

University of Alberta

Verification of the Weather Research and Forecasting Model for Alberta

by

Clark William Pennelly

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Earth and Atmospheric Sciences

©Clark William Pennelly

Spring 2013

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission

Abstract:

We simulated three heavy summer precipitation events in Alberta, Canada, using the Weather Research and Forecasting (WRF) model and compared the output precipitation data with observations to determine model accuracy. Each storm was simulated multiple times, using five different cumulus parameterization schemes and three grid resolutions. The explicit scheme, when simulated using 6 km grid resolution, was found to be the most accurate scheme when simulating these heavy precipitation events.

We also used the WRF model to perform daily forecasts during 2011-2012. The forecasts include 2-meter maximum and minimum temperatures, 10-meter wind speed, sea-level surface pressure, and daily accumulated precipitation. When compared against observations, the WRF forecasts showed seasonal differences and tendencies such as forecasting a smaller range of diurnal temperatures than observed. WRF was found to have high skill scores when forecasting maximum temperature against climatology and persistence forecasts, but was less skillful when forecasting minimum temperatures.

Table of Contents

	Page
Chapter 1: Introduction.....	1
1.1. The problem of weather forecasting.....	1
1.2. Research Objectives.....	2
1.3. Numerical weather prediction for Alberta.....	3
1.4. The Numerical Weather Prediction Model.....	3
1.5. Outline of content of thesis.....	5
1.6. Figures.....	7
1.7. References.....	8
Chapter 2: Verification of the WRF Model for Simulating Heavy Precipitation in Alberta.....	 10
2.1. Introduction.....	10
2.2. Methodology and model description.....	12
2.2.1. Cumulus parameterization schemes.....	13
2.2.2. Model Verification.....	15
2.3. Simulation Results.....	17
2.3.1. WRF 48 hour simulations using 15 km grid resolution... 17	
2.3.1.1. Rain storm A.....	17
2.3.1.2. Rain storm B.....	20
2.3.1.3. Rain storm C.....	22
2.3.2. WRF 48 hour forecasts using 30 km grid resolution.....	24
2.3.3. WRF 48 hour forecasts using 6 km grid resolution.....	25
2.3.3.1. Rain storm A.....	26
2.3.3.1. Rain storm B.....	27
2.3.3.3. Rain storm C.....	27
2.4. Discussion and Conclusions.....	28
2.5. Tables.....	30
2.6. Figures.....	35
2.7. References.....	46
Chapter 3: Verification of daily WRF forecasts for Alberta.....	50
3.1. Introduction.....	50

3.2.	Methodology.....	52
3.2.1.	Numerical experiments.....	52
3.2.2.	Numerical model.....	52
3.2.3.	Forecast verification.....	53
3.2.4.	Forecast skill.....	54
3.2.5.	Verification datasets.....	55
3.3.	Verification Results.....	55
3.3.1.	Seasonal Verification.....	55
3.3.1.1.	Summer season.....	57
3.3.1.2.	Winter season.....	58
3.3.1.3.	Fall season.....	59
3.3.1.4.	Spring season.....	59
3.3.1.5.	Averaged year.....	60
3.3.2.	Skill Score of temperature forecasts.....	62
3.3.2.1.	WRF forecast skill compared to climatology forecasts.....	63
3.3.2.2.	WRF forecast skill compared to persistence forecasts.....	64
3.3.2.3.	WRF forecast skill compared to Environment Canada's forecasts.....	64
3.4.	Discussion and Conclusions.....	65
3.5.	Tables.....	67
3.6.	Figures.....	72
3.7.	References.....	75
Chapter 4:	Summary and Conclusions.....	79
4.1.	Summary.....	79
4.2.	Conclusions of WRF precipitation simulations.....	79
4.3.	Conclusions of WRF daily forecasts.....	80
4.4.	Comments and suggestions for future work.....	81
4.5.	Concluding remarks.....	82
4.6.	References.....	84
	Appendix.....	85

List of Tables

	Page
2-1 WRF simulation initialization and end time for storms A, B, and C, using different cumulus parameterization schemes and grid resolutions.....	30
2-2 Domain averaged verification results for storm A, B, and C, when simulated at 15 km grid resolution.....	31
2-3 Verification results for storm A, B, and C, when simulated at 15 km grid resolution, for the river basin which received the most precipitation for the given storm.....	31
2-4 Precipitation scores for storm A when simulated at 15 km grid resolution, for three precipitation thresholds.....	32
2-5 Precipitation scores for storm A, B, and C, when simulated at 15 km grid resolution, at the highest precipitation threshold of 50 mm.....	32
2-6 Verification results between different cumulus parameterization schemes when simulating storm A at 30 km grid resolution.....	33
2-7 Verification results between the Explicit (EX) and Kain-Fritsch (KF) cumulus parameterization scheme for storm A, B, and C, when simulated at a grid resolution of 6 km.....	34
3-1 List of the nine observation stations used for the verification of the WRF model for Alberta.....	67
3-2 Characteristics of WRF forecasts used in this study.....	67
3-3 Forecast Day 1 errors, by station, during the summer season.....	68
3-4 Forecast Day 2 errors, by station, during the summer season.....	68

3-5	Seasonal forecast errors for Day 1 forecasts, averaged from 9 weather stations.....	69
3-6	Seasonal forecast errors for Day 2 forecasts, averaged from 9 weather stations.....	69
3-7	Precipitation verification for Day 1 and Day 2 forecasts during summer, averaged from 9 weather stations.....	70
3-8	Skill score values for Day 1 temperature forecasts for the year of study for nine weather stations, when compared against climatology, persistence, and Environment Canada.....	71
3-9	Skill score values for Day 1 and 2 temperature forecasts, averaged across the nine weather stations for each season, when compared against climatology, persistence, and Environment Canada.....	71
F-1	Precipitation contingency table to determine POD, FAR, BIAS, and ET scores.....	88

List of Figures

	Page
1-1 Numerical Weather Prediction domain setup with the parent domain and nested child domain over part of North America.....	7
2-1 Numerical Weather Prediction domain setup using an outer grid resolution of 45 km, and inner grid resolution of 15 km.....	35
2-2 Major river basins of Alberta, Canada.....	36
2-3 48 hour precipitation amounts for storm A.....	37
2-4 Precipitation verification statistics, for storm A, at each river basin, simulated at 15 km grid resolution.....	38
2-5 48 hour precipitation amounts for storm B	39
2-6 Precipitation verification statistics, for storm B, at each river basin, simulated at 15 km grid resolution.....	40
2-7 48 hour precipitation amounts for storm C	41
2-8 Precipitation verification statistics, for storm C, at each river basin, simulated at 15 km grid resolution.....	42
2-9 48 hour precipitation amounts for storm A when simulated at three different grid resolutions	43
2-10 Precipitation verification statistics, for storm A, at each river basin, simulated at 30 km grid resolution.....	44
2-11 48 hour precipitation amounts for storm B and C when simulated using 6 km grid resolution.....	45
3-1 Numerical Weather Prediction domain setup for WRF forecasts.....	72

3-2	Time series of daily WRF forecasts and observed meteorological variables at Edmonton International Airport, from 1 June 2011 to 31 August 2011.....	73
3-3	Comparison between Day 1 WRF forecasts and observed meteorological variables at Edmonton International Airport during the summer season...	74

Chapter 1: Introduction

1.1. The problem of weather forecasting

Accurate prediction is the goal of many scientists. But few scientists have a more complex and challenging medium to work with than the meteorologist. With a variety of air circulation sizes present in the atmosphere, numerous different materials on the Earth's surface, and the fact that water vapor can vary considerably in space as well as in time, it is clear that the interactions between the Earth and the atmosphere are very complex. With such intricate and ever-changing weather patterns that are often inadequately observed, it is not surprising that improvements in the accuracy of weather predictions have been painfully slow when compared against other fields of science.

Meteorologists have long hoped that they would be able to compute tomorrow's state of atmosphere that would compare closely to tomorrow's observations, just as astronomers are able to compute future eclipses. The basic physical equations governing the behavior of fluids have been known for over 100 years. Thus, in principle, at a first glance it might seem meteorologists should be able to solve these equations to provide weather forecasts, but only for the last 60 years have practical, although imperfect, computed forecasts been made on a daily basis. And even today, some weather elements, such as heavy precipitation, fog, and local storms, are often poorly forecast with operational numerical weather prediction (NWP) models (e.g. Roebber and Bosart, 1998).

There exist several obstacles to obtaining an accurate prognosis using the physical-mathematical modeling approach. First, the mathematical equations that describe air motion, thermodynamic processes, and precipitation microphysics are complex non-linear partial differential equations, which cannot be solved simply by analytical methods. Instead, numerical approximation techniques, which are extremely laborious and require an immense number of computations, must be used. The numerical computations require a fine spatial and temporal resolution; however, there are always important smaller scale features such as turbulent air motion, which are unable to be resolved.

A second obstacle to accurate weather forecasts comes from the lack of observational data to provide adequate initial conditions for the input to the weather model. If the weather model is incorrectly initialized colder than reality at a certain location, advection will transport this faulty coldness downwind. In addition, small initial

errors tend to magnify in nonlinear differential equations. Third, most of the time the net force acting on any part of the atmosphere is very small compared to the magnitudes of each of the individual forces acting on it. Thus minute differences in vertical forcing cause significant variation in cloud condensation and precipitation formation.

1.2. Research Objectives

The process of determining the accuracy of a NWP model when compared against observations is called verification. This study will verify the Weather Research and Forecasting model (WRF) when simulating historic precipitation events which caused extensive damage and loss of life due to flooding. I will also verify the WRF model when forecasting various near-surface parameters, such as temperature and wind speed.

I will determine the accuracy of the WRF model when simulating heavy precipitation for severe flooding cases. I will perform sensitivity tests to determine how the grid resolution and cumulus parameterization schemes affect the accuracy of the simulated precipitation field. Since this objective will assess WRF's ability to simulate heavy precipitation storms which caused damaging flooding, the analysis will be done from the perspective of a flood forecaster and will verify precipitation over the major river basins in Alberta.

I will also be using the WRF model to forecast weather over central and southern Alberta, and verify the forecast data against observed data. This is performed to determine how accurately WRF forecasts in this region. I will determine the forecast accuracy for temperature, wind speed, precipitation and surface pressure. The evaluation will also include how WRF performs against other forecast methods. I propose to investigate two scientific questions:

1: What is the accuracy of the WRF model when simulating flooding events in Alberta caused by extensive precipitation, and how does the selection of grid resolution and cumulus parameterization schemes affect the precipitation?

2: How accurately can the WRF model forecast temperature, pressure, precipitation, and wind speed for the following two days? In addition, how skillful are WRF temperature forecasts when compared against other forecast methods?

1.3. Numerical weather prediction for Alberta

The province of Alberta (whose boundaries extend from 49° to 60° N and 120° to 110° W) is situated in a meteorologically active area with the mean position of the polar jet stream straddling southern Alberta during the summer. The Rocky Mountains form a northwest-southeast barrier along the western border of Alberta. The plains to the east provide a striking contrast to the mountainous west. The combination of these two geographic features aids in the development of storms throughout the year, particularly during the summer months when the mountains can cause a subsidence inversion that acts as a capping lid for the build-up of Convective Available Potential Energy (CAPE). Consequently, southern Alberta is highly susceptible to severe convection, having on average 52 days with hail each summer (Smith et al. 1998).

Numerical prediction of summertime convection over Alberta can be challenging, particularly due to the spatial distribution of the convective precipitation. Erfani et al. (2003) used the Global Environmental Multiscale (GEM) model in a non-hydrostatic configuration to simulate a convective storm over central Alberta. While the simulated storm structure and intensity resembled radar observations, the simulated storm track deviated significantly from the observed storm track. Milbrandt and Yau (2006) used a mesoscale model to simulate a severe convective storm forming over Alberta. Despite the fine spatial resolution of the nested grid, the simulation did not coincide with the observed radar echoes. Summer storms can result in extensive precipitation, and accurate simulation of these events is desirable.

1.4. The Numerical Weather Prediction Model

The Weather Research and Forecasting Environmental Modeling System (WRF-EMS, henceforth as WRF) is a state of the art, full physics Numerical Weather Prediction model (Michalakes et al. 1999, Skamarock et al. 2008). WRF version 3.0 was used for this project. The Environmental Modeling system (Rozumalski 2006) was used for ease of installation, configuration, and execution of WRF. WRF was used configured in an “off-the-shelf” manner, where most settings were not changed from their default values. This was done to assess the accuracy of the model without changing any default configurations, other than cumulus parameterization schemes; likely the case when operated by a user who does not have enough information regarding specific configurations to decide which is more useful for their work.

Most of the WRF model predictions presented in this thesis use the following grid nesting configuration: there is a parent domain covering some of the northwest section of North America with a 45 km grid resolution, and a nested inner domain covering southern Alberta with a 15 km resolution (Figure 1-1). The simulations used 2-way nesting; the coarse grid solution is interpolated into the fine grid to provide the needed boundary conditions, and the fine grid feeds back into the coarse grid after processing the meteorological information for some time steps.

WRF precipitation simulations of historic flash flooding were initialized from the North American Regional Reanalysis (NARR, Mesinger et al. 2006) dataset, and boundary conditions were updated every three simulation hours from NARR. WRF daily forecasts were initialized from the North American Mesoscale model using 40 km resolution (NAM212, Black 1994), and boundary conditions were also updated every 3 hours. WRF forecasts and simulations were initialized at 1200 UTC on the day prior to meteorological data collection, and performed 69 hours of simulation. Colle et al. (1998) found that after 12-18 hours the model precipitation was generally spun-up, when initialized from a cold start (no hydrometeors). 69 hours of simulation length was done to simulation of the atmosphere for the following two entire days. WRF processes simulations using 45 pressure levels, using sigma vertical coordinates (Skamarock et al. 2008), with the pressure top at 50 mb. The Lin et al. (1983) microphysics scheme was used. It uses six categories of water substance: water vapor, cloud water, rain, cloud ice, snow, and graupel. The model output from the nested domain was analyzed using the Grid Analysis and Display System (GrADS, <http://www.iges.org/grads/>) to determine the meteorological information. This output was verified with observational data to determine model accuracy.

There have been many verification studies of the WRF model. Davis et al. (2006) verified precipitation forecasts over the continental United States. WRF was used by Flesch and Reuter (2012) to simulate major flooding events in Alberta as well as to understand how WRF performs when the topography of the nearby Rocky Mountains was altered. Cheng and Steenburgh (2005) performed temperature, dew point, and wind speed verification during summer in the western United States, as well as sensitivity studies regarding WRF's slab-soil model.

Other research has been performed using WRF that was not focused on verification. Gilliland and Rowe (2007) performed high resolution simulations of the WRF super cell benchmark scenario, which is included with the WRF software, using

different cumulus parameterization schemes. Done et al. (2004) performed high resolution forecasts for summer days with Mesoscale Convective Systems using 10 km grid resolution simulations with a cumulus parameterization scheme, as well as 4 km resolution simulations without one. WRF has been used from the high latitudes (e.g. Powers 2007), mid-latitudes (e.g. Davis et al. 2006) as well as near the equator (e.g. Skok et al. 2010). WRF has been used for research in forecast mode (e.g. Cheng and Steenburgh 2005), simulation of past events (e.g. Flesch and Reuter 2012), climate simulations (e.g. Qian et al. 2010) and air quality applications (e.g. Borge et al. 2008). WRF can also be coupled with a fire module (WRF-Fire, Patton and Coen 2004) and a chemistry module (WRF/Chem, Grell et al. 2005). There is a large amount of versatility behind the WRF model, which researchers find beneficial when conducting their studies.

As part of this study, I will evaluate WRF forecasts for an entire year. These WRF forecasts will be compared against other forecast methods, such as forecasts issued from the Weather Office at Environment Canada, to determine which forecast is more accurate for predicting temperatures. Even if WRF is found to be less accurate than Environment Canada's temperature forecast, which is anticipated due to their massive infrastructure in predicting the weather, WRF still can be used in important ways. While Environment Canada's forecasts might be more accurate, which will be examined in Chapter 3, their daily forecasts may not align with the information required by many users. An example would be that Environment Canada releases a forecast for heavy precipitation amounts at certain locations during an intense rain period, but does not publically release the precipitation information regarding the entire river basin. This information would be very valuable to a hydrologist, or a flood forecaster, and would be easy to produce using the WRF model. The ability to create your own powerful scripts to be used on model output from WRF can be very useful, from research applications to those who require detailed specifics about the weather for their own operational use.

1.5. Outline of content of thesis

This thesis was written in “paper” format and thereby abides by certain conventions. Chapter 2 and 3 are self-contained units with their own introduction, conclusion, and bibliography, and were written to be read independently of the other chapter. As a consequence, there is some duplicity when examining the entire thesis as a whole. Chapters intended for journal submission were written in the format required by the particular journal.

This thesis consists of 4 chapters. Chapter 1 provides some background of the thesis problem and reasons for the investigation as well as a description of the WRF model and some

Chapter 2 will be submitted for publication as a journal article in *Atmospheric Research*, a technical journal focused on precipitation. In this chapter an analysis of WRF's ability to simulate heavy precipitation storms which caused major damage due to floods is given. Three storms are presented in this chapter, as well as an analysis using different cumulus parameterization schemes to simulate each storm at varying resolutions. Background information regarding the purpose of cumulus parameterization schemes, as well as the differences between cumulus parameterization schemes available to WRF, is also presented. The model precipitation is assessed for accuracy at the scale of the major river basins of Alberta. A single cumulus parameterization scheme is selected as being most skillful across the Alberta region for these heavy precipitation events. An evaluation of the storms when simulated with different grid resolutions is also performed.

Chapter 3 will be submitted for publication as a journal article in *Weather and Forecasting*, a technical journal for original advances in weather forecasts and evaluations of numerical weather prediction. In this chapter an analysis of WRF's ability to forecast various parameters across central and southern Alberta is performed. Background information regarding forecast verification and model skill is presented. An analysis regarding WRF's forecast skill relative to other forecast methods is also performed. Seasonal forecast differences, such as biases and errors, are determined, as well as general findings regarding the model when assessed across the year of study.

Chapter 4 presents the summary of the research project as well as conclusions drawn from the results of the simulations and forecasts. This chapter also presents suggestions for future research.

1.6. Figures

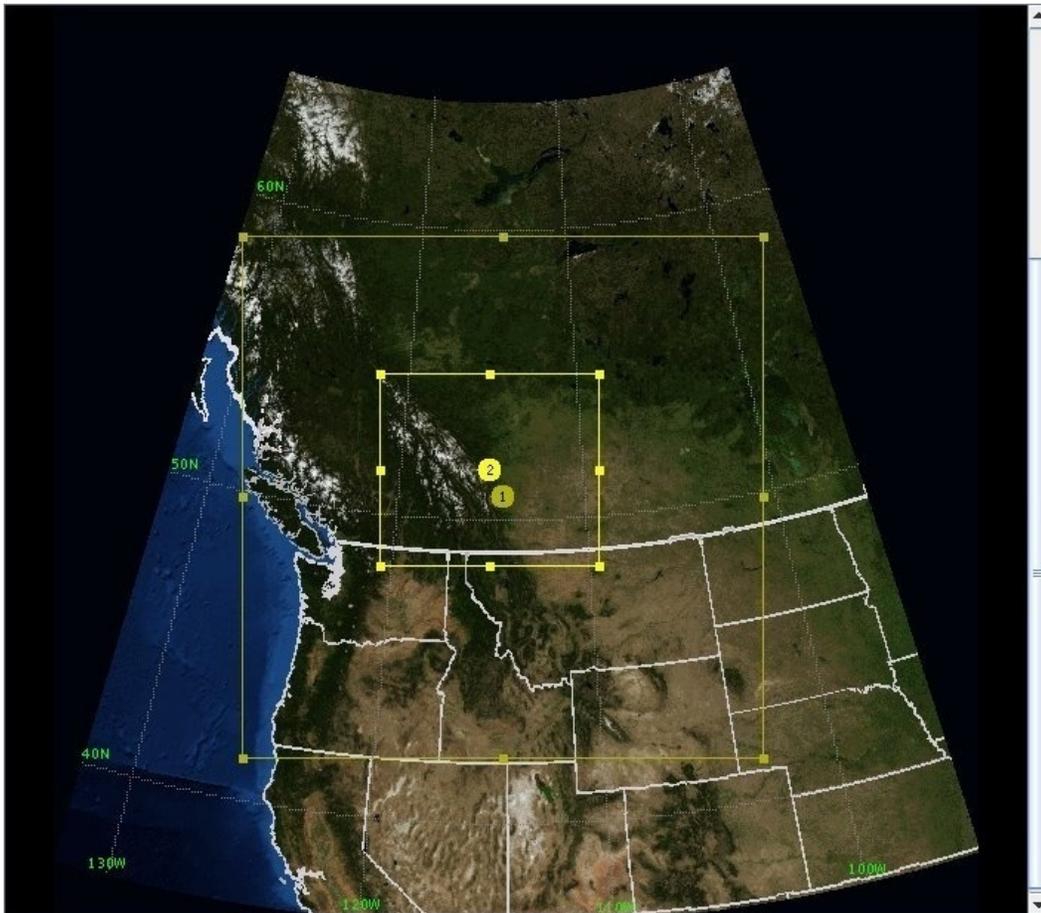


Figure 1-1: Numerical Weather Prediction domain setup with the parent domain (larger rectangle) and nested child domain over part of North America. The center point of each domain is indicated by the yellow circle while the number refers to the domain; the larger domain is 1, and the nested domain is 2.

1.7. References:

- Black, T. L., 1994: The New NMC Mesoscale Eta Model: Description and Forecast Examples, *Wea. Forecasting*, **9**, 265-278
- Borge, R., V. Alexandrov, J.J. Vas, J. Lumberras, E. Rodriguez, 2008: A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula, *Atmos. Envir.*, **42**, 8560-8574
- Cheng, W.Y.Y., and W. J. Steenburgh, 2005: Evaluation of Surface Sensible Weather Forecasts by the WRF and the Eta Models over the Western United States, *Wea. Forecasting*, **20**, 812-821
- Davis, C., B. Brown, R. Bullock, 2006: Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas, *Mon. Wea. Rev.*, **134**, 1772-1784
- Done, J., C.A. Davis, M. Weisman, 2004: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model, *Atmos. Sci. Letters*, **5**, 110-117
- Erfani, A., A. Méthot, R. Goodson, S. Bélair, K.-S. Yeh, J. Côté, and R. Moffet, 2003: Synoptic and mesoscale study of a severe convective outbreak with the nonhydrostatic Global Environmental Multiscale (GEM) model. *Meteor. Atmos. Phy.*, **82**, 31-53
- Flesch, T.K., and G.W. Reuter, 2012: WRF Model Simulation of Two Alberta Flooding Events and the Impact of Topography, *J. of Hydrometeor.*, **13**, 695-708
- Gilliland, E.K. and C.M. Rowe, 2007: A Comparison of Cumulus Parameterization Schemes in the WRF Model. Preprints, *21st Conference on Hydrology*, San Antonio, TX, Amer. Meteor. Soc., P2.16
- Grell, G.A., S.E. Peckham, R. Schmitz, S.A. McKeen, G. Frost, W.C. Skamarock, B. Eder, 2005: Fully coupled "online" chemistry within the WRF model, *Atmos. Envir.*, **39**, 6957-6975
- Lin, Y-L., R.D. Farley, H.D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model, *J. of Clim. and Appl. Meteor.*, **22**, 1065-1093
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi W., 2006: North American Regional Reanalysis, *Bull. Amer. Meteor. Soc.*, **87**, 343-360

- Michalakes, J., D. Dudhia, D. Gill, J. Klemp, W. Skamarock, 1999: Design of a next-generation regional Weather Research and Forecast Model, *Towards Teracomputing*, World Scientific, River Edge, New Jersey, 117-124
- Milbrandt, J.A. and M.K. Yau, 2006: A Multimoment Bulk Microphysics Parameterization. Part III: Control Simulation of a Hailstorm, *J. Atmos. Sci.* **63**, 3114-3136
- Patton, E.G. and J.L. Coen, 2004: WRF-Fire: A coupled atmosphere-fire module for WRF. Preprints of Joint MM5/Weather Research and Forecasting Model Users' Workshop, Boulder, CO, June 22-25. NCAR, 221-223
- Powers, J.G., 2007: Numerical Prediction of an Antarctic Severe Wind Event with the Weather Research and Forecasting (WRF) Model, *Mon. Wea. Rev.* **135**, 3134-3157
- Qian, Y., S.J. Ghan, L.R. Leung, 2010: Downscaling hydroclimatic changes over the Western US based on CAM subgrid scheme and WRF regional climate simulations, *International J. of Climatology*, **30**, 675-693
- Roebber, P.J. and L.F. Bosart, 1998: The Sensitivity of Precipitation to Circulation Details. Part I: An analysis of Regional Analogs, *Mon. Wea. Rev.*, **126**, 437-455.
- Rozumalski, R. A., 2006: SOO/STRC WRF EMS User's Guide. Available online at <http://strc.comet.ucar.edu/index.htm>
- Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.Y. Huang, W. Wang, J.G. Powers, 2008: A Description of the Advanced Research WRF Version 3, *NCAR technical note NCAR/TN-475+STR*, 113 pp
- Skok, G., J. Tribbia, J. Rakovec, 2010: Object-Based analysis and Verification of WRF Model Precipitation in the Low- and Midlatitude Pacific Ocean, *Mon. Wea. Rev.*, **138**, 4561-4575
- Smith, S.B., G.W. Reuter, M.K. Yau, 1998: The Episodic Occurrence of Hail in Central Alberta and the Highveld of South Africa, *Atmos.- Ocean*, **36**, 169-178

Chapter 2: Verification of the WRF Model for Simulating Heavy Precipitation in Alberta

2.1. Introduction

The Rocky Mountains form the Continental Divide extending some 2500 km from northern Canada to southern Texas. This mountain barrier strongly affects the weather and precipitation of the province of Alberta, Canada. The orographic effects are particularly evident during the summer due to differential slope heating which gives rise to convergence and often trigger convective outbreaks (Smith and Yau 1987). The summer season can experience extreme rainfall events associated with the passage of an upper air cutoff low and rapid lee cyclogenesis over the Alberta foothills region (Reuter and Nguyen 1993). The transport of water vapor to Alberta occurs in moist warm conveyor belts often originating from the Gulf of Mexico (Brimelow and Reuter 2005). These extreme rainfall events can lead to flash flooding in southern Alberta.

In June 2005, extensive rainfall caused flooding in southern Alberta (Ou 2008). Sixteen municipalities declared states of emergency. Thousands of people were forced to leave their homes along the rivers. The floods claimed four casualties and the estimated damage was 400 million Canadian dollars. The precipitation fell from four distinct storms with similar tracks. The dates and recorded rainfall amounts were 1-5 June (140 mm), 5-9 June (248 mm) 16-19 June (152 mm), and 27-29 June (90 mm). This paper focuses on numerical simulation of two of these extreme events, 5-9 June and 16-19 June. A third modeling case of the extreme rainfall event of 12-13 July 2010, with maximum recorded rainfall of 110 mm, was added to have an example of highly convective event. On 12 July 2010 the metropolitan city of Calgary suffered the most damaging hailstorm in Canada's recent history. The maximum hail size was 4 cm in diameter, and damages were assessed at 400 million Canadian Dollars in insurance claims.

Hydrological models estimating water flow for rivers in Alberta need a high spatial and temporal resolution of precipitation data. Rain gauge measurements alone do not provide adequate resolution, particularly in the orographic regions of south west Alberta. Weather radar imagery can estimate rainfall rate, but not over mountainous terrain because ground clutter distorts radar echoes. In addition, radar images have limited forecast skill, as they cannot be produced prior to the precipitation event. In recent years there have been efforts to use precipitation estimates from Numerical Weather Prediction (NWP) models as input for hydrological models.

With the advance of computing power and data assimilation it is possible to run NWP models as a tool for flood forecasters. An important issue is to assess the skillfulness of these models in predicting the spatial distribution of rainfall to obtain reliable estimates of the total water mass falling over the catchment areas of the river systems. One of the standalone NWP models used for mesoscale precipitation forecasting is the Weather Research and Forecasting Model (WRF). Flesch and Reuter (2012) used WRF to simulate heavy precipitation events over Alberta and examined the role of the topography in simulating and organizing the precipitation. Specifically, they performed simulations using the actual topographic grid and other simulations with a reduced mountain elevations. They concluded that a reduction of mountain elevation decreases maximum precipitation by about 50% over the mountains and foothills.

Due to the difference between the grid resolution of contemporary NWP models, typically between 15 and 30 km, and the scale of cumulus cells (1-10 km), NWP models often use cumulus parameterization schemes (CPS) to mimic the effects of cumulus clouds which are not resolved as they are smaller than individual model grid cells. These schemes attempt to trigger the convection and modify the temperature and moisture profiles within a model column based on the grid-scale (i.e. resolved) meteorological information. Common cumulus parameterization schemes are those of Betts and Miller (1986), Kain and Fritsch (1990), and Grell (1993). How cumulus parameterization schemes operate in NWP models is particularly important for hydrological applications, because the total volume of rainwater is sensitive to the cumulus parameterization scheme (Wang and Seaman, 1997). Kerkhoven et al. (2006) compared different cumulus parameterization schemes for an intense monsoon rainfall event in China and Japan and found that the Grell scheme was the most robust, performing well at different rainfall intensities.

The results of a NWP model can be quite dependent on the spatial resolution of the numerical grid. Intuitively, one would expect that simulations using the highest spatial resolution would provide the most accurate model simulation. Wang and Seaman (1997) and Done et al. (2004) indeed found that the finest grid resolution yielded the most accurate results, but Grubisic et al. (2005) and Roberts and Lean (2008) showed cases for which the finer grid spacing did not improve simulation accuracy. Furthermore, the finer grid spacing requires significantly more computation time and resources when performing simulations.

The purpose of this paper is to simulate intense Alberta summer rainstorms with the emphasis of evaluating the skillfulness of the model to accurately predict the spatial distribution of rainfall. Secondary objectives are to determine the optimum choice of cumulus parameterization schemes and the model's grid resolution. An inter-comparison between model precipitation and rain gauge observations will be performed on the model grid and also integrated across the watershed basins. Three storms will be simulated using the Weather Research and Forecasting model. Each storm case will be simulated five times at 15 km grid resolution, corresponding to each of the 5 cumulus parameterization schemes available. A single storm will be simulated at a grid scale of 30 km using all 5 cumulus schemes. All three storms will also be simulated at a grid resolution of 6 km, however only the two cumulus schemes found most accurate at 15 km resolution will be used for simulations at 6 km. This is to test whether increased resolution will produce a more accurate simulation over the two most accurate cumulus schemes at 15 km resolution. The model output will be examined for accuracy of location and amounts of precipitation, by comparing the simulated 48-hour output with the observed 48-hour precipitation amounts.

2.2. Methodology and model description

To test how well we can predict precipitation for heavy rainfall events in Alberta, we use the Weather Research and Forecasting (WRF) model (Michalakes et al. 1999). WRF features non-hydrostatic dynamics, multi-nest capability, and several physics options for boundary layer processes, radiation schemes, cloud microphysics, and cumulus parameterization schemes. Figure 2-1 shows our domain setup using an interactive nested domain inside the parent domain. The inner grid covers southern Alberta, and only the meteorological information from the inner grid was used in this study. Table 2-1 lists the simulations performed with the model start and end times for the three storms when simulated using different grid resolution and cumulus parameterization schemes. The model simulations were initialized at 1200 UTC from the North American Regional Reanalysis (NARR, Mesinger et al. 2006) dataset the day before the first day for which precipitation was to be collected, to allow for 18 hours of model spin-up. Boundary conditions for the outer domain were updated every 3 hours from the NARR dataset.

The WRF model was used in an off-the-shelf manner, using the Environmental Modeling System's (Rozumalski 2006) default configurations. The default configuration of WRF uses the Lin et al. (1983) bulk water microphysics scheme, the Noah (Skamarock et al. 2008) land surface scheme, and the Yonsei University planetary boundary layer scheme (Hong et al. 2006). The Lin et al. microphysics scheme resolves water vapor, cloud and precipitation processes using 6 hydrometeors: water vapor, cloud water, rain, cloud ice, snow, and graupel. The Noah land surface model uses atmospheric information from the surface layer to provide heat and moisture fluxes for 4-layers of soil. The Yonsei University planetary boundary layer scheme is responsible for the vertical sub-grid fluxes due to eddy transport in the entire vertical column. We selected the Advanced Research WRF dynamic model core to perform our simulations. The model has 45 vertical levels, with the top level at 50 mb. For more details of model choices we refer to Flesch and Reuter (2012). We used a 3-1 nesting option between outer and inner domains, as well as the same cumulus parameterization scheme for both domains for a given model simulation in an attempt to minimize inconsistencies at the interface of the computation grids (Warner et al. 1997). Three different spatial resolutions were used for the inner domain: 6 km, 15 km and 30 km. Other than slightly different domain setup and cumulus parameterization schemes, the simulations with different spatial resolutions used identical configurations and initialization data. Appendix A-D present brief descriptions of some WRF processes, such as the governing equations and grid nesting.

2.2.1. Cumulus parameterization schemes

By releasing latent heat and transporting water vapor and sensible heat, cumulus clouds modify the vertical profile of the environment. This takes place from subsidence of the environmental air, induced by the convection of mass upwards, as well as from detrainment of water substance from clouds (Ooyama 1971). Deep convection results in warming and drying of the environmental air as it is forced to sink (Yanai and Johnson 1993). Shallow convection will moisten and cool the environmental air from the detrainment of water vapor that evaporates.

Mesoscale models generally are unable to explicitly resolve convection. This is due to the difference in size between a cumulus cell (1-10 km) and the grid resolution of the numerical model (10-30 km). This results with cumulus clouds to be parameterized. The goal of cumulus parameterization is to determine the collective effects of cumulus clouds, rather than to resolve how a single cloud affects the vertical profile (Arakawa and

Schubert 1974). Different cumulus parameterization schemes have been developed based on different assumptions regarding how convection is triggered and how intense and deep the resulting convection.

All cumulus schemes have a trigger mechanism, which are the requirements for activation of the cumulus scheme. Cumulus schemes, once activated, will change the vertical profile of the grid column, often by modifying the moisture and temperature values. The modification will continue until the closure assumptions are met, which are a set of requirements to deactivate the parameterized convection inside the model grid. Each cumulus scheme has a different trigger assumption, modification process, and closure assumption, though cumulus schemes may have some aspects in common.

In this study we will perform some simulations without using a cumulus parameterization scheme. We will term these simulations explicit, identifying the simulations by the notation EX. We will compare those simulations with simulations using the Kain-Fritsch (KF), Betts-Miller-Janjic (BMJ), Grell-Devinji (GD) and Grell Three-Dimensional (G3D) cumulus parameterization schemes.

The KF scheme is a mass-flux parameterization scheme which determines the strength of convection from Convective Available Potential Energy (CAPE) when deep convection is triggered (Kain and Fritsch 1990). It is an extension of the earlier Fritsch-Chappell scheme (Fritsch and Chappell 1980) that modulates updraft/downdraft as well as entrainment and detrainment rates. The KF scheme triggers deep convection when a mixed parcel has positive vertical velocity over a depth that exceeds a specified cloud depth, typically 3 to 4 km (Kain 2003). KF mixes the air due to convection as well as related updrafts and downdrafts, as well as rigorously conserved mass, thermal energy, total moisture and momentum (Kain and Fritsch 1993). Convection triggered using the KF scheme will eliminate at least 90% of the CAPE over a certain amount of time, between 0.5 and 1 hour (Kain et al. 2003), which is the closure assumption. The removal of CAPE is performed by rearranging the mass in a column using the updraft, downdraft, and environmental mass fluxes. Once the CAPE has been removed, convective precipitation has been introduced into the model column.

The Betts-Miller-Janjic scheme is the extension of the Betts-Miller scheme (Betts and Miller 1986). Similarly to the KF scheme, BMJ is triggered when a parcel of air ascends a certain distance, as well as positive CAPE. The Betts-Miller scheme then adjusts the atmospheric temperature and moisture structure towards a reference structure by using deep and shallow convection. The reference structures are pre-determined

profiles of temperature and moisture inside the cumulus scheme. Janjic (1994) introduced a cloud efficiency factor into the Betts-Miller scheme to avoid spurious deep convection over warm oceans, a problem which happened using the older Betts-Miller scheme. Vaidya and Singh (2000) compared the previous Betts-Miller scheme to that of the BMJ and found that BMJ was more useful for forecasting precipitation over land. Janjic (2000) also introduced a variable relaxation time as well as more moisture profiles. Gilliland and Rowe (2007) verified WRF cumulus parameterization schemes at high resolution and found that the BMJ scheme had difficulty producing precipitation under some environments, such as a warm and dry environment. The BMJ scheme has been found favorable in other settings, and was used operationally in the WRF-NMM at NCEP.

The Grell-Dévényi (GD) scheme and the Grell three dimensional (G3D) scheme make use of ensemble parameterization with different closure assumptions and parameters (Grell and Dévényi 2002). The goal of the two Grell schemes are to determine the best configuration of an ensemble of parameters and closure schemes to feed back into the NWP model to statistically arrive at more accurate amount of precipitation than using only a single cumulus scheme. Since the Grell schemes are ensemble in nature, they feature numerous triggering mechanisms, adjustment processes, and closure assumptions, many of which are found in other parameterization schemes. From comparing to observed work, an optimal mixture of the subs-ensembles can be found. The G3D scheme was designed to be suitable for grid sizes less than 10 km (Skamarock et al. 2008), in addition to coarser resolutions. While there are configurations for the Grell schemes which the user can change to fine tune model simulations, there was no attempt at using anything other than the default settings during this study.

2.2.2. Model Verification

The observations used to assess precipitation were weather station archived data available from Environment Canada's Climate website (http://climate.weatheroffice.gc.ca/climateData/canada_e.html). Other observations were taken from the Government of Alberta; Agriculture and Rural Development AgroClimatic Information Service website (<http://www.agric.gov.ab.ca/app116/stationview.jsp>). These two sources of observed precipitation measurements account for approximately 120 observation stations used in this study. However, due to the low density of observation stations in Alberta, this results with a true observed grid resolution of approximately 60 km. Automatically recorded

precipitation observations can suffer from under-catchment during windy conditions (Colle et al. 2000), though there was no attempt to modify the observed data. The observation station data was interpolated to grid points of the same resolution as the model simulation by using the default Cressman (1959) function, Oacres, as part of the Grid Analysis and Display System (GrADS) mapping software (<http://www.iges.org/grads/>). While the Cressman function can suffer from producing more precipitation than other interpolation methods (5.7%, Hewitson and Crane 2005) other studies have used the Cressman function within the GrADS software (e.g. Davolio et al. 2009) with success.

Evaluation of the simulated precipitation was performed in two different ways; at the grid point and at the observation point.

Grid point: Each simulated grid point's 48 hour precipitation amount was compared to the corresponding observed grid point's interpolated 48 hour precipitation amount, to find the root-mean-squared-error (RMSE, Appendix E) at that grid point. The total RMSE was then found for each of the seven major river basins of southern Alberta (Fig. 2-2) below 54° N latitude. Analysis also determined the percent simulated precipitation compared to observed precipitation which was also performed for each river basin. This allowed for a spatial evaluation of each cumulus scheme in terms of over simulating or under simulating the precipitation amounts, while the RMSE gave information regarding how accurately WRF simulated the precipitation across the Alberta river basins.

Observation point: The second method of evaluation compared the observed amount of precipitation from over 120 observation stations with the simulated precipitation at those observation stations. This method used known values for observed precipitation, while the grid-point method used interpolated values determined by the Cressman function. The observation point analysis was performed at three 48-hour precipitation accumulated thresholds; above 10 mm, 25 mm, and 50 mm. The observed data and the simulated data were evaluated to determine the Probability of Detection (POD), False Alarm Ratio (FAR), BIAS, and Equitable Threat (ET, Schaefer 1990) for each threshold across the entire domain, by using a 2x2 contingency table (Wilks 1995). POD determined the percentage of stations at which WRF correctly simulated precipitation when precipitation was observed. FAR determines the percentage of falsely simulated events of precipitation when compared to the total number of simulated precipitation events. BIAS determined whether WRF simulated precipitation at more

stations than observed, or fewer than observed. A BIAS value of 1 indicates the same number of stations had simulated precipitation as were observed with precipitation, whereas a BIAS value of 2 would imply WRF simulated twice as many stations with precipitation as were observed. ET determines overall skill when simulating precipitation, and includes a correction term which reduces the effect of a correct precipitation simulation by chance. An ET score of 0 indicates the same accuracy as a random precipitation simulation, and positive ET scores indicate some level of accuracy, while a perfect precipitation simulation would have ET equal to 1. Appendix F contains formulas for the above precipitation statistics. We focused primarily on the higher threshold of precipitation due to the flooding consequences of high precipitation rates.

2.3. Simulation Results

2.3.1. WRF 48 hour simulations using 15 km grid resolution

In this section we present the model results for the three storm cases using different cumulus parameterization schemes with a 15 km resolution for the inner grid. The results for the most intense flooding event (storm A, 5-9 June 2005) will be thoroughly presented first, followed with the findings for storm B (16-19 June 2005) and storm C (11-14 July 2010).

2.3.1.1. Rain storm A

Storm A was part of a series of flooding events in Alberta which was analyzed by Ou (2008), who details the synoptic conditions of this storm in great detail; we present a brief synopsis of the event. On 5 June 2005 an upper-air blocking high was stationed over Alberta. With an upper-air trough approaching from the west, a surface low-pressure centre developed over Montana, forming a trough of low pressure extending into Alberta. A secondary low formed in this trough in south-eastern Alberta late on 6 June. This low moved slowly northwestwards on 8 June causing heavy precipitation across southern Alberta (Fig. 2-3a). Ou (2008) found that the heaviest precipitation in southern Alberta fell from 00 UTC 06 June to 12 UTC on 08 June, for which our analysis captured the 48 hours of precipitation from 06 UTC 06 June through 06 UTC 08 June. The intense radar echoes were organized in a precipitation band that approached Alberta from the southwest pushing northeastwards across the province. The Oldman River basin received an average precipitation amount of 107 mm during the 48 hour storm. Heavy

precipitation fell over the foothills of the Rocky Mountains, while lighter precipitation occurred throughout the southern part of Alberta. The southeastern border between Alberta and Saskatchewan had precipitation amounts around 50 mm, considerably smaller than the maximum over the Oldman River basin. The northern part of the domain, above 52°N, received low amounts of precipitation.

The WRF model was used to simulate storm A with five different cumulus parameterization schemes (EX, KF, BMJ, GD, G3D). Figure 2-3 illustrates the simulated 48 hour rainfall amounts with the observed precipitation (Fig. 2-3a) during the same time period. All schemes were skillful in reproducing the major features of the spatial distribution of precipitation, yet there were differences in the spatial amounts of precipitation between simulations. The differences between the simulations and the observed precipitation indicate that the cumulus schemes can give very different amounts of precipitation compared to one another. The EX scheme (Fig. 2-3b) produced the maximum precipitation furthest south compared against the other two cumulus schemes in Figure 2-3, as well as a distinct secondary location of precipitation near the southeastern border with Saskatchewan. However, the EX scheme placed the zone of maximum precipitation (136 mm) incorrectly inside the river basin to the north, the Bow River, 75 km from the observed maximum. The KF scheme (Fig. 2-3c) produced more widespread precipitation than the EX, with precipitation further north as well as heavier precipitation in the southeastern part of Alberta. The zone of maximum precipitation (140 mm) when simulated using the KF scheme is noticeably north compared to EX, inside the Red Deer River basin, 170 km from the observed maximum. The BMJ scheme (Fig. 2-3d) simulated more intense precipitation along the foothills, with a larger maximum precipitation amount of 163 mm. However, the location was inside the Red Deer River basin, 170 km from the observed maximum. These three cumulus schemes produced most of their precipitation within a similar range of latitudes, between 50°N and 53°N, and all schemes produced the precipitation further north when compared to the observed precipitation. However, the simulations must be compared against the observed precipitation (Fig. 2-3a) which had a maximum precipitation amount of 224 mm within the Oldman River basin. While the BMJ scheme best simulated the maximum precipitation amount, the EX scheme simulated the location of maximum precipitation better than other schemes. It is very difficult to visually determine which cumulus scheme produced the most accurate precipitation when evaluated across the entire domain of study, so we present the precipitation statistics next.

Each cumulus parameterization scheme simulated different amounts of precipitation for the domain of study during storm A (Fig. 2-4). The BMJ and KF schemes simulated more domain averaged precipitation than what was observed, with BMJ simulating 122% of the observed volume and KF simulating 132%. The KF scheme has been noted to over-simulate precipitation amounts from an overactive triggering of the cumulus scheme (e.g. Colle et al. 2003; Gochis et al. 2003; Liang et al. 2004; Wang and Seaman 1997). The EX scheme simulated the closest amount of precipitation at 101% of the observed amount. GD simulated 106% while G3D simulated 109%. The river basins with low amounts of observed precipitation, such as the Athabasca (4 mm) and North Saskatchewan received more simulated precipitation than was observed. When assessed over the entire domain, all cumulus schemes either properly simulated the amount of precipitation, or simulated more precipitation than was observed.

The domain average analysis (Table 2-2) shows that EX had the lowest RMSE at 29 mm with KF the next closest at 32 mm. BMJ had the highest error with 39 mm. KF simulated 32% more precipitation than was observed, yet had lower RMSE values than some cumulus schemes which better simulated the total amount of precipitation. The KF scheme was able to simulate precipitation more accurately compared to some of the other cumulus schemes in order to accomplish this, though the EX scheme was more accurate than KF when analyzed across the grid points.

The Oldman River received the most precipitation during this storm, at 107 mm when averaged across all grid points inside the river basin. While simulating only 58% of the observed basin precipitation (Table 2-3), KF simulated more precipitation than any other cumulus scheme for this river basin. The next closest was EX at 42%. KF had the lowest RMSE value (60 mm), with EX the next closest at 73 mm. GD, G3D, and BMJ performed similarly for this river basin, with 30%, 29%, and 29% respective simulated precipitation when compared to observed. Their RMSE values were also similar, with GD with 83 mm, G3D with 84 mm, and BMJ with 88 mm. The KF cumulus parameterization scheme simulated this river basin more accurately than the remaining cumulus schemes, by having the lowest RMSE value.

Table 2-4 shows the results from the observation point analysis for each cumulus scheme with the three different precipitation accumulation thresholds. The table shows similar results as the grid point analysis above. Due to KF and BMJ both simulating more precipitation than observed, both of these schemes would be expected to have a higher POD and BIAS values than the other schemes, which is observed in Table 2-4. It would

also be expected that KF and BMJ would have a higher FAR value than other schemes, due to simulating more precipitation than being observed, and thus falsely simulating more precipitation as well. This was the case for BMJ, but not for KF. KF was able to maintain a low FAR, at all thresholds, while simulating 32% more precipitation than was observed. Though KF had high POD (0.77) and low FAR (0.41) at the highest precipitation threshold, the EX scheme had slightly higher ET scores across the 25 mm (0.43) and 50 mm (0.39) threshold analysis than KF (0.39, 0.38), thus EX was slightly more accurate when simulating the precipitation from this storm when analyzed across the observation points. The BMJ, GD, and G3D schemes all had similar ET scores at 0.23, 0.21, and 0.23 respectively, at the highest precipitation threshold. The observation point analysis using POD, FAR, BIAS, and ET shows that EX and KF were able to simulate this storm more accurately than the other cumulus schemes.

Table 2-4 illustrates some general results when precipitation is analyzed using different precipitation thresholds. ET scores were generally higher at the 25 mm threshold than at both the 10 mm and 50 mm thresholds for all schemes. This is consistent with other studies (e.g. Cherubini et al. 2002), though Colle et al. (1999 and 2000) had the highest ET values at the lowest threshold when simulating precipitation over the Pacific Northwest. In addition, cumulus schemes generally had lower POD and higher FAR at higher thresholds which lead to lower ET scores. Other studies have found that numerical weather prediction models have higher FAR and lower POD and ET scores at higher precipitation thresholds (e.g. McBride and Ebert 2000)

2.3.1.2. Rain storm B

Storm B followed a common pattern for heavy rainfall over Alberta. A cutoff cold low supported a well developed surface low pressure centre. The vertically stacked system slowly moved northwards from Montana into southern Alberta. During the early stage, the system was quite convective and contained lightning, hail and squall lines across southern Alberta (Ou 2008). This storm produced an observed 48 hour maximum rainfall accumulation of 152 mm at Spring bank, about 25 km northwest of Calgary, and it was estimated that an area of about 50000 km² received ≥ 50 mm of rain (Ou 2008).

Storm B was simulated with the WRF model using five different cumulus parameterization schemes at a 15 km grid resolution. Figure 2-5 illustrates the simulated 48 hour precipitation field with the observed precipitation (Fig. 2-5a). The most intense rainfall from storm B occurred over the Red Deer River basin, with a basin-averaged

amount of 56 mm. This is about half the maximum rainfall observed for storm A, which averaged 107 mm across the Oldman River basin, though the Red Deer River basin is approximately twice as large as the Oldman River basin. Storm B had a larger amount of precipitation over the entire domain (36 mm) than storm A (30 mm), for the same length of time, 48 hours.

As for storm A, with the exception of the EX scheme (97% of observed), each cumulus parameterization scheme simulated storm B with more precipitation than was observed. Figure 2-6 shows the differences between each cumulus scheme and the river basin receiving the precipitation, as well as the RMSE values. For storm B, the EX and BMJ schemes had the largest domain averaged RMSE, at 30 mm and 25 mm (Table 2-2). The RMSE values were generally lower than for storm A. The EX scheme correctly simulated the amount of precipitation but was not able to simulate the precipitation in the correct location, gaining the highest (i.e. poorest) RMSE values, and was the least accurate scheme to simulate storm B when analyzed across the grid points. BMJ, which simulated 29% more precipitation than was observed, was slightly more accurate than EX. KF, G3D, and GD all performed similarly to one another; GD had the lowest RMSE value (21 mm), while G3D simulated a more correct amount of precipitation between these three schemes. The RMSE values for KF (23 mm) and G3D (22 mm) were close to each other in value, and lower than EX and BMJ. KF simulated the most precipitation of these three schemes, but had errors approximately the same as G3D. The GD scheme simulated this storm the most accurately when analyzed using the grid point method over the entire domain, with the lowest RMSE and a relatively correct amount of precipitation (119%) compared against the observed amount.

The Red Deer River basin received the greatest amount of precipitation during storm B (Table 2-3). GD closely simulated the amount of observed precipitation at 102% and had the lowest RMSE for this basin (22 mm). EX has the highest RMSE (34 mm), further indicating EX simulating this storm with lower accuracy than the other schemes. KF simulated this river basin with the second lowest RMSE (25 mm), with G3D close behind (26 mm). The BMJ scheme over simulated (123%) the precipitation for this basin, and had a high RMSE of 29 mm. GD performed the most accurately at the river basin which received the most precipitation, with the lowest RMSE and most accurate amount of precipitation when compared to the observed amount.

For storm B, the observation point precipitation analysis shows that KF and GD simulated this storm the most accurately. Table 2-5 shows the results for each scheme, at

the highest precipitation threshold. KF had the highest POD at the 50 mm threshold (0.92), and also the lowest FAR compared to the other schemes at 0.49. Even with a high BIAS of 1.81, KF had the highest ET score of 0.37 and was the most accurate scheme when evaluated using the observation point analysis. GD was the second most accurate scheme, with an ET of 0.33. EX simulated the storm with the least accuracy among the five cumulus schemes, with a lower POD (0.54) than FAR (0.59), which led to the lowest ET score for this storm (0.19). BMJ had a relatively average POD (0.81) and high FAR (0.58) the highest BIAS (1.92) which led to a low ET of 0.25. G3D performed similarly to BMJ, with lower POD but more accurate BIAS.

2.3.1.3. Rain storm C

Of the three storms in this study, storm C was the most convective in nature. We present a brief synopsis of the storm; for an in-depth description of this event we refer to Smith 2011. The upper air analysis of storm C showed diffluence at 250 and 500 mb, implying upwards movement of air. There is an exit region of the 250 mb jet in southern Alberta, further showing ascent of air. A cold front aligned northeast to southwest produced numerous thunderstorms across central and southern Alberta, producing significant damage. The Strathmore Radar recorded radar reflectivity values above 55 dBZ passing over the metropolitan city of Calgary, which indicates heavy precipitation with large hail. The observed hails sizes were up to 4 cm in diameter. These large hail stones produced damage to structures, vehicles, trees, and crops. The total insurance claims by Calgarians were \$400 million (Phillips 2010). This storm also produced heavy precipitation over the North Saskatchewan River basin, with an average of 47 mm of rain. While this amount is far less than the precipitation which storms A and B produced for the river basin with the heaviest precipitation, the North Saskatchewan River basin was the largest basin we studied, and a lower precipitation value would be expected when sampled over a much larger area.

Storm C was simulated with the WRF model using five different cumulus parameterization schemes and a grid resolution of 15 km. Figure 2-7 illustrates the simulated 48 hour precipitation field with the observed precipitation (Fig. 2-7a). Storm C had the largest differences in precipitation when simulated using the five cumulus schemes compared to storm A and B. The large differences in precipitation by these cumulus parameterization schemes may have been caused by the triggering mechanism for each scheme when stabilizing the atmosphere. This storm was

convectively unstable, with nearly 400 J/kg of Convective Available Potential Energy determined from the 1200 UTC 12 July 2010 sounding from Stony Plain near Edmonton. Since each cumulus parameterization scheme uses different triggering mechanism as well as closure assumptions, environments which favor convective days may be simulated very differently by different cumulus parameterization schemes.

From the domain average values (Fig. 2-8), both the EX and KF schemes simulated the precipitation close to the observed amount (97% and 109%). BMJ over-simulated with 115% of the observed precipitation amount when averaged over the entire domain (Table 2-2), while GD and G3D both simulated less than observed, at 73% and 78%. EX averaged the lowest RMSE of 20 mm, while the other four cumulus schemes averaged 23 mm. Storm C had a similar amount of precipitation when averaged across all grid points inside the domain (31 mm) as storm A (30 mm), while storm C had lower RMSE values when analyzed across the domain. The EX scheme simulated this storm the most accurately when analyzed using the grid point analysis across the domain of study, with the KF and BMJ scheme slightly less accurate.

The North Saskatchewan river basin received the most precipitation during this storm, with an average of 47 mm of precipitation across all grid points inside the river basin (Table 2-3). EX (KF) simulated 100% (102%) of the observed precipitation, and had a RMSE value of 24 mm (26 mm). The other three schemes simulated the heaviest precipitation river basin with similar RMSE values between 29 mm and 30 mm. However, the BMJ scheme simulated more precipitation than was observed (118%) while GD and G3D both simulated much less precipitation (61% and 67%). With low RMSE and accurate simulation of the accumulated precipitation, the North Saskatchewan River basin was simulated the most accurately using the EX scheme, with the KF having slightly less accuracy using the grid point analysis.

Storm C was analyzed for accuracy by using the observation point analysis for a precipitation threshold of 50 mm (Table 2-5). The BMJ scheme had the greatest POD value (0.57), with KF and EX the next closest at 0.43. Even while simulating more precipitation than observed, the BMJ scheme managed a relatively low FAR (0.50) compared against the other cumulus scheme simulations of this storm. With the greatest POD and relatively low FAR, BMJ had the highest ET score for this storm (0.24), with EX and KF slightly less accurate (0.23 and 0.20). The GD and G3D had very low BIAS scores, indicating very little simulated precipitation, which resulted in low POD and high

FAR scores. The observation point analysis shows that the EX, BMJ, and KF scheme all simulated this storm with similar skill.

Numerical simulations of storm C revealed some weakness in capturing the spatial distribution of the observed rainfall. The POD values for storm C were relatively low in comparison to storm A and storm B. This led to cumulus schemes generally having lower ET scores for storm C than for A and B. In addition, the GD and G3D schemes simulated this storm poorly, simulating much less precipitation than was observed indicated from small BIAS values. The lack of precipitation resulted in ET scores only marginally above zero for the GD and G3D schemes. While it has been noted that the Grell cumulus parameterization scheme has skillfully simulated heavy precipitation (e.g. Kerkhoven et al. 2006; Yang and Tung 2003), the Grell scheme has also been noted to underperform (e.g. Ratnam and Kumar 2005), often simulating a drier and colder atmosphere than observed (Gochis et al. 2000).

2.3.2. WRF 48 hour forecasts using 30 km grid resolution

In the previous section, different cumulus parameterization schemes were compared using a numerical grid resolution of 15 km. It is of interest to determine whether simulations performed using different grid resolutions in result with similar findings. This section presents our findings for storm A when simulated using the course grid resolution of 30 km.

Figure 2-9 shows the spatial distribution of the 48 hour simulated precipitation amounts for storm A using the KF scheme at different grid resolutions. The 30 km resolution (Fig. 2-9d) had skill when simulating this storm, though comparing with the observations (Fig. 2-9a) it is clear that the 30 km simulation is lacking areas of heavy precipitation. Figure 2-10a shows the percentage of simulated precipitation compared to observed precipitation. Comparing this figure, with storm A15 on Figure 2-4a, river basins which were simulated at 15 km grid resolution with a high percentage of simulated precipitation, such as the Athabasca and North Saskatchewan River basin, were simulated with even more of a difference at a resolution of 30 km. At 30 km resolution, each cumulus scheme simulated more precipitation across the domain than was observed. The values range from the EX scheme with the closest value at 116% of observed, to KF with 137% of observed (Table 2-6). No scheme showed improvement in simulating the total amount of precipitation over the river basins when simulated with a coarse resolution of 30 km compared to a grid resolution of 15 km.

The A30-EX simulation had a lower RMSE value (Table 2-6) than the A15-EX simulation (Table 2-2) when averaged across the domain, while the remaining schemes had higher errors for the lower resolution simulation. KF and BMJ had marginally higher errors at 30 km resolution than for 15 km resolution, but were lower in the magnitude of difference than GD and G3D. Generally, the cumulus schemes are more accurate at the grid resolution of 15 km rather than 30 km when evaluated using the grid-point analysis. In addition, the 15 km simulations better simulated the domain averaged precipitation amounts, whereas the 30 km simulations all over-simulated the precipitation.

The Oldman River basin had the heaviest precipitation for this storm, and EX had simulated this basin with high accuracy at 15 km resolution. When analyzed across the Oldman River basin, the EX and BMJ schemes were slightly more accurate when simulated at a resolution of 30 km, with lower RMSE values (Table 2-6), than for a resolution of 15 km (Table 2-3). However the remaining three cumulus parameterization schemes (KF, GD, G3D) had less accuracy when simulated at 30 km resolution across the Oldman River basin.

There were differences between the POD, FAR, BIAS, and ET when simulating storm A at 15 km (Table 2-5) and at 30 km resolution (Table 2-6). The EX scheme was the only simulation at 30 km resolution to have a higher POD (0.73) than for the same simulation using 15 km grid resolution (0.63). This may have been due to the larger BIAS value for the 30 km simulation (1.37) than at 15 km (0.93). An increase in BIAS indicates more precipitation which could increase the POD value if the precipitation is simulated where it is also observed. However, other schemes experienced a larger BIAS at 30 km resolution without increasing the POD; G3D had a BIAS of 0.87 at 15 km resolution and a BIAS of 1.43 at 30 km resolution, though experienced a decrease in POD. The higher BIAS values also caused much greater FAR values at the grid resolution of 30 km than at 15 km for all schemes. Generally, the 30 km grid resolution simulations had lower POD and ET, and higher BIAS, FAR, and RMSE than the same simulation using a grid resolution of 15 km. All schemes showed a decrease in accuracy when simulating this storm at the coarse resolution.

2.3.3. WRF 48 hour forecasts using 6 km grid resolution

2.3.3.1. Rain storm A

Storm A was simulated at different resolutions (Fig. 2-9) to determine how grid resolution can affect the output precipitation. The EX and KF schemes were simulated at a grid resolution of 6 km and compared against the same storm event at 15 km resolution. While WRF simulated this storm at 6 km resolution successfully (Fig. 2-9b), it was not known if a refinement of resolution results in more accurate precipitation coverage. The EX scheme had a much larger change in simulated precipitation than the KF scheme when the grid resolution was changed from 15 km to 6 km. A6-EX (Table 2-7) simulated 118% of the observed precipitation, while A15-EX (Table 2-2) simulated 101%. A6-KF simulated 133% of the observed precipitation, while A15-KF simulated 132%. For storm A, increasing the resolution resulted with a simulation with more precipitation, and the EX scheme was more accurate simulating the total precipitation across the grid points inside the domain than the KF scheme.

The high resolution 6 km simulations had higher errors than the 15 km simulations. The domain averaged RMSE for A6-EX was 35 mm (Table 2-7), compared to A15-EX with 29 mm (Table 2-2). The KF simulations show similar behavior. Higher resolution simulations often suffer from larger errors when evaluated at the individual grid points (Mass et al. 2002). This is because higher resolution corresponds to smaller grid points, and small displacement differences in the precipitation field will yield a larger error than compared to a coarser resolution.

For the observation point analysis, A6-EX had lower POD at 6 km (0.57, Table 2-7) than at 15 km resolution (0.63, Table 2-5) and higher FAR at 6 km (0.47) than 15 km (0.32). This led to a lower ET score at 6 km resolution (0.26) than 15 km (0.39). The KF simulation experienced similar changes between the resolutions, though was less accurate than EX, with higher RMSE, FAR, and BIAS. While KF also had higher POD than EX, the KF scheme had a lower ET value at 6 km resolution. This storm was simulated with less accuracy at 6 km resolution for both schemes than at 15 km resolution, though EX simulated this storm with higher accuracy than KF using 6 km grid resolution. Among the three different grid resolutions and five cumulus parameterization schemes, this storm was best simulated at 15 km resolution by using the EX scheme, with the KF scheme being slightly less accurate at a 15 km resolution.

2.3.3.2. Rain storm B

Storm B was simulated at 6 km grid resolution using the KF and EX cumulus parameterization scheme. Figure 2-11 illustrates the observed precipitation amount as well as when simulated using the KF scheme. Both schemes successfully simulated this storm, though it is difficult to visually determine if the storm was more accurately simulated at 6 km resolution than at 15 km resolution (Fig. 2-5).

B6-EX (B15-EX) simulated 103% (97%) of the total observed precipitation that occurred over the entire domain, while B6-KF (B15-KF) simulated 110% (122%), as shown in Table 2-7 (Table 2-2). EX had RMSE values of 30 mm at 15 km resolution, which was reduced to 25 mm for 6 km. KF showed a much smaller improvement between the two simulations, decreasing from 23 mm at 15 km resolution to 22 mm at 6 km resolution. Storm B, when simulated using a grid resolution of 6 km, resulted with more accuracy for the total precipitation than for a simulation using 15 km resolution when evaluated for the grid points inside the domain.

The EX scheme, at 6 km grid resolution, had a higher POD (0.73, Table 2-7) when compared to the 15 km grid resolution simulation (0.54, Table 2-5). The EX scheme also had lower FAR and higher ET when simulated at the finer resolution. While the KF simulations at 15 km resolution had higher POD values, it also had higher FAR values than was simulated at 6 km resolution. Consequently, the ET score for the KF simulations did not change between the 6 km resolution and 15 km resolution, both resulting with 0.37. The observation point precipitation analysis shows that a simulation of storm B at 6 km grid resolution would be more accurate when using the EX scheme, than at 15 km grid resolution. The KF scheme showed minor improvement in accuracy when the resolution was changed from 15 km to 6 km. While both KF and EX simulated this storm with the same ET score using 6 km grid resolution, the KF scheme was more accurate with a lower RMSE of 22 mm compared to EX with 25 mm, thus having more accuracy.

2.3.3.3. Rain Storm C

Storm C was simulated at 6 km grid resolution using the KF and EX scheme. While each scheme was able to simulate the precipitation (Fig. 2-11), it is difficult visually determine if increased resolution more accurately simulated the precipitation when compared against simulations of 15 km grid resolution (Fig. 2-7). At 6 km (15 km) resolution, the EX scheme simulated 115% (97%) of the observed precipitation, as shown

in Table 2-7 (Table 2-2). Simulating storm C at 6 km using the EX scheme resulted with a less accurate total precipitation when evaluated across the entire basin. The KF scheme showed a slight improvement when simulated at higher resolution, simulating 99% (109%) at 6 km (15 km) grid resolution. However, there were higher errors for the higher resolution simulations. C6-EX had a RMSE of 25 mm, compared to C15-EX with RMSE of 20 mm. KF had similar results, with a RMSE value of 26 mm at 6 km resolution and 23 mm at 15 km resolution.

For the highest precipitation threshold, at 6 km grid resolution, both EX and KF had nearly identical ET scores, at 0.27, which were higher than the ET score for the same storm when simulated at 15 km resolution (Table 2-5). When compared against the 15 km simulations, the EX scheme had higher POD and FAR values at 6 km resolution, whereas KF had higher POD and lower FAR at the higher resolution. While both schemes had higher ET score at the 6 km grid resolution, KF had a larger increase than EX when compared to the ET score of the 15 km simulations. Both schemes simulated storm C with very similar levels of accuracy at 6 km simulations, with nearly identical ET and RMSE values. The observation point and grid point analysis show that the EX scheme was slightly more accurate than KF at the higher resolution.

2.4. Discussion and Conclusions

The WRF model was used to simulate three heavy rainfall events over southern Alberta using 5 different cumulus parameterization schemes and three different grid resolution. The model results show that using a grid resolution of 15 km provides a compromise between computational efficiency and accurate resolution of the observed precipitation field. Refining the grid resolution from 15 km to 6 km drastically increases the computation time, while the accuracy of the precipitation distribution is only slightly better than simulations of 15 km grid resolution, with the difference in accuracy depending on the cumulus scheme. The EX scheme showed a slight improvement in accuracy for higher resolution simulations, while the KF showed less improvement.

Our finding was that for heavy rainfall events it is best to either use the Kain-Fritsch (KF) cumulus parameterization scheme, or alternatively, use the explicit scheme which does not parameterize convection. The KF scheme and the explicit scheme consistently had low errors, high probability of detection and equitable threat. GD and G3D simulated very similar to each other and had less skill than KF and EX for these three storms. BMJ simulations were also less accurate than using the KF and EX.

When comparing the 6 km resolution simulations against 15 km, some interesting findings emerged. Storm A was better simulated using the EX or KF scheme at 15 km resolution rather than at 6 km. Storm B was simulated more accurately using the EX scheme at 6 km resolution than at 15 km, while the KF scheme only showed a marginal increase in accuracy when simulated at higher resolution. Storm C, when simulated using the KF scheme at 6 km resolution, had more accuracy than at 15 km resolution. The EX scheme also showed some improvement for this storm, but not as much improvement as KF. However, the EX scheme appeared to be more sensitive to changes in horizontal resolutions for the other storms than the KF scheme, which did not exhibit as many changes between the 6 km resolution simulations and those at 15 km.

Of the true cumulus parameterization schemes (non-Explicit), the KF cumulus parameterization scheme best simulated the heavy precipitation events across Alberta during the summer season. The strength of the KF scheme has been observed in other work (e.g. Gochis et al. 2002). The KF scheme is thought to simulate convective precipitation more accurately because the scheme conserves mass while using the parameterization of convective downdrafts (Gochis et al. 2002) as well as using CAPE as part of the closure assumptions (Wang and Seaman 1997).

While the results indicate that these three heavy precipitation storms were generally better simulated when using the finer grid resolution of 6 km, the increase in accuracy was marginal. An objective of this project was to determine whether finer resolution would result in higher accuracy for precipitation events. However, the user of the model needs to consider the additional resources required to process a higher resolution simulation. The 15 km resolution simulations required slightly longer than 1 hour of computational time to perform a simulation on a new computer from 2010, from start to finish. The computer was a 32-bit machine, using 4 CPU's, and 4 GB of ram. The 6 km resolution simulations required a different computer, which used 8 processors and 8 GB of ram, and consumed over 18 hour of computational time to start and finish each storm simulation. While a researcher may prefer a higher resolution simulation that takes more than half a day to perform, an operational flood forecaster may choose the coarser resolution with similar accuracy to give them more time to prepare any warnings that may be required.

2.5. Tables

Table 2-1: WRF simulation initialization and end time for storms A, B, and C, using different cumulus parameterization schemes and grid resolutions for the nested domain. The model simulation (e.x. A6) is the combination of the grid resolution (6 km) for the particular storm (A). Model precipitation was accumulated for the nested domain starting from 06 UTC the day following initialization, and ended 48 hours later.

Model simulation	Initialization time	End time	Grid spacing (km)	Cumulus Parameterization schemes used
A6	12 UTC 5 Jun 2005	9 UTC 8 Jun 2005	6	EX, KF
A15	12 UTC 5 Jun 2005	9 UTC 8 Jun 2005	15	EX, KF, BMJ, GD, G3D
A30	12 UTC 5 Jun 2005	9 UTC 8 Jun 2005	30	EX, KF, BMJ, GD, G3D
B6	12 UTC 16 Jun 2005	9 UTC 19 Jun 2005	6	EX, KF
B15	12 UTC 16 Jun 2005	9 UTC 19 Jun 2005	15	EX, KF, BMJ, GD, G3D
C6	12 UTC 11 Jul 2010	9 UTC 14 Jul 2010	6	EX, KF
C15	12 UTC 11 Jul 2010	9 UTC 14 Jul 2010	15	EX, KF, BMJ, GD, G3D

Table 2-2: Domain averaged verification results for storms A, B, and C, when simulated at 15 km grid resolution. Storm A had an observed domain average precipitation accumulation of 30 mm, Storm B 36 mm, and Storm C 31 mm. Bolded values indicate the two most accurate cumulus options for the given analysis.

Storm	Statistic	EX	BMJ	KF	GD	G3D
Storm A15	Average precipitation (mm)	30	37	39	32	33
	RMSE (mm)	29	39	32	34	34
	Percent simulated of observed (%)	101	122	132	106	109
Storm B15	Average precipitation (mm)	35	47	44	43	42
	RMSE (mm)	30	25	23	21	22
	Percent simulated of observed (%)	97	129	122	119	116
Storm C15	Average precipitation (mm)	30	35	34	22	24
	RMSE (mm)	20	23	23	23	23
	Percent simulated of observed (%)	97	115	109	73	78

Table 2-3: Verification results for Storm A, B, and C, when simulated at 15 km grid resolution, for the river basin which received the most precipitation for the given storm. Storm A was over the Oldman with observed average precipitation of 107 mm, Storm B for Red Deer with 56 mm, and Storm C for North Saskatchewan with 47 mm. Bolded values indicate the two most accurate cumulus option for the given analysis.

Storm	Statistic	EX	BMJ	KF	GD	G3D
Storm A15: Oldman River	Average precipitation (mm)	45	31	61	33	31
	RMSE (mm)	73	88	60	83	84
	Percent simulated of observed (%)	42	29	58	30	29
Storm B15: Red Deer	Average precipitation (mm)	40	69	66	57	49
	RMSE (mm)	34	29	25	22	26
	Percent simulated of observed (%)	72	123	118	102	87
Storm C15: North Saskatchewan	Average precipitation (mm)	47	55	48	28	31
	RMSE (mm)	24	29	26	30	30
	Percent simulated of observed (%)	100	118	102	61	67

Table 2-4: Precipitation scores for storm A when simulated at 15 km grid resolution, for three precipitation thresholds. Bolded values indicate the two most accurate cumulus schemes for the given analysis.

	Threshold (mm)	EX	BMJ	KF	GD	G3D
POD	10	0.90	0.93	0.95	0.86	0.81
	25	0.77	0.80	0.88	0.71	0.73
	50	0.63	0.53	0.77	0.53	0.47
FAR	10	0.27	0.35	0.28	0.31	0.34
	25	0.23	0.35	0.32	0.38	0.39
	50	0.32	0.50	0.41	0.53	0.46
BIAS	10	1.25	1.42	1.32	1.25	1.23
	25	1.00	1.23	1.29	1.14	1.20
	50	0.93	1.07	1.30	1.13	0.87
ET	10	0.34	0.22	0.35	0.26	0.18
	25	0.43	0.32	0.39	0.25	0.24
	50	0.39	0.23	0.38	0.21	0.23

Table 2-5: Precipitation scores for storm A, B, and C, when simulated at 15 km grid resolution, at the highest precipitation threshold of 50 mm. Bolded values indicate the two most accurate cumulus schemes for the given analysis.

Storm	Statistic	EX	BMJ	KF	GD	G3D
A15	POD	0.63	0.53	0.77	0.53	0.47
	FAR	0.32	0.50	0.41	0.53	0.46
	BIAS	0.93	1.07	1.30	1.13	0.87
	ET	0.39	0.23	0.38	0.21	0.23
B15	POD	0.54	0.81	0.92	0.85	0.77
	FAR	0.59	0.58	0.49	0.51	0.58
	BIAS	1.31	1.92	1.81	1.73	1.85
	ET	0.19	0.25	0.37	0.33	0.24
C15	POD	0.43	0.57	0.43	0.03	0.10
	FAR	0.43	0.50	0.50	0.80	0.70
	BIAS	0.77	1.13	0.87	0.17	0.33
	ET	0.23	0.24	0.20	0.00	0.02

Table 2-6: Verification results between different cumulus parameterization schemes when simulating storm A at 30 km grid resolution. POD, FAR, BIAS, and ET values were calculated at a precipitation threshold of 50 mm. Bolded values indicate an increase in accuracy from the same storm simulated at 15 km grid resolution.

Statistic	EX	BMJ	KF	GD	G3D
Domain averaged precipitation accumulation (mm)	33	37	39	36	38
Domain averaged precipitation RMSE (mm)	27	41	37	44	46
Domain averaged percent simulated precipitation of observed	116	129	137	127	133
Oldman River basin precipitation (mm)	60	32	57	20	21
Oldman River basin RMSE (mm)	62	86	63	98	97
Oldman River basin Percent simulated precipitation of observed	54	29	51	18	19
Probability of Detection (POD)	0.73	0.43	0.67	0.33	0.30
False Alarm Ratio (FAR)	0.46	0.66	0.57	0.76	0.79
BIAS	1.37	1.27	1.57	1.37	1.43
Equitable Threat (ET)	0.32	0.10	0.20	0.02	-0.01

Table 2-7: Verification results between the Explicit (EX) and Kain-Fritsch (KF) cumulus parameterization scheme for storm A, B, and C, when simulated at a grid resolution of 6 km. POD, FAR, BIAS, and ET were calculated at a precipitation threshold of 50 mm. The 48 hour domain averaged observed precipitation amounts were 32 mm for Storm A, 39 mm for storm B, and 33 mm for storm C. Bolded values indicate an increase in accuracy from the same storm simulated at 15 km grid resolution.

Statistic	A6EX	A6KF	B6EX	B6KF	C6EX	C6KF
Domain averaged precipitation accumulation (mm)	38	42	40	43	38	33
Domain averaged precipitation RMSE (mm)	35	41	25	22	25	26
Domain averaged percent simulated precipitation of observed	118	133	103	110	115	99
Probability of Detection (POD)	0.57	0.67	0.73	0.85	0.63	0.50
False Alarm Ratio (FAR)	0.47	0.52	0.42	0.46	0.49	0.40
BIAS	1.07	1.40	1.27	1.58	1.23	0.83
Equitable Threat (ET)	0.26	0.24	0.37	0.37	0.27	0.27

2.6. Figures

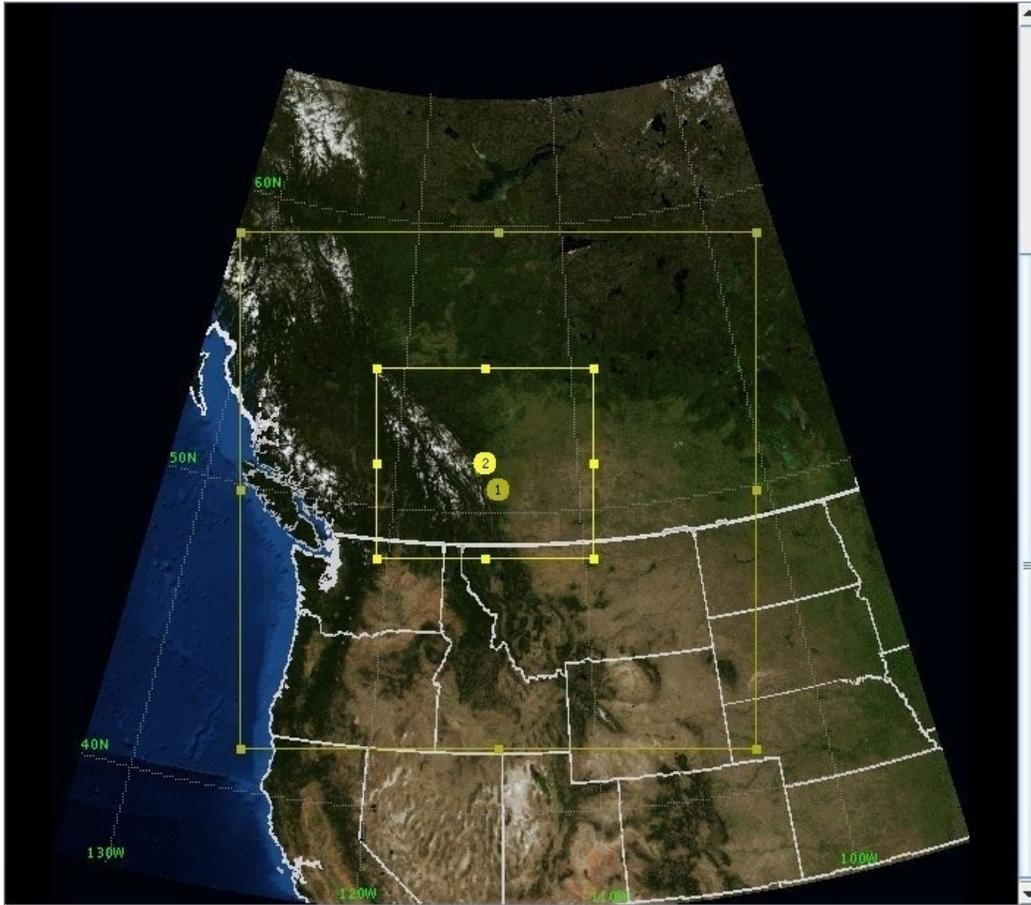


Figure 2-1: Numerical Weather Prediction domain setup using an outer grid resolution of 45 km, and inner grid resolution of 15 km. All precipitation data used for analysis is from Alberta land south of 54° N latitude inside the 15 km grid.

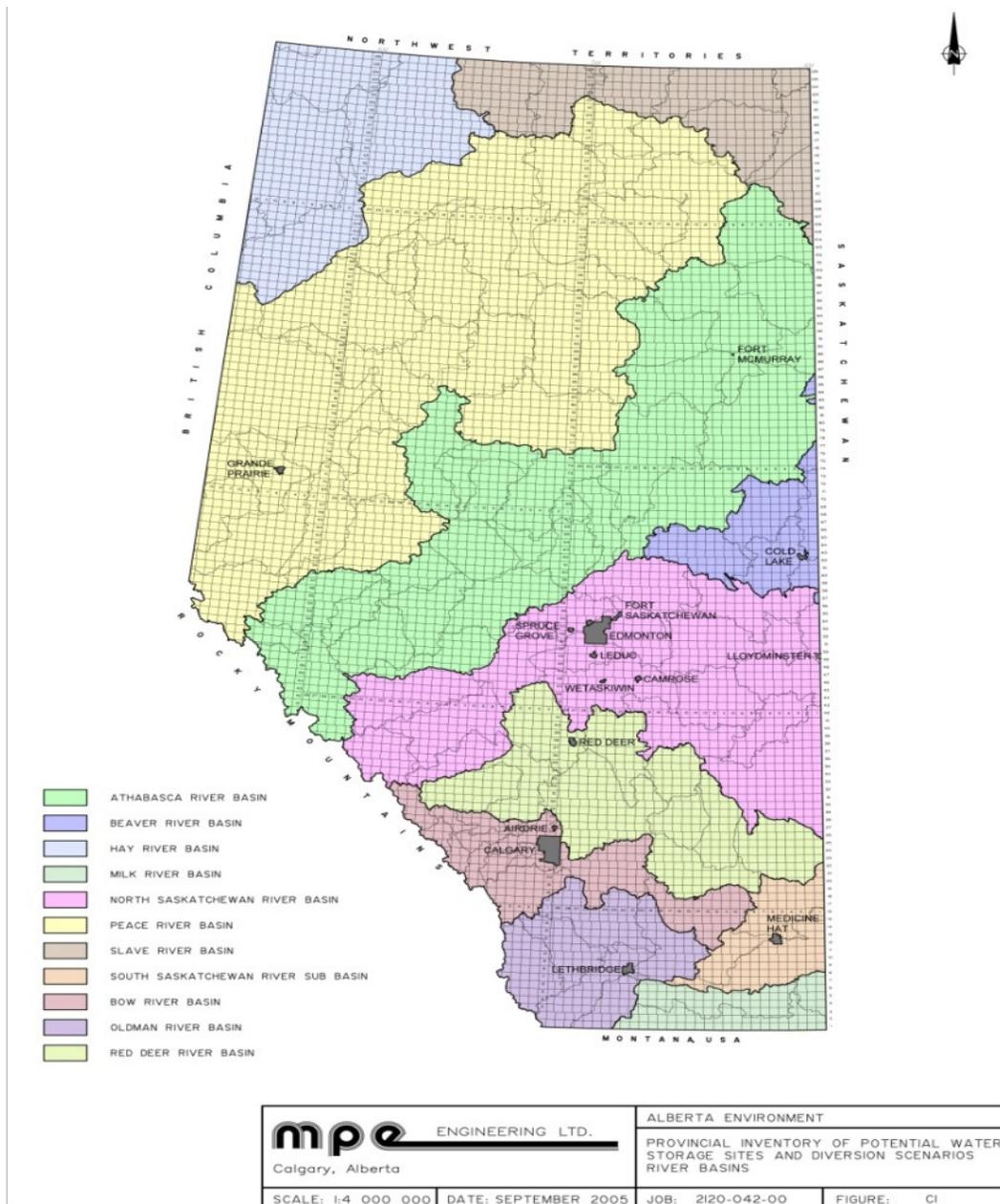


Figure 2-2: Major river basins of Alberta, Canada. The seven used in this study, ordered from the largest area to smallest, are the North Saskatchewan, Red Deer, Bow River, Athabasca, Oldman, South Saskatchewan, and Milk River.

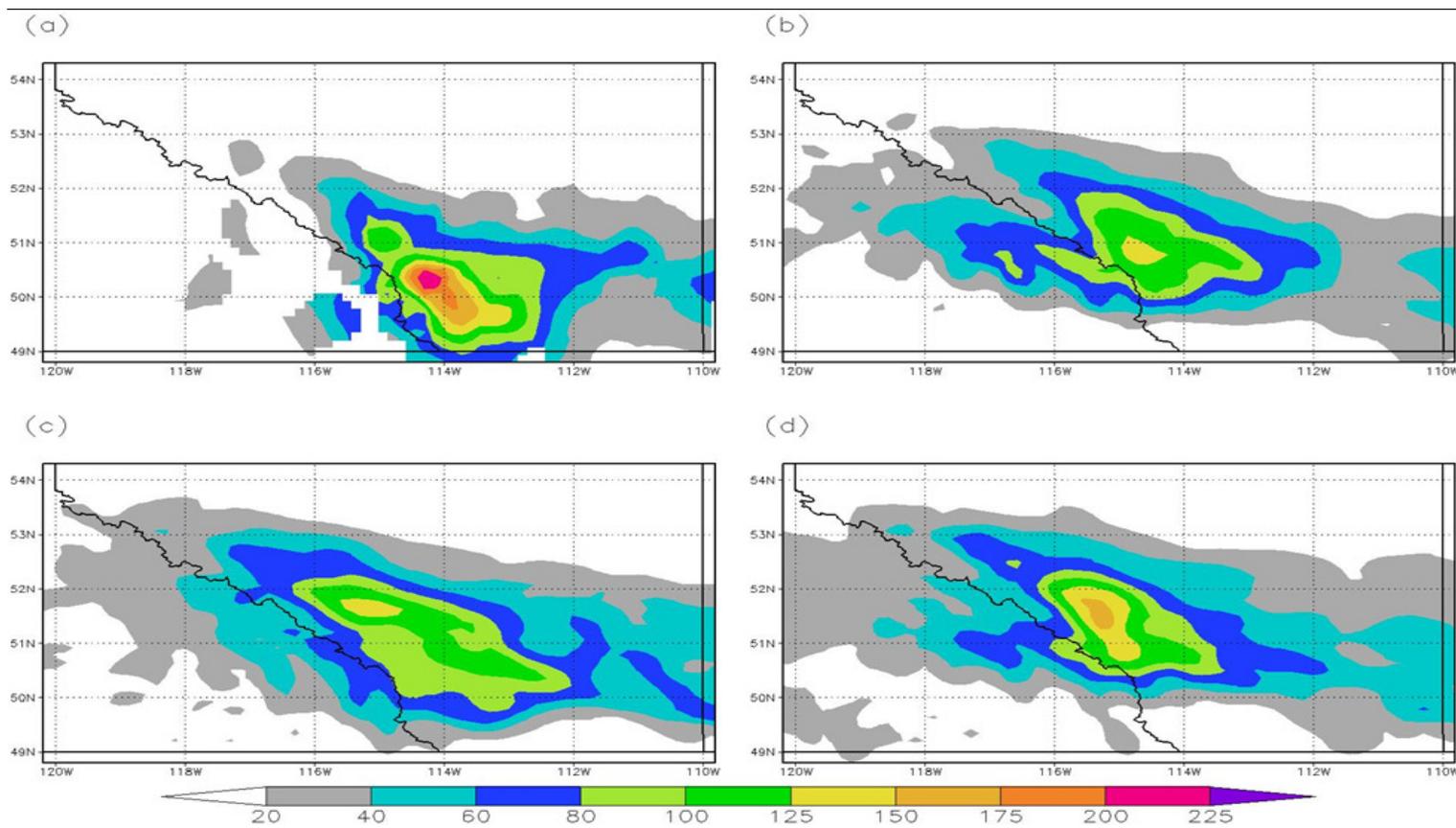


Figure 2-3: 48 hour precipitation amounts for Storm A for (a) the observed precipitation interpolated to 15 km resolution, and the simulations using 15 km grid resolution with the following cumulus parameterization scheme: (b) Explicit scheme, (c) Kain-Fritsch, (d) Betts-Miller-Janjic.

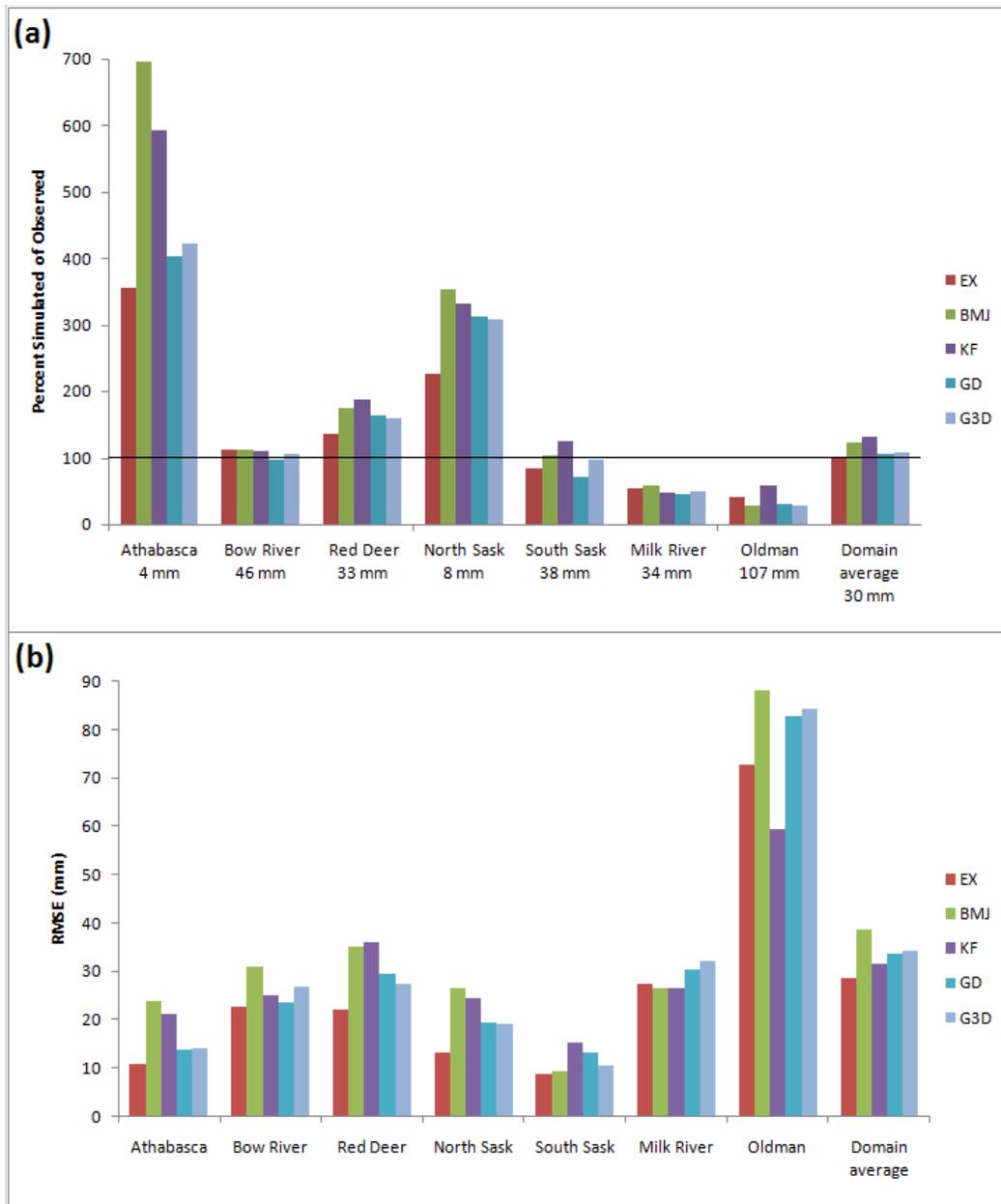


Figure 2-4: Precipitation verification statistics, for storm A, at each river basin, simulated at 15 km grid resolution. Percent simulated precipitation of observed (a) and RMSE values (b) for each cumulus scheme. Observed basin precipitation values are indicated below the basin name in (a).

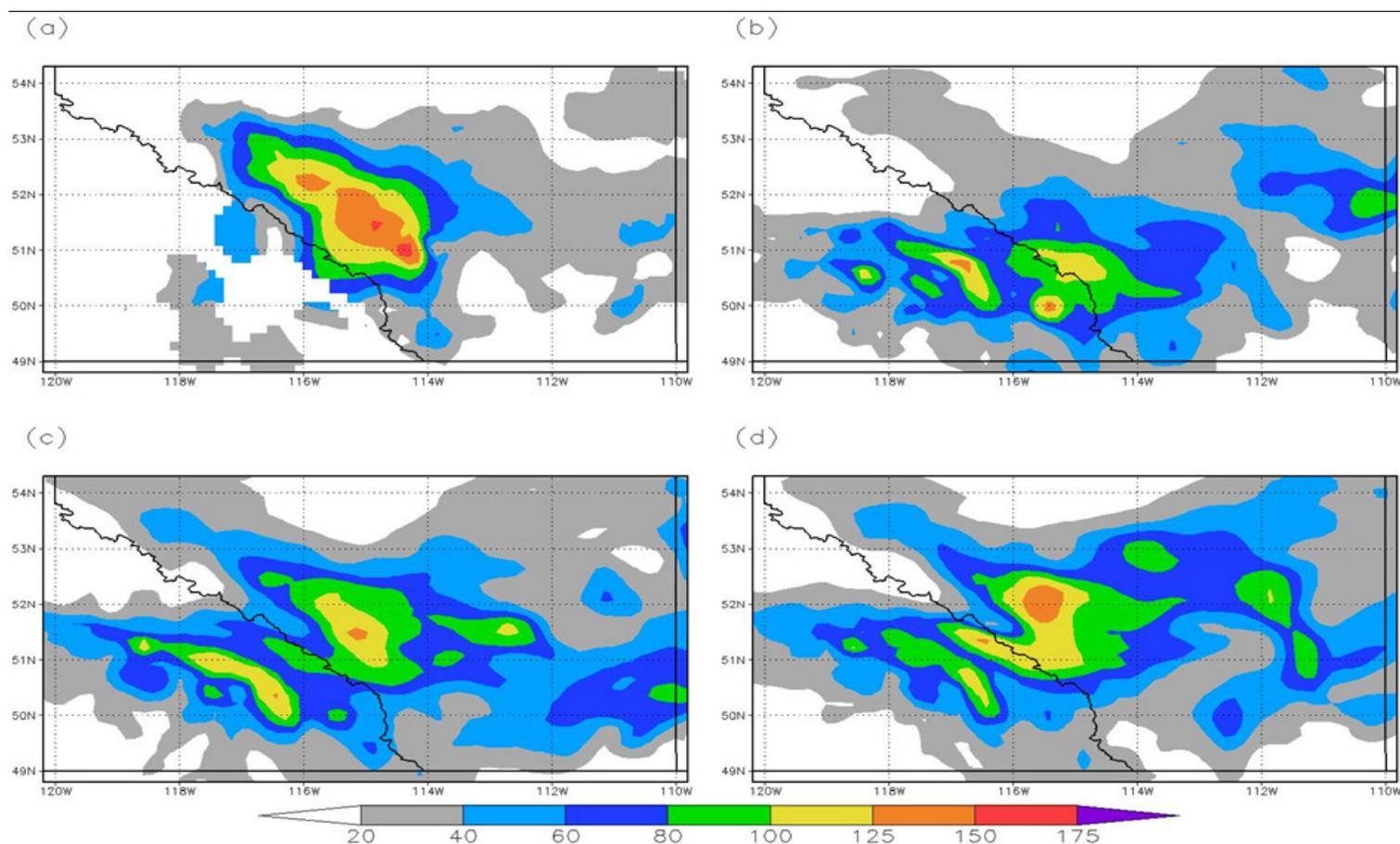


Figure 2-5: 48 hour precipitation amounts for Storm B for (a) the observed precipitation interpolated to 15 km resolution, and the simulations using 15 km grid resolution with the following cumulus parameterization scheme: (b) Explicit scheme, (c) Kain-Fritsch, (d) Betts-Miller-Janjic.

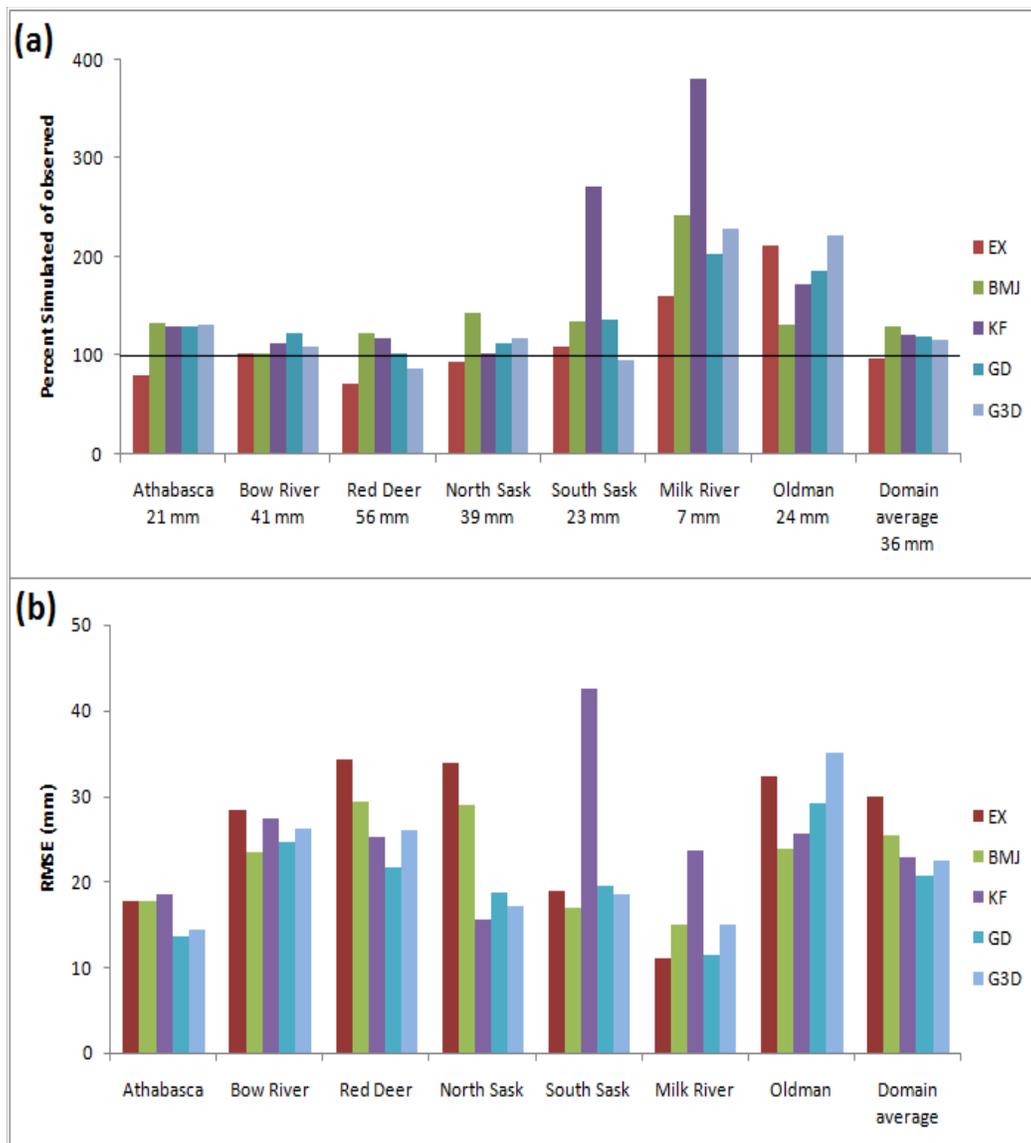


Figure 2-6: Precipitation verification statistics, for storm B, at each river basin, simulated at 15 km grid resolution. Percent simulated precipitation of observed (a) and RMSE values (b) for each cumulus. Observed basin precipitation values are indicated below the basin name in (a).

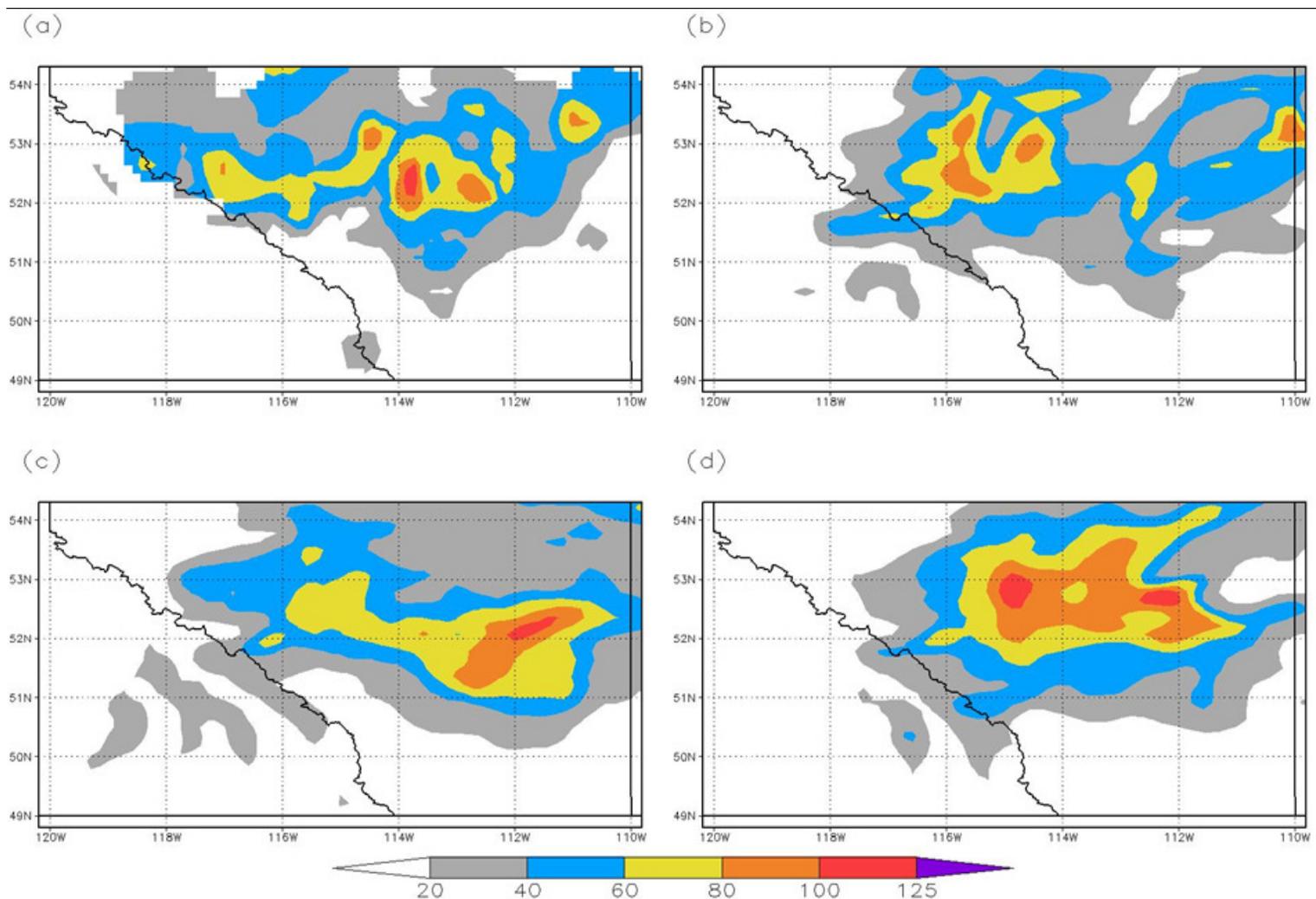


Figure 2-7: 48 hour precipitation amounts for Storm C for (a) the observed precipitation interpolated to 15 km resolution, and the simulations using 15 km grid resolution with the following cumulus parameterization scheme: (b) Explicit scheme, (c) Kain-Fritsch, (d) Betts-Miller-Janjic.

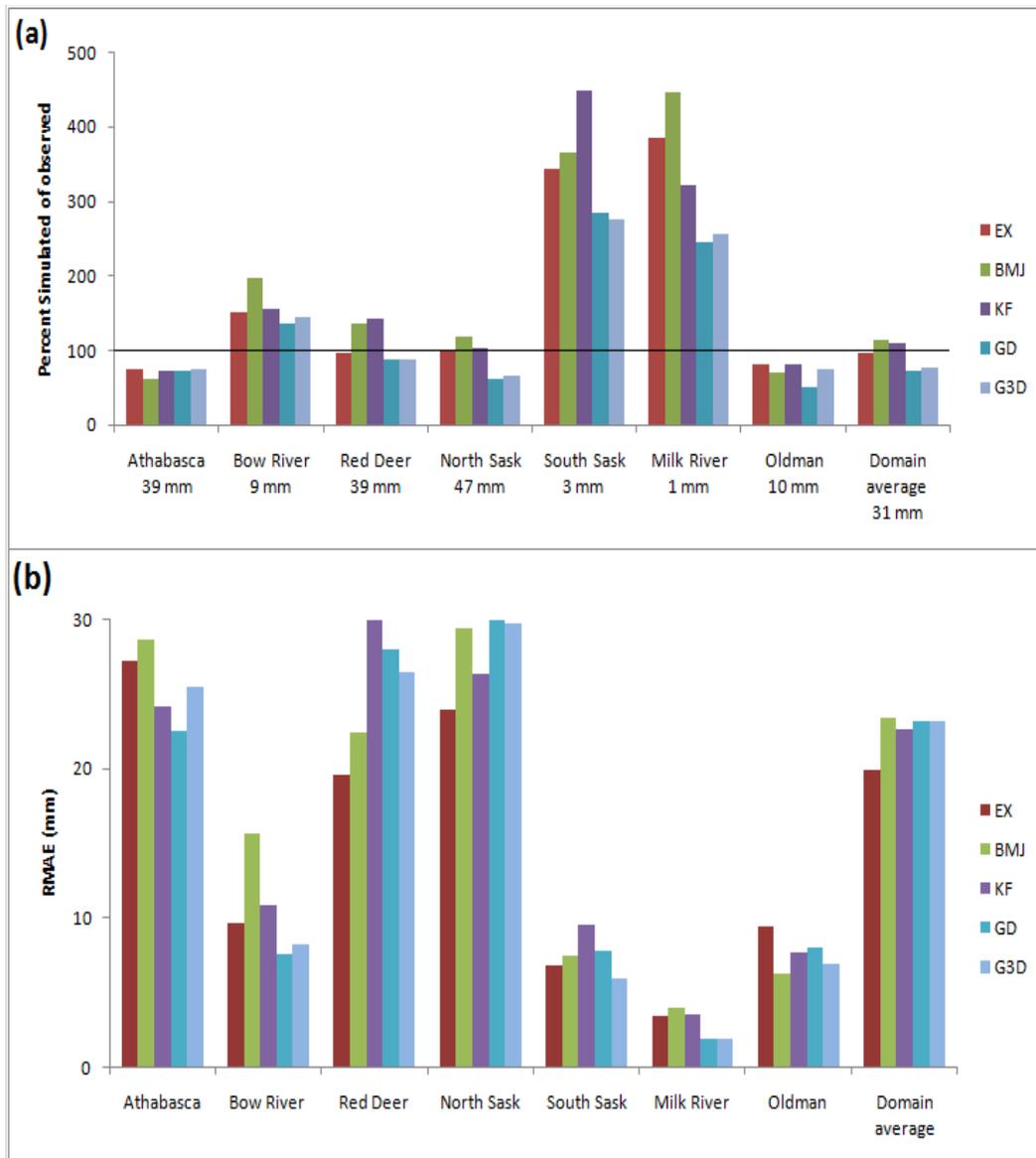


Figure 2-8: Precipitation verification statistics, for storm C, at each river basin, simulated at 15 km grid resolution. Percent simulated precipitation of observed (a) and RMSE values (b) for each cumulus scheme. Observed basin precipitation values are indicated below the basin name in (a).

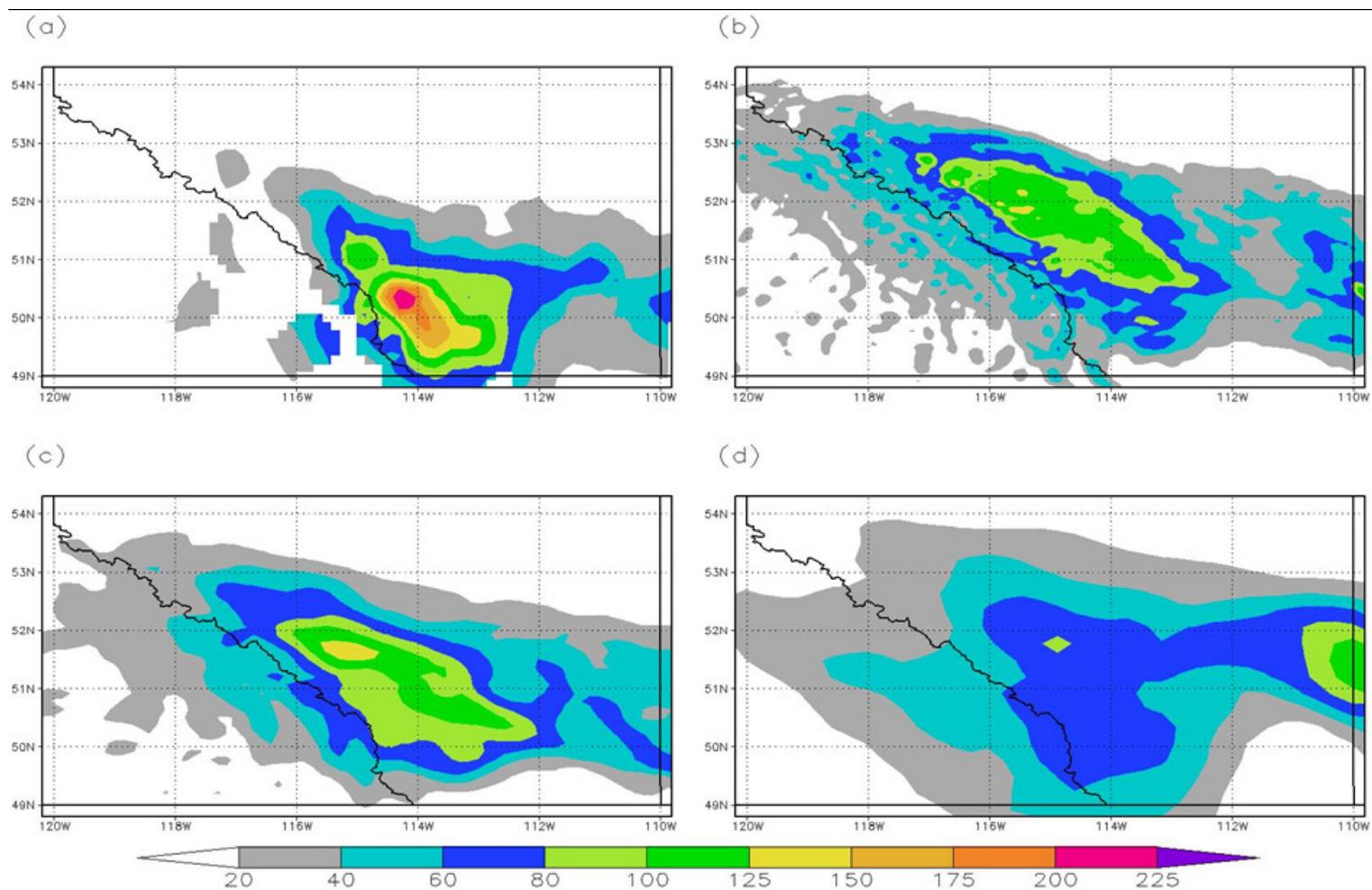


Figure 2-9: 48 hour precipitation amounts for Storm A when simulated at three different grid resolutions. The observed precipitation interpolated to 15 km resolution (a), and simulations using the Kain-Fritsch cumulus parameterization scheme at the following resolutions: (b) 6 km, (c) 15 km, (d) 30 km.

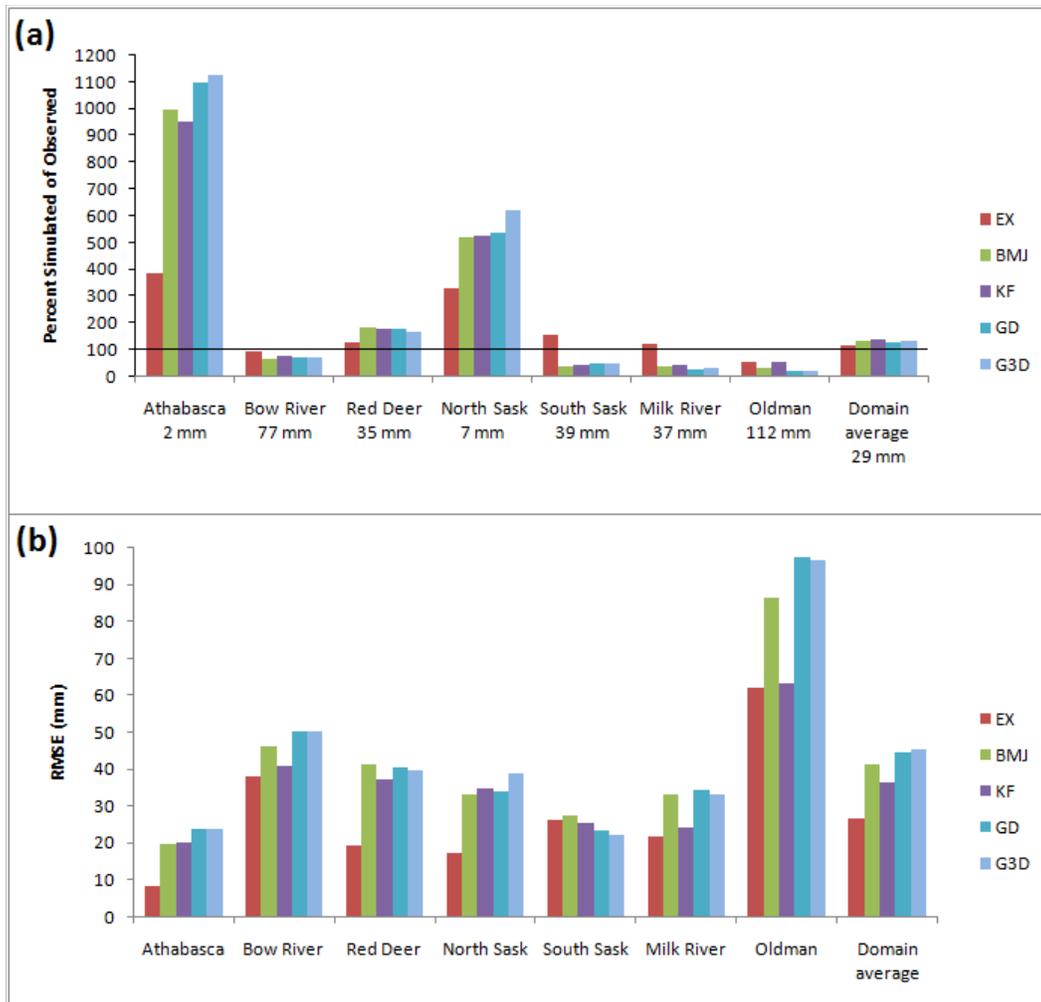


Figure 2-10: Precipitation verification statistics, for storm A, at each river basin, simulated at 30 km grid resolution. Percent simulated precipitation of observed (a) and RMSE values (b) for each cumulus scheme across all river.

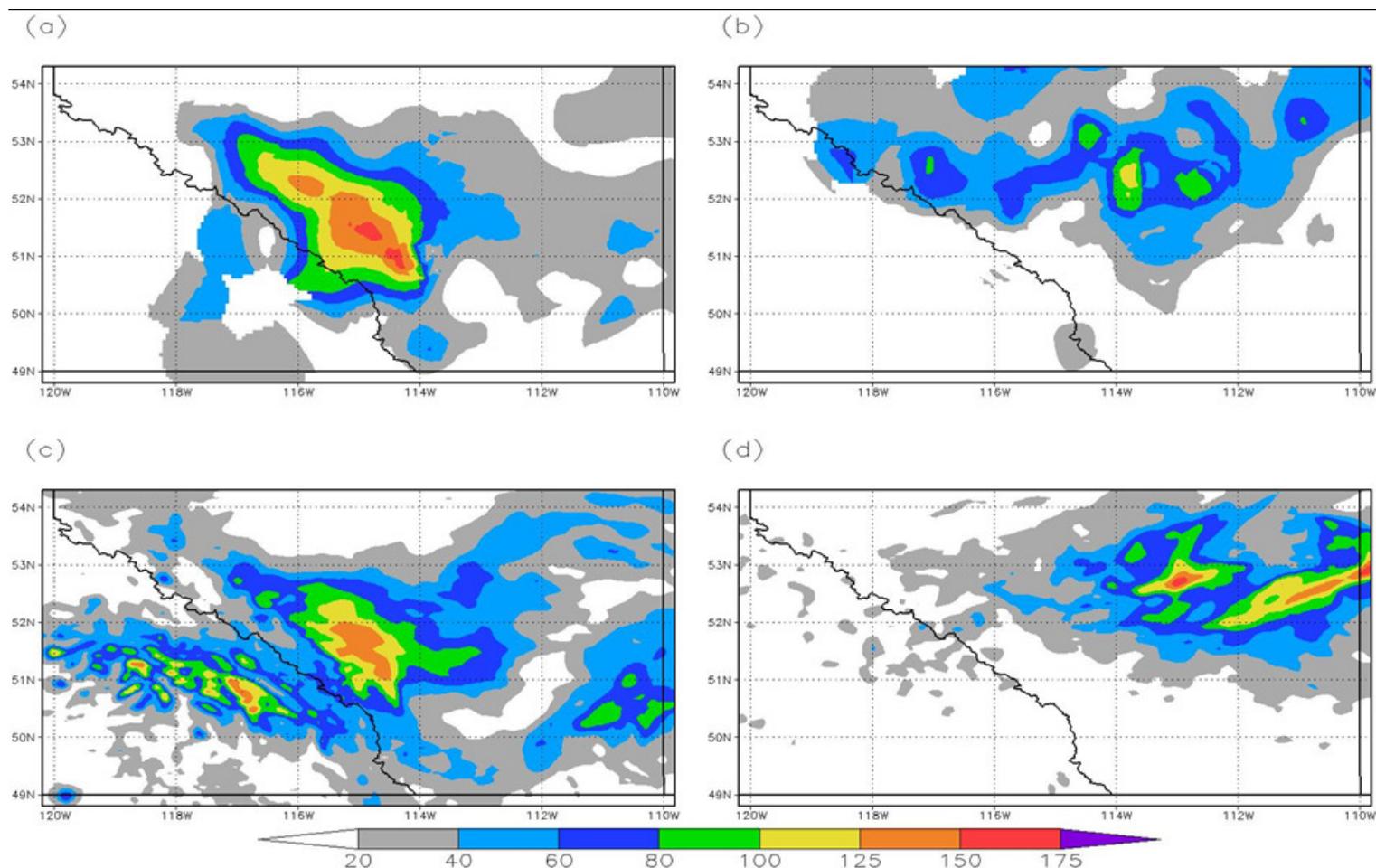


Figure 2-11: 48 hour precipitation amounts for storm B and storm C when simulated using 6 km grid resolution. Observed precipitation for storm B (a) and storm C (b) are above the simulated precipitation amounts when using the Kain-Fritsch cumulus scheme for storm B (c) and storm C (d).

2.7. References

- Arakawa, A. and H.W. Schubert, W.H., 1974: Interaction of a Cumulus Cloud Ensemble with the Large-Scale Environment, Part 1, *J. Atmos. Sci.*, **31**, 674-701
- Betts, A.K. and M.J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets, *Quart. J. R. Met. Soc.*, **112**, 693-709
- Brimelow, J.C. and G.W. Reuter, 2005: Transport of Atmospheric Moisture during Three Extreme Rainfall Events over the Mackenzie River Basin, *J. of Hydrol.*, **6**, 423-440
- Cherubini, T., A. Ghellia, F. Lalaurette, 2002: Verification of Precipitation Forecasts over the Alpine Region Using a High-Density Observing Network, *Wea. and Forecasting*, **17**, 238-249
- Colle, B.A., K.J. Westrick, C.F. Mass, 1999: Evaluation of MM5 and Eta-10 Precipitation Forecasts over the Pacific Northwest during the Cool Season, *Wea. and Forecasting*, **14**, 137-154
- Colle, B.A., C.F. Mass, K.J. Westrick, 2000: MM5 Precipitation Verification over the Pacific Northwest during the 1997-99 Cool Seasons*, *Wea. and Forecasting*, **15**, 730-744
- Colle, B.A., J.B. Olson, J.S. Tongue, 2003: Multiseason Verification of the MM5. Part II: Evaluation of High-Resolution Precipitation Forecasts over the Northeastern United States, *Wea. and Forecasting*, **18**, 458-480
- Cressman, G.P., 1959: An Operational Objective Analysis System, *Mon. Wea. Rev.*, **87**, 367-374
- Davolio, S., D. Mastrangelo, M.M. Miglietta, O. Drofa, A. Buzzi, P. Malguzzi, 2009: High resolution simulations of a flash flood near Venice, *Nat. Hazards Earth Sys. Sci.*, **9**, 1671-1678
- Done, J., C.A. Davis, M. Weisman, 2004: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model, *Atmos. Sci. Let.*, **5**, 110-117
- Flesch, T.K. and G.W. Reuter, 2012: WRF Model Simulation of Two Alberta Flooding Events and the Impact of Topography, *J. of Hydrometeor.*, **13**, 695-708
- Fritsch, J.M. and C.F. Chappell, 1980: Numerical Prediction of Convectively Driven Mesoscale Pressure Systems. Part I: Convective Parameterization, *J. Atmos. Sci.*, **37**, 1722-1733

- Gilliland, E.K. and C.M. Rowe, 2007: A Comparison of Cumulus Parameterization Schemes in the WRF model, Preprints, *21st Conference on Hydrology*, San Antonio, TX, Amer. Meteor. Soc., P2.16
- Gochis, D. J., Z-L. Yang, W.J. Shuttleworth, R.A. Maddox, 2000: Quantification of error in a pseudo-regional climate simulation of the North American Monsoon. *Proc. Second Southwest Weather Symp.*, Tucson, AZ, National Weather Service.
- Gochis, D.J., W.J. Shuttleworth, Z-L. Yang, 2003: Hydrometeorological Response of the Modeled North American Monsoon to Convective Parameterization, *J. of Hydrometeor.*, **4**, 235-250
- Gochis, D.J., W.J. Shuttleworth, Z-L. Yang, 2002: Sensitivity of the Modeled North American Monsoon Regional Climate to Convective Parameterization, *Mon. Wea. Rev.*, **130**, 1282-1298
- Grell, G.S., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations, *Mon. Wea. Rev.*, **121**, 764-787
- Grell, G.A. and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, *Geophys. Res. Letters*, **29**, 1693-1696
- Grubišić, V., R.K. Vellore, A.W. Huggins, 2005: Quantitative Precipitation Forecasting of Wintertime Storms in the Sierra Nevada: Sensitivity to the Microphysical Parameterization and Horizontal Resolution, *Mon. Wea. Rev.*, **133**, 2834-2859
- Hewitson, B.C. and R.G. Crane, 2005: Gridded Area-Averaged Daily Precipitation via Conditional Interpolation, *J. of Climate*, **18**, 41-57
- Hong. S.-Y., Y. Noh, J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318-2341
- Janjić, Z.I., 2000: Comments of "Development and Evaluation of a Convection Scheme for Use in Climate Models", *J. Atmos. Sci.*, **57**, 3686-3687
- Janjić, Z.I., 1994: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure schemes, *Mon. Wea. Rev.*, **12**, 927-945
- Kain, J.S. 2003: The Kain-Fritsch Convective Parameterization: An Update, *J. of Applied Meteor.*, **43**, 170-181
- Kain, J.S., M.E. Baldwin, S.J. Weiss, 2003: Parameterized Updraft Mass Flux as a predictor of Convective Intensity, *Wea. and Forecasting*, **18**, 106-116

- Kain, J.S. and J.M. Fritsch, 1990: A One-Dimensional Entraining/Detraining Plume Model and Its Application in Convection Parameterization, *J. of the Atmos. Sci.*, **47**, 2784-2802
- Kain, J.S. and J.M. Fritsch, 1993: Convective Parameterization for Mesoscale Models: The Kain-Fritsch Scheme, *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.*, No 46, Amer. Meteor. Soc., 165-170
- Kerkhoven, E., T.Y. Gan, M. Shiiba, G. Reuter, K. Tanaka, 2006: A comparison of cumulus parameterization schemes in a numerical weather prediction model for a monsoon rainfall event, *Hydrol Process.*, **20**, 1961-1978
- Liang, X-Z., L. Li, K.E. Kunkel, M. Ting, J.X.L. Wang, 2004: Regional Climate Model Simulation of U.S. Precipitation during 1982-2002. Part I: Annual Cycle, *J. of Clim.*, **17**, 3510-3529
- Lin, Y-L., R.D. Farley, H.D. Orville, 1983: Bulk Parameterization of the Snow Field in a Cloud Model, *J. of Clim. and Appl. Meteor.*, **22**, 1065-1093
- Mass C. F., D. Ovens K. Westrick, B.A. Colle, 2002: Does Increasing Horizontal Resolution Produce More skillful Forecasts?, *Bull. Amer. Meteor. Soc.*, **83**, 407-430
- McBride, J.L. and E.E. Ebert, 2000: Verification of Quantitative Precipitation Forecasts from Operational Numerical Weather Prediction Models over Australia, *Wea. Forecasting*, **15**, 103-121
- Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, W. Shi, 2006: North American Regional Reanalysis, *Bull. Amer. Meteor. Soc.*, **87**, 343-360
- Michalakes, J., J. Dudhia, D. Gill, J. Klemp, W. Skamarock, 1999: Design of a Next-Generation Regional Weather Research and Forecast Model, *Towards Teracomputing*, World Scientific, River Edge, New Jersey, 117-124
- Ooyama, V.K., 1971: A theory on parameterization of cumulus convection, *J. Meteor Soc. Japan*, **49**, 744-756
- Ou, A.A., 2008: Meteorological analysis of four rainstorms that caused severe flooding in Alberta during June 2005. MSc. Thesis. Univ. of Alberta, Edmonton, Canada
- Phillips, D., 2010: Canada's top 10 weather stories for 2010. Available at <http://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=53E29740-1>

- Ratnam, J.V. and K.K. Kumar, 2005: Sensitivity of the Simulated Monsoons of 1987 and 1988 to Convective Parameterization Schemes in MM5, *J. of Clim.*, **18**, 2724-2743
- Reuter, G.W. and C.D. Nguyen, 1993: Organization of cloud and precipitation in an Alberta storm, *Atmos. Res.*, **30**, 127-141
- Roberts, N.M. and H.W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events, *Mon. Wea. Rev.*, **136**, 78-97
- Rozumalski, R.A., 2006: SOO/STRC WRF EMS User's Guide. Available online at <http://strc.comet.ucar.edu/wrf>
- Schaefer, J.T., 1990: The Critical Success Index as an Indicator of Warning Skill, *Wea. Forecasting*, **5**, 570-575
- Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X-Y. Huang, W. Wang, J.G. Powers, 2008: A Description of the Advanced Research WRF Version 3, *NCAR technical note* NCAR/TN-475+STR, 113 pp
- Smith, H.D., 2011: Diagnosis of a Large Hail Event over Calgary, Alberta, 12 July 2010. Internal report. Prairie and Northern Region. Environment Canada.
- Smith, S.B. and M.K. Yau, 1987: The mesoscale effect of topography on the genesis of Alberta hailstorms, *Beirt. Phys. Atmos.* **60**, 371-392
- Vaidya, S.S. and S.S. Singh, 2000: Applying the Betts-Miller-Janjic Scheme of Convection in Prediction of the Indian Monsoon, *Wea. Forecasting*, **15**, 349-356
- Wang, W. and N.L. Seaman, 1997: A Comparison Study of Convective Parameterization Schemes in a Mesoscale Model, *Mon. Wea. Rev.*, **125**, 252-278
- Warner, T.T., R.A. Petersen, R.E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction, *Bull. Amer. Meteor. Soc.*, **78**, 2599-2617
- Wilks, D. S., 1995: Statistical Methods in the Atmospheric Sciences. Academic Press, 467 pp.
- Yanai, M. and R.H. Johnson, 1993: Impact of Cumulus Convection on the Thermodynamic Fields, *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.*, No 46, Amer. Meteor. Soc., 39-62
- Yang, M-J. and Q-C. Tung, 2003: Evaluation of Rainfall Forecasts over Taiwan by Four Cumulus Parameterization Schemes, *J. of the Met. Soc. Of Japan*, **18**, 1163-1183

Chapter 3: Verification of daily WRF forecasts for Alberta

3.1. Introduction

Diverse geography and the potential effects resulting from it characterize Alberta, Canada. The Rocky Mountains along the western boundary have a direct impact on weather due to their large elevation profiles (e.g. Reuter and Nguyen 1993). Plains in the east provide a different geography from the foothills and mountains in the western regions. The juxtaposition of these regions promotes storm development, resulting in convective summers often capable of producing severe conditions. Alberta summer temperatures reach above 30 °C, while winter temperatures occasionally drop below -40 °C. Sustained heavy winds in the lee of the Rockies are harvested for electricity by wind farms. With such diverse weather conditions, accurate forecasts are important; both economically and for quality of life. Verification of any numerical weather prediction (NWP) model being used in Alberta is therefore important.

A NWP model can produce a forecast for atmospheric conditions including winds, temperatures, and precipitation. However, until any model forecast has been compared statistically against daily observations, a process called verification, the accuracy of its forecast is unknown. While the equations that govern fluid dynamics have long been known, early attempts to use these equations for producing a forecast failed (Richardson 1922). Today, computational power can solve these equations with greater precision; however, the resulting numerical forecasts are often different than what is observed.

Several factors affect the forecast accuracy of NWP models. The 2-meter temperature is often verified for accuracy (e.g. Dallavalle and Dagostaro 2004; Mass et al. 2002; Murphy et al. 1989); however, studies indicate some difficulties in verifying the 2-meter temperature, including elevation and season. Large changes in elevation over short horizontal distances are difficult to represent accurately in a NWP model due to the limited horizontal and vertical resolution of the model, resulting with the model approximating the terrain poorly. This often will result with the smoothing of a valley into the nearby mountains, simulating a higher ground elevation of the valley than is observed. These elevation differences often result in larger temperature errors due to the effect which the lapse rate has on temperatures (e.g. Frauenfeld et al. 2005). Poor interpolation from surface temperatures can be compounded by geography and even season (Galanis and Anadranistakis 2002) resulting in further inaccuracies in the 2-meter

temperatures. Daily precipitation is often verified, although skill in forecasting precipitation is increasing at a slower rate than other parameters, such as the 2-meter temperature (e.g. Roebber and Bosart 1998; Dallavalle and Dagostaro 2004). Verification of precipitation requires a dense network of precipitation gauges, due to the highly variable spatial distribution of precipitation, which makes accurate precipitation forecasts difficult to produce. These precipitation gauges suffer from under-catchment during windy conditions, with as much as 10-40% lower amounts of precipitation than the true observed amounts (Colle et al. 2000), causing potential biases in observed data, which is then compared against precipitation values from a forecast to perform verification.

The verification process determines model biases such as forecasting colder temperatures than are typically observed, as well as the errors when forecasting temperatures. Error analysis may be performed using the mean absolute error, mean error, and/or root mean squared error. We will verify the forecasts produced by the Weather Research and Forecasting (WRF) model (Michalakes et al. 2001; Skamarock et al. 2008). While WRF has been verified at other locations (e.g. Chen et al. 2009; Cheng and Steenburgh 2005) it has not been verified over Alberta. WRF is available to the research community without cost, and forecast verification over Alberta could be useful to other researchers using WRF or other variants, such as WRF-Chem (Grell et al. 2005) or WRF-Fire (Patton and Coen 2004), in this region.

The forecast skill of a model is determined by comparing its forecast against other forecast methods, such as climatology, persistence, or forecasts produced from meteorologists using NWP models (Murphy 1988). If a model is found to have more skill when measuring one parameter, such as wind speed, when compared to another forecast method, then the model was able to forecast wind speed more accurately. Determining model skill compared to a climate forecast is useful, as skill scores can be used to compare forecasting performance across different geographical regions throughout the seasons (Murphy 1988).

This project verified forecasts produced by the WRF model and determined model skill compared to climatology, persistence, and the forecasts produced by Environment Canada. All forecasts were performed for one full year, divided into four seasons to determine differences in WRF's ability to forecast each season. Verification was evaluated for daily maximum and minimum temperature, wind speed, sea-level pressure, and daily precipitation accumulations from nine weather stations across Alberta.

3.2. Methodology

3.2.1. Numerical experiments

From 15 March 2011 through 29 February 2012, the WRF model performed a forecast each day which provided information regarding the temperature, pressure, wind speed, and precipitation for the following two days. The first day in the forecast is designated as "Day 1", immediately following the day the forecast was initialized, while the second day is designated as "Day 2". Each model forecast provided the three-dimensional fields of temperature, pressure, wind speed, water vapor, precipitation, etc. For the purpose of verifying the model forecasts, we compared the near-surface conditions at nine weather stations located in southern and central Alberta (Table 3-1). For each forecast day, all nine station reported maximum daily temperature, minimum daily temperature, 1200 UTC wind speed, 1200 UTC sea level pressure, and 24 hour liquid precipitation accumulation. Any snowfall was recorded in liquid-equivalent amounts of precipitation. There was no attempt to adjust any of the forecast values as this project focuses on verifying the model output.

Four seasons fell within the span of data collection: spring was from 15 March 2011 to 31 May 2011; summer was from 1 June 2011 to 31 August 2011; fall was from 1 September 2011 to 30 November 2011; winter was from 1 December 2011 to 29 February 2012. Verification and skill were also analyzed by integrating all four seasons into a year-long average. The spring season started late due to computational problems, and days from 1 March 2011 through 14 March 2011 were not used.

3.2.1. Numerical model

The modeling system used in this study is the Weather Research and Forecasting Environmental Modeling System, version 3.0 (Rozumalski 2006). WRF is a limited area model that uses 45 vertical levels with the pressure level top at 50 mb; other characteristics of the model are shown in Table 3-2. Figure 3-1 shows the domain setup using a grid resolution of 15 km in the nested domain, and forecast output is only taken from the nested 15 km domain, for which the output field is saved every 30 minutes for analysis. There is a 3:1 ratio between the nested domain and parent domain and two-way feedback between each domain to help prevent errors from the lateral boundary conditions from working their way into the nested domain (Warner et al. 1997). The

forecasts were initialized at 1200 UTC from the NAM212 dataset (Black 1994) and boundary conditions of the larger domain were updated every 3 hours.

3.2.3. Forecast verification

The forecasts were evaluated in two ways. The first evaluation determined the forecast accuracy of WRF which verified daily forecasts from WRF directly against observed data. Root mean square error (RMSE) and mean error (ME) were calculated for temperature, wind, and pressure forecasts for the whole year and separated into the four seasons. The errors were evaluated at each of the nine stations in Table 3-1, and were averaged across all stations to provide a single value of the seasonal error. RMSE was used to determine accuracy of the forecasts, as opposed to mean absolute error, to penalize large deviations from observed values due to squaring the error between observed and forecast values (Appendix G). ME was used to determine if the forecasts over-predicted values (positive ME), under-predicted (negative ME), or had neutral bias (ME= 0). Together, both RMSE and ME allow for analysis of the accuracy between forecast and observed values, as well as understanding of any over/under prediction of the weather model.

Brown et al. (2004) reviews several methods to verify the forecast precipitation amounts. In this study we have adopted the 2x2 contingency table method (Wilks 1995) and computed the Probability of Detection (POD), False Alarm Ratio (FAR), and Equitable Threat (ET) scores. Appendix F contains the formulas for these precipitation verification statistics. POD determined the percentage of events which WRF correctly forecast precipitation when precipitation was observed. FAR determined the percentage of events which WRF incorrectly forecast precipitation when precipitation was forecast. BIAS determined if WRF forecast more precipitation than observed, or less than observed. A BIAS value of 1 indicates the same number of precipitation events were forecast as were observed, whereas a BIAS value of 2 would result for WRF forecasting twice as many events of precipitation than observed. ET determines overall skill when forecasting precipitation, and includes a correction term which reduces the effect of a correct forecast by chance. An ET score of 0 indicates the same accuracy as a random precipitation forecast, and positive ET scores indicate some level of accuracy, while a perfect forecast would have ET equal to 1. Precipitation thresholds are values to determine the accuracy of precipitation forecasts for various amounts of precipitation, and three were used: 0.2 mm, 1.0 mm, 5.0 mm. Unlike error analysis that uses the

difference between observed and forecast amounts, precipitation thresholds break each forecast and observation down into either a rain or no-rain event, and where each rain/no-rain event is contingent upon reaching the threshold being analyzed. For example, a forecast for 0.5 mm of precipitation would meet the 0.2 mm threshold, but would not reach the 1.0 mm or 5.0 mm threshold, and only result with a rain forecast at the 0.2 mm threshold. Verification of precipitation was only performed during the summer season, as each of the other seasons had too few days with precipitation. The summer season was subsequently compared against the entire year to determine any differences in accuracy for precipitation. For other methods to verify precipitation we refer to Brown et al. (2004), Davis et al. (2006), Ebert and McBride (2000), or McBride and Ebert (2000).

3.2.4. Forecast skill

The statistics above were used to evaluate the accuracy of the WRF model with respect to observed conditions. WRF was also compared against other common forecasting methods to determine the model's forecast skill (Appendix H). The forecast methods compared against WRF were a climatology based forecast, persistence forecast, and Environment Canada's Weather Office forecast. Skill scores can only be found for parameters which are common among all forecast methods. As a consequence, skill was only calculated for temperature forecasts. The climatology-based forecast used temperature history from the last 30 years at a particular location to forecast the average daily maximum and daily minimum that would occur at that location each day.

Climatology forecasts have identical Day 1 and Day 2 forecast values. A persistence-based forecast assumes the weather will be the "same as yesterday" for Day 1 forecasts (and "same as the day before yesterday" for Day 2 forecasts) by forecasting the same maximum and minimum temperatures recorded from one day as the forecast for the following day (or two days). Environment Canada's Weather Office uses numerical weather prediction models, meteorologists, climate and satellite data, and other resources to make their forecasts.

WRF's Day 1 and Day 2 forecasts were compared against these forecast methods, as well as observed values, to determine the relative skill between WRF and the various forecast methods each season. Skill scores which compare the above forecast methods to WRF's forecast determine which forecast method had higher accuracy. A positive skill score indicates that WRF has greater accuracy than the forecast method being compared, and a negative skill score indicates the opposite. The values of skill score range from

positive one, which would be a perfect WRF forecast, to negative infinity. A skill score of zero would result from two forecasts of the same accuracy. Assessing skill provides no information about the errors in accuracy; those errors are determined from the verification. Both climatology and persistence forecasts require fewer resources when producing a forecast than WRF does, and it is anticipated that the additional costs to produce WRF forecasts should come with greater skill than both of these methods.

3.2.5. Verification datasets

Data used for verification of the WRF model consisted of observed sea level pressure and 10-meter wind speed (both measured at 1200 UTC), 2-meter maximum temperature, 2-meter minimum temperature, and liquid precipitation accumulation, recorded each day from each observation station (Table 3-1), from the Weather Office website of Environment Canada (<http://climate.weatheroffice.gc.ca>). There was no attempt to modify any of the observational data.

Day 1 and Day 2 temperature forecasts were also recorded to evaluate the relative skill of the WRF model compared to climatology, persistence, and Environment Canada's forecast. The normal climate temperatures, averaged across the same day of the year for the past 30 years, were recorded for each station (<http://climate.weatheroffice.gc.ca>) for every day in the study. Persistence data were recorded once observed values were available. Environment Canada provides four forecasts each day for each weather station, available at <http://www.weatheroffice.gc.ca>. The temperature forecasts released at 11 AM (local time) were recorded for Day 1 and Day 2 forecasts because they aligned most closely with same initialization data as the WRF forecast. Forecasts from Environment Canada, climatology, and persistence will provide information about the relative skill of the WRF model by having other forecasts methods to compare with.

3.3. Verification Results

3.3.1. Seasonal Verification

The verification was performed across 9 observation stations in Alberta. A time series of WRF forecasts and observations was performed for Edmonton International Airport (Fig. 3-2). The WRF maximum temperature forecasts (Fig. 3-2a) are shown with the observations, and while some days show large deviations from the observations, the

maximum temperature agrees fairly well with the observations. The WRF minimum temperature forecast (Fig. 3-2b) shows more days with larger deviations from the observations, with the observed minimum temperature often colder than the forecast values. While the pressure is often forecast fairly similar to the observations (Fig. 3-2c), both forecasts for wind speed and precipitation contain large deviations from observations (Fig. 3-2d and 3-2e). These figures also show that Day 1 and Day 2 forecasts are different, though it is difficult to determine their difference in accuracy.

Table 3-3 shows the errors for each observation station during the summer season for Day 1 forecasts, while Table 3-4 shows Day 2 forecasts errors. Comparing these two tables together illustrates the differences in forecasting Day 1 and Day 2 weather at each station. For example, WRF forecasts at ZPC had a similar Day 1 maximum temperature RMSE (2.26°C) as Day 2 (2.57°C). ZPC also had similar ME values when forecasting maximum temperature for Day 1 (-1.13°C) as Day 2 (-1.12°C) indicating both forecast days at Pincher Creek had a cold-bias when forecasting minimum temperatures during the summer season. Most stations show a lower, more accurate, RMSE value for a Day 1 forecast than a Day 2 forecast. Day 2 forecasts were considered similar, but less accurate, than Day 1 forecasts due to having higher RMSE values and similar ME values across all observation stations during this summer season. Ideally, Day 2 forecasts errors would be the same as Day 1 forecasts. However, a Day 2 forecast results from a longer simulation length, as well as uses initial conditions further removed from observations, both of which can allow initial errors to grow and propagate. This can lead to Day 1 and Day 2 forecasts being similar, yet with Day 2 forecasts of lower accuracy. Similar results were found across the remaining seasons, and further discussion of results will be primarily about Day 1 forecasts except for anything significant regarding Day 2 forecasts.

Figure 3-3 illustrates Day 1 forecast values matched with observations during the summer season at Edmonton International Airport. The diagonal line present on each graph is a line where the forecast value is the same as the observed value; the perfect forecasting line. The maximum temperature forecasts (Fig. 3-3a) has a relatively tight fit around the perfect forecasting line, which is not the case for minimum temperature forecasts (Fig. 3-3b). The minimum temperature forecasts have a larger deviation from the perfect forecasting line, and also show a warm bias, as they are generally above the perfect forecasting line, indicating a higher forecast for minimum temperature than observed. WRF pressure forecasts (Fig. 3-3c) matched well with observations, but typically are forecast slightly lower than are observed. The forecasts for wind speed (Fig.

3-3d) show the opposite; large variation between forecast value and observed value, as well as a tendency to forecast higher speeds than were observed. Precipitation forecasts (Fig. 3-3e) were often much different than the observations, with slightly more days forecast with precipitation than were observed. While WRF forecasts produced for Edmonton International Airport had some differences when compared against observations, it is important to verify the model over many locations.

We verified the WRF model's forecast for nine observation stations in Alberta. We previously showed that there was variability for Edmonton International Airport forecasts. We would expect that each station being used to verify the WRF model would show some level of variation between the forecast values and observed. Table 3-3 shows the spatial differences in forecast errors found at these nine stations across the summer season. For example, YQL shows a warm bias in maximum temperature with a positive ME value of 0.53°C, as well as forecasting faster wind speeds than was observed, with a ME value of 0.85 km/h. ZPC shows the opposite; a cold bias in maximum temperature (ME= -1.13°C) and often forecasts wind speeds slower than observed (ME= -1.13 km/h). While all stations experienced a warm bias when forecasting the minimum temperatures during this summer season, the ME values ranged from 0.51°C at YQF to 3.9°C at YXH. The maximum temperature ME values ranged from -1.13°C at ZPC to 0.6°C at YYC, and most stations had a negative value, indicating a cold bias when forecasting maximum temperatures. The RMSE values also show variability between each station. YET had the largest RMSE for wind speed at 8.27 km/h while YLL had the lowest at 5.86 km/h. The maximum temperature RMSE values ranged from 1.79°C at YEG to 2.4°C at YET, while the minimum temperature values were from 1.94°C at YQF to 4.59°C at YXH. The summer season had less variability between stations when forecasting maximum temperature than for minimum temperature. Each station had different RMSE and ME values than other stations, indicating that WRF forecasts during the summer season have some degree of spatial variability. While the data were analyzed for each station across each forecast variable, the stations were averaged together to determine a single value of RMSE and ME for each variable for the four season.

3.3.1.1. Summer season

The summer season forecasts produced by the WRF model had ME for maximum temperatures (Table 3-5) at -0.24 °C while the minimum temperature forecast ME was much higher in value, at 2.5°C. With a slightly negative ME for maximum

temperature (under-prediction) and a positive ME for minimum temperature (over-prediction), WRF forecast a smaller diurnal temperature range than was observed during this season. Day 2 forecasts showed this smaller diurnal temperature range as well (Table 3-6). The summer season had the lowest RMSE for maximum temperature forecasts at 2.08°C, while the RMSE for minimum temperature was 3.39°C. The summer season had the most accurate maximum temperature forecast while the forecasts for minimum temperature were not the most accurate, being slightly less accurate than the spring season.

While summer wind forecasts were predicted at higher speeds than observed (ME= 2.27 km/h), the RMSE value of 7.17 km/h is low compared to the other seasons. Other studies using WRF found similar wind RMSE values during the summer season (e.g. Chen et al. 2009).

The summer pressure forecasts had a -1.4 mb ME value, indicating summer forecasts often under-predicted the pressure value. The RMSE value was 2.32 mb, which was slightly more accurate than the averaged year.

3.3.1.2. Winter season

The ME values for the winter season's temperature showed similarities to that of the summer: negative ME for the maximum temperature (-1.63°C) and positive ME for the minimum temperature (1.98°C) (Table 3-5). However, the RMSE values were far higher during winter. Winter season had a RMSE of 3.48°C for the maximum temperature and 4.54°C for minimum temperature. While winter performed the least accurate for maximum temperature forecasts, the winter season had lower RMSE values than the fall season for minimum temperatures.

The winter season forecast greater wind speed than was observed, with the highest ME of all seasons at 3.44 km/h. The RMSE values for wind speed were also the highest for winter season, at 10.22 km/h. Winter wind forecasts were the least accurate when compared to the other seasons. Forecasting wind speeds during the winter has been found to be more inaccurate than other seasons, as Vincent et al. (2008) found that winter had higher wind RMSE values than the summer season using their model.

The pressure forecasts produced during the winter season had a RMSE of 3.13 mb, which was higher than the other three seasons. With a ME value of 1.62 mb, the winter season also forecast higher pressure than was observed. Compared to summer, which had a negative ME in the pressure at -1.40 mb, WRF showed greater mean errors

when forecasting pressure during the extreme seasons, while spring and fall were more neutral with ME close to 0 mb. Reasons why WRF forecasts had lower pressure during summer and higher pressure during winter when compared to observations are not known. Winter was the least accurate season when forecasting pressure due to consistently over-predicting pressure and the greatest RMSE value.

The winter season experienced the largest range of temperatures as well as more days with strong winds than any other season. The range of observed temperatures at YEG during the winter was from -35°C to $+9^{\circ}\text{C}$, while during summer the range was 0°C to 29°C . The season with the widest range of temperatures could be expected to have large errors, due to a wider range of values for the mathematical computations performed by the numerical model. Winter also experienced more days with strong winds caused by large horizontal temperature gradients, and had large wind speed errors as a result.

3.3.1.3. Fall season

WRF fall temperature forecasts showed the same diurnal temperature problems as mentioned earlier, with a maximum temperature ME of -1.2°C , and a minimum temperature ME of 3.98°C (Table 3-5). The maximum temperature RMSE (2.73°C) was more accurate than the averaged year. However, the minimum temperature was forecast with the least accuracy of all seasons, with a RMSE of 5.07°C .

With a positive ME value (2.76 km/h), the forecast wind speed in the fall was often faster than the observed speed. The RMSE were not as large, at 9.18 km/h , compared to the winter season. WRF forecasts for fall wind speed were slightly more accurate than forecasts made during the winter season.

While having a slight high pressure bias (ME= 0.22 mb), the fall pressure forecast RMSE value was the lowest of all seasons, at 2.12 mb . Pressure forecasts made during the fall season were more accurate than any other season.

3.3.1.4. Spring season

Spring had a larger cold-bias for the maximum temperature forecast (ME= -1.31°C) than summer (-0.24°C) while the RMSE for maximum temperature forecasts for spring (2.74°C) was higher than summer (2.08°C) (Table 3-5). However, minimum temperatures were best forecast during this season, with the lowest RMSE of 3.24°C and the most neutral warm bias (ME= 1.08°C) of all seasons.

The spring season had a positive bias when forecasting wind speed, with a ME of 0.77 km/h, which was lower than the summer season. The RMSE value during spring (7.12 km/h) was slightly lower than the summer season (7.17 km/h). WRF had approximately the same wind forecast RMSE values between both spring and summer, but more often forecast faster winds during summer, making spring the most accurate season when forecasting wind speeds.

With a neutral ME of 0.02 mb and a similar RMSE as summer (2.32 mb), WRF pressure forecasts produced during the spring season were of the same accuracy as summer forecasts, but also showed no bias in forecasting pressure compared against observed values, unlike the summer season.

3.3.1.5. Averaged year

The averaged year results were an average across the four seasons. The following averaged year results will indicate typical model bias, while the above results indicate specific seasonal biases.

The RMSE in maximum temperature for the average year was 2.82°C while the minimum temperature was 4.19°C (Table 3-5). The averaged year showed a negative ME (-1.08°C) when forecasting maximum temperature, and a positive ME (2.43°C) when forecasting minimum temperatures. For this year of study, WRF forecast were found to forecast maximum temperature more accurately than minimum temperatures. The WRF forecast errors also indicate that WRF typically has a cold bias when forecasting maximum temperatures and warm bias for minimum temperatures. WRF forecasts during this year of study were found to have a smaller range in diurnal temperature than what was observed, during all four seasons. Hu et. al (2010) evaluated three planetary boundary layer schemes in the WRF model and also found a cold bias in the maximum temperature. They suggested that the cold bias could be due to errors simulating cloud cover and/or soil moisture.

With a ME value of 2.37 km/h, WRF forecasts for wind speed during this year of study often over-predicted wind speeds. Cheng and Steenburgh (2005) found WRF forecast stronger winds than observed during the summer season, over the western United States. The RMSE values for wind speed were averaged to 8.59 km/h for our year of study. This result was lower than those found by Rife and Davis (2005) when verifying wind over complex terrain over New Mexico using the MM5 model; however, verification of WRF forecasts in this project did not happen at stations with complex

terrain, as most stations used were in the foothills or in the prairie. Borge et al (2008) used the WRF model over the Iberian Peninsula, using many different model configurations, and found the “best case” simulation, using their optimal configuration, had wind RMSE values very similar to this study.

With a positive ME value of 0.11 mb, WRF pressure forecasts during this year of study were found to be slightly higher in pressure than observed. However, WRF showed large ME values during summer (-1.40 mb) and winter (1.62 mb) indicating that WRF has some difficulties during the extreme seasons which are not present during fall and spring. The yearly RMSE was averaged to 2.51 mb, with only the winter season having lower accuracy, at 3.13 mb.

The precipitation analysis compared precipitation during the summer season and the precipitation during the entire year of study. Alberta has a convective summer, with many days of precipitation. Of the days with precipitation, 42% took place during the summer season. The POD values (Table 3-7) are greater during summer than the averaged year for each threshold and forecast day, indicating that summer forecasts were able to more correctly forecast days with precipitation. The FAR values were less during the summer, indicating that summer forecasts were less likely to falsely forecast days with precipitation. The ET scores were slightly lower during the summer season than the averaged year, though the values were fairly close suggesting that both seasons were of similar accuracy. The BIAS values for precipitation accumulation during both seasons were higher than 1.00 for all thresholds, indicating that WRF forecast more precipitation than was observed for these seasons. This is likely due to the Kain Fritsch cumulus parameterization scheme, which often simulates more precipitation than was observed (e.g. Colle et al. 2003; Gochis et al. 2002; Liang et al. 2004; Wang and Seaman 1997). However, the summer season had lower BIAS values on average, thus more accurately forecast the amounts of precipitation. There were some differences between the forecast days; Day 2 forecasts typically had lower BIAS values and lower forecast accuracy, indicated by lower ET values, than Day 1 forecasts. While having slightly lower ET scores, the summer season was found to forecast precipitation slightly more accurate than the averaged year.

The model does exhibit some biases such as forecasting a smaller range of diurnal temperatures which was mentioned earlier. The Day 2 forecasts (Table 3-6) were less accurate than Day 1 forecasts (Table 3-5), for most parameters, indicated by larger RMSE for Day 2 forecasts. The ME values are different between Day 1 and Day 2

forecasts. For example, the Day 1 ME for minimum temperature during the summer was 2.50°C, while it was 2.67°C for Day 2 forecasts for the same season. This small change, positive in direction, was found to occur for all seasons except the winter season for minimum temperature forecast. This indicates that WRF typically forecasts warmer minimum temperatures further into the forecast length; a warmer minimum temperature is forecast for Day 2 forecast than for Day 1. Similar results exist for pressure and wind speed, as WRF forecasts stronger winds and lower pressure further into the forecast length. The winter season did not show the same behavior, possible due to the unseasonably warm winter 2011-2012 season, which will be discussed in the skill score section below.

3.3.2. Skill Scores of WRF temperature forecasts

We evaluated the skill of the WRF model compared to a climatology, persistence, and Environment Canada forecast. The meteorological variables in these comparisons were maximum and minimum temperatures. This was done to determine relative accuracy of the WRF model compared to other forecast methods. Skill scores were calculated at each observation station for the four seasons, as well as the entire year. Table 3-8 shows WRF skill scores when calculated for the entire year of study at each observation station. The variation in skill when compared against climatology forecasts of maximum temperature, from 0.75 at YET to 0.86 at WCT, is much smaller than for minimum temperature, which was from 0.09 at WCT to 0.61 at YLL and ZPC. Skill scores for WRF minimum temperature forecasts compared to persistence forecasts show a large variation, from -0.73 at WCT to 0.36 at ZPC. This indicates that some stations were more accurately forecast using the WRF forecast for minimum temperatures than the persistence forecast, while others showed the opposite. When WRF is compared to the maximum and minimum temperature forecasts issued by Environment Canada, most stations resulted with negative skill score values, though a few had slightly positive skill scores for maximum temperature. While there was variability in skill across the different stations, resulting with some forecast methods being more accurate than others at a single observation station, the goal of this project was to determine WRF accuracy across central and southern Alberta, and the skill from the nine stations will be averaged to determine a single value for model accuracy against other forecast methods.

3.3.2.1. WRF forecast skill compared to climatology forecasts

During the winter season, WRF had the highest Day 1 skill (0.84) when compared to a climatology based forecast for maximum temperature (Table 3-9). The fall season was slightly lower (0.81), with summer (0.72) and spring (0.66) even lower. A skill score of 0.81 indicates that WRF's MSE values were 19% of the climatology based MSE. Day 2 forecasts result with similar skill scores, though they are slightly lower. All seasons had positive skill compared to the maximum temperature climatology forecasts, and thus WRF was more accurate forecasting maximum temperatures across all seasons during this year of study than a forecast based on climatology.

WRF minimum temperature forecasts were also compared against climatology forecasts. The winter season, with a skill score of 0.70, was more accurate than a climatology based forecast. The remaining seasons all had negative skill, with spring and fall slightly less than zero (-0.02 and -0.07), while the summer season was -0.72. A negative skill score of -0.72 indicated that WRF's MSE values were 172% of the MSE for climatology. This indicates that WRF was less accurate than climatology for determining minimum temperatures for three seasons. However, averaging the seasons together results with an overall positive skill for the year of study, at 0.39.

The maximum temperature skill scores were typically positive, and larger, than the minimum temperature skill scores. The lower skill scores for minimum temperature were due to lower accuracy of the WRF model when forecasting minimum temperatures. This is indicated by greater RMSE values for minimum temperatures than for maximum temperatures (Table 3-5). WRF was also earlier shown to forecast temperatures less accurately during the winter season (see section 3.3.1.2.). The greater skill scores, for maximum and minimum temperature, during winter resulted from a lengthy unseasonably warm period during this season. This caused greater climatology errors when compared against WRF forecast errors, resulting in greater skill scores during winter. The fall season experienced a similar effect to a lesser extent. Winter was expected to have lower, yet positive, skill when forecasting maximum temperatures due to greater RMSE between forecast value and observed value, than the other seasons. For minimum temperatures, fall and winter were expected to have lower, and negative, skill scores when compared against climatology.

3.3.2.2. *WRF forecast skill compared to persistence forecasts*

WRF forecast maximum temperature during the summer season with the highest skill, at 0.71 (Table 3-9), when compared against persistence forecasts. The remaining seasons all had positive skill, with fall at 0.69, winter at 0.60, and spring at 0.26. The fall season had similar RMSE as spring, though the spring season had lower skill. The lower skill during spring resulted from low errors between the persistence forecast and observations, which came about from slowly changing day to day temperatures during this season. The average of the four seasons give a value of 0.61 for the entire year. With positive skill scores for all seasons, WRF forecast maximum temperatures more accurately than a persistence based forecast during this year of study.

WRF forecast minimum temperature during the winter season with the highest skill, at 0.35, when compared against persistence forecasts. The remaining three seasons all had negative skill, with spring at -0.18, fall at -0.43, and summer at -0.59. These negative scores result from high RMSE between minimum temperatures forecast by WRF and the observed minimum (Table 3-5). WRF skill scores compared against climatology show the same three seasons with negative skill scores for minimum temperatures as compared to persistence forecasts, indicating that WRF does not perform accurately for minimum temperatures.

The relatively warmer winter season had positive skill due to the errors between persistence temperatures and observed. Persistence forecasting will have low errors in periods of time where the temperature changes slowly. However, during the winter season, the temperature changed quickly from day to day, and resulted with large errors between the persistence forecast and observed values. The WRF skill scores for minimum and maximum temperatures were greater for a Day 2 forecast than Day 1 forecast when compared to persistence.

3.3.2.3. *WRF forecast skill compared to Environment Canada's forecasts*

WRF forecast maximum temperature during the fall season with the highest skill, at 0.00 (Table 3-9), when compared against Environment Canada's forecasts. A skill of 0.00 indicates that maximum temperature forecasts issued from Environment Canada, for the nine observation stations used in this study had the same accuracy as WRF maximum temperature forecasts for the same nine stations during the fall season. The remaining seasons had negative skill, with winter at -0.01, summer at -0.29, and spring at -0.84. With three seasons having negative skill, and the fourth season having identical skill, the

WRF model had lower accuracy when forecasting maximum temperatures when compared against forecast issued by Environment Canada's weather office during this year of study.

WRF forecast minimum temperature during the winter season with the highest skill, at -0.20, when compared again Environment Canada's forecasts. With the most accurate season forecast by WRF having negative skill, the WRF model had lower accuracy when forecasting minimum temperatures when compared against forecasts issued by Environment Canada's weather office during this year of study.

3.4. Discussion and Conclusions

The WRF model showed seasonal differences in forecast accuracy during this year of study. WRF forecast summer and spring with greater accuracy than the remaining seasons, with maximum temperatures the most accurate during the summer season and minimum temperature and wind speed forecasts were the most accurate during the spring season. The fall season was the most accurate for pressure forecasts, but had the least accuracy when forecasting minimum temperatures. The winter season had the least accuracy for maximum temperate, pressure and wind speed forecasts.

The Day 2 forecast was always slightly less accurate than the Day 1 forecast. In addition, WRF had certain biases forecasting further into the forecast length. Day 2 forecasts tended to have warmer minimum temperatures, higher wind speed, and lower pressure than compared to a Day 1 forecast, for all seasons except winter. WRF forecasts during the winter (summer) over-predicted (under-predicted) the 1200 UTC pressure value, while both spring and fall had relatively low ME. WRF had some difficulties forecasting the diurnal temperature range, typically forecasting a colder maximum temperature and warmer minimum temperature than observed. WRF forecasts for minimum temperature had greater RMSE values than maximum temperature, resulting with less accuracy when forecasting minimum temperatures.

WRF had slightly greater accuracy when forecasting precipitation across the summer season than compared to the averaged year. This finding is beneficial due to the heavy precipitation that usually occurs during summer in Alberta, often with severe convection. While WRF produces more precipitation than observed, likely due to the overactive Kain-Fritsch cumulus scheme, the ET scores remained positive, indicating the model had more skill than random forecasts. ET scores were also lower for light and heavy thresholds than for medium thresholds, for Day 1 forecasts, as other studies have

found (e.g. Cherubini et al. 2001; Colle et al. 1999). The greater FAR values (40-61%) are partially attributed to the overactive Kain-Fritsch cumulus scheme.

WRF had greater accuracy than climatology and persistence forecasts when forecasting the daily maximum temperature. WRF had greater errors associated with forecasting the daily minimum temperature, and, as a result, had less skill when forecasting minimum temperatures. While the WRF daily maximum temperature forecast was skillful, WRF's minimum temperature forecast was inaccurate compared to all other forecast methods. Using the minimum temperature forecast produced by WRF within the same region of study, and for similar configurations as this project, is not recommend without using other adjustments to improve accuracy. WRF was less accurate at forecasting temperatures, both maximum and minimum, when compared against forecasts issued by Environment Canada's Weather Office.

3.5. Tables

Table 3-1: List of the nine observation stations used for the verification of the WRF model for Alberta.

Location	Station	Latitude	Longitude	Elevation (m)
Edmonton International Airport	YEG	53.317°N	113.583°W	723
Calgary International Airport	YYC	51.117°N	114.017°W	1084
City of Pincher Creek	ZPC	49.517°N	113.983°W	1189
Lethbridge County Airport	YQL	49.633°N	112.800°W	929
Medicine Hat Airport	YXH	50.033°N	110.717°W	716
Lloydminster Airport	YLL	53.317°N	110.067°W	668
Edson Airport	YET	53.579°N	116.465°W	926
City of Coronation	WCT	52.067°N	111.450°W	791
Red Deer Regional Airport	YQF	52.183°N	113.900°W	904

Table 3-2: Characteristics of WRF forecasts used in this study.

Model Parameter:	Option selected for WRF forecast:
Nesting Option	Nest: 15 km inside a 45 km parent domain
Land Surface Model	MODIS
Projection	Polar Stereographic
WRF Core	Advanced Research WRF (ARW), Non hydrostatic
Simulation Length	69 hours
Data set: Boundary and Initial Conditions	NAM212
Initialization Time	1200 UTC
Microphysics	Lin et al. (1983)
Cumulus Parameterization Scheme	Inner and outer domain: Kain Fritsch (Kain and Fritsch 1990)
Planetary Boundary Layer Scheme	Yonsei University (Hong et al. 2006)
Land Surface Physics Scheme	Noah

Table 3-3: WRF root mean squared errors (RMSE) and mean errors (ME) for Day 1 forecasts during the summer season, evaluated at each observation station used in the verification process.

Station	RMSE Maximum Temperature (°C)	RMSE Minimum Temperature (°C)	RMSE Wind Speed (km/h)	ME Maximum Temperature (°C)	ME Minimum Temperature (°C)	ME Wind Speed (km/h)
YEG	1.79	3.19	7.02	0.07	2.37	2.71
YYC	2.15	3.93	6.6	0.6	3.45	1.62
ZPC	2.26	3.31	7.92	-1.13	2.25	-1.13
YQL	1.92	4.16	6.29	0.53	3.34	0.85
YXH	2.33	4.59	7.94	-0.89	3.9	4.66
YLL	1.81	2.29	5.86	0.19	1.01	1.35
YET	2.4	3.21	8.27	-0.81	2.34	5.79
WCT	2.22	3.93	6.44	-0.23	3.32	0.15
YQF	1.81	1.94	8.24	-0.48	0.51	4.46
Average	2.08	3.39	7.18	-0.24	2.50	2.27

Table 3-4: WRF root mean squared errors (RMSE) and mean errors (ME) for Day 2 forecasts during the summer season, evaluated at each observation station used in the verification process.

Station	RMSE Maximum Temperature (°C)	RMSE Minimum Temperature (°C)	RMSE Wind Speed (km/h)	ME Maximum Temperature (°C)	ME Minimum Temperature (°C)	ME Wind Speed (km/h)
YEG	1.92	3.32	8.16	-0.11	2.39	3.69
YYC	2.40	4.35	6.79	0.64	3.82	1.60
ZPC	2.57	3.21	8.14	-1.12	2.16	-0.88
YQL	2.43	4.25	6.37	0.58	3.49	0.93
YXH	2.51	4.75	8.26	-1.11	4.00	4.13
YLL	2.07	2.31	7.97	0.03	1.14	0.27
YET	2.70	3.30	7.92	-1.14	2.40	5.49
WCT	2.51	4.31	8.16	-0.19	3.74	1.56
YQF	2.24	2.34	8.79	-0.40	0.92	5.36
Average	2.37	3.57	7.84	-0.31	2.67	2.46

Table 3-5: WRF mean errors (ME) and root mean squared errors (RMSE) for Day 1 forecasts, averaged from 9 weather stations, for all seasons. A positive ME indicates a forecast with a higher value than observed, while negative numbers indicate a larger observed value than forecast. Seasons with the ME closest to zero were forecast with the least bias, and are indicated in bold. Seasons with the lowest RMSE were forecast with the greatest accuracy, and are indicated in bold.

	Day 1	Spring	Summer	Fall	Winter	Year
Maximum Temperature (°C)	ME	-1.31	-0.24	-1.20	-1.63	-1.08
	RMSE	2.74	2.08	2.73	3.48	2.82
Minimum Temperature (°C)	ME	1.08	2.50	3.98	1.98	2.43
	RMSE	3.24	3.39	5.07	4.54	4.19
Wind speed (km/h)	ME	0.77	2.27	2.76	3.44	2.37
	RMSE	7.12	7.17	9.18	10.22	8.59
Pressure (mb)	ME	0.02	-1.40	0.22	1.62	0.11
	RMSE	2.32	2.32	2.12	3.13	2.51

Table 3-6: WRF mean errors (ME) and root mean squared errors (RMSE) for Day 2 forecasts, averaged from 9 weather stations, for all seasons. A positive ME indicates a forecast with a higher value than observed, while negative numbers indicate a larger observed value than forecast. Seasons with the ME closest to zero were forecast with the least bias, and are indicated in bold. Seasons with the lowest RMSE were forecast with the greatest accuracy, and are indicated in bold.

	Day 2	Spring	Summer	Fall	Winter	Year
Maximum Temperature (°C)	ME	-1.24	-0.31	-1.20	-1.94	-1.17
	RMSE	3.03	2.37	2.94	3.63	3.04
Minimum Temperature (°C)	ME	1.38	2.67	4.06	1.76	2.51
	RMSE	3.54	3.57	5.21	4.48	4.31
Wind speed (km/h)	ME	0.82	2.46	3.05	2.79	2.33
	RMSE	7.54	7.84	9.81	10.76	9.17
Pressure (mb)	ME	-0.28	-1.85	0.05	1.93	-0.03
	RMSE	3.13	3.18	2.59	3.86	3.23

Table 3-7: Precipitation verification for WRF forecasts during summer 2011 and the entire year of data for both Day 1 and Day 2 forecasts, averaged from 9 weather stations. Three different precipitation accumulation thresholds are used: 0.2 mm, 1.0 mm, and 5.0 mm. The most accurate value for POD (Probability of Detection), FAR (False Alarm Ratio), BIAS, and ET (Equitable Threat) for the given threshold is in bold.

Forecast Day	Threshold	POD		FAR		BIAS		ET	
		Summer	Year	Summer	Year	Summer	Year	Summer	Year
Day 1	0.2 mm	0.87	0.78	0.40	0.49	1.47	1.57	0.27	0.28
Day 1	1.0 mm	0.81	0.70	0.43	0.52	1.46	1.50	0.30	0.30
Day 1	5.0 mm	0.58	0.54	0.60	0.61	1.51	1.42	0.22	0.26
Day 2	0.2 mm	0.82	0.75	0.40	0.51	1.41	1.58	0.23	0.25
Day 2	1.0 mm	0.71	0.64	0.46	0.55	1.35	1.48	0.23	0.25
Day 2	5.0 mm	0.57	0.53	0.58	0.61	1.42	1.41	0.22	0.26

Table 3-8: Day 1 WRF skill scores when compared against three other forecast methods: climatology, persistence, and Environment Canada's (EC) forecast. Skill was found for maximum and minimum temperatures for each station for the entire year of study.

Positive skill score indicates that WRF forecasts were more accurate, while negative skill score indicates the opposite.

Station	Climatology Maximum	Climatology Minimum	Persistence Maximum	Persistence Minimum	EC Maximum	EC Minimum
YEG	0.80	0.28	0.59	-0.08	0.00	-0.73
YYC	0.78	0.42	0.65	-0.14	-0.17	-0.71
ZPC	0.81	0.61	0.53	0.36	-0.34	-0.32
YQL	0.82	0.46	0.69	0.22	0.04	-0.60
YXH	0.80	0.41	0.59	-0.09	-0.57	-1.17
YLL	0.82	0.61	0.64	0.16	0.06	-0.76
YET	0.75	0.18	0.58	-0.16	-0.38	-1.10
WCT	0.86	0.09	0.66	-0.73	0.14	-1.40
YQF	0.79	0.49	0.60	0.16	-0.07	-0.44
Average	0.80	0.39	0.61	-0.03	-0.14	-0.81

Table 3-9: WRF skill score values for maximum and minimum temperature were compared to three other forecast methods: Climatology, persistence, and Environment Canada's (EC) forecast. Positive skill score values indicate that WRF was more accurate than the forecast method being compared against WRF, while negative skill indicates the opposite. Values are averaged from nine weather stations. Note: the winter of 2011-2012 was very mild, indicated by high skill scores for WRF when compared to the climatology forecast.

Season	Climatology Maximum	Climatology Minimum	Persistence Maximum	Persistence Minimum	EC Maximum	EC Minimum
Day 1						
Spring	0.66	-0.02	0.26	-0.18	-0.84	-0.57
Summer	0.72	-0.72	0.71	-0.59	-0.29	-2.06
Fall	0.81	-0.07	0.69	-0.43	0.00	-1.69
Winter	0.84	0.70	0.60	0.35	-0.01	-0.20
Year 2011-2012	0.80	0.39	0.61	-0.03	-0.14	-0.81
Day 2						
Spring	0.59	-0.21	0.53	0.06	-0.42	-0.62
Summer	0.64	-0.94	0.78	-0.17	-0.46	-2.12
Fall	0.79	-0.14	0.80	0.12	-0.08	-1.44
Winter	0.82	0.71	0.77	0.66	0.02	-0.15
Year 2011-2012	0.77	0.36	0.76	0.37	-0.13	-0.76

3.6. Figures

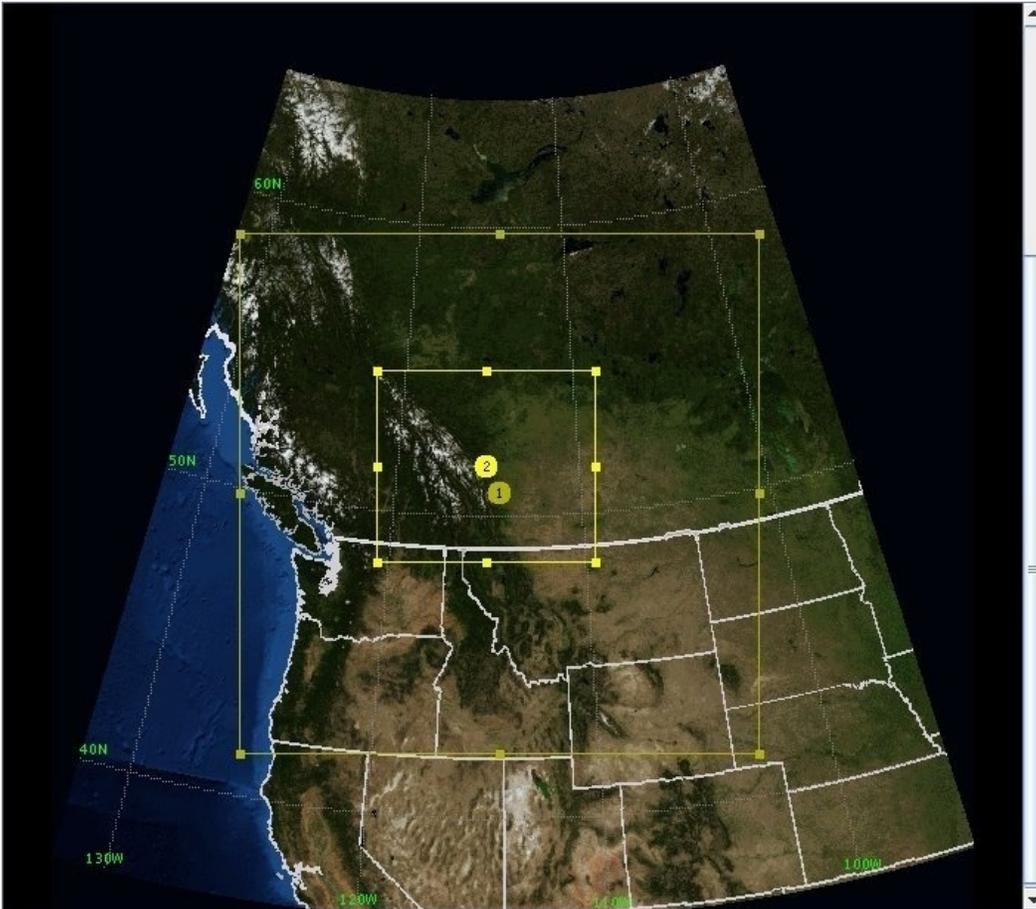


Figure 3-1: Numerical Weather Prediction domain setup for WRF forecasts using a grid resolution of 45 km for the outer domain and 15 km for the inner (nested) domain. The center point of each domain is indicated by the yellow circle while the number refers to the domain; the larger domain is 1, and the nested domain is 2.

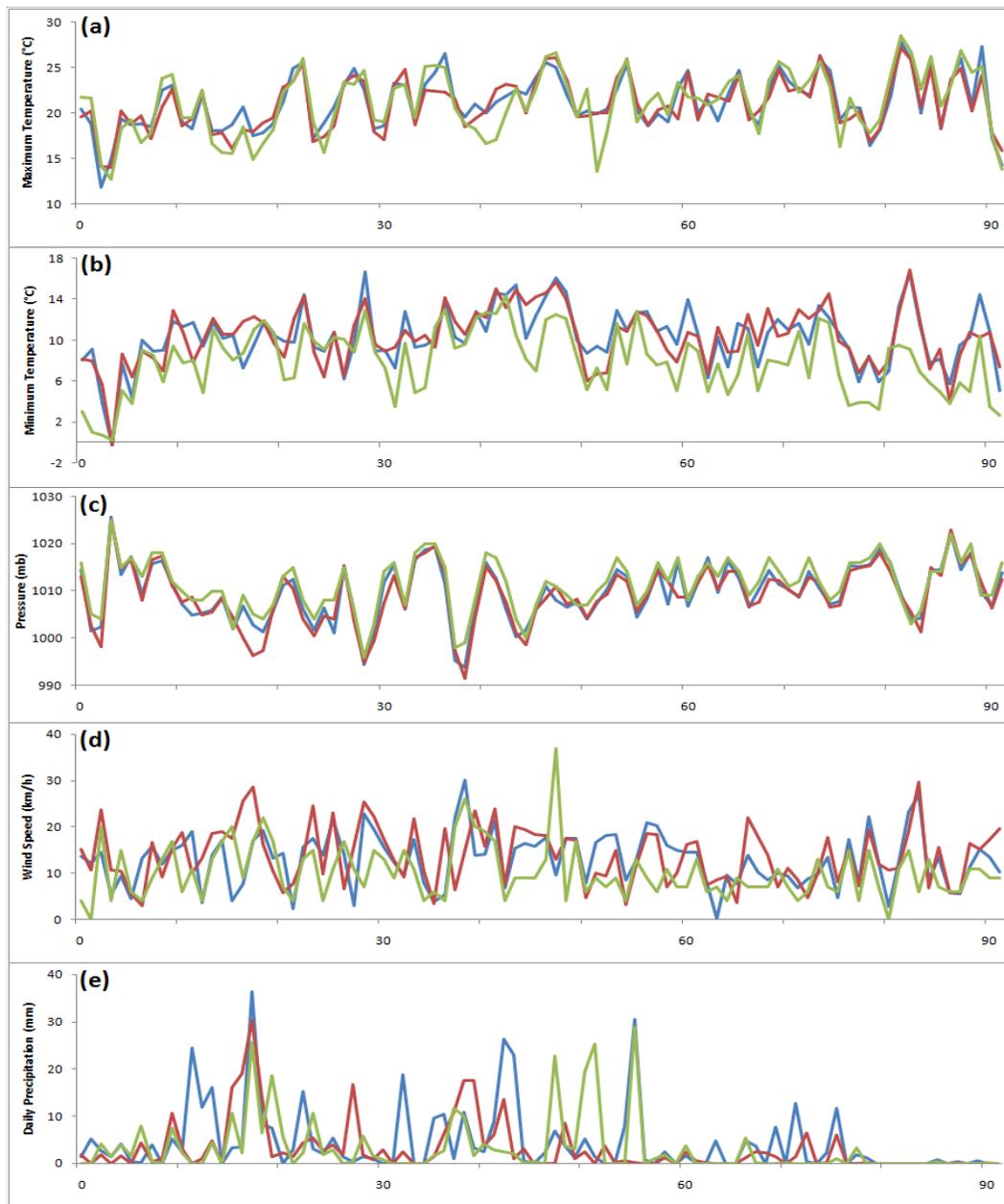


Figure 3-2: Time series of daily WRF forecasts and observed meteorological variables at Edmonton International Airport, from 1 June 2011 to 31 August 2011. Day 1 forecast is blue, Day 2 forecast is red, and the observations are in green. The variables are: (a) maximum 2-meter temperature in °C, (b) minimum 2-meter temperature in °C, (c) 1200 UTC sea level pressure in millibars, (d) 1200 UTC wind speed in km/h, (e) and precipitation accumulation in millimeters. The horizontal axis is the day in the series, starting from 1 June 2011.

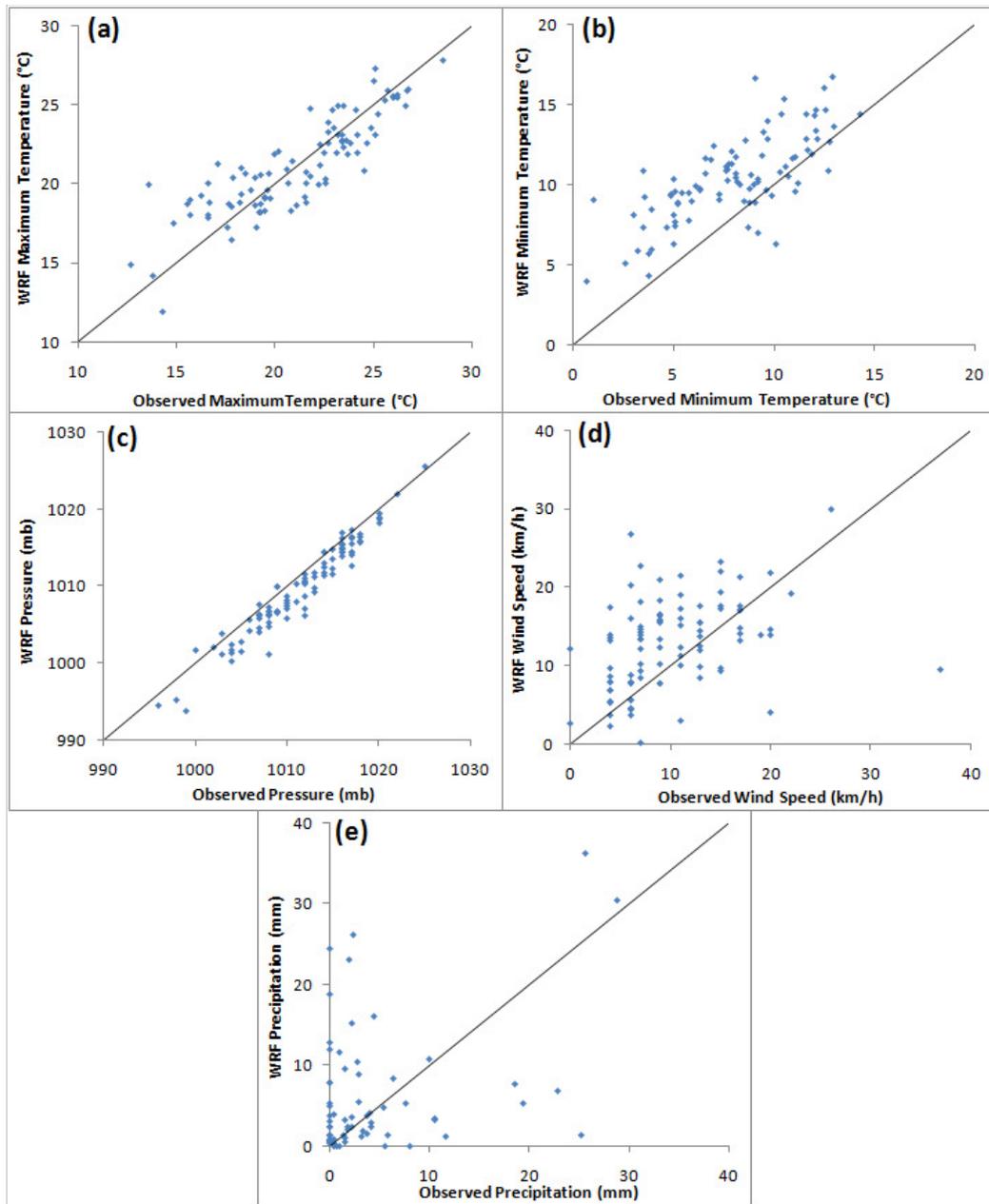


Figure 3-3: Comparison between WRF Day 1 forecast variables and observed values, at Edmonton International Airport during the summer season. Variables compared are: (a) maximum 2-meter temperature in °C, (b) minimum 2-meter temperature in °C, (c) 1200 UTC sea level pressure in millibars, (d) 1200 UTC wind speed in km/h, and (e) precipitation accumulation in millimeters for a threshold above 0.2 mm. The solid diagonal line indicates where forecast values equal observed values: a perfect forecast.

3.6. References

- Black, T. L., 1994: The New NMC Mesoscale Eta Model: Description and Forecast Examples, *Wea. Forecasting*, **9**, 265-278
- Borge, R., V. Alexandrov, J.J. Vas, J. Lumberras, E. Rodriguez, 2008: A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula, *Atmos. Envir.*, **42**, 8560-8574
- Brown B. G., R. R. Bullock, C. A. Davis, J. H. Gotway, M. B. Chapman, A. Takacs, E. Gilleland, K. Manning, 2004: New Verification Approaches for Convective Weather Forecasts, Preparings, *11th conference on Aviation, Range, and Aerospace*, Hyannis, MA, American Meteorological Society, 9.4
- Cherubini T., A. Ghellia, F. Lalaurette, 2001: Verification of Precipitation Forecasts over the Alpine Region Using a High-Density Observing Network, *Wea. Forecasting*, **17**, 238-249
- Colle, B. A., K. J. Westrick, C. F. Mass, 1999: Evaluation of MM5 and Eta-10 Precipitation Forecasts over the Pacific Northwest during the Cool Season, *Wea. Forecasting*, **14**, 137-154
- Colle B. A., C. F. Mass, K. J. Westrick, 2000: MM5 Precipitation Verification over the Pacific Northwest during the 1997-99 Cool Seasons*, *Wea. Forecasting*, **15**, 730-744
- Colle B. A., J. B. Olson, J. S. Tongue, 2003: Multiseason Verification of the MM5. Part II: Evaluation of High- Resolution Precipitation Forecasts over the Northeastern United States, *Wea. Forecasting*, **18**, 458-480
- Chen M., S. Fan, J. Zhong, X. Huang, Y. Guo, W. Wang, Y. Wang, B. Kuo, 2009: A WRF-based rapid updating cycling forecast system of BMB and its performance during the summer and Olympic Games 2008, Proceedings, *World Meteorological Organization Symposium on Nowcasting and Very Short Range Forecasting*, WMO, Whistler, British Columbia, Canada, 31 Aug-4 Sep
- Cheng W. Y. Y. and W. J. Steenburgh, 2005: Evaluation of Surface Sensible Weather Forecasts by the WRF and the Eta Models over the Western United States, *Wea. Forecasting*, **20**, 812-821
- Dallavalle, J. P. and V. J. Dagostaro, 2004: Objective Interpretation of Numerical Weather Prediction Model Output - A Perspective Based on Verification of Temperature and Precipitation Guidance, *Symposium 50th Anniversary of Operational Numerical Weather Prediction*, College Park, Md, NCEP, 5.8

- Davis, C., B. Brown, R. Bullock, 2006: Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas, *Mon. Wea. Rev.*, **134**, 1772-1784
- Ebert, E. E. and J. L. McBride, 2000: Verification of Precipitation in Weather Systems: Determination of systematic errors, *J. of Hydrol.*, **239**, 179-202
- Frauenfeld O W., T. Zhang, M. C. Serreze, 2005: Climate change and variability using European Centre for Medium-Range Weather Forecasts reanalysis (ERA-40) temperatures on the Tibetan Plateau, *J. of Geo. Research*, **110**, D02101, doi: 10.1029/2004JD005230
- Galanis G. and M. Anadranistakis, 2002: A one-dimensional Kalman filter for the correction of near surface temperature forecasts, *Meteor. App.*, **9**, 437-441
- Gochis D. J., W. J. Shuttleworth, Z. Yang, 2002: Hydrometeorological Response of the Modeled North American Monsoon to Convective Parameterization, *J. of Hydrometeor.*, **4**, 235-250
- Grell, G.A., S.E. Peckham, R. Schmitz, S.A. McKeen, G. Frost, W.C. Skamarock, B. Eder, 2005: Fully coupled "online" chemistry within the WRF model, *Atmos. Envir.*, **39**, 6957-6975
- Hong. S.-Y., Y. Noh, J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318-2341
- Hu X-M., J.W. Nielsen-Gammon, Z. Fuqing, 2010: Evaluation of Three Planetary Boundary Layer Schemes in the WRF model, *J. of Applied Meteorology and Climatology*, **49**, 1831-1844
- Kain, J.S. and J.M. Fritsch, 1990. A One-Dimensional Entraining/Detraining Plume Model and Its Application in Convection Parameterization, *J. of the Atmos. Sci.*, **47**, 2784-2802
- Liang, X., L. Li, K. E. Kunkel, M. Ting, J. X. L. Wang, 2004: Regional Climate Model Simulation of U.S. Precipitation during 1982-2002. Part I: Annual Cycle, *J. of Climat.*, **17**, 3510-3529
- Lin, Y-L., R.D. Farley, H.D. Orville, 1983. Bulk Parameterization of the Snow Field in a Cloud Model, *J. of Clim. and Appl. Meteor.*, **22**, 1065-1093
- Mass C. F., D. Ovens, K. Westrick, B. A. Colle, 2002: Does Increasing Horizontal Resolution Produce More skillful Forecasts?, *Bull. Amer. Meteor. Soc.*, **83**, 407-430

- McBride, J. L. and E. E. Ebert, 2000: Verification of Quantitative Precipitation Forecasts from Operational Numerical Weather Prediction Models over Australia, *Wea. Forecasting*, **15**, 103-121
- Michalakes, J., S. Chen, J. Dudhia, L. Hart, J. Klemp, J. Middlecoff, W. Skamarock 2001: "Development of a Next Generation Regional Weather Research and Forecast Model" in *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the Use of High performance Computing in Meteorology*. Eds. Walter Zwiefelhofer and Norbert Kreitz. World Scientific, Singapore. 269-276
- Murphy, A. H., 1988: Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient, *Mon. Wea. Rev.*, **116**, 2417-2424
- Murphy, A. H., B. G. Brown, Y. Chen, 1989: Diagnostic Verification of Temperature Forecasts, *Wea. Forecasting*, **4**, 485-501
- Patton, E.G., and J.L. Coen, 2004: WRF-Fire: A coupled atmosphere-fire module for WRF. Preprints of Joint MM5/Weather Research and Forecasting Model Users' Workshop, Boulder, CO, June 22-25. NCAR, 221-223
- Reuter, G. W. and C. D. Nguyen, 1993: Organization of the cloud and precipitation in an Alberta storm. *Atmospheric Research*, **30**, 127-141
- Richardson, L.F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, London.
- Rife, D. L. and C. A. Davis, 2005, Verification of Temporal Variations in Mesoscale Numerical Wind Forecasts, *Mon. Wea. Rev.*, **113**, 3368-3381
- Roebber, P.J. and L.F. Bosart, 1998: The sensitivity precipitation to circulation details. *Mon. Wea. Rev.*, **126**, 437-455.
- Rozumalski, R. A., 2006: SOO/STRC WRF EMS User's Guide. Available online at <http://strc.comet.ucar.edu/wrf>
- Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Baker, M. G. Duda, X. G. Huang, W. Wang, J. G. Powers, 2008: A Description of the Advanced Research WRF Version 3, *NCAR technical note NCAR/TN-475+STR*, 113 pp
- Vincent C. L., W. Bourke, J. D. Kepert, M. Chattopadhyay, Y. Ma, P. J. Steinle, C. I.W. Tingwell, 2008: Verification of a high-resolution Mesoscale NWP System, *Australian Meteorological Magazine*, **57**, 213-233
- Wang, W. and N. S. Seaman, 1997: A Comparison Study of Convective Parameterization Schemes in a Mesoscale Model, *Mon. Wea. Rev.*, **125**, 252-278

Warner T. T., R. A. Petersen, R. E. Treadon, 1997: A tutorial on land boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction, *Bull. Amer. Meteor. Soc.*, **78**, 2599-2617

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Chapter 4: Summary and Conclusions

4.1. Summary

This study verified the Weather Research and Forecasting model when simulating heavy precipitation events and determined the accuracy when forecasting weather for southern Alberta. The model output was verified against observations from weather stations to determine the differences between simulations and observations to identify the accuracy of WRF's forecasts.

WRF was used to simulate three historic events in central and southern Alberta which caused extensive damage due to heavy precipitation and related flooding. Each storm was simulated using different cumulus parameterization schemes and varying grid resolution. The simulated precipitation was compared against the observed precipitation. An evaluation was performed across the major river basins in Alberta to determine which cumulus parameterization scheme simulated the precipitation the most accurately. This was performed for three grid resolutions: 30 km, 15 km, and 6 km.

In addition to the simulated precipitation analysis, WRF's forecast accuracy was analyzed. WRF forecasts were produced for each day from 15 March 2011 to 29 Feb 2012. Each forecast provided the maximum and minimum temperature, daily precipitation accumulation, 1200 UTC wind speed and surface pressure each day for the following 2 days. The forecasts were verified against observations from nine weather stations to determine the seasonal differences in forecast accuracy. In addition, WRF temperature forecasts were compared against other forecast methods to determine how skillful the WRF model was in predicting temperatures.

4.2. Conclusions of WRF precipitation simulations

Three summer-time heavy precipitation events were simulated using the WRF model. These simulations were performed using the Kain-Fritsch, Betts-Miller-Janjic, Grell-Devenyi, Grell-Devenyi 3D ensemble, and the explicit cumulus parameterization schemes. The simulations were performed at 15 km grid resolution, a coarse resolution of 30 km, and a high resolution of 6 km. The results show that the Kain-Fritsch cumulus parameterization scheme was the most accurate scheme when simulating the heavy precipitation from these three events when evaluated over the major river basins of Alberta, and was slightly more accurate at the highest grid resolution of 6 km than at 15 km. However, the gain in accuracy was very small for the large increase in computational

time to perform the 6 km simulations. Simulations at 30 km were generally less accurate than simulations with higher resolution using the same cumulus schemes. The performance of the explicit scheme was comparable to that of the KF scheme, although with slightly less accuracy when simulated using a grid resolution of 15 km. The strengths of the KF cumulus parameterization scheme are thought to be due to the parameterization of entrainment/detrainment to and from convective clouds (Gochis 2002) as well as the closure assumption around convective available potential energy (CAPE) values (Wang 1997).

4.3. Conclusions of WRF daily forecasts

Daily forecasts were performed from 15 March 2011 through 29 February 2012 using the WRF model over central and southern Alberta. The year-long analysis for various near-surface parameters indicates that the spring and summer seasons were forecast more accurately than winter and fall during this period. Spring and summer had the lowest temperature and wind errors compared to the remaining seasons. The winter season was the least accurate, indicated by higher error values. The precipitation results indicate that WRF precipitation forecasts had skill during all seasons, with ET values greater than 0.0 for all thresholds. WRF precipitation forecasts were slightly more accurate during the summer season compared to the rest of the year; a useful finding considering the potential for severe storms and significant precipitation during summer.

Some model biases were found regarding these forecasts. At a longer forecast length (Day 2), WRF tended to forecast warmer daily minimum temperatures, higher wind speeds, and lower pressure than compared to the same day when forecast at a shorter length (Day 1). This was observed for all seasons except winter. In addition, WRF forecast a smaller range of diurnal temperatures than was observed, forecasting a warmer minimum temperature and colder maximum temperature, for each season during the year of study. WRF forecasts for maximum temperature were much more accurate than minimum temperature forecasts.

The WRF temperature forecasts were compared against other methods of forecasting: climatology, persistence, and Environment Canada's Weather Office. The WRF skill scores indicate that WRF maximum temperature forecasts were generally more accurate than a forecast using climatology or persistence, while the opposite is true for WRF minimum temperature forecasts. WRF temperature forecasts, both maximum and minimum, were shown to be less accurate than forecasts issued by Environment

Canada's Weather Office. The skill score results were consistent across all seasons analyzed; although skill scores did vary from season to season.

4.4. Comments and suggestions for future work

This study was designed to verify the accuracy of the Weather Research and Forecasting model to simulate heavy summertime precipitation storms, as well as daily forecasts for near-surface parameters. The following are some suggestions for further research which would complement this project.

The Kain-Fritsch cumulus scheme was found to be the most accurate cumulus parameterization scheme when simulating three heavy precipitation events. However, it is not known how accurately KF, or the other CPS, will simulate other important aspects of severe summer storms, such as storm tracks, damaging hail, or strong downbursts. Further research into the verification of cumulus parameterization schemes would provide useful information regarding the strengths and weaknesses of each scheme when simulating severe storms using the WRF model.

The heavy storms were verified by using different cumulus parameterization schemes to determine which scheme was the most accurate. There are a variety of other methods to fine tune a model to increase accuracy. For example, changing the microphysics schemes could result in a different accumulated precipitation amount across the area of study. Changing the microphysics schemes may result in larger differences in precipitation than changing the cumulus schemes, potentially more accurate as well. A study to examine how the various microphysics schemes of WRF can simulate heavy precipitation events would be a useful follow-up to the results presented in Chapter 2. In addition to the microphysics schemes, a verification study on how the various planetary boundary layer schemes produce precipitation and other forecast parameters would follow the work of this thesis rather well. Borge et al (2008) performed many different simulations using different configuration settings, including planetary boundary layer and microphysics schemes, using the WRF model and examined wind speed and direction, temperature, and humidity, but did not examine how the configurations affected precipitation amounts.

The daily forecast results provided information regarding how accurate WRF was across the different seasons for near-surface parameters. There was no verification of any upper-air meteorological parameters in this study. While near-surface parameters are far

easier to observe, and thus verify, verification using WRF forecasts for upper-air levels would complement Chapter 3 of this study.

Lastly, as we noted earlier, WRF had difficulty forecasting near-surface parameters for areas that were inside a mountain valley. Due to terrain smoothing from using a grid resolution of 15 km, some valley locations had their simulated elevation much higher than their observed elevation, some by approximately 500 meters. The difference in heights caused their meteorological data to be rather inaccurate. For example, a severe cold bias in the temperature forecast was common, which prompted the removal of 2 weather stations (Jasper and Banff) from the initial analysis of 11. A study that determines a proper grid resolution for accurately simulating weather inside a mountain valley could be useful, as WRF has the capability to easily reconfigure a domain with embedded high resolution nested domains over mountain areas. I feel this suggestion is important because mountainous locations often have difficulty with proper radar coverage, as well as can suffer from severe meteorological conditions (flash flooding, heavy snowfall, avalanches induced from weather, etc.) which might go undetected due to being outside the radar coverage. Roeger et al. (2003) produced forecasts with different resolutions over the Rocky Mountains in British Columbia which experience avalanches and found that a coarse resolution (30 km) forecast produced precipitation with similar skill compared against high resolution (2km, 10km) forecasts. However, the 30 km forecast had the worst errors when evaluated for wind speed and was not very accurate for temperature forecasts, when compared against the higher resolution forecasts. This shows that high resolution forecasts have potential to be more useful in mountainous regions.

4.5. Concluding remarks

While it was noted that temperature forecasts produced by the WRF model were not as accurate as those from Environment Canada, it should be noted that the WRF model still can perform a valuable role as a numerical weather prediction model. Environment Canada produces easy to read forecasts for the general public, but does not produce spatial precipitation forecasts which are easily accessible by the public. It is possible, and very useful, to make computer scripts to easily create meaningful information from the WRF model output. It is not difficult to produce surface charts which show where a possible frost might occur, as well areas with heavy winds, or

predicting the amount of precipitation which will fall in each river basin. The combination of easily configuring the WRF model to your particular interests and the ability to create automatic scripts to perform meaningful tasks can be very productive. The WRF model is currently used by both researchers and in operations, showing how useful a NWP model can become once specifically tailored to your objectives.

4.6. References

- Gochis D. J., W. J. Shuttleworth, Z. Yang, 2002: Hydrometeorological Response of the Modeled North American Monsoon to Convective Parameterization, *J. of Hydrometeor.*, **4**, 235-250
- Roeger C, R. Stull, D McClung, J. Hacker, X. Deng, H. Modzelewski, 2003: Verification of Mesoscale Numerical Weather Forecasts in Mountainous Terrain for Application to Avalanche Prediction, *Wea. Fore.*, **18**, 1140-1160
- Wang, W. and N. S. Seaman, 1997: A Comparison Study of Convective Parameterization Schemes in a Mesoscale Model, *Mon. Wea. Rev.*, **125**, 252-278

Appendix

A: WRF Vertical Coordinate

The WRF equations use a terrain-following hydrostatic pressure vertical coordinate, η , defined as:

$$\eta = (\rho_h - \rho_{ht}) / \mu \quad \text{where } \mu = \rho_{hs} - \rho_{ht}$$

ρ_h is the hydrostatic component of the pressure, and ρ_{hs} and ρ_{ht} refer to the surface and top boundary (Skamarock et al. 2008). The vertical coordinate is a traditional sigma-coordinate, and varies from 0 at the upper boundary, to 1 at the surface.

B: Governing equations of the WRF model:

The WRF model uses the following flux-form Euler governing equations (Skamarock et al. 2008):

$$\partial_t U + (\nabla \cdot \mathbf{V}u) - \partial_x(\rho\phi_x) + \partial_\eta(\rho\phi_x) = F_U \quad (1)$$

$$\partial_t V + (\nabla \cdot \mathbf{V}v) - \partial_y(\rho\phi_y) + \partial_\eta(\rho\phi_y) = F_V \quad (2)$$

$$\partial_t W + (\nabla \cdot \mathbf{V}w) - g(\partial_\eta \rho - \mu) = F_W \quad (3)$$

$$\partial_t \Theta + (\nabla \cdot \mathbf{V}\theta) = F_\Theta \quad (4)$$

$$\partial_t \mu + (\nabla \cdot \mathbf{V}) = 0 \quad (5)$$

$$\partial_t \phi + \mu^{-1}[(\mathbf{V} \cdot \nabla \phi) - gW] = 0 \quad (6)$$

$$\partial_\eta \phi = -\alpha\mu \quad (7)$$

And the equation of state:

$$\rho = \rho_0(R_d\theta/\rho_0\alpha)^{\gamma} \quad (8)$$

Where the variables are:

\mathbf{V}	three dimensional coupled vector velocities
F	forcing term for U , V , W , and Θ
ρ	pressure
ρ_0	reference sea-level pressure
α	inverse density of air
μ	hydrostatic pressure difference between surface and top of the model
ϕ	geopotential
Θ	coupled potential temperature
g	acceleration due to gravity

η	terrain-following hydrostatic-pressure vertical coordinate
R_d	gas constant for dry air
Υ	ratio of heat capacities for dry air at constant pressure
U	coupled horizontal component of velocity in the x-direction
V	coupled horizontal component of velocity in the y-direction
W	coupled vertical component of velocity
u	horizontal component of velocity in x-direction
v	horizontal component of velocity in the y-direction
w	vertical component of velocity

Equations 1 and 2 are the horizontal momentum equations. Equation 3 is the vertical momentum equation. Equation 4 is the conservation equation for potential temperature. Equation 5 is the conservation equation for dry air. Equation 6 determines the geopotential, and is the only non-conservative equation for equations 1 through 6. Equation 7 determines the inverse density. Equation 8 is the equation of state.

The WRF model allows the use of four projections: the Lambert conformal, polar stereographic, Mercator, and latitude-longitude. Each grid point inside the WRF model has the same horizontal dimension. However, the grid points are then projected onto the Earth and are allowed to have slightly different sizes, to be fit properly to the map projection. This is performed by using map scale factors, which require the governing equations (1-7), and momentum variables, to be redefined (Skamarock et al. 2008).

The governing equations are further rewritten using perturbation variables to reduce truncation errors in the horizontal pressure gradient calculation, as well as errors in the vertical pressure gradient. The new variables are perturbations from a hydrostatically-balanced reference state. This causes changes to the momentum equations, the equation for mass conservation and the geopotential equation (Skamarock et al. 2008). These equations, as well as the equation of state, are solved by the WRF model, and contain a Coriolis term, mixing terms, as well as parameterized physics.

C: WRF boundary conditions

WRF simulations were initialized from a global model. WRF used the global model to produce boundary conditions for the outer domain, which was updated every 3 hours during the simulations. There are two lateral boundary regions in the coarse grid; the specified region and the relaxation region. The specified region is the outermost row

and column on each side of the outer domain. This region is set entirely by the interpolation from the global model. The relaxation region is the columns and rows adjacent to the specified region, and penetrates a certain amount of user-set grid columns/rows into the outer domain. The relaxation region is where WRF is relaxed towards the global model, with certain weighing values (Skamarock et al. 2008). This is performed every time the boundary conditions are updated, and produces a smooth mixture between the freshly updated boundary condition and the simulation.

D: Nesting

WRF simulations use two-way nesting between the coarse grid and the fine grid. The fine grid boundary conditions are interpolated from the coarse grid simulation. Two-way nesting also set the coarse grid solution as the mean of the fine grids inside each coarse grid. The mass, thermodynamic, scalar, and chemistry variables are averaged throughout the fine grid cells when interpolating back to the coarse grid. Horizontal momentum components are averaged from the fine grid cells which boarder the coarse grid's cell-face.

The nested domain generates boundary conditions in a similar method as the coarse domain. Whereas the coarse domain received boundary conditions from a global model, the nested boundary conditions are produced by the coarse domain. The coarse grid points which contain the fine grid perimeter are interpolated to the outer rows and columns of the nested domain. Unlike the coarse grid boundary conditions, there is no relaxation region inside the nested domain. The fine grid receives updated boundary conditions after each time-step of the coarse domain.

E. Grid point analysis:

The grid point precipitation analysis is performed by using the root-mean-squared-error (RMSE) between the observed grid point precipitation value, O , and the simulated grid point precipitation value, S , for every grid point, N , in the river basin:

$$RMSE = \sqrt{\frac{1}{N} \sum_{1}^{N} (S - O)^2}$$

F. Observation point analysis:

Table F-1: Precipitation contingency table, where each element (HIT, MISS, FALSE, and NONE) hold the number of observation stations in which the observation and simulation exceed or fail to exceed a precipitation threshold. For example, for a precipitation threshold of 5 mm, if there was 7 mm of observed precipitation (Rain) at an observation station and 3 mm of simulated precipitation (Non-event) at the same observation station, this would be a MISS, increasing the MISS counter by one.

		Observed:	
		Rain	Non-event
Simulated:	Rain	HIT	FALSE
	Non-event	MISS	NONE

For a given precipitation threshold, the observation point statistics can be found for each storm using the known number of hits, misses, and false events from the contingency table above. Hits will be designated in formula form as H, misses with M, false as F, and N will be the total number of observation stations which meet or exceed the threshold being evaluated.

The Probability of Detection (POD) is calculated as:

$$POD = \frac{H}{(H + M)}$$

The False Alarm Ratio (FAR) is calculated as:

$$FAR = \frac{F}{(H + F)}$$

The BIAS is calculated as:

$$BIAS = \frac{(H + F)}{(H + M)}$$

The Equitable Threat (ET) is calculated as:

$$ET = \frac{(H - E)}{(H + M + F - E)}$$

Where the variable E is:

$$E = \frac{(H + F) * (H + M)}{N}$$

G. Error analysis:

The error analysis uses the observation value, O_i for the i th station, and a forecast value, F_i for the i th station, to determine the Mean Error (ME) and Root Mean Squared Error (RMSE) at the i th station when performed for a period of time in days, N :

$$ME = \frac{1}{N} \sum_1^N (F_i - O_i)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_1^N (F_i - O_i)^2}$$

H. Skill score analysis:

Skill score, S , is found by using the ratio between the mean-squared-error (MSE) of WRF forecasts, MSE_{WRF} , and the MSE of a reference forecast, MSE_{REF} , where the reference forecast is climatology, persistence, or Environment Canada. MSE is found from the WRF temperature forecast value, F_{WRF_i} , at an i th station, and the reference temperature forecast value, F_{REF_i} , at the i th station, where O_i is the observed value at the i th station, when performed for a period of time in days, N :

$$MSE_{REF} = \frac{1}{N} \sum_1^N (F_{REF_i} - O_i)^2$$

$$MSE_{WRF} = \frac{1}{N} \sum_1^N (F_{WRF_i} - O_i)^2$$

$$S = 1 - \frac{MSE_{WRF}}{MSE_{REF}}$$