Bayesian networks, Petri nets, machine learning, and text mining emerging fields of accident analysis studies. These newer methods have been employed standalone or in combination with the classical accident and risk analysis methods such as fault tree analysis (FTA) and HFACS.

6 REFERENCES

- Aguirre, F., Sallak, M., Schön, W. & Belmonte, F. (2013). Application of evidential networks in quantitative analysis of railway accidents. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 227(4), 368-384.
- Al-Shanini, A., Ahmad, A. & Khan, F. (2014). Accident modelling and analysis in process industries. Journal of Loss Prevention in the Process Industries, 32, 319-334.
- Alawad, H., Kaewunruen, S. & An, M. (2019). Learning from accidents: machine learning for safety at railway stations. IEEE Access, 8, 633-648.
- Bearfield, G. & Marsh, W. (2010). Causal modelling of lower consequence rail safety incidents. European Safety and Reliability Conference (ESREL). London, UK.
- Bridgelall, R. & Tolliver, D. D. (2021). Railroad accident analysis using extreme gradient boosting. Accident Analysis & Prevention, 156, 106126.
- Brown, D. E. (2016). Text mining the contributors to rail accidents. IEEE Transactions on Intelligent Transportation Systems, 17(2), 346-355.
- Dindar, S., Kaewunruen, S., An, M. & Sussman, J. M. (2018). Bayesian Network-based probability analysis of train derailments caused by various extreme weather patterns on railway turnouts. Safety science, 110, 20-30.
- Dindar, S., Kaewunruen, S. & An, M. (2019). Rail accident analysis using large-scale investigations of train derailments on switches and crossings: Comparing the performances of a novel stochastic mathematical prediction and various assumptions. Engineering failure analysis, 103, 203-216.
- Dindar, S., Kaewunruen, S. & An, M. (2020). Bayesian Network-based Human Error Reliability Assessment of Derailments. Reliability Engineering & System Safety, 197, 106825.
- Dirk, S., Sebastian, H. R., Welte, J. & Schnieder, E. (2013). Integration of Petri nets into STAMP/CAST on

- the example of Wenzhou 7.23 accident. *IFAC Proceedings Volumes*, 46(25), 65-70.
- George, P. G. & Renjith, V. (2021). Evolution of Safety and Security Risk Assessment methodologies to use of Bayesian Networks in Process Industries. *Process Safety and Environmental Protection*, 149, 758–775.
- Ghofrani, F., He, Q., Goverde, R. M. & Liu, X. (2018). Recent applications of big data analytics in railway transportation systems: A survey. *Transportation Research Part C: Emerging Technologies*, 90, 226-246.
- Ghomi, H., Bagheri, M., Fu, L. & Miranda-Moreno, L. F. (2016). Analyzing injury severity factors at highway railway grade crossing accidents involving vulnerable road users: A comparative study. *Traffic injury prevention*, 17(8), 833-841.
- Gutierrez-Osorio, C. & Pedraza, C. (2020). Modern data sources and techniques for analysis and forecast of road accidents: a review. *Journal of Traffic and Transportation Engineering (English Edition)*, 7(4), 432-446.
- Halim, Z., Kalsoom, R., Bashir, S. & Abbas, G. (2016). Artificial intelligence techniques for driving safety and vehicle crash prediction. *Artificial Intelligence Review*, 46(3), 351-387.
- Hegde, J. & Rokseth, B. (2020). Applications of machine learning methods for engineering risk assessment—A review. *Safety science*, 122, 104492.
- Heidarysafa, M., Kowsari, K., Barnes, L. & Brown, D. (2018). Analysis of railway accidents' narratives using deep learning. In 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, USA, IEEE.
- Heinrich, H. W. (1931). *Industrial Accident Prevention. A Scientific Approach*. New York, McGraw-Hill.
- Hua, L., Zheng, W. & Gao, S. (2019). Extraction and Analysis of Risk Factors from Chinese Railway Accident Reports. In *IEEE Intelligent Transportation Systems Conference (ITSC)*, Auckland, New Zealand, IEEE.
- Huang, L., Wu, C., Wang, B. & Ouyang, Q. (2018). A new paradigm for accident investigation and analysis in the era of big data. *Process safety progress*, 37(1), 42-48.
- Huang, W., Zhang, Y., Kou, X., Yin, D., Mi, R. & Li, L. (2020). Railway dangerous goods transportation system risk analysis: An Interpretive Structural Modeling and Bayesian Network combining approach. *Reliability Engineering & System Safety*, 204, 107220.
- Ismail, S. N., Ramli, A. & Aziz, H. A. (2021). Influencing factors on safety culture in mining industry: A systematic literature review approach. *Resources Policy*, 74, 102250.
- Kabir, S. & Papadopoulos, Y. (2019). Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: A review. *Safety science*, 115, 154-175.
- Kamerkar, N., Patil, K., Kale, A. & Patil, G.-M. N. (2018). Text Mining Applied to Rail Accidents. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 4(3), 383-386.
- Karthi, M. & Priscilla, R. (2018). The patrons for anticipating veracity of rail mishaps using text mining and ID3 algorithm. *International Journal of Pure and Applied Mathematics*, 119(15), 1753-1759.
- Klockner, K. (2015). *Human factors in rail regulation: Modelling a theory of non-linear rail accident and incident networks using the Contributing Factors Framework*. Ph.D. Thesis, Central Queensland University.
- Lam, C. Y. & Tai, K. (2020). Network topological approach to modeling accident causations and characteristics: Analysis of railway incidents in Japan. *Reliability Engineering & System Safety*, 193, 106626.
- Li, K. & Wang, S. (2018). A network accident causation model for monitoring railway safety. *Safety science*, 109, 398-402.
- Li, X., Shi, T., Li, P., Gao, F. & Xiang, W. (2018). Application of Text Mining Techniques in Railway Safety Supervision System. In *IOP Conference Series: Earth and Environmental Science*, IOP Publishing.
- Liang, C., Ghazel, M., Cazier, O. & El-Koursi, E.-M. (2017). Risk analysis on level crossings using a causal Bayesian network based approach. *Transportation research procedia*, 25, 2167-2181.
- Liang, C. & Ghazel, M. (2018). A risk assessment study on accidents at French level crossings using Bayesian belief networks. *International journal of injury control and safety promotion*, 25(2), 162-172.
- Liang, C., Ghazel, M. & Cazier, O. (2018). Using Bayesian Networks for the Purpose of Risk Analysis at Railway Level Crossings. *IFAC-PapersOnLine*, 51(9), 142-149.
- Liang, C., Ghazel, M., Cazier, O. & Bouillaut, L. (2020). Advanced model-based risk reasoning on automatic railway level crossings. *Safety science*, 124, 104592.
- Liu, J., Schmid, F., Zheng, W. & Zhu, J. (2019). Understanding railway operational accidents using network theory. *Reliability Engineering & System Safety*, 189, 218-231.
- Liu, J., Schmid, F., Li, K. & Zheng, W. (2021). A knowledge graph-based approach for exploring railway operational accidents. *Reliability Engineering & System Safety*, 207, 107352.
- Luo, Z., Li, K., Ma, X. & Zhou, J. (2014). Analyzing railway accidents based on complex network and cascading failure. In *Proceedings of the International Conference on Electrical and Information Technologies for Rail Transportation (EITRT2013)*, Springer.
- Marcot, B. G. & Penman, T. D. (2019). Advances in Bayesian network modelling: Integration of modelling technologies. *Environmental modelling & software*, 111, 386-393.
- Marsh, W. & Bearfield, G. (2004). Using Bayesian networks to model accident causation in the UK railway

- industry. In Probabilistic safety assessment and management, Springer.
- Mirabadi, A. & Sharifian, S. (2010). Application of association rules in Iranian Railways (RAI) accident data analysis. *Safety Science*, 48(10), 1427-1435.
- Mkrtychyan, L., Podofilini, L. & Dang, V. N. (2015). Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability engineering & system safety*, 139, 1-16.
- Qiao, W., Liu, Y., Ma, X. & Liu, Y. (2020). Human Factors Analysis for Maritime Accidents Based on a Dynamic Fuzzy Bayesian Network. *Risk Analysis*, 40(5), 957-980.
- Rungskunroch, P., Jack, A. & Kaewunruen, S. (2021). Benchmarking on railway safety performance using Bayesian inference, decision tree and petri-net techniques based on long-term accidental data sets. *Reliability Engineering & System Safety*, 213, 107684.
- Sigurdsson, J., Walls, L. & Quigley, J. (2001). Bayesian belief nets for managing expert judgement and modelling reliability. *Quality and Reliability Engineering International*, 17(3), 181-190.
- Soleimani, S., Mohammadi, A., Chen, J. & Leitner, M. (2019a). Mining the Highway-Rail Grade Crossing Crash Data: A Text Mining Approach. In 18th IEEE International Conference on Machine Learning and Applications (ICMLA), Boca Raton, USA, IEEE.
- Soleimani, S., Mousa, S. R., Codjoe, J. & Leitner, M. (2019b). A comprehensive railroad-highway grade crossing consolidation model: a machine learning approach. *Accident Analysis & Prevention*, 128, 65-77.
- Soleimani, S., Leitner, M. & Codjoe, J. (2021). Applying machine learning, text mining, and spatial analysis techniques to develop a highway-railroad grade crossing consolidation model. *Accident Analysis & Prevention*, 152, 105985.
- Song, H. & Schnieder, E. (2018). Evaluating Fault Tree by means of Colored Petri nets to analyze the railway system dependability. *Safety Science*, 110, 313-323.
- Syeda, K. N., Shirazi, S. N., Naqvi, S. A. A. & Parkinson, H. J. (2017). Exploiting Natural Language Processing for Analysing Railway Incident Reports. IGI Global.
- Underwood, P. & Waterson, P. (2013). Accident analysis models and methods: guidance for safety professionals. Loughborough: Loughborough University, 28.
- Vernez, D., Buchs, D. & Pierrehumbert, G. (2003). Perspectives in the use of coloured Petri nets for risk analysis and accident modelling. *Safety science*, 41(5), 445-463.
- Weber, P., Medina-Oliva, G., Simon, C. & lung, B. (2012). Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. *Engineering Applications of Artificial Intelligence*, 25(4), 671-682.
- Williams, T., Betak, J. & Findley, B. (2016). Text mining analysis of railroad accident investigation reports. In Joint Rail Conference, Columbia, USA, ASME.
- Williams, T. & Betak, J. (2018). A comparison of LSA and LDA for the analysis of railroad accident text. *Procedia computer science*, 130, 98-102.
- Williams, T. P. & Betak, J. F. (2016). Identifying Themes in Railroad Equipment Accidents Using Text Mining and Text Visualization. In International Conference on Transportation and Development Houston, USA.
- Wu, C., Cao, C., Sun, Y. & Li, K. (2015). Modeling and analysis of train rear-end collision accidents based on stochastic petri nets. *Mathematical Problems in Engineering*, 2015.
- Xin, M., Ke-Ping, L., Zi-Yan, L. & Jin, Z. (2013). Analyzing the causation of a railway accident based on a complex network. *Chinese Physics B*, 23(2), 028904.
- Yan, H., Yang, N., Peng, Y. & Ren, Y. (2020). Data mining in the construction industry: Present status, opportunities, and future trends. *Automation in Construction*, 119, 103331.
- Ye, H. & Zheng, W. (2016a). A human reliability analysis method based on cognitive process model for risk assessment. In IEEE International Conference on Intelligent Rail Transportation (ICIRT), Birmingham, UK, IEEE.
- Ye, H. & Zheng, W. (2016b). An improved method for building conditional probability tables of Bayesian networks in human reliability analysis and its application for railway. In IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Xi'an, China, IEEE.
- Yu, G., Zheng, W., Wang, L. & Zhang, Z. (2018). Identification of significant factors contributing to multi-attribute railway accidents dataset (mara-d) using som data mining. In 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, USA, IEEE.
- Zhang, G., Zheng, W. & Wang, L. (2018). A human and organizational analysis method for chinese high-speed railway accidents/incidents based on bayesian network. In 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, USA, IEEE.
- Zhang, Q., Zhuang, Y., Wei, Y., Jiang, H. & Yang, H. (2020). Railway safety risk assessment and control optimization method based on FTA-FPN: a case study of Chinese high-speed railway station. *Journal of advanced transportation*, 2020.

With great power comes great responsibility: Organisational factors influencing power disparities between network controllers and other operational groups

Bride Luva and Anjum Naweed

Central Queensland University, Appleton Institute for Behavioural Science, Adelaide, Australia

ABSTRACT

The rail system relies on the effective coordination of multidisciplinary functional groups to meet a single operational objective — the safe and timely movement of rail traffic. Problems in the way these groups interact manifest in system error, including trains exceeding movement authorities, traffic encroaching into active worksites, and workers entering the rail corridor without adequate protection. Understanding how power imbalances within these groups impact how they communicate is under-researched and provides a new dimension for establishing the contributing factors in accidents and incidents. The aim of this study was to identify what organisational factors shape any power disparities between network controllers and other operational groups. Data were collected from network controllers ($N = 55$) across eight rail operating organisations in Australia and New Zealand using scenario-based interviewing. Preliminary findings derived through thematic analysis revealed that power disparities were perceived to be influenced by various accountability mechanisms associated with error-investigation and an apportioning of blame. This included finding fault, a fear of being at fault, the threat of performance management (upon error discovery), deflection of blame, and unjust treatment. Power was also perceived to be imbued in the role through responsibility and operational imperatives, which created a sense of empowerment and a notion of privileged knowledge. Findings point to the clear existence of “authority gradients” between the network controller and other operational groups and illustrate how such power disparities may result in breakdown of safeworking and implications for training.

1 INTRODUCTION

As a multidisciplinary team, track maintenance workers, traincrew and network controllers share a common goal, ensuring the safe and timely movement of rail traffic (Dobson et al. 2014). While remotely located, each team operate as specialists in their respective domains, but hierarchical structures across multidisciplinary teams may result in power imbalances that impact safety (Gordon & O'Connor 2012). Rail organisations necessarily operate using hierarchical structures (Cheng 2019), however, organisations that are designed this way are known to be less effective at collaborative communication (Speroff et al. 2010). It is therefore important to understand how the organisational influence of power imbalances within these groups may impact on how they communicate.

While the term “network controller” varies across countries and organisations, (e.g. train controller, train dispatcher, rail traffic controller), it is used here to identify workers responsible for directing rail traffic movement and authorising activities on the network. The role can also be distributed across other discrete functions (e.g. signaller, area controller). The network controller controls the signalling system that indicates the paths train crew must navigate. The network controller also negotiates and implements worksite protections for work groups undertaking maintenance work (Cheng & Tsai 2011).

Failed communication between these operational groups manifest in various ways and include: trains exceeding their authority and passing a signal at stop (i.e., a “SPAD” or Signal Passed at Danger); errors in the placement or removal of protections for worksites, which may result in rail traffic entering an active worksite; and workers entering the rail corridor without adequate protection from rail traffic movement.

Communication may be adversely impacted and impeded by the relationships between the parties exchanging information (Dunbar 2015) including an unequal power distribution (Streeton et al. 2016). In relationships where there is an unequal distribution of power, those that perceive themselves to hold less powerful positions are less likely to challenge the decisions of those they perceive to be in positions of more power — this phenomenon is referred to as an ‘authority gradient’ (Dobson et al. 2014).

In the Australian context, current major review papers have identified evidence of authority gradients in rail and the associated impact on safety (Luva & Naweed 2021). However, there is a paucity of rail research looking at the influence that power disparity has on miscommunication. Understanding how organisational factors influence power imbalances within rail operational groups and how they communicate, provides a new dimension for establishing the contributing factors in accidents and incidents.

1.1 Research Aim

This study set out to explore how the structural hierarchy of rail organisations and their operational imperatives influence distributions of power across multidisciplinary groups, impacting on their communication.

2 METHODOLOGY

2.1 Research Design

A qualitative research orientation was used, focusing on data collection from network controllers. As network controllers sit at the top of the operational hierarchy with a requirement to communicate with all other operational groupings, they were the focus of this study. To develop insights into their perceptions and experiences, semi-structured, one-to-one interviews were conducted and incorporated application of the Scenario Invention Task Technique. This technique elicits expert knowledge by posing questions to create hypothetical scenarios that allow participants to share their insights and experiences. The technique has been used to elicit knowledge across various industries, including healthcare (Naweed et al. 2021), aviation (Kingshott & Naweed 2018), maritime (Pabel et al. 2020) and rail (Naweed 2020).

2.2 Participants and Recruitment

Participants were recruited from eight rail organisations across Australia ($n = 7$) and New Zealand ($n = 1$) using a maximum variation sampling approach. A total of 55 network controllers and signallers (4 female, 51 male) were interviewed, ranging from 22 to 67 years of age ($M_{age} = 45.5$, $SD = 11.1$).

2.3 Procedure

Following provision of participant consent, a combination of closed and open-ended questions were initially posed, allowing participants to delve into the substance of their roles within the operational hierarchy and their relationships with other operational working groups (e.g., *“what is it about the job that you enjoy?”* *“If you could choose one word to describe your relationship with drivers, what would that be?”*). Having provided an opportunity to build rapport with each participant and establish the context, the scenario invention task technique was initiated. Participants were asked to consider a situation in which a network controller might inadvertently increase safety-risk, such as leading a train driver into a SPAD. To support their examples, participants created illustrations of their scenarios with felt-tip markers and A3-sized paper. All interviews were voice recorded and ranged from 90 to 120 minutes in length.

2.4 Data Analysis

Voice recordings were transcribed and analysed using thematic analysis (Braun & Clarke 2012) concentrated

at the “latent level” (Bengtsson 2016), indicating that it sought to interpret and identify the underlying meaning being expressed through concepts and implicit ideas. A framework was developed for understanding how the hierarchy of power was distributed across the multidisciplinary working groups and underpinned by operational activities.

3 RESULTS AND DISCUSSION

Preliminary analysis of data identified three broad themes (i.e., organisational factors) that contributed to the power disparities between network controllers and other operational groups. These were: (1) the accountability mechanisms; (2) the power vested in roles; and (3) the status attached to roles. Quotes and excerpts from data are provided to support points where relevant

3.1 Accountability Mechanisms

Participants attributed various high risk error scenarios solely to the responsibility of traincrew, for example *“if drivers drove to signals, they wouldn’t have SPADs”* and the *“driver not paying attention.”* However, traincrew were also held accountable under circumstances where drivers adhered to the signal without dispute. For example, errors in holding a train for too long at a signal and causing a scheduling delay were also attributed to traincrew, even when the network controller was distracted with other tasks, *“the driver didn’t challenge the signal, he didn’t call me up and say – I’m sitting at the signal, what am I sitting here for?”*

Participants placed effort into finding the causes of error and apportioning blame because they were afraid of being at fault themselves, *“oh, shit have I done something wrong?”*, *“has there been something that I could have done or is it the driver?”* The finding of a fault and fear of being at fault had two facets. The first was error as a reflection of the abilities of the participant, which called into question their capability, expertise or performance. The second was a sense of inequity in the network controller retaining accountability for incidence of error when traincrew ultimately operated the train:

One of the things that is a big bugbear of mine is when we do have a SPAD where a driver has passed the signal, where it’s the driver’s fault, the first thing they ask is – why wasn’t the signal at proceed?

The extent to which participants felt negative or fearful about their contribution to incidence of error was influenced directly by the organisation, for example, how staff performance was managed upon error discovery. While participants roundly felt that they were being unfairly blamed, the accountability assigned to the network controller by the organisations was indicative of the correlating position of power in the operational hierarchy, *“the buck stops at train control”*.

3.2 Power Vested in Roles

The network controller role was considered to possess power through: (1) knowledge; (2) system controls; and

(3) regulatory principles of authority. Holding knowledge was powerful in the sense that not having access to relevant knowledge was dangerous. Maintaining safety relied on effective communication between members of a multidisciplinary team. One participant expressed an awareness of the danger of creating a situation in which other operational staff might be hesitant to share or withhold relevant information:

So, you've got to have the ability to talk not just with a person but have them feel like they want to talk back to you at the same time. That's something that I've always found if you do that, you don't have somebody withholding information from you that they could very easily wish to share, but they feel suddenly like they shouldn't because apparently you're talking and they should listen and not tell you anything, you know?

An important dimension of the power wielded by the network controller role was in the knowledge of the operational plan, also referred to as the *"bigger plan."* Knowledge of the operational plan and designated authority to enact the plan was perceived to empower the network controller:

The idea is to come through the controller because they know exactly what's happening in that area at that time and allow everything to happen or not to happen.

Authority and power to make autonomous decisions sat within the purview of the network control role, with some crossover between responsibilities and roles across participant organisations. Participation of both network controller and signallers was included in this study. The difference between roles was summarised by one participant as, *"a signaller works at a location and controls a location – a train controller controls the whole territory."* A perceived attitude of superiority (in the network controller) and deference (from the signaller) between these operatives was revealed in the findings.

Providing permission or authorising activities on the network was:

Based on the original fundamentals of a train authority where the controller will dictate what he wants to occur to either the track worker or the train driver.

Signals were called and routes were set by network controllers, and when the signalling system was not operational, a *"special proceed authority"* was issued by the network controller to authorise the movement of rail traffic. For track workers, the network controller issued a *"track occupancy authority"* or *"track warrant"* granting them permission to undertake maintenance or repair work on the network. Any person entering the rail corridor (rail worker or member of the public) would be required to obtain authorisation from the network controller responsible for the relevant portion of the rail corridor. In some organisations, shunting that happened outside a yard required a shunt plan to be authorised by the network controller. In short, any activity that occurred on the network or in the rail corridor did so within a system controlled by the network controller, and by virtue of the power vested in their role.

3.3 Status Attached to Roles

As a multidisciplinary team, rail workers engaged in activities equally important to the function and operability of the network, but the contribution of some members of the team were perceived to be more highly valued than others:

Control's this ivory tower where some sort of deity or something and you're working together for the same goal, aren't you?

Participants with significant experience within the rail industry described the network controller as a career progression aspiration, *"I always thought [network control] was too far to reach for"* and *"I thought it was a little out of my league"*. Participants referred to working their way *"through the ranks"* and passing exams to access a new promotion or grading within operational streams, such as signalling, before progression to network controller. In many instances, multiple applications needed to be made before successfully attaining a network control position and/or considerable time was invested to progress further.

The hierarchical structure of rail organisations was illustrated by participants describing having worked their way *"up the ladder"*, and the organisational positioning of the network controller at the apex of the operational hierarchy, *"train controllers are up the top. That was the pinnacle of where you wanted to be at"*.

The status elevation of the network controller to the top of the operational hierarchy was reinforced by nine participants from across five organisations who all referred to the network controller as *"God"*. This label had a number of connotations, with participants attributing it to their holistic view of the network and ability to access various systems from their control location, but also a sense of being more capable than their multidisciplinary peers, with their word as law.

Further to the *"God"* label, other metaphors used to describe and characterise the status attached to the controller role were: *"the hub' and everyone else as being 'the wheel'"*, the *"backbone' of the network"*, the *"king of the castle"* and the *"masters of the universe"*. Along with these self-perceptions of status attached to their role, there was also an awareness of the negative perceptions held about network controllers from others in the multidisciplinary team. Drivers, for example, were postulated as perceiving network controllers to be *"arrogant"* or *"obnoxious pricks."*

4 GENERAL DISCUSSION

While the network controller authorises movement and activities on the network, the actual responsibility for undertaking tasks is retained by the relevant working group. However, it is evident through the mechanisms that organisations used to performance manage and/or investigate occasions of error, that accountability is attributed to network controllers for activities they are not responsible for. This generated animosity on the part of the network controller and was evident in attitudes towards interactions with other workers.

Power over disseminating information translated to power over individuals and the activities of the individual working groups. Knowledge boundaries were created between working groups, engendering insular teams with distinct operational expertise for their knowledge domains. This gave rise to “expertise-based status and power” differences across teams, influencing their interactions (Jackson 1996, p. 61).

Status and power are intrinsically linked. Power is an advantage of status, borne out of positioning high within a hierarchy and is held by those with access to, or the ability to confer or withhold knowledge or resources. The network control role was acknowledged as being positioned at the “highest level of the management system” (Cheng 2019, p. 213). Having operational oversight of the entirety of the rail network further contributed to a perception of omniscience.

Taken together, both the power vested in roles and the status attached to them fostered a culture of superiority in those further up the hierarchy, and resentment from those further down. This supported the notion that perceptions of status of network controllers influence attitudes and behaviours towards other members of the operational team.

4 CONCLUSIONS

There is a correlation between railway culture, unequal power distributions within organisational hierarchies and the existence of authority gradients between workers within and across multidisciplinary operational groups in rail. The generation of an authority gradient is not exclusively the domain of organisation influence, yet its contribution is significant. The findings in this study may be used to understand how the authority gradient is generated and how unequal power dispersal across groups impacts on whether they are likely to act upon instructions that they know to be erroneous, volunteer information that may be relevant to a decision being made or question a directive that they know to be unsound.

5 ACKNOWLEDGEMENT

The authors are grateful to the contacts and participants across all eight organisations for their valuable input and participation in this research. This work was funded by an Australian Research Council Discovery Early Career Research Award (DE160101137).

6 REFERENCES

Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, 2, 8-14.

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology*, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological (pp. 57–71). American Psychological Association.

Cheng, Y.-H. (2019). Railway safety climate: a study on organizational development. *International Journal of Occupational Safety and Ergonomics*, 25(2), 200-216.

Cheng, Y.-H., & Tsai, Y.-C. (2011). Railway-controller-perceived competence in incidents and accidents. *Ergonomics*, 54(12), 1130-1146.

Dunbar, N. E. (2015). A Review of Theoretical Approaches to Interpersonal Power. *Review of Communication*, 15(1), 1-18.

Dobson, K., Moors, A., & Norris, B. J. (2014). Literature Review of Safety Critical Communication Methodologies. Retrieved 23 Aug 2018, from Interfleet Transport Advisory Ltd. <https://www.era.europa.eu>

Gordon, S., & O'Connor, B. (2012). What goes without saying in patient safety. In R. Koppel & S. Gordon (Eds.), *First, Do Less Harm: Confronting the Inconvenient Problems of Patient Safety* (pp. 41-61). Ithaca: ILR Press.

Jackson, S. (1996). The consequences of diversity in multidisciplinary work teams. In (pp. 53-76).

Kingshott, K., & Naweed, A. (2018). “Taxiing down the runway with half a bolt hanging out the bottom”: Affective influences on decision making in General Aviation maintenance engineers. In R. Charles and J. Wilkinson (Eds.), *Contemporary Ergonomics and Human Factors 2018* (pp. 21-29). *Proceedings of Ergonomics & Human Factors 2018 Conference*. Birmingham, UK (23-25 April)

Luva, B., & Naweed, A. (2021). Authority gradients between team workers in the rail environment: a critical research gap. *Theoretical issues in ergonomics science*, 1-29.

Naweed, A. (2020). Getting mixed signals: Connotations of teamwork as performance shaping factors in network controller and rail driver relationship dynamics. *Applied Ergonomics*, 82, 102976

Naweed, A., Stahlut, J., & O’Keeffe, V. (2021). The Essence of Care: Versatility as an Adaptive Response to Challenges in the Delivery of Quality Aged Care by Personal Care Attendants. *Human Factors*, 00187208211010962.

Pabel, A., Naweed, A., Ferguson, S. A., & Reynolds, A. (2020). Crack a smile: the causes and consequences of emotional labour dysregulation in Australian reef tourism. *Current Issues in Tourism*, 23(13), 1598-1612.

Speroff, T., Nwosu, S., Greevy, R., Weinger, M., Talbot, T., Wall, R., . . . Burgess, H. (2010). Organisational culture: variation across hospitals and connection to patient safety climate. *BMJ Quality & Safety*, 19(6), 592-596.

Streeton, A., Bisbey, C., O’Neill, C., Allen, D., O’Hara, S., Weinhold, M., . . . Grubbs, P. (2016). Improving Nurse-Physician Teamwork: A Multidisciplinary Collaboration. *Medsurg Nursing*, 25(1), 31-34,66. Retrieved from <https://www.proquest.com>

Canadian & Cold Regions Rail Research Conference 2021

With 48,000 route kilometres of track, Canada has one of the most extensive rail networks in the world. The importance of rail transportation to both the Canadian economy and the quality of life enjoyed by Canadians is significant. The system moves people, goods and services to their destinations, provides jobs and sustains economic growth.

The Canadian rail sector faces many challenges, such as high safety and security standards, long travel distances, sparse population, harsh environments and extreme climatic changes. In addition, industry experts predict that 30% of the skilled work force in the rail sector will be eligible for retirement within the next five to ten years, and therefore the need to attract high caliber candidates and provide sector specific university training at the undergraduate and graduate level is greater than ever.

For Canada to continue to grow and prosper, it is critical that the transportation system evolve to become more innovative, more efficient and more resilient to our harsh and changing environment. The Canadian Rail Research Laboratory (CaRRL) was established to be Canada's premier education and research program in railway engineering; CaRRL is an important step towards addressing these industry challenges.

Visit carrl.ca to learn more about our innovative research!

Proceedings cover photo courtesy of Dr. Parisa Haji Abdulrazagh