Forecasting Short-Term Road Surface Temperatures – A Neural Network-based Approach

by

Tasnia Nowrin

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

TRANSPORTATION ENGINEERING

Department of Civil and Environmental Engineering

University of Alberta

ABSTRACT

The ability to forecast road surface temperatures (RST) in advance is an important asset for winter maintenance (WRM) operations as it effectively reduces WRM costs through more efficient use of their maintenance resources, which in turn would improve road safety and mobility of traveling public during winter months. However, considering the number of road networks that must be monitored and the extent to which RST varies over time, WRM agencies have constantly been looking for better ways to generate accurate road weather forecasts as they strive to optimize their WRM services and provide safe travel conditions. Despite its significance, there exist significant gaps in knowledge and methods in current literature for generating reliable short-term RST forecasts. In addition, forecasting itself is a challenging task that often involves many unknown variables creating uncertainties through their interactions. As a result, forecasting accuracies may vary significantly across different regions – an important aspect that has been often neglected in the existing literature.

To tackle these issues, this thesis aims to develop highly reliable RST forecasting models using one of the most sophisticated and successful methods; namely, neural networks. In particular, this thesis attempts to assess the performances of the two uniquely developed neural network-based short-term forecasting models and investigate the hypothesis that geographical (e.g., latitude and altitude) attributes affect forecasting accuracy for improved generalization potentials.

RST measurements collected by six selected stationary road weather information systems (RWIS) stations in Alberta, Canada, were utilized to validate the feasibility and applicability of the proposed method developed herein. The findings indicate that the two developed models were found to generate highly accurate results. The first method using conventional artificial neural

network (ANN) models yielded forecasts with mean absolute error (MAE) values of 0.74, 1.34, 1.82, 2.40, 2.84, and 3.30°C for 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h ahead forecasts, respectively. The second method using long short-term memory (LSTM) models, on the other hand, generated even more promising results, with average MAE values of 0.63, 1.17, 1.79, 2.28, 2.74, 3.14°C for the 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h forecasting horizons, respectively.

The novelty of this study lies in investigating the probable effect of several external factors on model performance, where it was revealed that forecasting horizon and geographical attributes influenced forecasting accuracies. Upon investigating the hypothesis that locational attributes affect forecasting accuracies, the results showed that accuracy improved with increasing latitude and decreasing elevations, which are worthwhile findings that can potentially lead to developing more refined models for generating highly accurate location-specific RST forecasts.

PREFACE

The work presented in this thesis is either published or is under review for publication.

Conference paper

Nowrin, T., and Kwon, T.J. (2021). Forecasting Short-Term Road Surface Temperatures using Artificial Neural Network. *Presented at and Proceedings of the 101st Transportation Research Board Annual Conference,* Washington, D.C., United States, January 2022.

Journal paper (Submitted)

Nowrin, T., and Kwon, T.J. (2022). Forecasting Short-Term Road Surface Temperatures considering Forecasting Horizon and Geographical Attributes – An ANN-based Approach. Under review in the *Cold Regions Science and Technology.*

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Tae J. Kwon for his unwavering support, guidance, and encouragement during my studies. It has been a great privilege to find such a supportive supervisor, as his guidance has been invaluable. I also want to thank Dr. Juliana Leung, Dr. Tony Z. Qiu, and Dr. Mustafa Gul for spending their valuable time reviewing my work and being committee members at my final defense.

Secondly, I would like to acknowledge the financial and operational assistance from both Alberta Transportation and the Natural Sciences and Engineering Research Council of Canada (NSERC). I would like to thank Beata Bielkiewicz of Alberta Transportation in particular for sharing RWIS data, which was crucial to our research.

Thirdly, I want to express my appreciation to my wonderful colleagues, Simita Biswas, Mingjian Wu, Andy Wong, Shuoyan Xu, Davesh Sharma, Qian Xie, Xueru Ding, and other lab mates, for their insightful recommendations and support during my research.

Finally, I want to express my deepest gratitude to my beloved parents for their unconditional love, understanding, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

WRM	Winter Road Maintenance
ANN	Artificial Neural Network
LSTM	Long Short-Term Memory
RNN	Recurrent Neural Network
RWIS	Road Weather Information System
RST	Road Surface Temperature
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
RPU	Remote Processing Unit
LOS	Level of Service
FPD	Freezing Point Depressant
ASTM	American Society for Testing and Materials
GEM	Global Environmental Multiscale
MLP	Multilayer perceptron
ReLU	Rectified Linear Unit
ITS	Intelligent Transportation Systems
FHWA	Federal Highway Association
FFS	Free Flow Speed
RSC	Road Surface Conditions

Chapter 1 INTRODUCTION

1.1 Background

The performance of a country's transportation systems directly impacts the state of its economic development (Blum, 1997). A safe, efficient, and reliable transportation network not only facilitates a sustainable economy but also optimizes road users' experience in terms of safety and mobility. Nevertheless, the provision of such a system heavily depends on the effectiveness of the implemented road maintenances strategies. This is particularly true for countries that experience long winter seasons, where road maintenance personnel face the more challenging task of dealing with snow and ice, making it more difficult to ensure safety and mobility. As a result, the impacts of winter weather conditions have been actively studied. Nilsson and Obrenovic (1998) found that an individual is twice as likely to be involved in an accident during the winter (Nilsson and Obrenovic, 1998). According to the US Federal Highway Association (FHWA), 24% of weather-related vehicle crashes occur on snowy, slushy, or icy pavement, and 15% happen during snowfall or sleet each year (FHWA 2021). Similarly, a gazette published by Royal Canadian Mounted Police stated that nearly 30% of collisions reported to the National Collision Database occurred on wet, snowy, or icy roads in 2017 (R.C.M.P., 2019). Qiu and Nixon (2008) observed about an 84% increase in the crash rate, 75% increase in injury rate, and 9% increase in fatality rate due to snow precipitation. In addition to the adverse road safety issues caused by winter weather, researchers have found that winter road conditions can also significantly reduce accessibility and mobility. Agarwal et al. (2005) found that both capacity and speed could be reduced by 19.0%-27.0% and 11.0-15.0%, respectively. Another study conducted by Kwon et al. (2013) for an urban freeway in

Canada suggested that the free flow speed (FFS) and capacity reductions were 17.0% and 44.2%, respectively, for snow-covered road surface conditions.

Therefore, to ensure the safety and mobility of road network users, countries that experience severe winter weather place heavy emphasis on road maintenance strategies and spend billions each year. North America as a whole spends over $3 billion annually (Usman et al., 2010; Tsapakis et al., 2013; Kwon and Fu, 2021) on WRM, with Canada alone accounting for CAD 1.3 billion (Suggett et al., 2006). For this reason, the winter road maintenance (WRM) authorities and researchers are constantly looking for ways to minimize the cost and resources of WRM operations. The term WRM refers to a group of activities that serve to keep roads contaminant free. WRM includes mechanical operations such as sand spreading and snow plowing, chemical application to combat icing, and thermal techniques to prevent snow and ice from bonding to the pavement. Other than having an extensive toolkit to combat winter weather, it is also important for WRM personnel to determine where WRM is needed and what treatments are required. Therefore, every operation requires certain key decisions to be made beforehand. For this purpose, collecting and analyzing real-time read weather information is crucial. One such tool is the road weather information system (RWIS), which plays a vital role in decision-making for a wide variety of purposes, including measuring, recording, and communicating road weather information. RWIS is a combination of advanced sensor technologies embedded near a road segment of interest to capture road weather information such as atmospheric and pavement conditions. Atmospheric data include air temperature, humidity, visibility distance, wind speed, and direction. Pavement data include pavement temperature, pavement freezing point, pavement condition (e.g., wet, icy, flooded), pavement chemical concentration, and subsurface conditions (e.g., soil temperature).

Among the many conditions recorded, road surface temperature (RST) is considered one of the most important road weather variables for deciding between the different types of operations (e.g., anti-icing, de-icing, plowing, and sanding). Frequently applied freezing point depressants (FPDs) for roadway winter operations such as sodium chloride (NaCl), magnesium chloride ($MgCl_2$), and calcium chloride ($CaCl_2$) have distinct temperature ranges where they function best. (Ketcham et al., 1996 and Minsk, 1998). Anti-icing is an effective preventative technique used before the start of a winter storm or prior to a precipitation event to prevent the formation of ice or the development of compacted snow and it requires knowledge of RST ahead of time (Ketcham et al., 1996). Having accurate RST ultimately leads to better WRM planning, optimized resource-allocation, resulting in cost reductions. A report published by Iowa Department of Transportation in 2009 quantified the benefits and costs associated with road weather information forecasts and reported that the percentages of benefits over total winter maintenance costs were 5.6 percent in Iowa, 6.5 percent in Nevada, and 0.9 percent in Michigan (Iowa DOT, 2009). Furthermore, another study revealed that timely winter road maintenance focused on obtaining bare pavement conditions during heavy snowfall might minimise overall highway traffic delays by 5 to 36 percent, depending on the demand level (Shahdah and Fu, 2010). As a result, authorities have been exploring different ways to improve their ability to make WRM-related decisions based on real-time road weather information. One such techniques is forecasting of RST, which allows maintenance employees to forecast road surface conditions (RSC) ahead of time, enabling the deployment of more advanced and proactive activities, such as anti-icing operations prior to snowstorms. By doing so, a higher level of service can be provided, and at the same time reduce overall maintenance cost and minimize environmental damages from over-salting (Ramakrishna and Viraraghavan, 2005). Ultimately, efficient planning and decision support with real-time and near-future road weather information including RST can not only improve cost-efficiency and resource-optimization but also reduce pavement damage.

1.2 Problem Statement and Motivation

Maintenance authorities undertake WRM intending to control the adverse effects brought forth by winter weather. While appropriate treatments can reduce cost and environmental damage, timely operations are necessary to improve safety and mobility, which is why many authorities implement RWIS as it provides real-time road weather information. Among the data collected by RWIS, RST is the most influential variable on WRM operations; it provides information about road surfaces such as susceptibility to black ice formation, glazing, or frost, which is vital for implementing ice and snow removal operations. Using only real-time information may not be enough for WRM operations to be truly proactive. If RST can be forecasted ahead of time, it will enable maintenance personnel to develop mitigation strategies well before the weather event, which will indubitably increase the level of service.

As mentioned previously, most jurisdiction uses road weather information systems (RWIS) to make road surface measurements. Based on the meteorological variables collected by RWIS, site-specific road weather forecasts can be made. Hence, numerous models for predicting RST have been developed in the past. Depending on the forecasting objectives and parameters included in the model, the complexity and accuracy level of the models generated tend to differ significantly. Some of these models include simple meteorological elements like air temperature and humidity. In contrast, others employ extensive physical and mathematical modeling of solar radiation, heat flow rate, and other aspects. As early as 1957, Barber determined the maximum RST using infinite surface temperature medium periodic variations of heat conduction equations. Since then, many similar models have been developed that employ the same energy balance equation to predict RST using weather measurements from RWIS stations (Barber, 1957). Such models require many variables with complex equations and a significant amount of data to maintain. Moreover, due to the inability of these models to simulate the complex interactions between the model parameters at many problematic sites, a large degree of error is still present (Kršmanc et al., 2013).

Machine learning algorithms and neural networks have garnered attention in recent years due to recent advances in computer science. Neural networks are a relatively new type of statistical mapping that distinguishes itself by their capacity to generalize patterns that the model hasn't seen before. It is capable of learning and simulating nonlinear connections between input and target variables without requiring knowledge of the underlying mechanics, which is an edge over other parametric approaches in dealing with complicated situations. For this reason, assessing the applicability of such techniques has become an active field of research in many different sectors. Nevertheless, there is a lack of adequate research to evaluate the feasibility of these techniques to make road weather forecasts. Although prior studies have helped shape the current practices in forecasting RST, there is a resurgent need to further investigate and validate the feasibility of newer and more advanced technologies to overcome the limitations of current practices and utilize the power of computing advancements. Equally important, there is a lack of research on identifying the underlying external variables such as local weather and geographical characteristics; hence, it is crucial to examine how these factors affect forecasting accuracy.

1.3 Research Objectives

Despite the current widespread use of numerical models in forecasting road surface temperatures, it is also critical that the applicability of newer and more advanced technologies get explored to break the status quo and overcome current limitations. Furthermore, it is equally important to validate the performances of these techniques over varied geographic locations to detect the presence of spatial variations in forecasting accuracies, which will aid in developing more refined models in the future. To tackle these issues, this thesis develops neural network models using historical road surface measurement collected by RWIS. In particular, this thesis has the following objectives:

1) Conduct an extensive literature review on winter road maintenance operations and common practices for forecasting road surface temperatures;
2) Develop neural networks-based models for forecasting RST at various forecasting horizons for six spatially and geographically diverse areas in Alberta, Canada;
3) Assess and compare the overall forecasting accuracies of the developed models; and
4) Examine if locational attributes have any effect on forecasting accuracies across the six distinct stations and identify probable factors

1.4 Organization of Thesis

This thesis is comprised of five chapters. The rest of the content is organized as follows:

Chapter 2 reviews existing literature, covering current winter road monitoring techniques and numerous factors that may impact RST. Finally, it highlights the current state-of-the-art in estimating and predicting road surface temperatures (RST).

Chapter 3 initially describes the overall methodological approach for constructing neural network models. The discussion is then expanded to include the underlying theory of artificial neural networks (ANN) and long short-term memory (LSTM) algorithms.

Chapter 4 first presents real-world case studies to describe the study area, data collection, and processing techniques. Then, it illustrates the development of the ANN model and discusses the probable factors affecting the accuracies. After that, it presents the development of LSTM models and discussions on the results regarding the impact of probable factors. Finally, a comparative analysis between ANN and LSTM model performance is conducted to determine which of the two is the better model for forecasting RST.

Chapter 5 highlights the main findings and contributions of this research and potential extensions for future research.

Chapter 2 LITERATURE REVIEW

This chapter starts with an overview of the winter road maintenance operations and then describes the possible factors affecting road surface temperatures. After that, it provides an in-depth review of past efforts and current practices in applying, modeling, and forecasting road surface temperatures. A summary of this chapter is presented at the end of this section, along with a discussion on the limitations present in existing research.

2.1 Winter Road Maintenance

Winter road maintenance (WRM) operations refer to a combination of systems-based approaches to ensure safe access to infrastructures and the provision of an adequate level of service (LOS) in adverse weather conditions. Commonly used snow and ice control techniques are mechanical, chemical, and thermal methods. Mechanical removal methods involve the physical removal of snow or ice without any application of chemicals. On the other hand, chemical methods include the application of chemicals for de-icing and anti-icing purposes. In addition, there are thermal techniques that involve the application of heat to remove snow and ice or to prevent ice formation. Whether a road segment should be treated and which treatment method to use under a given set of conditions are difficult decisions to make. They must be made while considering budgetary constraints and the negative environmental impacts caused by the method chosen. Thus, the decision-making process is always facilitated by various road weather variables such as climatic conditions, site-specific factors, local rules of practice, available resources, target LOS, etc. Among the deciding factors, road surface temperature (RST) is considered one of the most critical factors for determining the most suitable maintenance operation under the given condition. This section provides an overview of the different WRM operations available and how they relate to RST.

Mechanical methods include strategies such as the physical removal of snow by plowing, brooming, gritting, blowing, or sand spreading. The purpose of mechanical removal is to pick up the loose or unbonded snow, shear it from the road if necessary, and relocate it to an area off the road (Perrier et al., 2006a). Spreading and plowing operations are commonly performed in rural and urban regions experiencing frozen precipitation or snowfall. In rural regions, snow is often moved to the edge of the roadway instead of disposing it to a separate storage location (Perrier et al., 2006b). However, in urban areas with high traffic volume and limited spacing at the edges of the roadway, the large volumes of plowed snow must be transported elsewhere for disposal (Hashemloo, 2008). Mechanical methods can be very efficient and fast for low priority roads, low LOS routes, or unpaved roads. Furthermore, to facilitate de-icing or anti-icing operations, snow plowing is performed first to remove as much snow and loose ice as possible before applying chemicals (Perrier et al., 2006a). In areas with low traffic, brooming can significantly reduce the need for chemical methods.

Thermal methods provide heat to the roadway surface to remove or prevent snow and ice buildup. They help reduce traffic delays and damages from accidents caused by black ice, glaze ice, or packed snow (Perrier et al., 2006a). However, the expense of installing and running a fixed or mobile heating system makes them unsuitable for widespread usage. They are usually installed in selected critical locations, including bridge decks, toll plazas, on and off-ramps, and steep grades (Perrier et al., 2006a).

Chemical methods use chemicals as freezing point depressants to melt ice. Chemicals play a significant role in WRM due to their high effectiveness and low cost compared to their alternatives.

The most frequently applied freezing point depressants (FPDs) for roadway winter operations are sodium chloride (NaCl), magnesium chloride ($MgCl_2$), and calcium chloride ($CaCl_2$), as well as calcium magnesium acetate (CMA) and potassium acetate (KAc) to a lesser extent (Ketcham et al., 1996). Chloride salts are one of the most extensively used and readily available salts, available in solid or liquid form (Shi et al., 2009). Among the chlorides, sodium chloride (NaCl) is the most widely used material for WRM due to its abundance and affordability (Fischel, 2001). Different chemicals have different temperature ranges where they perform best. The eutectic temperature and the lowest effective temperature are mainly used to assess the performance of certain materials at low temperatures. The eutectic temperature is defined as the lowest freezing point for a corresponding concentration of a eutectic mixture (Hashemloo, 2008; Keep et al., 2000). It is generally determined by laboratory tests, such as ASTM D1177. The miscibility or solubility of a chemical with a low eutectic temperature is likely to be high (Du et al., 2021). Moreover, the greater the difference between eutectic and ambient temperatures, the faster the snow and ice melt (Fischel, 2001). Therefore, the eutectic temperature of the chemical should be significantly below the expected ambient temperature to achieve a quick chemical dissolution and a low freezing point for the brine solution (Du et al., 2021). However, the eutectic temperature alone is insufficient to assess the material's performance, as most chemicals lose their effectiveness before the eutectic temperature is reached (Du et al., 2021). As a result, the effective temperature, which indicates the lowest temperature for compounds in practical usage, must be introduced (Fischel, 2001). The current practice in finding effective temperature involves field observations and anecdotal evidence of performance (Du et al., 2021). Typically, the lowest effective temperature is much greater than the eutectic temperature (Nixon et al., 2001). Table 2.1 shows the lowest effective temperatures and eutectic temperatures for the five frequently used chemicals mentioned above.

Table 2-1 Lowest effective temperature for different chemicals compared to the eutectic temperature (Source: Data obtained from Minsk, 1998 and Ketcham et al., 1996)

Chemical Agent	Lowest Effective Temperature (°C)	Eutectic Temperature (°C)	Eutectic Concentration (%)
Sodium Chloride (NaCl)	-7	-21	23.3
Calcium Chloride ($CaCl_2$)	-29	-51	29.8
Magnesium Chloride ($MgCl_2$)	-23	-33	21.6
CMA	-7	-27.5	32.5
Potassium Acetate (KAc)	-25	-60	49

There are two different approaches for applying these chemical freezing-point depressants on the road surface, and they are: de-icing and anti-icing, for which the main difference is their objective. Anti-icing is a preventive measure to inhibit the formation of ice or prevent the development of compacted snow tightly bonded to the pavement, making it easier to remove. Chemicals are applied at the commencement of a winter storm or previous to a precipitation event (Ketcham et al., 1996). Anti-icing is usually done using liquid ingredients, but dry or pre-wet solid granular materials can also be used (Nixon et al., 2001). In practice, liquid anti-icing operations should be conducted at pavement temperatures of about -5℃ (23°F) and above (Ketcham et al., 1996). Chemicals applied on the pavement will remain until they are dissolved by the precipitation during a snow storm event. As a result, anti-icing improves LOS, reduces the need for chemicals, saves money, and improves safety and mobility when compared to de-icing and sanding (Ketcham et al., 1996). Although anti-icing is extremely beneficial, it heavily depends on accurate weather forecasts. If inaccurate information is given to winter maintenance personnel, it can lead to excess material waste, resulting in financial and environmental damages. Therefore, effective anti-icing activities require reliable and accurate weather and pavement information, such as those provided by RWIS. In comparison, de-icing is a reactive operation conducted to break or melt the ice that has already formed on the pavement surface. A popular de-icing strategy is to wait until two centimeters or more of snow are present on the pavement surface before de-icing treatments (Ketcham et al., 1996). Due to its reactive nature, de-icing often provide less safety when compared to anti-icing (Du et al., 2021). Nevertheless, de-icing is still a viable option in many instances, for example, roads with a lower priority service level, instances where weather forecasts are erroneous, and when anti-icing activities are not possible (Du et al., 2021; Ketcham et al., 1996). Besides these two strategies, pre-wetting dry salt with a liquid chemical before spreading it on the pavement surface is another efficient strategy. Pre-wetting is the process of wetting salt aggregates before spreading them on a road surface. This strategy is more economical than applying solid salt alone because it reduces salt usage by starting the salt solution with a primary liquid and guarantees that the salt will stick to the road surface (Minsk, 1998; Burtwell et al., 2001; Blackburn et al., 2004).

Although these chemicals allow many roads to remain operable during winter storms, various studies have found that they can have negative consequences such as motor vehicle corrosion, infrastructural damage, degradation of roadside vegetation, and sodium intrusion of drinking water (Perrier et al., 2006a).

The use of sand and other abrasives is another common technique in WRM operations, one that is considered relatively inexpensive and effective in improving pavement surface traction. However, they can not prevent nor break the bond between pavement and ice, rendering them only appropriate in low-temperature scenarios where chemical treatments are ineffective due to the strong bond between contaminant and pavement. (Perrier et al., 2006a). An example use case is when the temperature drops below -12°C, at which point sand is applied to provide additional road surface traction to make up for the inability of chemical treatments to maintain bare pavement conditions. (Hashemloo, 2008). Despite sand and other abrasives being effective yet inexpensive, they have several possible detrimental consequences. These include damages to cars, additional highway cleanup costs, clogging of sewers and drains, airborne dust problems, and so on. Furthermore, the performance of abrasives has received mixed reviews. According to Ketcham et al. (1996), abrasives deliver no appreciable increase in friction or improvement in pavement condition on roads getting appropriately scheduled anti-icing treatments with successful prevention or mitigation of the hazards of packed snow. Therefore, abrasives should not be used as a routine operations of an anti-icing program.

This section provided an overview of WRM operations as well as illustrated the importance of road surface temperature in choosing chemical treatment agents and treatment application methods. The dependence of chemical treatments on road surface temperature highlighted why knowing the local road temperature ahead of time is critical for successful winter material applications. The next section will discuss the possible effects of various factors on RST.

2.2 Factors Affecting Road Surface Temperatures

Pavement temperature is affected by many interacting factors, including meteorological (i.e., solar radiation, wind speed, cloud cover, air temperature, humidity), geographical (i.e., latitude, altitude, topography, land-use), and road structure (i.e., construction materials, traffic). Some of these factors vary over time and space (i.e., weather, traffic), while others are permanent for a location (i.e., geography). This section discusses some of the critical factors influencing the variations in road surface temperatures (RST).

Latitude

For any location, its climate and RST are heavily impacted by its latitude via the laws of radiation geography, which affects the amount of solar radiation a place receives (Chapman et al., 2001). Consequently, winter seasons are longer and colder at higher latitudes as they receive less sunlight than low latitude locations.

Altitude

The ambient lapse rate causes RST to drop with altitude: up to 9.8℃ every 1000 m, but more often 6.5℃ per 1000 m. (Tabony, 1985). Another study by Shao et al. (1997) showed that the relationship between altitude and RST is typically non-linear. Moreover, the impacts of altitude on RST should be most noticeable when atmospheric stability is low; however, as stability grows, the greater involvement of topography should diminish this link (Chapman et al., 2001).

Topography

The topography is often thought to be the dominant factor in creating fluctuations in RST during severe nights (Bogren & Gustavsson, 1991). Small variations in topography can create considerable changes in air temperature and RST over the mesoscale landscape (Chapman et al., 2001). This phenomenon can be explained by the widely accepted katabatic theory, which states that a layer of thick cold air accumulates at the surface during steady circumstances, causing a temperature inversion. If this occurs in an undulating topography, the layer of cold air becomes mobile and descends as a katabatic flow, following drainage lines until it reaches a topographic or thermal barrier (Chapman et al., 2001). As a result, valley bottoms often experience the lowest temperature caused by the topography-induced fluctuations in temperature, which, in turn, affects RST due to the linear relationship between air and road surface temperature (Gustavsson, 1990).

However, RST tends to be slightly higher than the air temperature due to the thermal inertia of the road construction (Bogren & Gustavsson, 1991).

Screening and Sky-view Factors

Screening obstructs surface exposure to incoming daytime shortwave radiation. As a result, topographically screened areas are likely to yield lower RST (Bogren, 1991). Furthermore, buildings and forests provide wind shelter, which lowers the RST (Gustavsson, 1995).

The sky-view factor (Ψ_s) is a dimensionless parameterization used to quantify the amount of visible sky at a location. It is commonly used as a substitute for screening in RST investigations (Chapman et al., 2001). The variable itself is represented by a number between one and zero, one being perfectly flat and open terrain, and zero for sites with significant number of barriers such as buildings and trees (Oke,1992). Several studies have labelled Ψ_s as a controlling factor in RST variations (e.g., Bärring et al., 1985; Elliasson, 1996; Postgård & Nunez, 2000) and have found Ψ_s to play a crucial role in the nocturnal radiation budget. Often, it is observed that places with a lower value of Ψ_s experience increased nocturnal air and surface temperatures (Chapman et al., 2001).

Land-use Patterns

A phenomenon known as the urban heat island causes urban regions to be several degrees warmer than adjacent rural areas. One explanation for this observation is the significant difference in building and vegetation densities, leading to surface geometry variations. As a result, the effects of factors such as screening and sky-view factor cause microclimatic changes in these locations. Furthermore, the thermal properties of construction materials, as well as anthropogenic heat from buildings and transportation, impact urban climates (Chapman et al., 2001). Therefore, the land-use pattern can directly or indirectly influence the variations in RST.

Pavement Materials and Traffic

The construction materials of the pavement are important contributors influencing RST as heat absorbance, storage, and reradiation abilities of materials vary significantly. For example, asphalt has a higher surface temperature, heat storage capacity, and heat emission than concrete and bare soil (Asaeda et al., 1996). These materials absorb solar radiation during the daytime and emit it at night.

The impact of traffic can also significantly modify RST. Vehicles cause mixing of hot and cold air layers and act as shadow limiting longwave radiation loss from the road surface (Thornes, 1991). Moreover, slower-moving vehicles can produce increased RST of up to 2°C (Parmenter and Thornes, 1986). The effects of traffic on RST are more noticeable during the night (Chapman et al., 2001).

As discussed in this section, the development and modeling of RST forecasting models largely depend on the factors that influence RST. The next section dives deeper into the current and historical practices of modeling road surface temperatures.

2.3 Current Practices in Forecasting Road Surface Temperatures

The discussions so far have demonstrated the need for road surface temperature (RST) forecasts. In the past, numerous models for predicting RST have been developed. The complexity and accuracy level of the models developed tend to vary significantly based on the forecasting objectives and factors included in the model. Some of these models use simple meteorological factors such as air temperature and humidity, while others incorporate complex physical and mathematical modeling of solar radiation, heat flow rate, etc. However, in most cases, forecast models for RST can be broadly divided into two categories, numerical forecast models and statistical-based models. More recently, machine learning and neural network models are being developed. This section illustrates the past and current efforts of modeling RST.

2.3.1 Models Developed Using Heat Transfer Theories

Numerical models have been studied extensively over the past decades. As early as 1957, Barber (1957) attempted to determine the maximum RST using thermal diffusion theory. The research connected pavement temperature with wind, precipitation, air temperature, and solar radiation. Using this found relationship, it is possible to predict the RST at different times and locations. The study concluded that actual values of a structure's thermal characteristics, as well as the ambient circumstances, must be known in order to compute the exact temperature for that structure (Barber, 1957).

Sass (1992) proposed a model based on the ground and the surface energy-balance equation to predict pavement temperature and ice. The model consisted of two primary phases: phase one involved computing the atmospheric surface heat fluxes, and phase two involved finding the road temperatures, road water, and ice using the flux values from phase one. The developed numerical model underwent sensitivity testing using Danish road station data. The assessment revealed considerable model sensitivity to the atmospheric input data. Therefore, the study concluded that high-quality road-temperature forecasts would require linking the model to a realistic atmospheric forecast model and careful handling of initial road and atmospheric conditions.

Shao et al. (1996) constructed a numerical model based on an unsteady one-dimensional heat conduction equation to forecast road surface temperature and state (i.e., dry, wet, frost, or ice). The model's development process required a significant number of complex equations and parameters.

The model first makes forecasts on standard meteorological variables such as air temperature, dewpoint temperature, mean wind speed, total cloud amount dominant cloud type, and several others using time series analysis. These forecasted parameters are then used as inputs to the numerical equations for computing heat fluxes and final forecasts.

Crevier and Delage (2001) developed the widely used Environment and Temperature of Roads model, also known as METRo, first implemented at the Ottawa Regional Centre in October 1999. It delivers 24-h road weather forecasts two times per day, at 3:00 am and 3:00 pm local time. METRo is a numerical model that calculates the temperature evolution by solving the energy balance equation at the road surface and the heat transmission in the road material. METRo forecasts road surface conditions in three stages: initialization of the road temperature profile using historical measurements, coupling of the forecast with observations during the overlap period when both the meteorological forecast and the roadside observations are available, and finally forecasting the pavement temperature itself. The model incorporates data from RWIS stations as well as meteorological forecasts from the Canadian Meteorological Centre's operational Global Environmental Multiscale (GEM) model. GEM model provides solar fluxes either in automatic mode or as a function of cloud cover and temperature. The energy balance equation used in METRo is highly complex. It considers seven types of fluxes: the incoming flux, the absorbed incoming infrared radiation flux, the emitted flux, the sensible turbulent heat flux, the latent heat flux, the flux associated with phase changes of precipitating water, and an anthropogenic flux. Despite the model’s popularity and the many advantages it provides, its complex nature and requirement for enormous volumes of data make it time and resource-intensive to maintain and develop. Furthermore, the study makes no indication of how accurate the forecasts are, nor does it compare the accuracy of forecasting short-term periods (less than 6 hours) to long-term periods

(Hashemloo, 2008). It has only been reported that about one-half of the time, the error in surface road temperature prediction is within ±2K.

2.3.2 Models Developed Using Statistical Methods

Researchers applied statistical analysis for estimating the pavement temperature to determine the most effective asphalt binders for low-temperature climates. Bosscher et al. (1998) applied statistical regression techniques to model pavement temperature as a function of air temperature and other meteorological factors. The goal of the study was to create statistical models that could estimate daily low and high pavement temperatures using meteorological data. Using linear regression analysis, the model defined the minimum pavement temperature as a function of minimum air temperature. Although the R^2 value of 0.963 showed a good correlation, the standard error was relatively high (2.71°C). The model was further improved by adding more predictor variables, including the average air temperature and average solar intensity calculated during the 24 hours preceding the time at which the minimum pavement temperature occurred and found a standard deviation of 1.73 °C. A similar analysis was conducted for the estimation of high pavement design temperature, and the best model yielded a standard error of 1.87 °C and R^2 of 0.916.[33]

Sherif and Hassan (2004) applied stepwise regression analysis to predict pavement temperature using air temperature, dew point, relative humidity, wind speed, wind gust, and wind direction data collected from six RWIS stations in the City of Ottawa (2001-2002 winter season). Six models were established for each station, out of which one model was selected for use in the Ottawa area, called the Ottawa-wide model. The study concluded that air temperature and dew point were the most significant independent variables regarding the model's goodness of fit and P-value. The regression model was further improved by incorporating time-lag-dependant variables.

One of the most widely used statistical technique for modeling time series data is the classical time series models such as Auto Regressive Moving Average (ARMA), Auto Regressive Integrated Moving Average (ARIMA), Seasonal Auto Regressive Integrated Moving Average (SARIMA), etc. These models are developed by considering the temporal variation of the variable. Researchers have used this approach to characterise seasonal fluctuations in pavement structural qualities in the literature (Ali & Parker 1996). Hashemloo (2008) adopted SARIMA to forecast road surface temperatures.

2.3.3 Models Developed Using Machine Learning and Neural Networks

Despite the extensive studies and widespread use of numerical models discussed above, they are highly complex and resource-intensive, making it difficult to a faster and flexible implementation. On the other hand, traditional statistical technique for time series modeling such as ARIMA is based on the assumption that the time series under investigation are created by linear processes (Zhang et al., 1998) thereby ignoring more intuitive pattern presented in the data. As a result, it may not be applicable for most nonlinear real-world cases as many applications, such as road surface temperature, is nonlinear in nature with uncertain behaviour that evolves over time, necessitating the solution of highly nonlinear, time-variant problems. Moreover, traditional time series models do not support missing values. This is when neural networks set themselves apart from the traditional practices.

With the advancements in computer science and the availability of unprecedented amounts of data, methods such as machine learning techniques and neural networks have started to gain momentum in recent years. One notable method—Artificial Neural Network (ANN)—has differentiated itself by its ability to generalize patterns it has not seen during training. It can learn and emulate nonlinear relationships between input and corresponding target variables without requiring insight into the underlying mechanisms. This ability gives ANN an advantage in dealing with complex problems over other parametric approaches. Furthermore, they are fast, flexible and capable of handling missing values. For this reason, similar studies are still being conducted to assess the applicability of such techniques to make road weather forecasts.

One of the earliest works combined ANN with a numerical method to improve the nowcasts generated by the numerical model. In that study, Shao (1998) forecasted the error of RST nowcasts provided by an automated numerical model using a three-layer neural network (NN). The goal was to increase the accuracy of the nowcasts by adjusting the errors. The network was trained by an error-backpropagation algorithm with numerous meteorological parameters, including time-lag-dependant forecast error, air, surface temperature, dew point, wind speed, wind direction, and surface state. Although the study reported promising results at forecasting errors, this approach added extra complexity to the already complicated numerical method for forecasting RST.

In another study, Horii et al. (2000) developed a 3 hours-ahead prediction model for RST using ANN. The model used the time-series changes of the variables, including air temperatures, RST, temperatures at 5 and 10 centimeters underground, wind velocities, and the amounts of rainfall and snowfall per hour as inputs to the three-layer neural network. The model's performance was evaluated for its ability to differentiate between two categories, whether the forecasted value was below zero or above zero. However, it did not report any evaluation of the accuracies of the forecasted results.

Hashemloo (2008) developed forecasting models (3-hour and 5-hour ahead) for RST using ANN and compared its performance with time series models. The study used time-lag-dependant hourly RST and current air temperature data collected from an RWIS station as inputs to the model. The study concluded that multi-variable Seasonal Autoregressive Integrated Moving-Average (SARIMA) was the most competitive technique and had the lowest degree of forecasting errors.

Besides ANN, support vector regression (SVR) has also been developed to predict the land surface temperature and has shown promising results (Mathew et al., 2018). Others have explored the use of similar methods to generate air temperature forecasts. Astsatryan et al. (2021) forecasted hourly air temperatures up to 24 hours and found a maximum error in the range of 8.97 to 11.63 for different forecasting horizons. Jakaria et al. (2020) found that multilayer perceptron regressors yielded high RMSE when forecasting air temperature at a particular hour. More advanced neural network architectures such as recurrent neural network (RNN) and long short-term memory (LSTM) have also been implemented to forecast air temperature (Tran et al., 2021).

2.4 Summary

In this chapter, various winter maintenance options and their dependency on road surface temperature have been reviewed. Following the demonstration of the significance of RST as a decision-making tool for selecting the best WRM program, numerous factors that may influence road temperatures were identified. Finally, a detailed overview of past and current methodologies for forecasting RST along with the limitations of these methods were presented.

Two fundamental limitations in prior studies on forecasting effectiveness have been recognized as a result of the foregoing discussions. Firstly, these approaches can only predict the current state of the road surface based on current environmental parameters, such as air temperature, wind speed, wind direction, humidity, solar radiation, and so on. They are unable to forecast future road surface conditions. Previous studies have also shown the existence of time-series aspects of road surface conditions (Kangas et al., 2015), but only a few research studies have provided insight into the evolution of time-series prediction models. Furthermore, despite their great popularity and applicability in other fields, few researchers utilized machine learning and neural network models to predict RST. Thus, there is still room for assessing the feasibility of newer and more advanced machine learning technologies to predict future road surface conditions utilizing the time-series property of the RST data. Secondly, the studies discussed above were mostly site-specific whereby failing to identify factors that cause forecasts to vary between sites or regions. These are factors such as local weather and geographical characteristics, all of which have the potential to influence forecasting accuracies. Therefore, it is crucial to examine how those factors affect model performance.

Chapter 3 METHODOLOGY

Recognizing the limitations in forecasting road surface temperatures (RST), transportation agencies are looking for better ways to forecast consistent and precise road temperature data. To tackle this challenge, this chapter provides an overview of the proposed methodology for forecasting RST and the underlying theory behind the implemented algorithms. Detailed descriptions of each processing procedure, including data processing, selection of features, and network architecture are also included.

3.1 Methodological Framework

This research aims at implementing neural networks to forecast hourly RST 6 hours into the future. Two highly popular neural networks, namely artificial neural network (ANN) and long short-term memory (LSTM), were considered for this study. The classic ANN has been studied in previous research to predict road surface conditions; however, the studies were limited to specific sites and small datasets. Furthermore, the LSTM has received less attention in forecasting RST. LSTM is a newer and more advanced type of neural network that is designed to capture data's time-series features. This thesis employs two types of neural networks in order to bridge the gap between previous research and current state-of-the-art algorithms while applying them to a greater spatial range. While the general development process is the same for both, they differ in their network architectures, input features, and other neural network-related parameters. Figure 3.1 illustrates a methodological framework for developing neural network-based forecasting models.

There are two phases: phase one involves data preparation, and phase two focuses on model development and hyperparameter tuning. These phases are discussed in detail in the following sections.

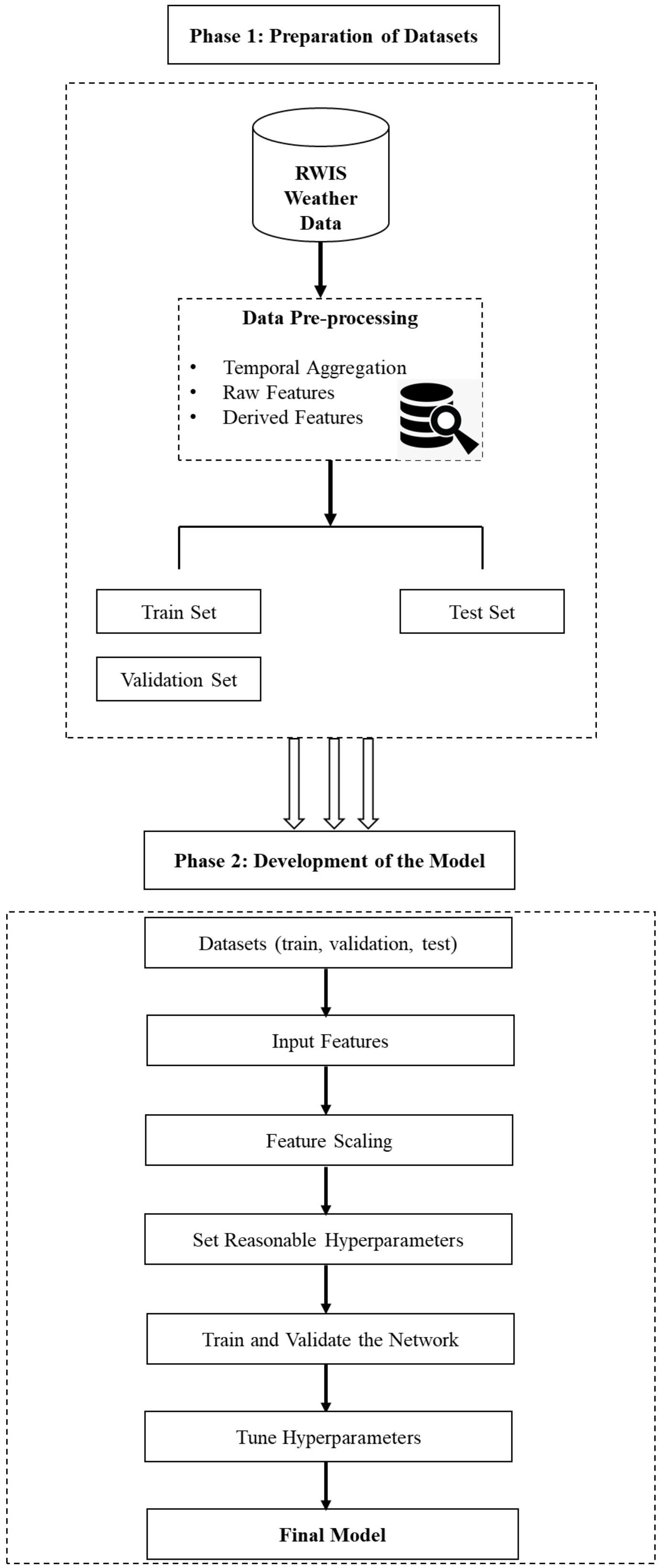


Figure 3-1 Methodological Framework for Forecasting RST using Neural Networks

3.1.1 Phase I: Preparation of Datasets

In Phase one, the raw data undergoes several transformations, including the identification and manipulation of missing values, selection of input variables, and the splitting of dataset into train-validation-test sets.

3.1.1.1 Data Collection and Processing

The primary information required for this study is the road surface temperature (RST) data. For most jurisdictions, road surface measurements are made using one of the most critical highway intelligent transportation systems (ITS) structures known as road weather information systems (RWIS). These are automated road weather reporting stations composed of different advanced sensor technologies embedded near the road. A typical RWIS station tower includes cameras, road surface sensors, remote processing units (RPU), and communication equipment for collecting, transmitting, and disseminating road weather and surface condition data to processing centers. The collected data are then processed and made available for further use. Several road weather and surface condition information can be collected with these sensors, one of which is RST. The data obtained offers the necessary inputs for WRM agencies to provide more efficient and cost-effective maintenance services to the right location at the right time with the appropriate treatment, which is one of the main reasons why RWIS stations have gained popularity. Approximately more than 3,000 RWIS stations are currently being utilized across North America, and the number will continue to rise as authorities strive to improve their existing WRM services and optimize resources (Kwon and Gu, 2017).

RWIS stations constantly collect road surface and weather conditions, reporting every few minutes (typically every 10-20 min). Raw observations obtained from RWIS were aggregated into hourly averages for the convenience of developing hourly forecasting models. Missing observations were replaced with the mean value of the previous and next available observations. Additional processing procedures required for neural networks are discussed in detail below.

3.1.1.2 Input Features

From a modeling perspective, one of the most important factors that may affect the forecast accuracy is the input variables. Researchers have explored a wide range of options in an effort to identify the most suitable input features, including univariate features such as prior time lags, derived features such as hour, days, months as functions of sine and cosine, meteorological variables like air temperature, humidity, wind speed, wind direction, dew point, and so on. Out of these candidate variables, their correlation with RST ultimately determines which feature gets chosen. The reason being highly correlated features are more likely to increase the model's predicting power. Regarding this study, prior time lags and the hour as a derived feature were chosen as inputs. Therefore, autocorrelation and partial autocorrelations plots will be used as tools to find the highly correlated lags.

3.1.1.3 Train, Validation, and Test dataset

Datasets are split into the train, validation, and test sets. The training dataset is a portion of the dataset used to fit the model, and the validation dataset provides an unbiased performance evaluation that can be used for model comparison and hyperparameter tuning purposes. Therefore, choosing the model that performs well in the validation set may help minimize the test dataset's forecasting error. At this stage, some performance metrics will be assessed to compare among the models. Finally, the testing dataset is utilized to offer an unbiased assessment of the final model's performance.

This study utilizes direct multi-step forecasting. This method involves developing separate models for each intended horizon. Therefore, the labels in the training, validation and test datasets are determined based on the forecasting horizon considered. The labels in this study contain only one variable, RST.

3.1.2 Phase II: Development of Neural Network Models

Some of the important aspects involved in model development are discussed below in detail.

3.1.2.1 Scaling of datasets

The preprocessed training dataset was used to train the neural network model. But before feeding the dataset into the network, the input features needed to be standardized to improve performance. In this study, the range of input features (i.e., RST, sines, and cosines of the hour) vary significantly. Therefore, the datasets are standardized via Equation (3.1):

$$x'_t = \frac{x_t - \mu_x}{s_x} \tag{3.1}$$

Where $x_t$and x'_t are the original and transformed explanatory variables, respectively and μ_x and s_x are the mean and standard deviation of the original variable x, respectively.

3.1.2.2 Network Architecture and Hyperparameters

The number of layers and its associated hyperparameters define the neural network architecture. Hyperparameters are user-defined values that govern the training process. Therefore, values selected for hyperparameters significantly impact model performance. Examples of hyperparameters include the number of hidden layers, batch size, epochs, activation function, etc. Hyperparameter tuning is the process of determining the right combination of hyperparameters that allows the model to maximize its performance. There are no hard and fast rules for selecting these parameters, so they are determined through trial and error. In this study, only the number of neurons in the hidden layer was tuned. Other parameters were selected through an extensive literature review, out of which commonly used values were chosen to minimize the complexity hyperparameter tuning process.

3.1.2.3 Training and validating the network

During the model training process, model performance was tracked using two metrics: mean absolute error (MAE) and root mean square error (RMSE). MAE measures the average of the residuals in the database, whereas RMSE measures the standard deviation of the residuals. The formulas for calculating MAE and RMSE are shown in Equations 3.2 and 3.3.

$$MAE = \frac{\sum_i^N |Predicted_i - Actual_i|}{N} \tag{3.2}$$

$$RMSE = \sqrt{\frac{\sum_i^N (Predcted_i - Actual_i)^2}{N}} \tag{3.3}$$

Where N is the number of samples, and i = 1, 2, 3…., N

The number of neurons in the hidden layer was determined by iteratively changing the number of neurons and evaluating the model performance as a result of the changes made. After each trial, the performance metrics were analyzed to find the one with the lowest error on the validation dataset. Thus, the model with the lowest error is the final model and is consequently used in the test phase.

3.2 Artificial Neural Network

ANN is a widely used technique to solve complex non-linear problems due to its profound ability to learn the underlying relationship found in a given dataset. Once fully developed, an ANN model is capable of generalizing patterns that were not presented to the model during development. This ability to learn and emulate nonlinear features without needing to consider the underlying mechanism of the problem gives the neural network a great advantage over traditional statistical mapping (Shao, 1998).

The basic concept of ANN was inspired by the human brain's synapse activity. These networks are a collection of highly interconnected processing units called neurons. Neurons are connected by edges that transmit signals between neurons. The edges themselves are assigned weights that determine the strength of the signal at a connection. Before the input dataset enters the neuron, the input data is multiplied by the edge's weight and added by a bias. And upon entering the neuron, this combined sum is then processed by a user-defined non-linear function before getting outputted.

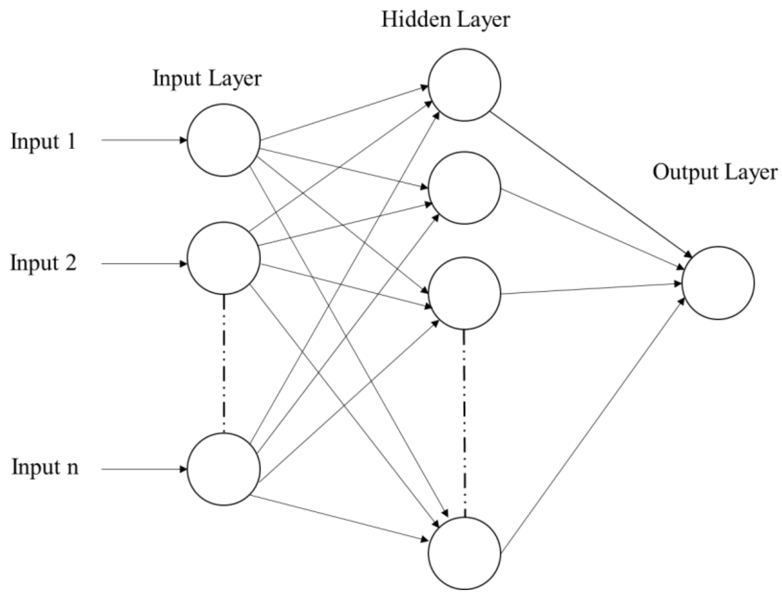


Figure 3-2 General Architecture of a Neural Network

An ANN network's basic architecture consists of three layers: input, hidden, and output. Each layer consists of one or more neurons, as represented in Figure 3.2. There are two types of neurons in an ANN network: passive nodes that do not change or modify the data and active nodes that do change and transform the data. Concerning our neural network, the input layer neurons are passive, while the hidden and output layer neurons are active (Hashemloo, 2008). As shown in Figure 3.2, excluding the input layer, each neuron receives input signals from other neurons in its previous layer.

Recall that the output of a neuron is a function of the weighted input, bias, and activation function (Figure 3.3).

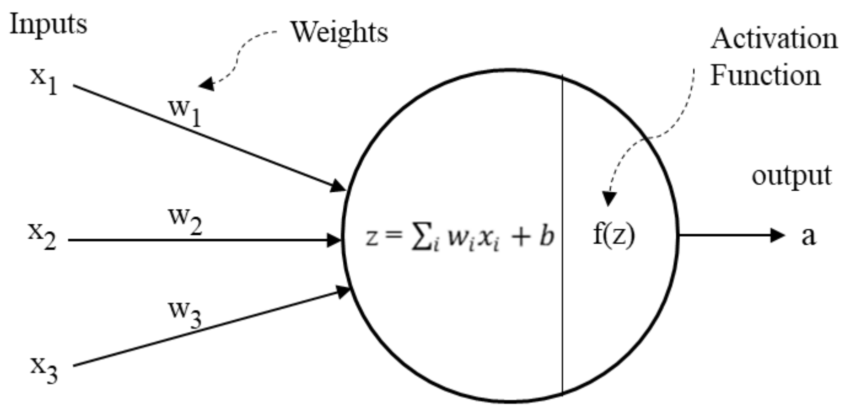


Figure 3-3 An Active Node of Neural Network

The operations in an active node can be represented as Equation 3.4:

$$z = \sum_{i=1}^{i=n} w_i x_i + b \tag{3.4}$$

where z is the weighted sum of inputs and a bias b, x_i is the i^{th} input to the neuron, w_i is the connection weight of the i^{th} input.

Activation function can be presented as Equation 3.5:

$$a = f(z) \tag{3.5}$$

where a is an output from the neuron and f is the activation function.

The activation function controls the activation level of a neuron (Karlik and Olgac, 2011), which helps in learning and mapping the complicated relationships between the inputs and outputs. There are several commonly used activation functions: Sigmoid, TanH, ReLU, and Softmax. In recent years, rectified linear unit (ReLU) has become the most widely adopted activation function because it overcomes the vanishing gradient problem and allows models to learn faster. This function works by changing all negative input values to 0, while leaving positive values untouched (Pauly et al., 2017). Figure 3.4 shows a ReLU function.

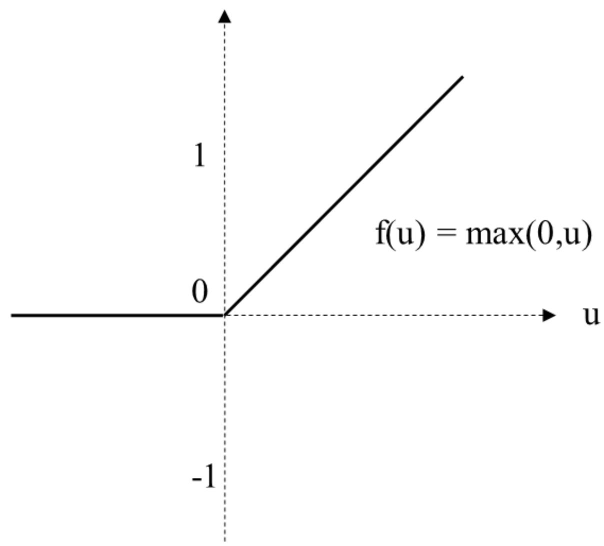


Figure 3-4 ReLU Activation Function

During training, the neural network's output is compared to actual sensor measurements to calculate error, which is then used to adjust the weight values through backpropagation. Backpropagation is the practice of fine-tuning the weights of the neural network based on error (i.e., loss). An optimization function is used to alter weights in a way that reduces error, making the model more accurate at the same time. The key concepts of the backpropagation neural network presented here are implemented in the following chapter.

Note that ANN is an overarching term covering many different types of neural networks. In this study, a three-layer fully connected feed-forward with the back-propagation neural network, known as a basic multi-layer perceptron, was used due to its proven ability to map inputs to target variables in prior research (Shao, 1998; Hashemloo, 2008; Horii and Fukuda, 2017). This model is the simplest form of ANN and thus has been referred to as ANN in this thesis.

3.3 Long Short-Term Memory

Long short-term memory (LSTM) is a popular neural network architecture developed based on the recurrent neural network (RNN) architecture. RNNs have hidden states, and they allow past outputs to be utilized as inputs (Schmidhuber, 2015). RNNs have a significant benefit over ANNs in that they can model data sequences, assuming each sample is dependent on preceding ones. As a result, RNN is normally a desired tool for sequence or time-series data. However, RNNs suffer from a major problem, which is vanishing and exploding gradient. This makes learning long data sequences difficult. The gradients contain information utilized in the RNN parameter update, and as the gradient shrinks, the parameter updates become minor, implying that no meaningful learning occurs. To overcome the limitations of RNN, Hochreiter and Schmidhuber (1997) proposed the LSTM model in 1997. LSTM fixes the vanishing gradient issue in RNN as well as can learn long-term dependencies of temporal sequences and remember information for extended periods. The LSTM includes units called memory cells which are the key to the LSTM framework as depicted in Figure 3.5.

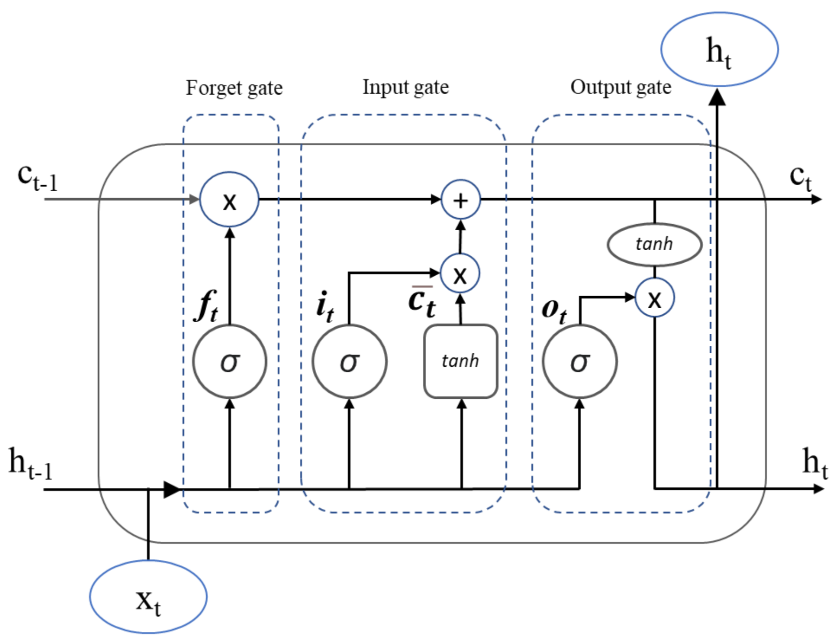


Figure 3-5 Architecture of LSTM cell

Where, x_t is the input vector, h_{t-1} is the previous cell output, c_{t-1} is the previous cell memory, h_t is the current cell output, c_t is the current cell memory, and σ is the sigmoid function. The addition and multiplication signs in the diagram refer to element-wise addition and element-wise multiplication, respectively.

Each memory cell is composed of three gates; input gate (*i*), forget gate (*f*), and output gate (*o*). The cell is tightly controlled by these gates. An input at time step x_t and the hidden state of the preceding step h_{t-1} enter into the LSTM cell, and then the hidden state h_t is calculated as follows:

$$\boldsymbol{f_t = \sigma(x_t W_f + h_{t-1} U_f + b_f)} \tag{3.6}$$

$$i_t = \sigma(x_t W_i + h_{t-1} U_i + b_i) \tag{3.7}$$

$$\overline{c_t} = tanh(x_t W_c + h_{t-1} U_c + b_c) \tag{3.8}$$

$$c_t = f_t \otimes c_{t-1} \oplus i_t \otimes \bar{c}_t \tag{3.9}$$

$$o_t = \sigma(x_t W_o + h_{t-1} U_o + b_o) \tag{3.10}$$

$$h_t = o_t \otimes \tanh(c_t) \tag{3.11}$$

In the equations above, *W*, *U* denote weight matrices and *b* bias vector parameters that must be learned during training.

The forget gate is the first LSTM gate, and it determines what information from the cell should be discarded in the cell state. It takes x_t and h_{t-1} into account and outputs a number between 0 and 1 for each number in the cell state c_{t-1}. If the output is 1, the information is saved, and if the output is 0, it is discarded.

The input gate is the second gate, and it calculates how much information should be added to and updated in the cell state. Input gate contains two components: a sigmoid layer and a tanh layer.

The overall process involves the sigmoid layer determining which values to update first. Then, tanh generates a vector of new candidate values, $\bar{c}_t$, that might be added to the state. Following this, the old state, c_{t-1}, is updated to new state c_t by performing an element-wise multiplication between f_t (output of the forget gate at time t) and the old state, at which point the output of this operation is added to the product of input (i_t) and candidate ($\bar{c}_t$) vectors.

The final gate, the output gate, determines the network output based on the cell state. First, the sigmoid function calculates what parts of the cell state will be outputted. Then, the cell state goes through tanh (to push the values to be between −1 and 1) and is multiplied by the output of the sigmoid gate so that the output contains only the intended parts of the cell state.

LSTM allows networks to learn, forget, and update hidden states when new data are presented with its gate activation mechanism. Thus, LSTM does not suffer from the vanishing and exploding gradient issues in RNN. Furthermore, their superior ability to connect previous information to the present task makes them suitable for sequential data, which in our case is road surface temperatures (RST). Therefore, the modeling of changes in RST with time using LSTM is expected to outperform traditional ANN.

3.4 Summary

This chapter presented an overview of a typical methodological framework for neural network-based models, a detailed explanation of data collection, processing, model development, and the theory behind the two neural networks used in this study, namely the ANN and LSTM.

The first part of the chapter discussed some key aspects such as data preparation, train-validation-test datasets, input features, network architecture and hyperparameters in detail. MAE and RMSE were chosen as performance evaluation methods for the developed model. This chapter then proceeds with the theory of ANN and LSTM to explore the predictability of these algorithms for forecasting sequential data such as RST. Although classic ANN has been explored in prior research, they used spatially limited datasets. Therefore, this thesis employs both algorithms to bridge the gap in prior research as well as assess the feasibility of state-of-the-art LSTM in forecasting RST. The more advanced and complex internal mechanisms of the LSTM network distinguish it from basic ANN in their ability to retain previously learned information and use it as input in determining the new output. In the next chapter, these neural networks are implemented in a case study to forecast RST for analyzing and comparing the outcomes from different perspectives.

Chapter 4 CASE STUDY

The proposed methodology is applied in a case study that covers a large area in the province of Alberta. The climate of Alberta is characterized as a semi-arid continental climate with warm summer and cold winter (Jiang et al., 2017). Winter air temperatures typically vary between −25.1°C and −9. 6°C, summer temperatures between 8.7°C and 18.5°C (Sekhon et al., 2010). Alberta has an area of 661,848 km^2 (255,541 sq mi) that extends 1,223 km (760 mi) north to south and 660 km (410 mi) east to west at its maximum width. As a result, the climate of Alberta varies greatly which makes it an ideal study area for spatial analysis.

A total of six stationary RWIS sites were chosen, covering southern to northern parts of the province based on the geographic locations and the availability of required data points. The selected stations cover area from latitude of 49.00N to 55.00N and altitude of 711m to 1044m. The careful selection of those six sites ensures a geographically diverse set of stations and thus will aid in analyzing the impact of geographical characteristics on forecasting abilities. Alberta Transportation provided the RWIS datasets required for this study. Figure 4.1 illustrates the study sites, and their longitude, latitude, and elevations are tabulated in Table 4.1.

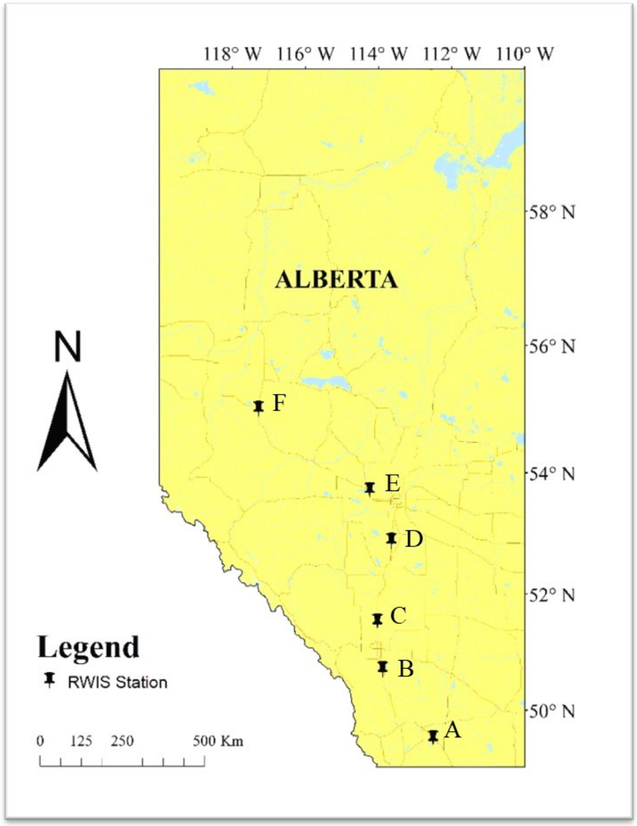


Figure 4-1 Study Sites in Alberta, Canada

Table 4-1 Study sites in Alberta, Canada

Station ID	Latitude	Longitude	Elevation(m)
A	49.502	-112.503	936
B	50.717	-113.883	1044
C	51.533	-114.026	1040
D	52.889	-113.643	797
E	53.716	-114.238	714
F	55.012	-117.282	711

4.1 Data Preparation and Processing

The RWIS data points collected in this study were logged every couple of minutes. An hourly aggregation is performed on the raw data obtained from RWIS stations. Temperature values are in degrees Celsius. Datasets and their associated weather stations are chosen based on the representativeness of winter months, the relative distance between two stations, and the number of missing data points within the period considered to analyze the effects of geographical attributes on forecasting capabilities.

Datasets are split into the train, validation, and test sets. For consistency in splitting the datasets, each of these sets starts at 00:00 hour and ends at 23:00 hour. Training dataset extends from December 1, 2019, to February 10, 2020, validation set from February 11 to February 15, 2020, and finally test dataset from February 16 to February 20, 2020. As such, six sets of datasets are prepared for the six stations in the study area. A table summarizing the descriptive statistics of the datasets is presented in Table 4.2.

Throughout the research period, the ranges of RST at stations A through F vary significantly which dictated the selection of these stations for the spatial analysis. From Table 4.2, the mean RST for the training set is -3.68 to -9.54°C, -1.41 to -6.71°C for the validation set, and -3.83 to -10.70°C for the testing set. As the climate differs widely from southern to northern Alberta, the average RST decreases with increasing latitude. Similarly, there are considerable variations in the minimum and maximum RST across the stations. The minimum RST for the training set is -25.70 to -34.00°C, -10.45 to -13.59°C for the validation set, and -14.97 to -20.40°C for the testing set. Furthermore, the training, validation and testing set's maximum RST is 15.15 to 5.10°C, 10.15 to 1.28°C, and 10.87 to 4.45°C, respectively. It can also be seen that in the northern part of Alberta, the standard deviation of the RST is somewhat smaller for validation and testing sets.

Table 4-2 Descriptive statistics of the datasets

	Station	Minimum	Maximum	Mean	Standard Deviation	No. of Observations
Training Set	A	-25.70	15.15	-3.68	6.71	1728
	B	-30.50	11.47	-6.66	7.13	
	C	-30.83	11.18	-5.84	7.71	
	D	-34.17	6.08	-8.12	7.50	
	E	-33.17	5.27	-9.11	7.09	
	F	-34.00	5.10	-9.54	7.57	
Validation Set	A	-10.45	10.15	-1.41	5.31	120
	B	-11.52	9.63	-3.08	5.23	
	C	-10.68	6.75	-2.67	4.64	
	D	-13.59	3.64	-5.44	4.43	
	E	-13.45	1.87	-6.71	4.17	
	F	-12.82	1.28	-5.94	4.09	
Testing Set	A	-14.97	10.87	-3.83	6.41	120
	B	-15.46	12.78	-5.53	6.86	
	C	-15.65	12.80	-5.86	6.60	
	D	-16.63	6.88	-8.70	5.90	
	E	-20.53	4.45	-9.82	5.90	
	F	-20.40	5.57	-10.70	5.90	

This research relied heavily on open-source libraries under the Python environment. Python is a high-level, interpreted programming language that is used for a broad range of applications. In this study, the models were implemented using TensorFlow, an open-source software library provided by Google. Additionally, packages including NumPy, Pandas, and Matplotlib were also utilized for data processing, manipulation, and visualization.

4.2 Artificial Neural Network

In this study, the same ANN development process is followed to develop models for each site. To make a fair comparison of the forecasting accuracies of the developed models, the same train-validation-test period, input features, and network architectures were used. A total of six models were developed for each station to forecast RST from one hour ahead to six hours ahead.

4.2.1 Model Development

When developing ANN models, RST at prior time lags were used as input features. Autocorrelation plots and partial autocorrelation plots of up to 24 lags were carefully examined to select highly correlated input features to develop the forecasting models. Furthermore, hour of the day can reveal hidden patterns in the datasets. For example, certain hours of the day have similar patterns than others - during daytime the temperatures tend to be higher than nighttime. Moreover, because of the cyclical nature of hour, 23:59 hour is closer to 00:00 hour. Therefore, hour can be transformed into an interpretable feature to convey its cyclical nature to the model. This can be achieved with sine and cosine transformations to encode the cyclical time feature into two features. Therefore, the input vector consists of the prior time lags and the forecasting hour as the functions of sine and cosines. Three lag-dependent variables were used for the first two models (i.e., 1-h, 2-h) to provide the model with enough inputs for mapping the target. In terms of the remaining models, the input vector contained the same number of time lags as the forecasting horizon. For example, the 3-h model would have three lags, and the 4-h model would have four legs, etc. The features are scaled before training the neural network.

Studies have shown that ANN model performance is related to network complexity, meaning models are dependent on the number of layers and the number of neurons in each layer (Chaturvedi, 2008). In this thesis, only the number of neurons in the hidden layer was determined through a trial in an effort to further improve the model performance. The neuron number varied between 6 to 14, with an increment of 2. After each trial, the performance metrics (MAE and RMSE) were analyzed to find the configuration with the lowest error on the validation dataset, and the number of neurons that yielded the lowest error was chosen. Other hyperparameters for model training were chosen based on common values found in literature. As a result, the following configuration was chosen: hidden layer 1, batch size 32, epoch size of 1000 with early stopping, ReLU as the activation function, Adam optimizer with a learning rate of 0.001, and the mean squared error (MSE) loss function as the objective function (Hashemloo, 2008; Ustaoglu et al., 2008).

With these specifications, the number of neurons in the hidden layer was determined by iteratively altering the number of neurons and evaluating the performance metrics of the model. As a result, the final model is the one with the lowest error on the validation dataset, and it is utilized in the testing phase. The specifications of the final models developed for site A are shown in Table 4.3. All others are shown in Appendix A.

Table 4-3 Summary of the ANN models developed for site A

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	8	54	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = ReLU
2-h ahead	3	14	73	
3-h ahead	3	14	78	
4-h ahead	4	10	98	
5-h ahead	5	8	67	
6-h ahead	6	8	43	

4.2.2 Results and Discussion

The developed models were tested on a test dataset that the model had not seen during training. The resulting forecasts were compared with the actual RST values. Figure 4.2 shows the results for site A while all other sites are shown in Appendix B.

This study adopted six different models to forecast pavement surface temperatures 1, 2, 3, 4, 5, and 6 hours ahead for each station. A visual inspection of Figure 4.2 reveals that the forecasts are closer to the actual values for short forecasting horizons and deviates as the forecasting horizon increases. As the forecasting horizon increases, it becomes more difficult for the model to capture temporal changes in RST as one can anticipate. A thorough discussion on the effects of horizons on the forecasting accuracy is presented later in this section.

The performance of the proposed models was evaluated further using two statistical measures-MAE and RMSE. MAE is a linear score which means that all the individual differences are weighted equally in the average. On the other hand, the RMSE is a quadratic scoring rule where the errors are squared before they are averaged. Hence, the RMSE gives a relatively high weight to large errors, making it the most useful when large errors are particularly undesirable. The greater the difference between MAE and RMSE, the greater the variance in the individual errors in the sample. While analyzing the various factors that may affect the RST forecasts, the MAE values were primarily employed to compare the developed models while also examining the RMSE for any significant errors.

The proposed models' performance statistic results are presented in Table 4.4.

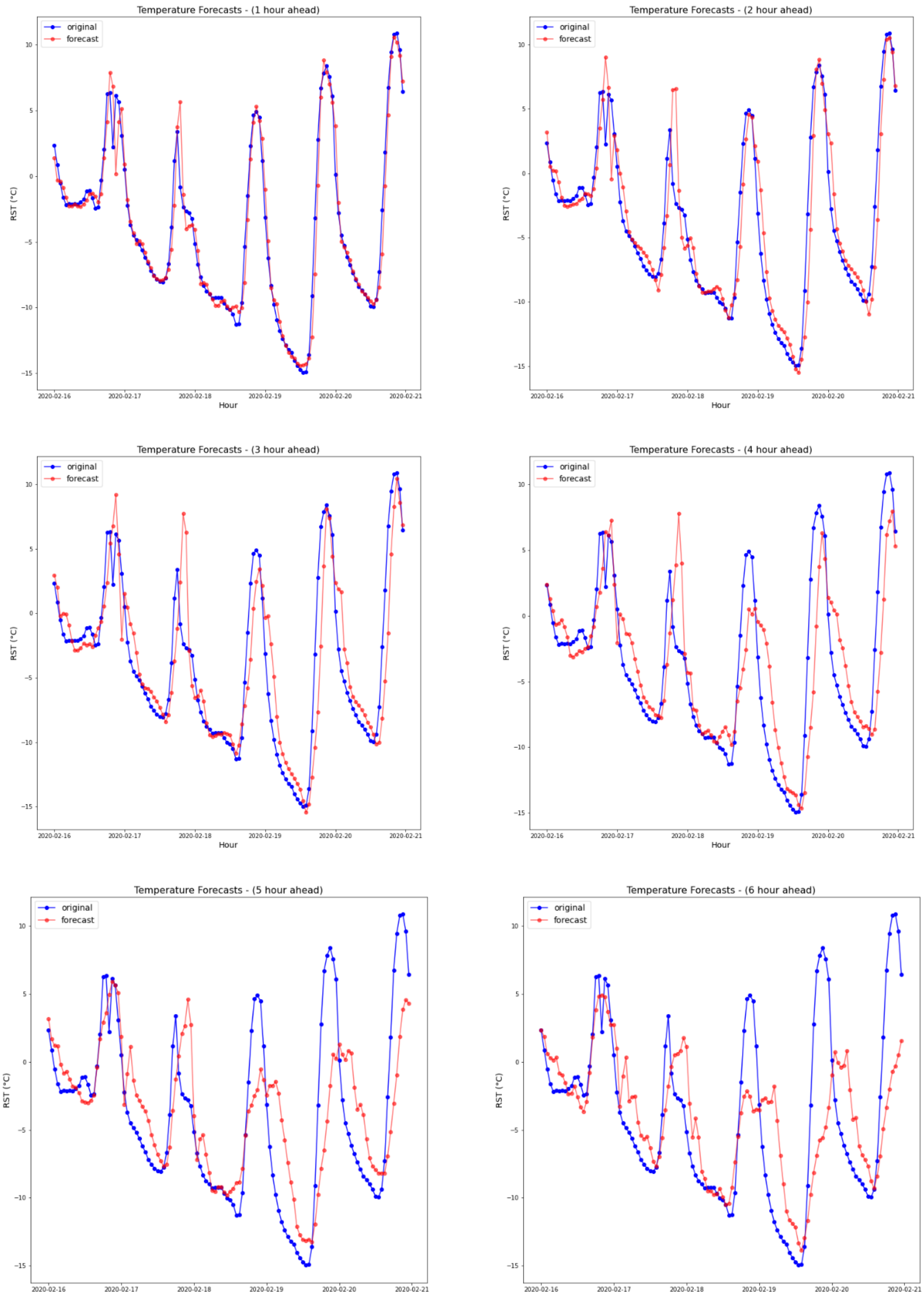


Figure 4-2 Forecasted results for station A (ANN)

Table 4-4 Summary of the proposed models' performance statistics (ANN)

Forecasting Horizon	Performance Metrics	Site A	Site B	Site C	Site D	Site E	Site F	Average
1	MAE (°C)	0.88	0.83	0.73	0.68	0.64	0.70	0.74
	RMSE (°C)	1.45	1.12	0.98	0.88	0.93	1.11	1.08
2	MAE (°C)	1.59	1.46	1.29	1.20	1.26	1.22	1.34
	RMSE (°C)	2.39	2.05	1.71	1.58	1.85	1.84	1.90
3	MAE (°C)	2.13	1.99	1.71	1.69	1.59	1.79	1.82
	RMSE (°C)	3.10	2.79	2.26	2.22	2.27	2.66	2.55
4	MAE (°C)	2.75	2.63	2.30	2.30	2.16	2.27	2.40
	RMSE (°C)	3.79	3.62	2.96	2.88	3.03	3.11	3.23
5	MAE (°C)	3.25	3.03	2.83	2.76	2.61	2.56	2.84
	RMSE (°C)	4.35	4.25	3.64	3.45	3.59	3.68	3.83
6	MAE (°C)	3.53	3.60	3.37	3.10	3.17	3.03	3.30
	RMSE (°C)	4.87	5.08	4.19	3.94	4.20	4.13	4.40

The results in Table 4-4 show that the average MAE value for the 1-hour forecasting horizon is as low as 0.74°C and does not exceed 3.30°C for the 6-hour forecasting horizon. Similarly, the average RMSE values range from 1.08°C to 4.40°C, as observed in Table 2. The lowest MAE values for 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h ahead forecasts are 0.64, 1.2, 1.59, 2.16, 2.56, and 3.03°C, respectively, whereas the lowest RMSE values are within the range of 0.88 to 3.94°C; the results are better or at least consistent with previous studies (Hashemloo, 2008).

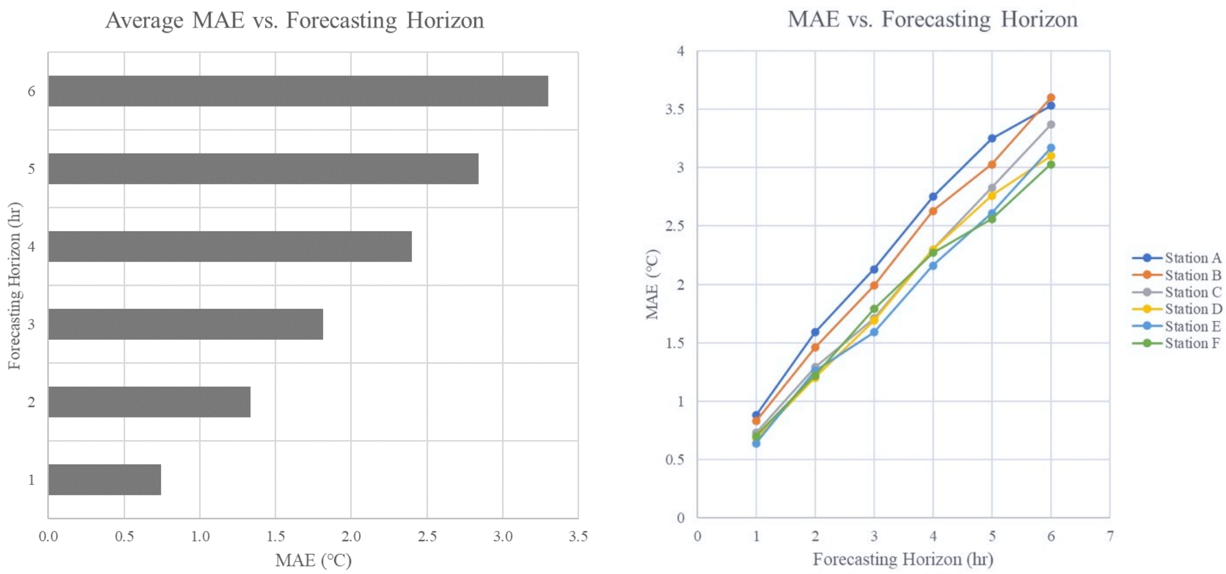
This section briefly highlighted the performances of the developed ANN models. In the following sub-sections, the qualitative and quantitative results for these different models are evaluated and discussed in more detail regarding forecasting horizons and the possible impact of geographical attributes on their forecasting accuracies.

4.2.2.1 Effect of Forecasting Horizons on Forecasting Accuracy

The forecasting horizon has a proportional influence on the accuracy of the forecasts. As illustrated in Figure 4.2, the forecasts deviate from the actual values of RST as the forecasting horizon gradually increases. An investigation on its impact on the forecasting accuracy becomes more relevant when considering diversely located forecasting sites as accuracy for a particular forecasting horizon may differ significantly from others.

To examine such phenomena, the MAE values for different forecasting horizons were calculated and shown in Figure 4.3. It can be seen in Figure 4.3(a) that the average forecasting errors increase with increasing forecasting horizons for all stations, which is expected. Similarly, the average RMSE values range from 1.08°C to 4.40°C, as observed in Table 4.4. The relatively minor differences between MAE and RMSE values reveal that very large errors were unlikely to have occurred although there exists some variation in the magnitude of the errors. Furthermore, what is discovered in Figure 4.3(b) is that the forecasting errors at some stations can be much higher for increased horizons while others may exhibit smaller errors. For instance, compared to other stations, the MAE values at higher forecasting horizons were found to be relatively larger at stations A and B. The MAE value for 6-h ahead forecasts at stations A and B were 3.53°C and 3.6°C, respectively, whereas the lowest value was found to be as low as 3.03°C at station F. Similar patterns can be observed for other forecasting horizons as well. The rise in errors in these two stations was steeper compared to other stations. This was likely due to the higher variability in temperature within the forecasting horizons for some stations over others.

The impact of forecasting horizon may dictate the preference of one model over another at a particular station to keep the error within the tolerance limit while maximizing the forecasting horizon.



(a) **(b)**

Figure 4-3 Relation of MAE with Forecasting Horizons (ANN)

4.2.2.2 Effect of Geographical Factors on Forecasting Accuracy

An investigation was conducted to examine the hypothesis that locational attributes would affect RST forecasting accuracies. Making such speculation becomes more reasonable when considering unique weather characteristics of Alberta, where southern parts typically experience less severe and varying weather patterns. In contrast, the opposite weather patterns are exhibited by the northern parts. As such, locational factors; namely, latitude and elevation, were considered to evaluate the extent to which each factor would affect the quality of the forecasts made by six different ANN models developed for each station in this study. The resulting relationships between each of the locational factors and their respective model performance (in MAE) are shown in Figure 4.4 and Figure 4.5.

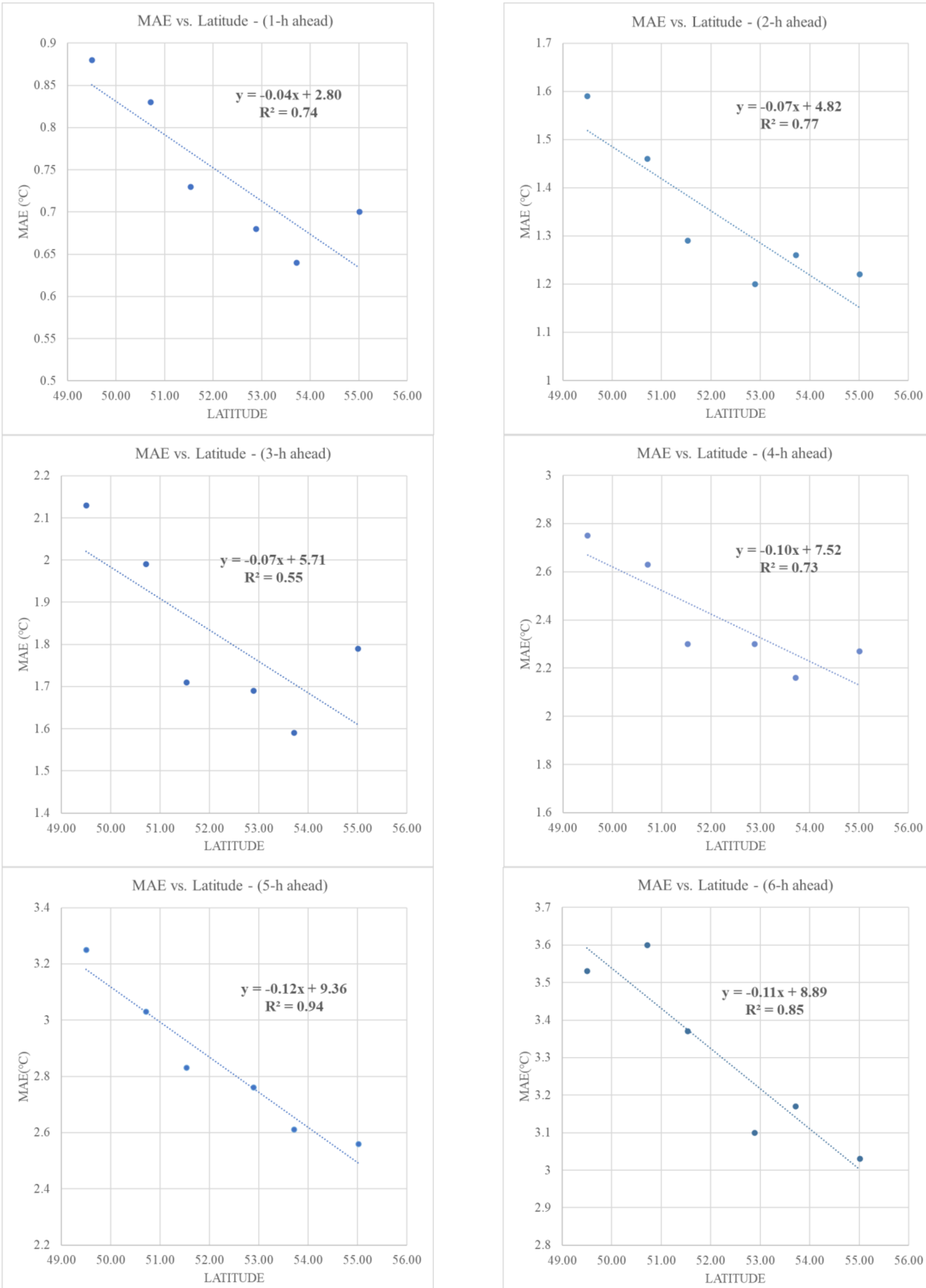


Figure 4-4 Relation of MAE with Latitude (ANN)

As can be seen from these figures in Figure 4.4, latitude was strongly correlated with model performance. The R-squared values for 1, 2, 3, 4, 5, and 6-h ahead forecasts from Figure 4.4 are 0.74, 0.77, 0.55, 0.73, 0.94 and 0.85, respectively whereby demonstrating a good correlation of forecasting error with latitude. One can visually confirm that the MAE values show a decreasing trend with increasing latitude, which means the accuracy improves with increasing latitude for all forecasting horizons. Although station F had somewhat greater error than neighboring stations in some cases, the errors were still much lower than stations A and B (stations located in the lower latitude zone). What is more relevant is that stations A and B, which are in the lower altitude zone, have far larger errors than stations E and F, which are located in the higher latitude zone. Similar trends can be found if corresponding RMSE values in Table 4.4 are carefully observed.

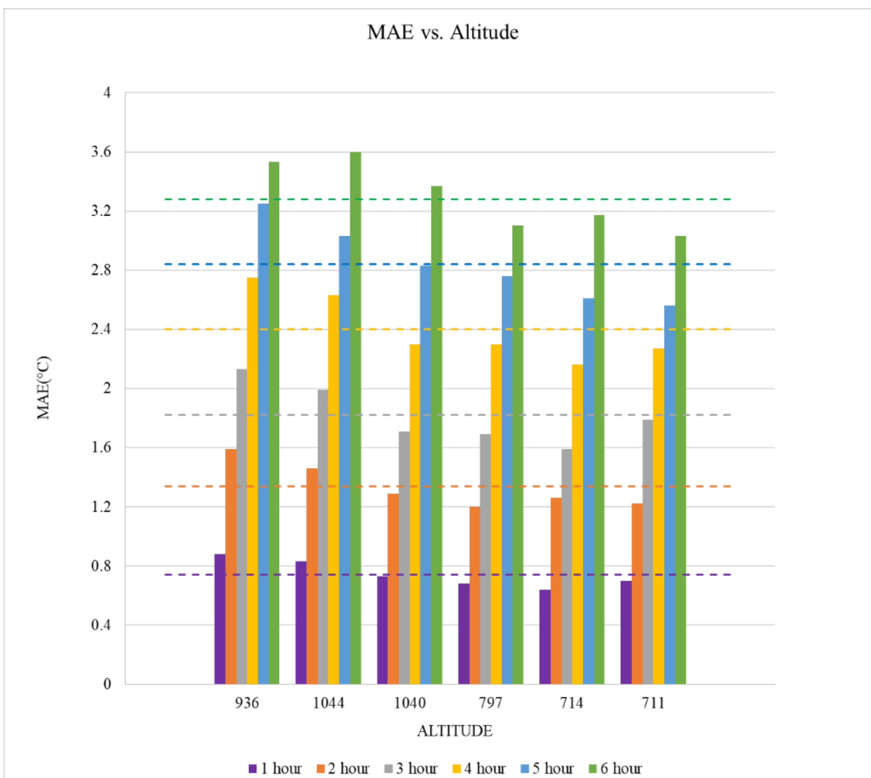


Figure 4-5 Relation of MAE and Altitude (ANN)

In Figure 4.5, stations with higher elevations generally experienced larger forecasting errors. For 1-hour ahead forecasts, stations with elevations greater than 900 meters had errors equal to or larger than the average, represented by the dashed line. Similarly, for other forecasting horizons, stations A(936m) and B(1044m) had errors above average. Interestingly, station C(1040m) had performed better than A and B despite having similar altitudes. However, the more relevant thing was that stations with similar elevations had shown similar forecasting errors for all forecasting horizons. Stations D(797m), E(714m), and F(711m), all had errors lower than the average in all cases.

The influence of geographical attributes can be explained by the pattern of RST in these regions. It can be noticed in Table 4.2 that the standard deviation of the test data at stations A to F are 6.41, 6.86, 6.60, 5.9, 5.89, and 5.9°C, respectively. Therefore, RST at stations D to F are less spread out than the others, which causes less variability in temperatures in these regions. Stations D to F are located at a greater latitude but lower altitude than stations A to C. As a result, less sunlight is experienced by D to F due to its high latitude, hence why these stations show less variability in temperature changes during the day. Such phenomena can be attributed to the fact that Northern Alberta is open to cold arctic weather systems carrying extremely cold air masses from the north, resulting in unusual winter conditions with very limited daily sunlight hours during winter. On the other hand, southwestern Alberta enjoys milder temperatures due to the proximity to the Rockies, which lets warm, dry air masses blow from the mountains to these regions.

The impact of altitude on forecasts may be explained by the relationship between RST and altitude, which is typically non-linear (Shao et al., 1997) and most noticeable when atmospheric stability is low (Chapman et al., 2001). Therefore, high-altitude areas are more likely to experience dramatic changes in RST than low-altitude areas, causing larger errors in forecasting.

4.3 Long Short-Term Memory

The same long short-term memory (LSTM) development process is followed to develop models for all sites using the same datasets as ANN. However, when developing an LSTM model, there are several network-specific factors to consider as discussed in detail below.

4.3.1 Model Development

The input features for LSTM were determined using the same technique as ANN. However, the LSTM algorithm demands specific input and output shapes to fit the model as discussed in Chapter 3. Recall that this network requires a three-dimensional array as an input, where the first dimension represents the batch size, the second dimension represents the number of time-steps in a sequence, and the third dimension represents the number of units in one input sequence. Therefore, a single sequence in this study contains three units: a lagged value of RST, the sine, and the cosine of the corresponding hour. The output contains only the forecasted RST value. The features were also scaled before training the neural network.

Similar to ANN, a trial was conducted to determine the number of units or neurons in the hidden layer. The number of units employed ranged from 6 to 14, with an increment of two. The performance parameters (MAE and RMSE) were examined after each trial to determine which had the lowest error on the validation dataset. The model with the lowest error was then employed in the test phase. Other hyperparameters for training the models include number of hidden layers 1, batch size 32, epoch size of 1000 with early stopping, Keras default hyperbolic tangent (tanh) as the activation function, Adam optimizer with a learning rate of 0.001, and the mean squared error (MSE) loss function as the objective function (Tabrizi et al., 2021, Tran et al., 2020).

The specifications of the models developed for site A are shown in Table 4.5, while all others are shown in Appendix A.

Table 4-5 Summary of the LSTM models developed for site A

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	12	99	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = tanh
2-h ahead	3	12	129	
3-h ahead	3	14	162	
4-h ahead	4	14	193	
5-h ahead	5	14	259	
6-h ahead	6	6	264	

4.3.2 Results and Discussion

This section highlights the performance of the developed models using LSTM. Six different models per station were developed to forecast pavement surface temperatures 1, 2, 3, 4, 5, and 6 hours ahead of time. The developed models are then put to the test on the test datasets to see how well they perform and what effects the forecasting horizon and geographical attributes have on forecasting accuracy.

The forecasted RST values were compared to the actual RST values. The outcomes for site A are given in Figure 4.6, while all other sites are listed in Appendix B. Figure 4.2 shows that for short forecasting horizons, the forecasts are closer to the actual values, but deviate as the forecasting horizon lengthens. This implies that the model had a harder time capturing temporal changes in RST as the forecasting horizon increased. Later in this section, there is a detailed examination of the implications of horizons on forecasting accuracy.

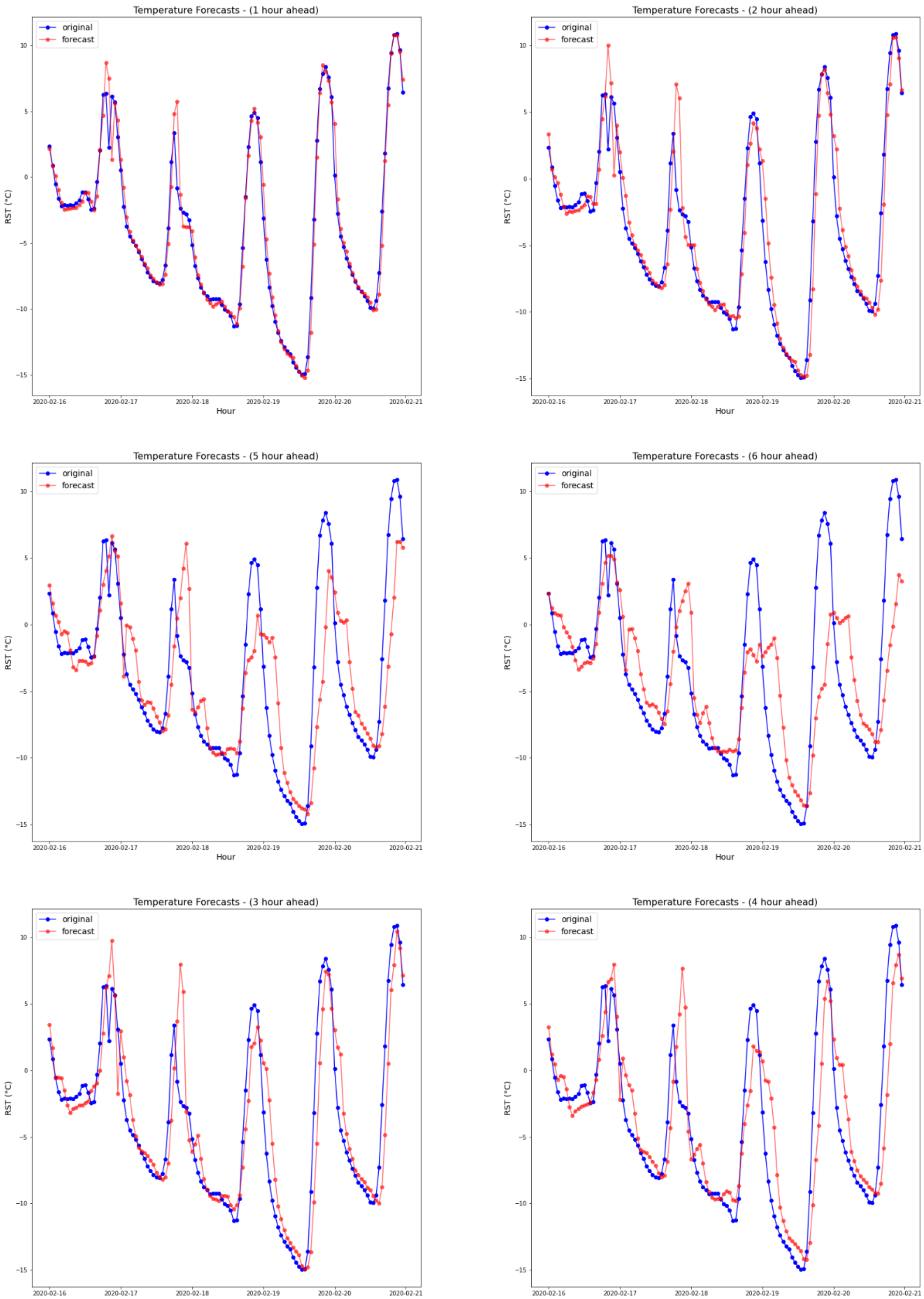


Figure 4-6 Forecasted Results for Station A (LSTM)

The performances of these models were further expressed in MAE and RMSE values. Table 4.6 shows the performance statistics for all the proposed models.

Table 4-6 Summary of the proposed models' performance statistics (LSTM)

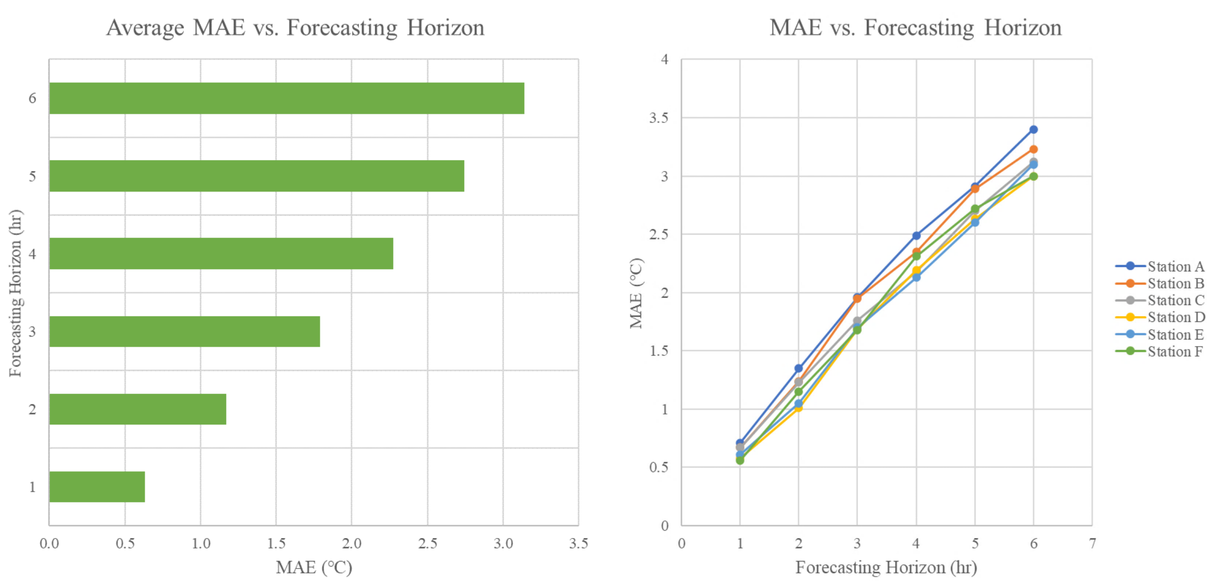
Forecasting Horizon	Performance Metrics	Site A	Site B	Site C	Site D	Site E	Site F	Average
1	MAE (°C)	0.71	0.67	0.67	0.58	0.61	0.56	0.63
	RMSE (°C)	1.25	0.96	0.89	0.78	0.83	0.90	0.94
2	MAE (°C)	1.35	1.24	1.23	1.01	1.05	1.15	1.17
	RMSE (°C)	2.13	1.74	1.59	1.32	1.47	1.68	1.66
3	MAE (°C)	1.96	1.95	1.76	1.69	1.70	1.68	1.79
	RMSE (°C)	2.86	2.71	2.30	2.17	2.37	2.41	2.47
4	MAE (°C)	2.49	2.35	2.18	2.19	2.13	2.31	2.28
	RMSE (°C)	3.48	3.26	2.91	2.73	2.92	3.16	3.08
5	MAE (°C)	2.91	2.89	2.70	2.63	2.60	2.72	2.74
	RMSE (°C)	4.06	4.02	3.51	3.26	3.51	3.61	3.66
6	MAE (°C)	3.40	3.23	3.12	3.00	3.10	3.00	3.14
	RMSE (°C)	4.58	4.57	3.96	3.78	4.03	4.06	4.16

According to the data in Table 4.6, the average MAE value for the 1-hour forecasting horizon is as low as 0.63°C and does not surpass 3.14°C for the 6-hour forecasting horizon. Similarly, as shown in Table 4.6, the average RMSE values vary from 0.94°C to 4.16°C, indicating some variability in the extent of the errors. For 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h ahead forecasts, the lowest MAE values are 0.56, 1.01, 1.68, 2.13, 2.60, and 3.00°C, respectively, while the lowest RMSE values are 0.78 to 3.78°C. Furthermore, it is noteworthy that these LSTM models have outperformed the ANN models shown in Table 4.4 in terms of their forecasting accuracies. A more detailed discussion comparing ANN and LSTM is presented later in this chapter.

This section assessed the overall performances of the developed LSTM models. To further investigate the probable impact of forecasting horizon and geographical attributes on the forecasting accuracy, the performances of models developed with LSTM are evaluated in greater detail in the following sub-sections.

4.3.3.1 Effect of Forecasting horizons on Forecasting Accuracy

Figure 4.7 illustrates the MAE values over different forecasting horizons. As seen in Figure 4.7(a), forecasting errors rise as forecasting horizons increase for all stations. Similarly, the average RMSE values range from 0.94°C to 4.16°C, as observed in Table 4.6. When each individual station was analyzed in terms of forecasting horizons, the MAE values at longer horizons were somewhat greater, with a slightly steeper rise for stations A and B compared to others. This further reinforces the findings in the previous section with ANN models.



(a) **(b)**

Figure 4-7 Relation of MAE with Forecasting Horizons (LSTM)

4.3.3.2 Effect of Geographical Factors on Forecasting Accuracy

The hypothesis that geographical attributes would affect RST forecasting accuracies was further tested on the LSTM models developed in this section. Figures 4.8 and 4.9 demonstrate the resultant correlations between each of the locational variables and their corresponding model performance (in MAE).

As shown in Figure 4.8, model performance was found to be influenced by latitude, where accuracy increased with increasing latitude across all forecasting horizons. The RMSE values provided in Table 4.6 show almost identical trends, demonstrating that latitude does affect forecasting accuracy. The R-squared values were also calculated to demonstrate the correlation. For 1, 2, 3, 4, 5, and 6-h ahead forecasts from Figure 4.8 are 0.89, 0.57, 0.80, 0.35, 0.57 and 0.78 respectively which indicates a considerably good correlation of forecasting error with latitude. As expected, the findings from the analysis with the LSTM models strongly reinforce the results found with the ANN models.

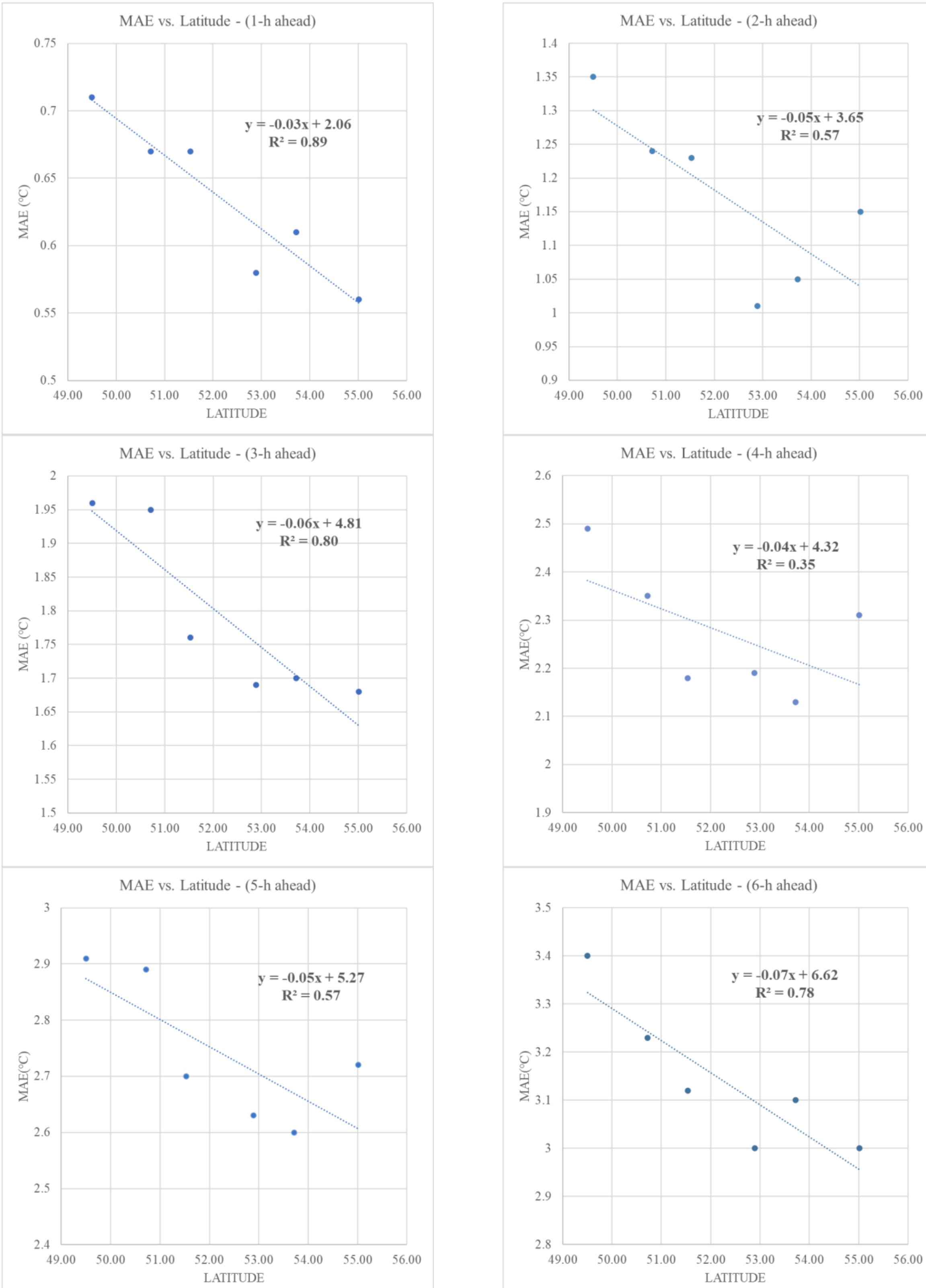


Figure 4-8 Relation of MAE with Latitude (LSTM)

Upon further investigating the influence of altitude on the LSTM models in Figure 4.9, it is revealed that the stations with higher elevations experienced larger forecasting errors. For 1-hour and 2-hour ahead forecasts, stations with elevations greater than 900 meters have errors larger than the average, represented by the dashed line. Similarly, for other forecasting horizons, stations A(936m) and B(1044m) have errors above average in all cases. Although stations C and D behaved somewhat differently than expected, the overall performance of stations with similar altitude ranges (>900m and <800m) were similar. As expected, the analysis of the influence of altitude on the forecasting performances closely resembles the findings from the ANN models.

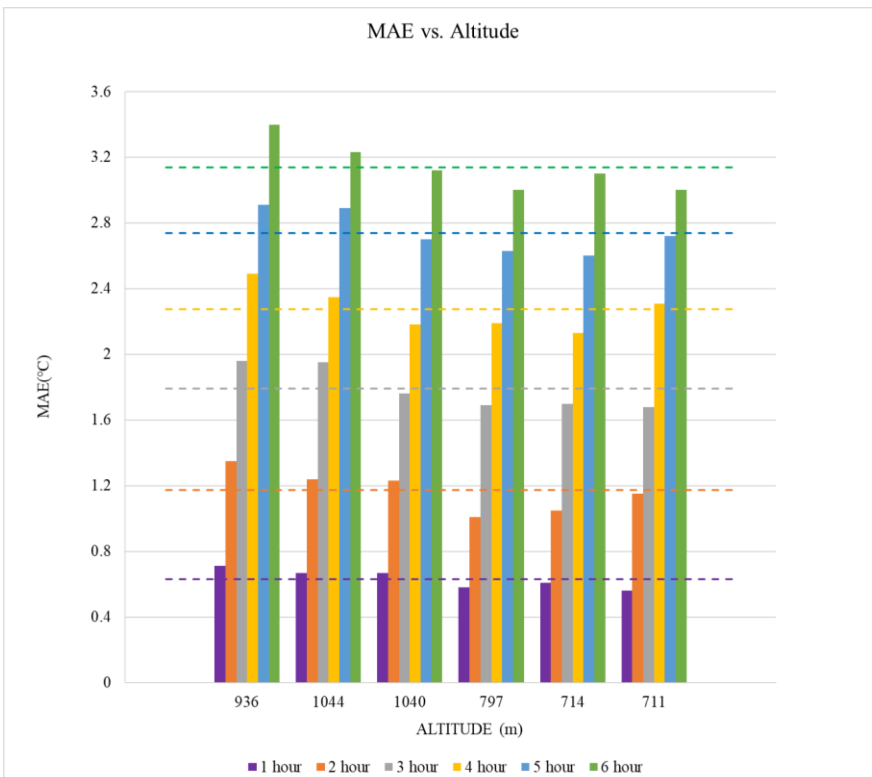


Figure 4-9 Relation of MAE with Altitudes (LSTM)

4.4 Comparison of Models Developed

The preceding sections employed two distinct neural networks: the ANN to bridge the gap in earlier research employing spatially large-scale datasets, and the LSTM to investigate the viability of a more sophisticated neural network model for RST forecasting. Both of these algorithms were employed to forecast hourly RST for 6 hours using data from six stations located in geographically distinct areas. Upon investigating their performance statistics displayed in Table 4.4 and Table 4.6, noticeable differences can be observed. Therefore, a comparative analysis of these two neural networks was conducted and is presented here. ANN models generated average MAE values of 0.74, 1.34, 1.82, 2.40, 2.84, 3.30°C for 1 to 6-hour forecasts, whereas LSTM produced 0.63, 1.17, 1.79, 2.28, 2.74, 3.14°C. Similarly, ANN models generated RMSE values of 1.08, 1.90, 2.55, 3.23, 3.83, 4.40°C, whereas LSTM models yielded RMSE values of 0.94, 1.66, 2.47, 3.08, 3.66, 4.16°C. Figures 4.10 and 4.11 illustrate the comparative performances of the ANN and LSTM in terms of MAE and RMSE, respectively.

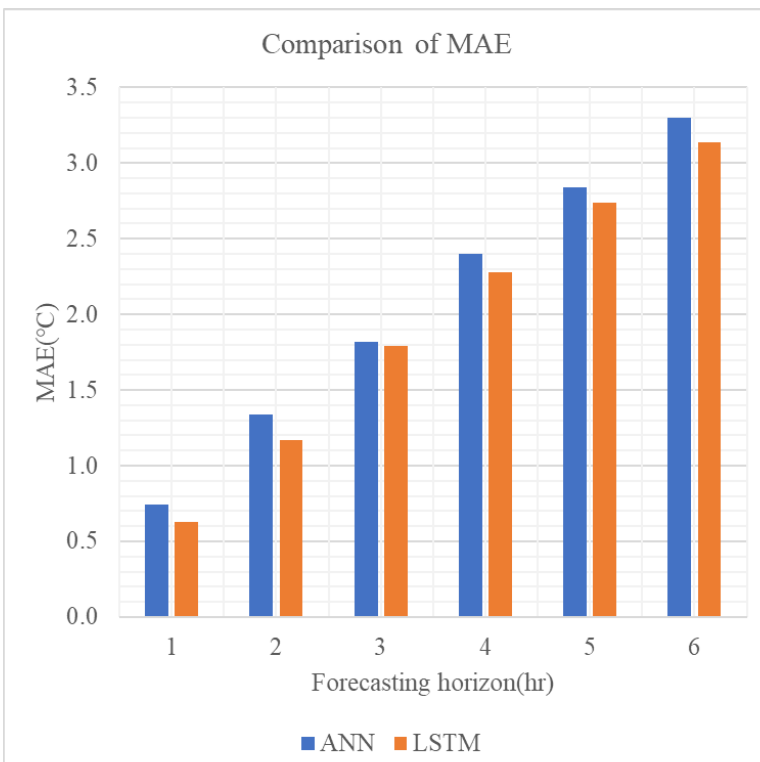


Figure 4-10 Comparison between ANN and LSTM in terms of Average MAE

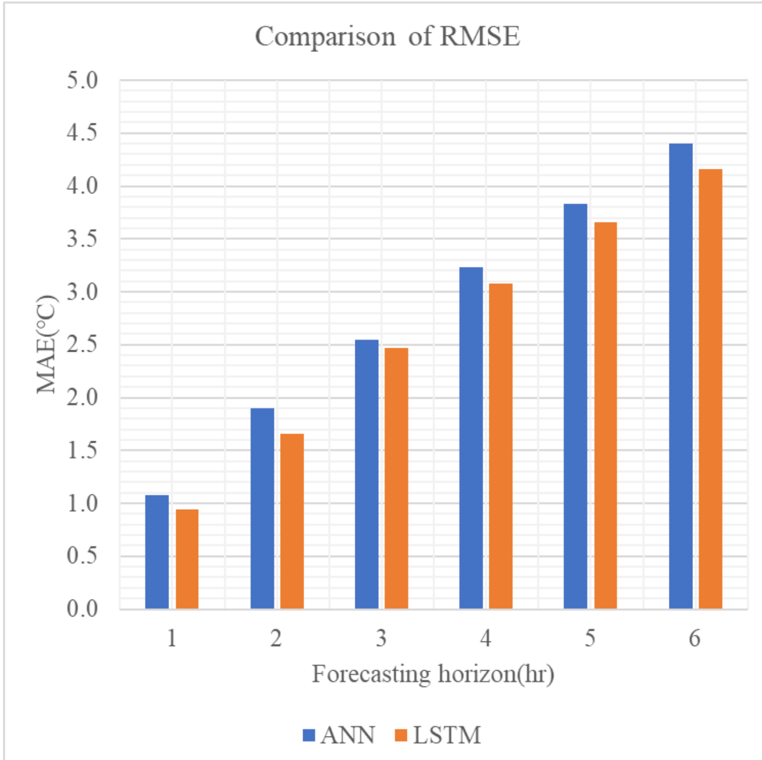


Figure 4-11 Comparison between ANN and LSTM in terms of Average RMSE

Based on Figures 4.10 and 4.11, LSTM performed better for all forecasting horizons. Temperatures exhibit the property of time-series sequences. It changes gradually with time, and the changes are closely related to its immediate past values. These patterns are utilized for predicting the next value in the series with respect to the previous information. The internal mechanism of the LSTM network discussed in Chapter 3 makes it capable of capturing the sequential or time-series property of RST, which makes LSTM more suitable for these tasks.

4.5 Summary

This chapter implemented ANN and LSTM networks at six different locations within Alberta for the winter months to forecast RST over the course of six hours. The impact of several factors such as forecasting horizon, latitude, and altitude on forecasting accuracies was also investigated.

Finally, a comparative analysis between ANN and LSTM was conducted in terms of their forecasting accuracies.

The models developed herein using both ANN and LSTM showed promising outcomes. Traditional ANN model forecasts yielded average MAE values of 0.74, 1.34, 1.82, 2.4, 2.84, 3.30°C for 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h forecasting horizons, respectively, and average RMSE values of 1.08 to 4.40°C. Furthermore, the lowest MAE values for 1-h to 6-h ahead forecasts were 0.64, 1.2, 1.59, 2.16, 2.56, and 3.03°C, respectively, while the lowest RMSE values ranged from 0.88 to 3.94°C. Similarly, the LSTM models generated average MAE values of 0.63, 1.17, 1.79, 2.28, 2.74, 3.14°C for the 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h forecasting horizons, respectively, and average RMSE values of 0.94 to 4.16°C. Upon analyzing the influence of factors affecting forecasting accuracies based on these outcomes, the study found accuracy increased with increasing latitudes and decreasing altitudes. Finally, the comparison between the neural networks revealed the superiority of the LSTM over multi-layer perceptron in modeling RST.

Chapter 5 CONCLUSION AND FUTURE WORK

This chapter summarises this thesis and highlights the significant findings and contributions made by the study presented here. Furthermore, the limitations of this thesis are discussed later in this chapter, along with the recommendations for future study.

5.1 Research Overview

Recent advancements in data storage and processing have encouraged researchers to assess the applicability of statistical models using neural networks to forecast weather parameters. Although prior studies have shown good forecasting accuracy, there is a lack of research in the developed model's variability of performance over space and time, as well as identifying factors that contribute to such variabilities.

In this research, a series of forecasting models were developed from short-term training data using multi-layer perceptron and LSTM neural network models at six geographically different RWIS sites within Alberta. Tuning of hyperparameters was applied to enhance the performance of the models. Although the forecasts from both neural network models showed promising results, LSTM outperformed ANN in terms of accuracy, possibly due to the inherent ability of LSTM architecture to model sequential data. Furthermore, a comparative analysis using primarily the MAE values over different forecasting horizons was done for different RWIS stations. And finally, various factors were identified to have a plausible impact on the performance of the models from both a modeling and geographical perspective. It was revealed that location influences model performance as forecasts made at different locations produced different patterns of RST. The forecasting accuracy was found to be sensitive to several factors such as latitude, altitude, and forecasting horizons.

5.2 Research Findings

5.2.1 Feasibility of Neural Networks for Forecasting RST

Although prior studies have contributed to developing and advancing methods for generating road weather forecasts, very few studies have implemented machine learning algorithms (i.e., ANN, RNN, LSTM) to create short-term RST forecasts. This study attempted to fill this research gap by analyzing the feasibility of neural networks for forecasting RST.

Forecasts by traditional ANN models showed promising results, with average MAE values of 0.74, 1.34, 1.82, 2.4, 2.84, 3.30°C for forecasting horizons 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h, respectively; and an average RMSE value within the range of 1.08 to 4.40°C. Furthermore, the lowest errors for 1-h to 6-h ahead forecasts in terms of MAE were reported to be 0.64, 1.2, 1.59, 2.16, 2.56, and 3.03°C, respectively, whereas the lowest RMSE values were between 0.88 to 3.94°C. The LSTM models, on the other hand, generated even more promising results, with average MAE values of 0.63, 1.17, 1.79, 2.28, 2.74, 3.14°C for the 1-h, 2-h, 3-h, 4-h, 5-h, and 6-h forecasting horizons, respectively, and average RMSE values of 0.94 to 4.16°C. Thus, the results revealed that neural networks have great potential for forecasting highly accurate RST with minimal effort and time.

5.2.2 Factors Affecting the Forecasting Accuracy of RST

At the beginning of this study, a hypothesis was made that locational attributes might influence forecasting accuracies. From both a modeling and a geographical standpoint, many elements were found to reasonably influence the models' performance. It was discovered that model performance is influenced by location since forecasts generated by developed models at various locations exhibited varied RST patterns. Several parameters, including latitude, altitude, and forecasting horizons, were shown to affect forecasting accuracy. Among these variables, latitude was found to have the highest impact, where accuracy improved as latitude increased for forecasting horizons. Furthermore, locations at higher altitudes exhibited larger errors in forecasting.

5.3 Research Contributions

There are four contributions of this research as outlined below:

- This study provided a comprehensive review of existing literature, including winter road maintenance strategies, factors affecting RST, and current RST forecasting practices, which will aid transportation authorities in better understanding the progression of RST forecasting techniques and provide new insights into RST estimation.

- Machine learning-based method for forecasting RST has not been sufficiently explored in the transportation field. Further investigation was conducted in this thesis to assess the ability of machine learning-based methods in capturing the temporal variations of RST.

- A comparative analysis between two popular neural networks showed that LSTM models are much better at capturing the time series property of RST than ANN.

- Using the models developed herein, this study attempted to relate several external factors with forecasting accuracy. The investigation suggested a plausible influence of geographical attributes (latitude and altitude) and forecasting horizon. Long term, it is expected to pave the way for a significant body of new information by exploring the impact of other locational factors on the forecasting accuracies, which will, in turn, lead to the development of highly accurate location-specific models.

5.4 Limitations and Recommendations for Future Work

Statistical modeling strategies using neural networks for forecasting RST have yielded promising outcomes. Traditional large-scale, resource-intensive numerical forecasting methods are substantially more complicated than these strategies. However, several challenges must be resolved before they can be used in the field.

The work performed in this thesis can be extended in several directions. It is worthwhile mentioning that this study represents an initial effort to relate geographical and meteorological parameters using neural network-based models. Further analyses over space and time using RWIS data from other stations and during other seasons will be required to further validate the conclusiveness of our findings, and to investigate other potential factors that may affect model performance. Although the developed models in this research only incorporated the RST values of the prior time lags to forecast RST into the future, the framework presented here could easily be replicated and expanded to include more variables as input features such as air temperature, humidity, wind speed and other meteorological factors to explore how they influence model performance at different locations. It can also be beneficial to explore more recent advancements in neural network models such as convolutional neural networks, other variations of long-short term memory, and evaluate their applicability and performance. Nevertheless, the findings reported in this thesis will ultimately provide RWIS forecasting vendors and winter maintenance personnel with added knowledge to help make more informed and proactive WRM operations-related decisions.

BIBLIOGRAPHY

Agarwal, M., Maze, T. H., & Souleyrette, R. (2005, August). Impacts of weather on urban freeway traffic flow characteristics and facility capacity. In *Proceedings of the 2005 mid-continent transportation research symposium* (pp. 18-19).

Ali, H. A., & Parker, N. A. (1996). Using time series to incorporate seasonal variations in pavement design. Transportation research record, 1539(1), 33-43.

Asaeda, T., Ca, V. T., & Wake, A. (1996). Heat storage of pavement and its effect on the lower atmosphere. *Atmospheric environment*, *30*(3), 413-427.

Astsatryan, H., Grigoryan, H., Poghosyan, A., Abrahamyan, R., Asmaryan, S., Muradyan, V., ... & Giuliani, G. (2021). Air temperature forecasting using artificial neural network for Ararat valley. *Earth Science Informatics*, *14*(2), 711-722.

Barber, E. S. (1957). Calculation of maximum pavement temperatures from weather reports. *Highway Research Board Bulletin*, (168).

Bärring, L., Mattsson, J. O., & Lindqvist, S. (1985). Canyon geometry, street temperatures and urban heat island in Malmö, Sweden. *Journal of Climatology*, *5*(4), 433-444.

Blackburn, R. R. (2004). *Snow and ice control: Guidelines for materials and methods* (Vol. 526). Transportation Research Board.

Blum, U. (1997). Benefits and external benefits of transport: A spatial view. In *The full costs and benefits of transportation* (pp. 219-243). Springer, Berlin, Heidelberg.

Bogren, J. (1991). Screening effects on road surface temperature and road slipperiness. *Theoretical and applied climatology*, *43*(1), 91-99.

Bogren, J., & Gustavsson, T. (1991). Nocturnal air and road surface temperature variations in complex terrain. *International journal of climatology*, *11*(4), 443-455.

Bosscher, P. J., Bahia, H. U., Thomas, S., & Russell, J. S. (1998). Relationship between pavement temperature and weather data: Wisconsin field study to verify superpave algorithm. *Transportation Research Record*, *1609*(1), 1-11.

Burtwell, M. (2001). Assessment of the performance of prewetted salt for snow removal and ice control. *Transportation research record*, *1741*(1), 68-74.

Chapman, L., Thornes, J. E., & Bradley, A. V. (2001). Statistical modelling of road surface temperature from a geographical parameter database. *Meteorol Appl*, *8*(4), 409-419.

Chaturvedi, D. K. (2008). Factors affecting the performance of artificial neural network models. *Soft Computing: Techniques and its Applications in Electrical Engineering*, 51-85.

Du, S., Akin, M., Bergner, D., Xu, G., & Shi, X. (2022). Material application methodologies for winter road maintenance: a renewed perspective. *Canadian Journal of Civil Engineering*, *49*(1), 1-10.

Eliasson, I. (1996). Urban nocturnal temperatures, street geometry and land use. *Atmospheric environment*, *30*(3), 379-392.

Fischel, M. (2001). Evaluation of selected deicers based on a review of the literature CDOT-DTD-R-2001-15. *Colorado Department of Transportation, Denver*, 9-13.

Government of Canada, R. C. M. P. (2019, December 17). *RCMP Gazette*. Royal Canadian Mounted Police. Retrieved February 9, 2022, from https://www.rcmp-grc.gc.ca/en/gazette/just-the-facts-winter-driving

Gustavsson, T. (1990). Variation in road surface temperature due to topography and wind. *Theoretical and Applied Climatology*, *41*(4), 227-236.

Gustavsson, T. (1995). A study of air and road-surface temperature variations during clear windy nights. *International journal of climatology*, *15*(8), 919-932.

Hashemloo, B. (2008). *Forecasting pavement surface temperature using time series and artificial neural networks* (Master's thesis, University of Waterloo).

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Horii, M., & Fukuda, T. (2000). Forecast model of road surface temperature in snowy areas using neural network. In *Snow Engineering: Recent Advances and Developments. Proceedings of the Fourth International Conference on Snow Engineering. Norwegian University of Science and Technology*.

Iowa Department of Transportation (DOT). (2009). *Cost Benefits of Weather Information for Winter Road Maintenance*. Retrieved April 14, 2022, from https://westerntransportationinstitute.org/wp-content/uploads/2016/08/4W1576_Final_Report.pdf

Jakaria, A. H. M., Hossain, M. M., & Rahman, M. A. (2020). Smart weather forecasting using machine learning: a case study in tennessee. *arXiv preprint arXiv:2008.10789*.

Jiang, R., Gan, T. Y., Xie, J., Wang, N., & Kuo, C. C. (2017). Historical and potential changes of precipitation and temperature of Alberta subjected to climate change impact: 1900–2100. *Theoretical and applied climatology*, *127*(3), 725-739.

Kangas, M., Heikinheimo, M., & Hippi, M. (2015). RoadSurf: a modelling system for predicting road weather and road surface conditions. *Meteorological applications*, *22*(3), 544-553.

Karlik, B., & Olgac, A. V. (2011). Performance analysis of various activation functions in generalized MLP architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems*, *1*(4), 111-122.

Keep, D., & Parker, D. (2000). Proactive guide to snow and ice control: a guide for highway winter maintenance personnel. *Prepared by Ice and Snow Technologies, LLC., Washington*.

Ketcham, S., Minsk, L. D., Blackburn, R. R., & Fleege, E. J. (1996). *Manual of practice for an effective anti-icing program: a guide for highway winter maintenance personnel* (No. FHWA-RD-95-202). United States. Federal Highway Administration.

Kršmanc, R., Slak, A. Š., & Demšar, J. (2013). Statistical approach for forecasting road surface temperature. *Meteorological applications*, *20*(4), 439-446.

Kwon, T. J., & Fu, L. (2021). Performance evaluation of road surface temperature forecasts. *Canadian Journal of Civil Engineering*, *48*(5), 532-539.

Kwon, T. J., & Gu, L. (2017, August). Modelling of winter road surface temperature (RST)—A GIS-based approach. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)* (pp. 551-556). IEEE.

Kwon, T. J., Fu, L., & Jiang, C. (2013). Effect of winter weather and road surface conditions on macroscopic traffic parameters. *Transportation research record*, *2329*(1), 54-62.

Mathew, A., Khandelwal, S., & Kaul, N. (2018). Analysis of diurnal surface temperature variations for the assessment of surface urban heat island effect over Indian cities. *Energy and Buildings*, *159*, 271-295.

Minsk, L. (1998). *Snow and ice control manual for transportation facilities*. New York: McGraw-Hill.

Nilsson, G., & Obrenovic, A. (1998). *Road accident risk twice as high in the winter. Swedish Road and Traffic Research Institute (VTI)* (No. 435). Report.

Nixon, W. A., & Williams, A. D. (2001). A guide for selecting anti-icing chemicals, version 1.0. *IIHR Technical Rep*, *420*.

Oke, T. (1987). *Boundary Layer Climates*. London [England]: Routledge.

Parmenter, B., & Thornes, J. (1986). *The use of a computer model to predict the formation of ice on road surfaces*. Crowthorne, Berkshire: Transport and Road Research Laboratory.

Pauly, L., Hogg, D., Fuentes, R., & Peel, H. (2017, July). Deeper networks for pavement crack detection. In *Proceedings of the 34th ISARC* (pp. 479-485). IAARC.

Perrier, N., Langevin, A., & Campbell, J. F. (2006a). A survey of models and algorithms for winter road maintenance. Part I: system design for spreading and plowing. *Computers & Operations Research*, *33*(1), 209-238.

Perrier, N., Langevin, A., & Campbell, J. F. (2006b). A survey of models and algorithms for winter road maintenance. Part II: system design for snow disposal. *Computers & Operations Research*, *33*(1), 239-262.

Postgård, U., & Nunez, M. (2000). Continuous measurements of sky-view factors along roads and their relationship to air and road surface temperature. *Postgård, U (2000) Road climate variations related to weather and topography. PhD Thesis, Earth Sciences Centre A*, *52*, 2000.

Qiu, L., & Nixon, W. A. (2008). Effects of adverse weather on traffic crashes: systematic review and meta-analysis. *Transportation Research Record*, *2055*(1), 139-146.

Ramakrishna, D. M., & Viraraghavan, T. (2005). Environmental impact of chemical deicers–a review. *Water, Air, and Soil Pollution*, *166*(1), 49-63.

Sass, B. H. (1992). A numerical model for prediction of road temperature and ice. *Journal of Applied Meteorology and Climatology*, *31*(12), 1499-1506.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85-117.

Sekhon, N. S., Hassan, Q. K., & Sleep, R. W. (2010). Evaluating potential of MODIS-based indices in determining "snow gone" stage over forest-dominant regions. *Remote Sensing*, *2*(5), 1348-1363.

Shahdah, U., & Fu, L. (2010, January). Quantifying the mobility benefits of winter road maintenance—A simulation based analysis. In *TRB 89th Annual Meeting Compendium of Papers* (Vol. 33984).

Shao, J. (1998). Improving nowcasts of road surface temperature by a backpropagation neural network. *Weather and Forecasting*, *13*(1), 164-171..

Shao, J., & Lister, P. J. (1996). An automated nowcasting model of road surface temperature and state for winter road maintenance. *Journal of Applied Meteorology and Climatology*, *35*(8), 1352-1361.

Shao, J., Swanson, J. C., Patterson, R., Lister, P. J., & McDonald, A. N. (1997). Variation of winter road surface temperature due to topography and application of thermal mapping. *Meteorological Applications*, *4*(2), 131-137.

Sherif, A., & Hassan, Y. (2004). Modelling pavement temperature for winter maintenance operations. *Canadian Journal of Civil Engineering*, *31*(2), 369-378.

Shi, X., Fay, L., Gallaway, C., Volkening, K., Peterson, M. M., Pan, T., ... & Nguyen, T. A. (2009). *Evaluation of alternative anti-icing and deicing compounds using sodium chloride and magnesium chloride as baseline deicers–Phase I* (No. CDOT-2009-1). Colorado. Dept. of Transportation. Applied Research and Innovation Branch.

Snow and ice. FHWA Road Weather Management. (n.d.). Retrieved February 9, 2022, from https://ops.fhwa.dot.gov/weather/weather_events/snow_ice.htm

Suggett, J. M., Hadayeghi, A., Mills, B., Andrey, J. C., & Leach, G. (2006). Development of winter severity indicator models for Canadian winter road maintenance. In *Annual Conference & Exhibition of the Transportation Association of Canada, 2006. Congres et exposition annuels de l'Association des transport du Canada, 2006. Transportation Association of Canada (TAC)*.

Tabony, R. C. (1985). Relations between minimum temperature and topography in Great Britain. *Journal of Climatology*, *5*(5), 503-520.

Tabrizi, S. E., Xiao, K., Thé, J. V. G., Saad, M., Farghaly, H., Yang, S. X., & Gharabaghi, B. (2021). Hourly road pavement surface temperature forecasting using deep learning models. *Journal of Hydrology*, *603*, 126877.

Thi Kieu Tran, T., Lee, T., Shin, J. Y., Kim, J. S., & Kamruzzaman, M. (2020). Deep learning-based maximum temperature forecasting assisted with meta-learning for hyperparameter optimization. *Atmosphere*, *11*(5), 487.

Thornes, J. E. (1991). Thermal mapping and road-weather information systems for highway engineers. *Highway meteorology*, *39*, 67.

Tran, T. T. K., Bateni, S. M., Ki, S. J., & Vosoughifar, H. (2021). A review of neural networks for air temperature forecasting. *Water*, *13*(9), 1294.

Tsapakis, I., Cheng, T., & Bolbol, A. (2013). Impact of weather conditions on macroscopic urban travel times. *Journal of Transport Geography*, *28*, 204-211.

Usman, T., Fu, L., & Miranda-Moreno, L. F. (2010). Quantifying safety benefit of winter road maintenance: Accident frequency modeling. *Accident Analysis & Prevention*, *42*(6), 1878-1887.

Ustaoglu, B., Cigizoglu, H. K., & Karaca, M. (2008). Forecast of daily mean, maximum and minimum temperature time series by three artificial neural network methods. *Meteorological*

Applications: A journal of forecasting, practical applications, training techniques and modelling, *15*(4), 431-445.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, *14*(1), 35-62.

APPENDIX A

Table A-1 Summary of the ANN models developed for site B

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	14	58	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = ReLU
2-h ahead	3	14	57	
3-h ahead	3	14	64	
4-h ahead	4	10	104	
5-h ahead	5	14	70	
6-h ahead	6	8	50	

Table A-2 Summary of the ANN models developed for site C

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	14	68	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = ReLU
2-h ahead	3	14	57	
3-h ahead	3	14	83	
4-h ahead	4	8	91	
5-h ahead	5	6	132	
6-h ahead	6	12	38	

Table A-3 Summary of the ANN models developed for site D

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	6	88	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = ReLU
2-h ahead	3	12	74	
3-h ahead	3	10	62	
4-h ahead	4	12	91	
5-h ahead	5	6	115	
6-h ahead	6	10	50	

Table A-4 Summary of the ANN models developed for site E

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	14	64	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = ReLU
2-h ahead	3	8	74	
3-h ahead	3	10	84	
4-h ahead	4	12	94	
5-h ahead	5	12	73	
6-h ahead	6	10	50	

Table A-5 Summary of the ANN models developed for site F

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	14	62	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = ReLU
2-h ahead	3	14	70	
3-h ahead	3	14	73	
4-h ahead	4	12	88	
5-h ahead	5	14	68	
6-h ahead	6	12	70	

Table A-6 Summary of the LSTM models developed for site B

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	12	109	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = tanh
2-h ahead	3	14	131	
3-h ahead	3	14	188	
4-h ahead	4	14	236	
5-h ahead	5	12	236	
6-h ahead	6	12	246	

Table A-7 Summary of the LSTM models developed for site C

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	10	108	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = tanh
2-h ahead	3	14	127	
3-h ahead	3	14	143	
4-h ahead	4	14	166	
5-h ahead	5	14	150	
6-h ahead	6	14	147	

Table A-8 Summary of the LSTM models developed for site D

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	12	100	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = tanh
2-h ahead	3	14	126	
3-h ahead	3	14	113	
4-h ahead	4	14	138	
5-h ahead	5	14	152	
6-h ahead	6	12	120	

Table A-9 Summary of the LSTM models developed for site E

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	12	102	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = tanh
2-h ahead	3	14	123	
3-h ahead	3	10	134	
4-h ahead	4	8	173	
5-h ahead	5	8	167	
6-h ahead	6	12	103	

Table A-10 Summary of the LSTM models developed for site F

Model	No. of Lags in Input	No. of neurons	Epochs	Others
1-h ahead	3	14	114	Hidden layer = 1 Batch size = 32 Optimizer = Adam Learning Rate = 0.001 Activation = tanh
2-h ahead	3	10	125	
3-h ahead	3	12	136	
4-h ahead	4	10	115	
5-h ahead	5	10	153	
6-h ahead	6	10	135	

APPENDIX B

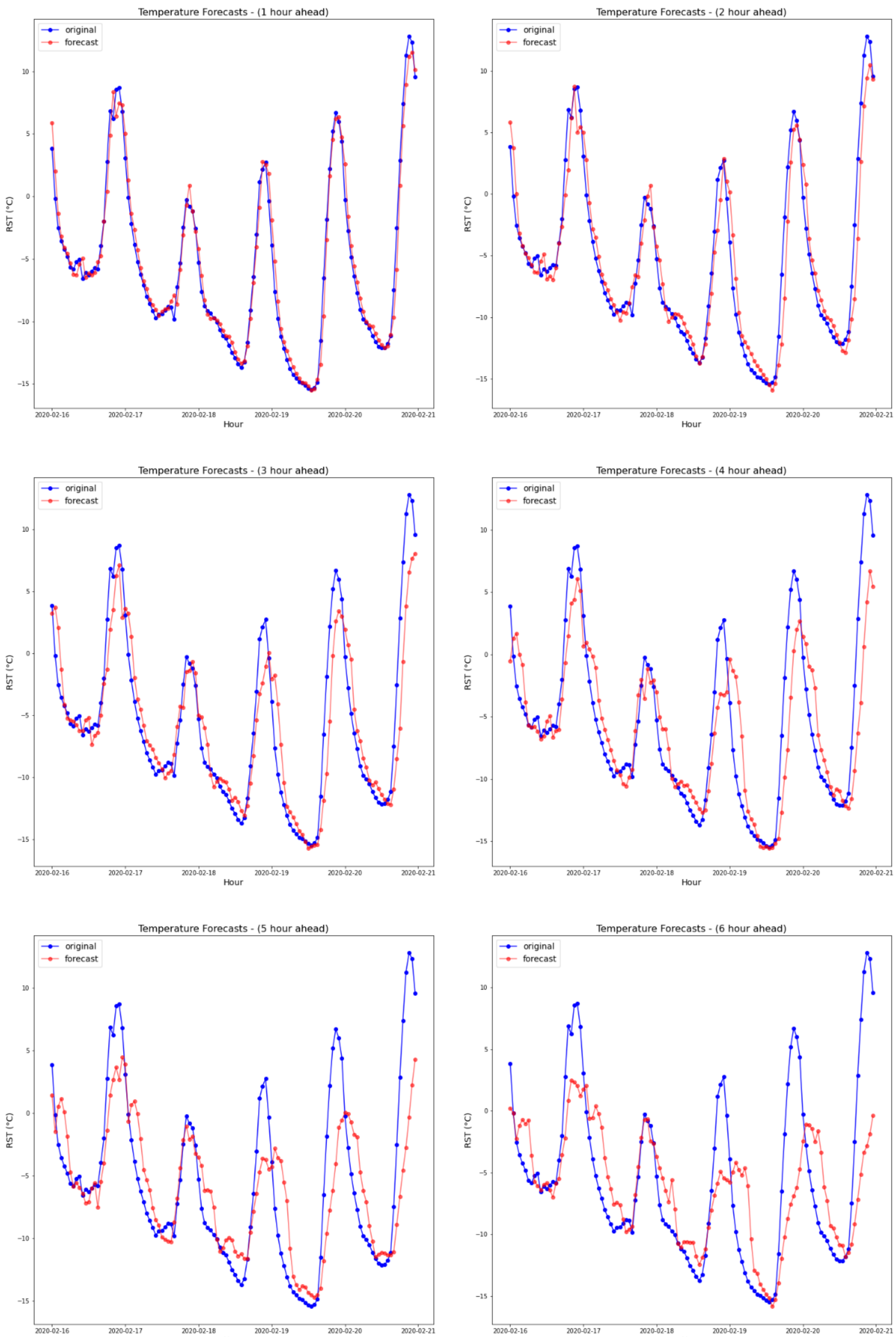


Figure B-1 Forecasted results for station B (ANN)

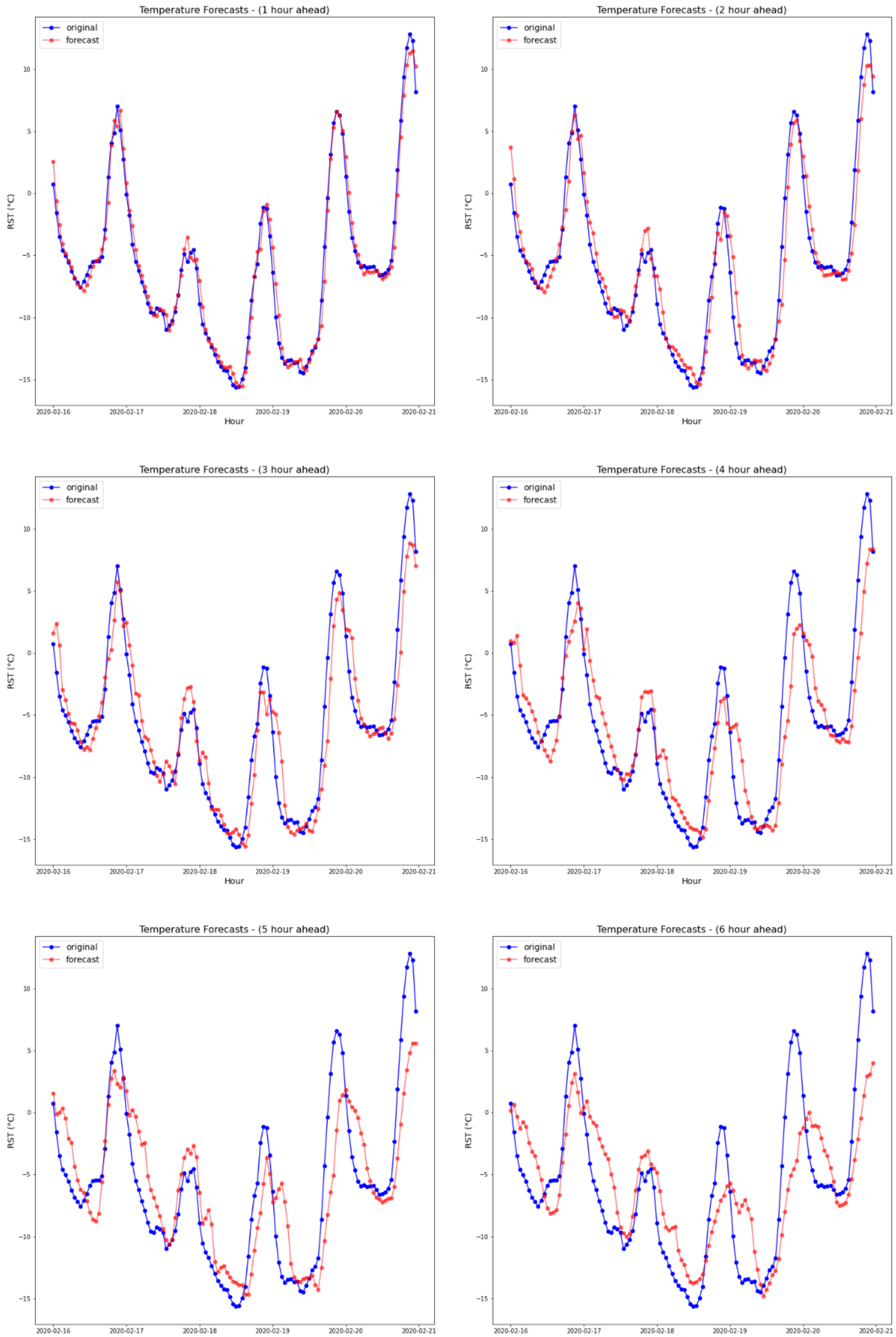


Figure B-2 Forecasted results for station C (ANN)

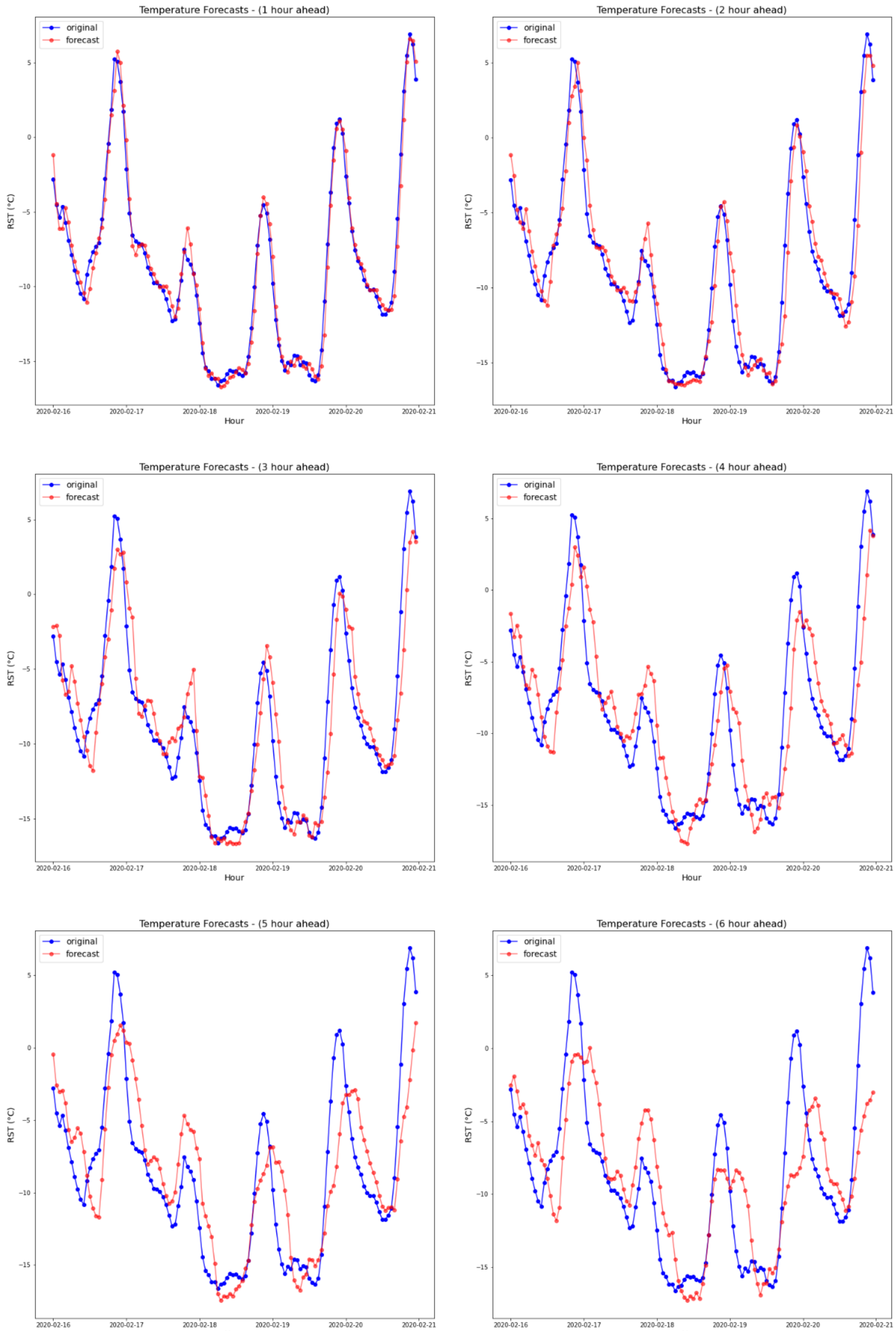


Figure B-3 Forecasted results for station D (ANN)

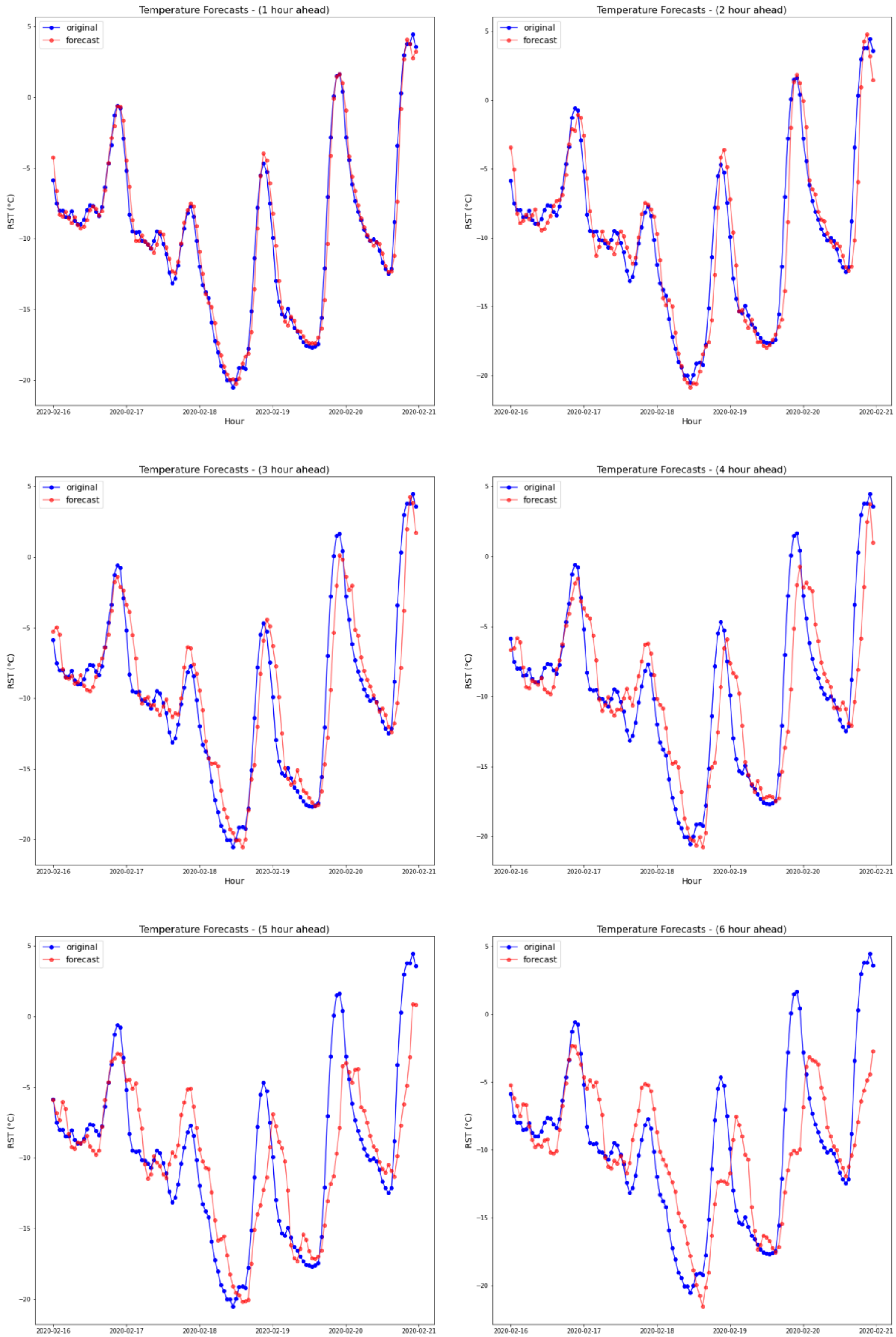


Figure B-4 Forecasted results for station E (ANN)

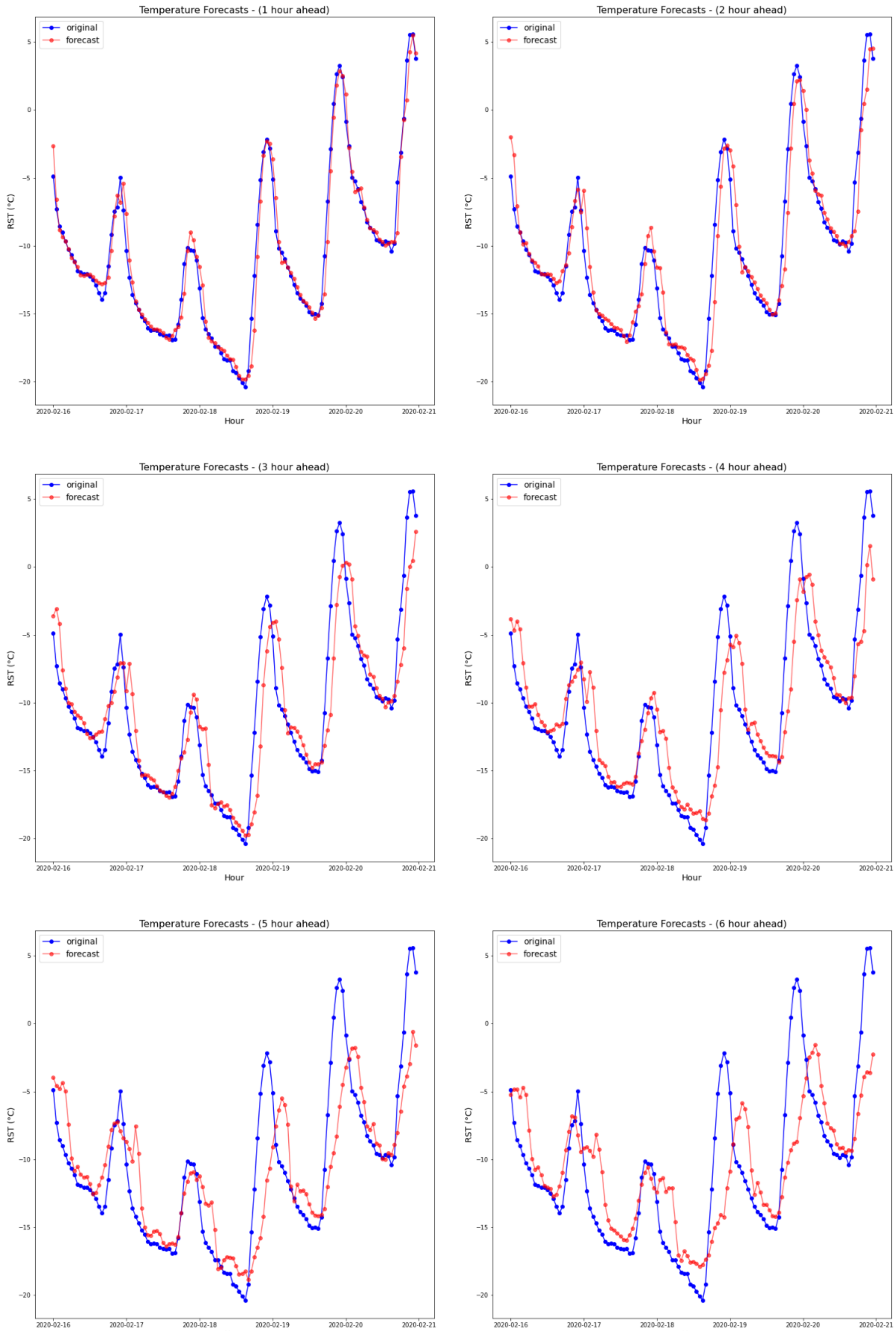


Figure B-5 Forecasted results for station F (ANN)

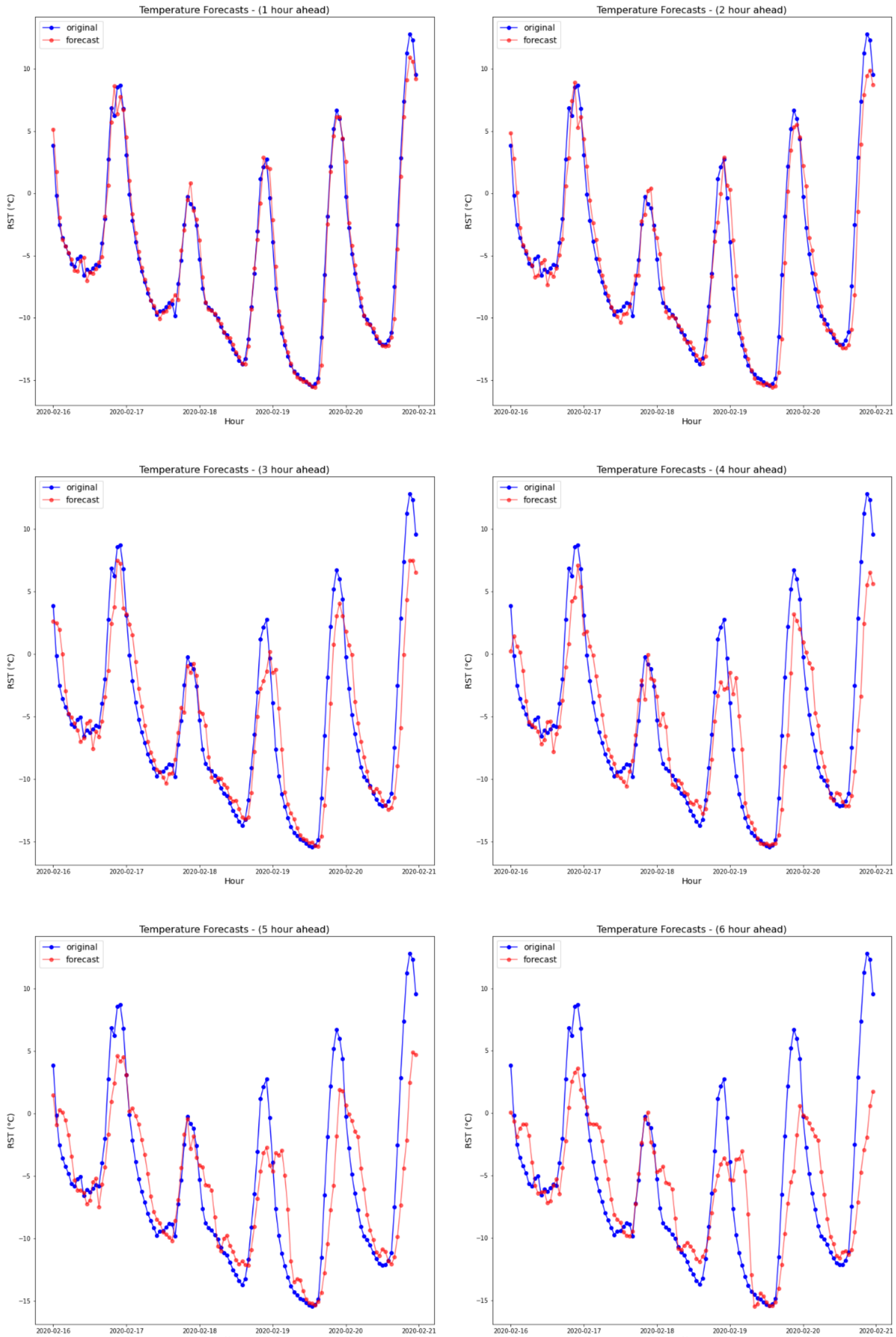


Figure B-6 Forecasted results for station B (LSTM)

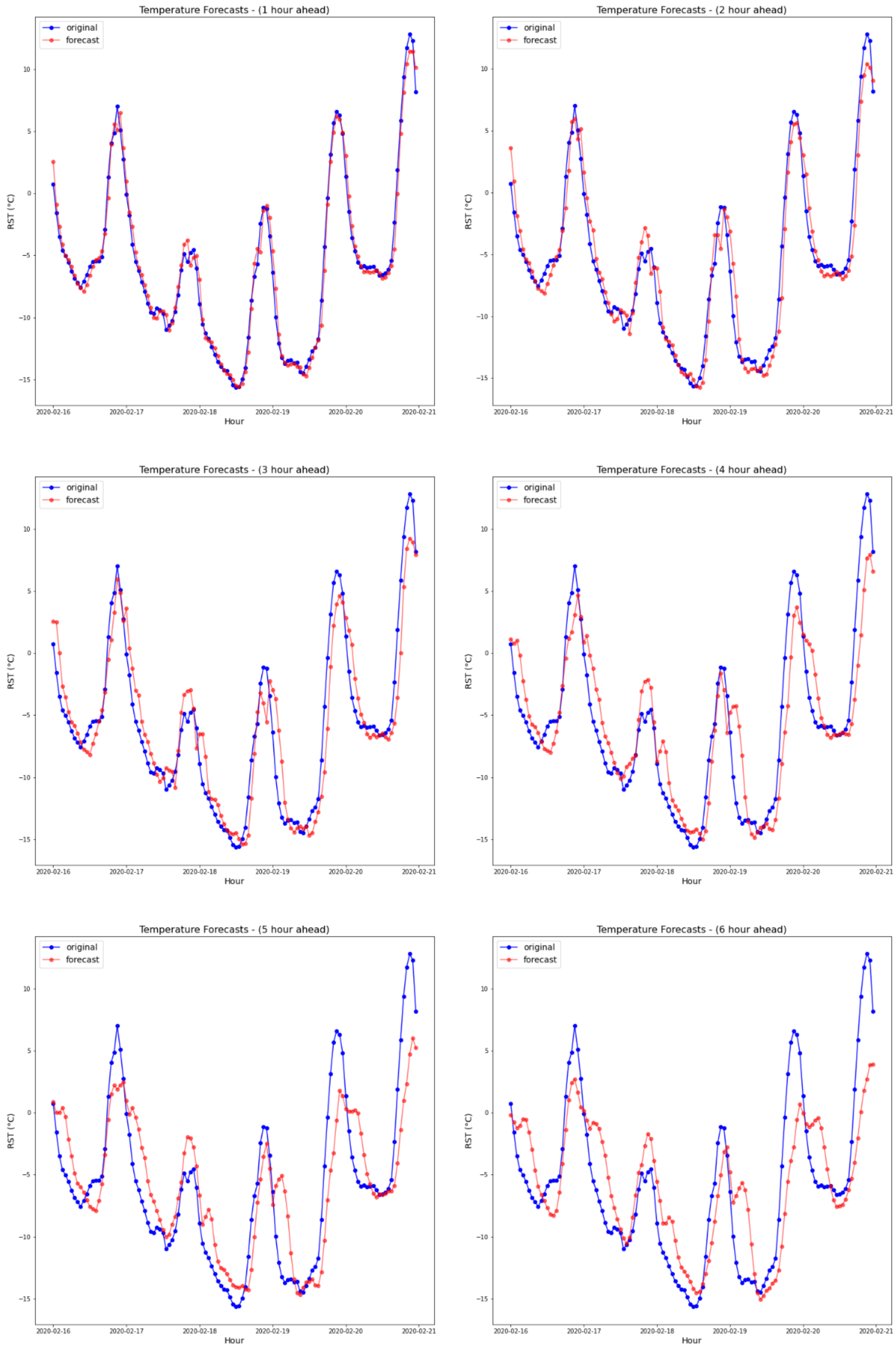


Figure B-7 Forecasted results for station C (LSTM)

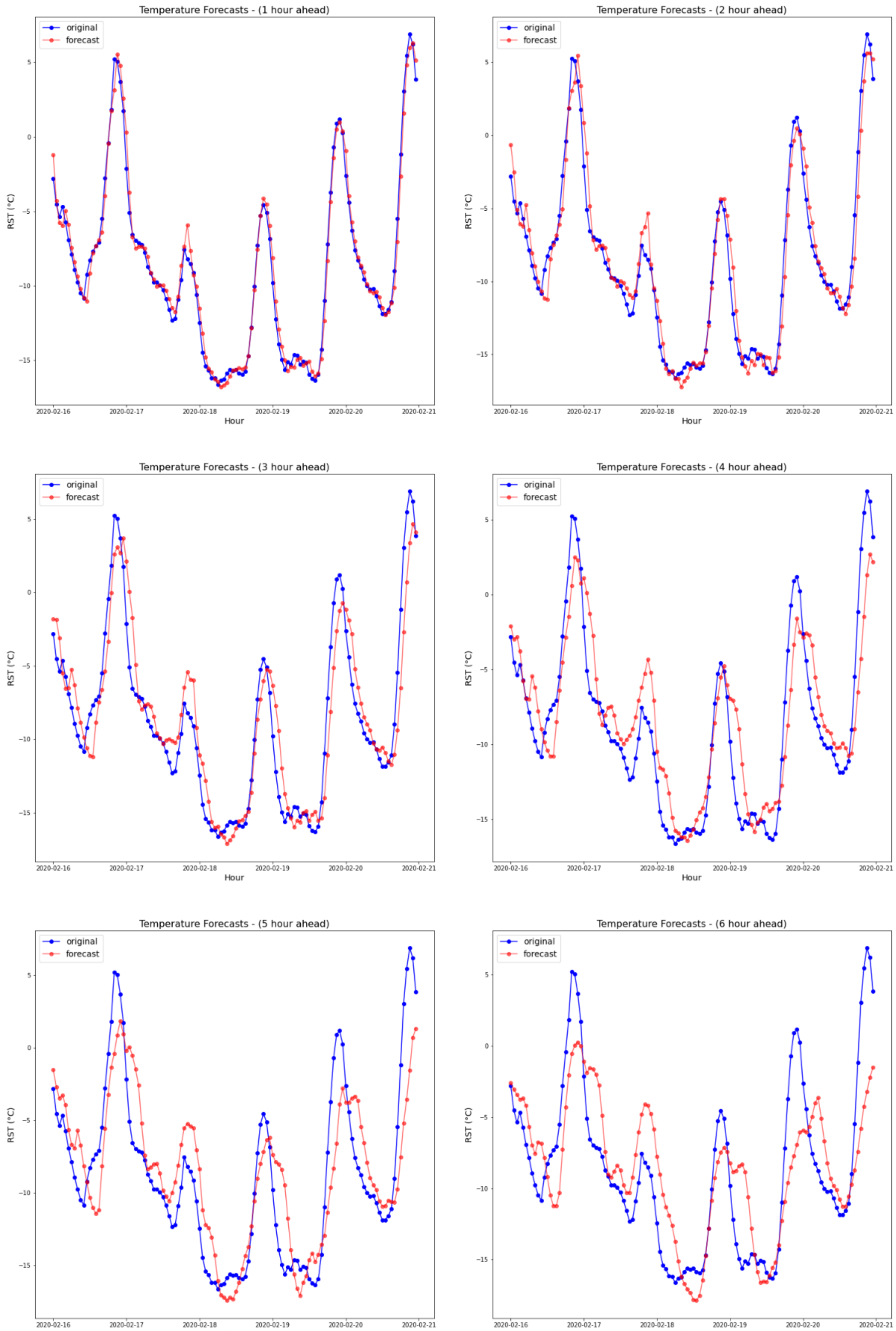


Figure B-8 Forecasted results for station D (LSTM)

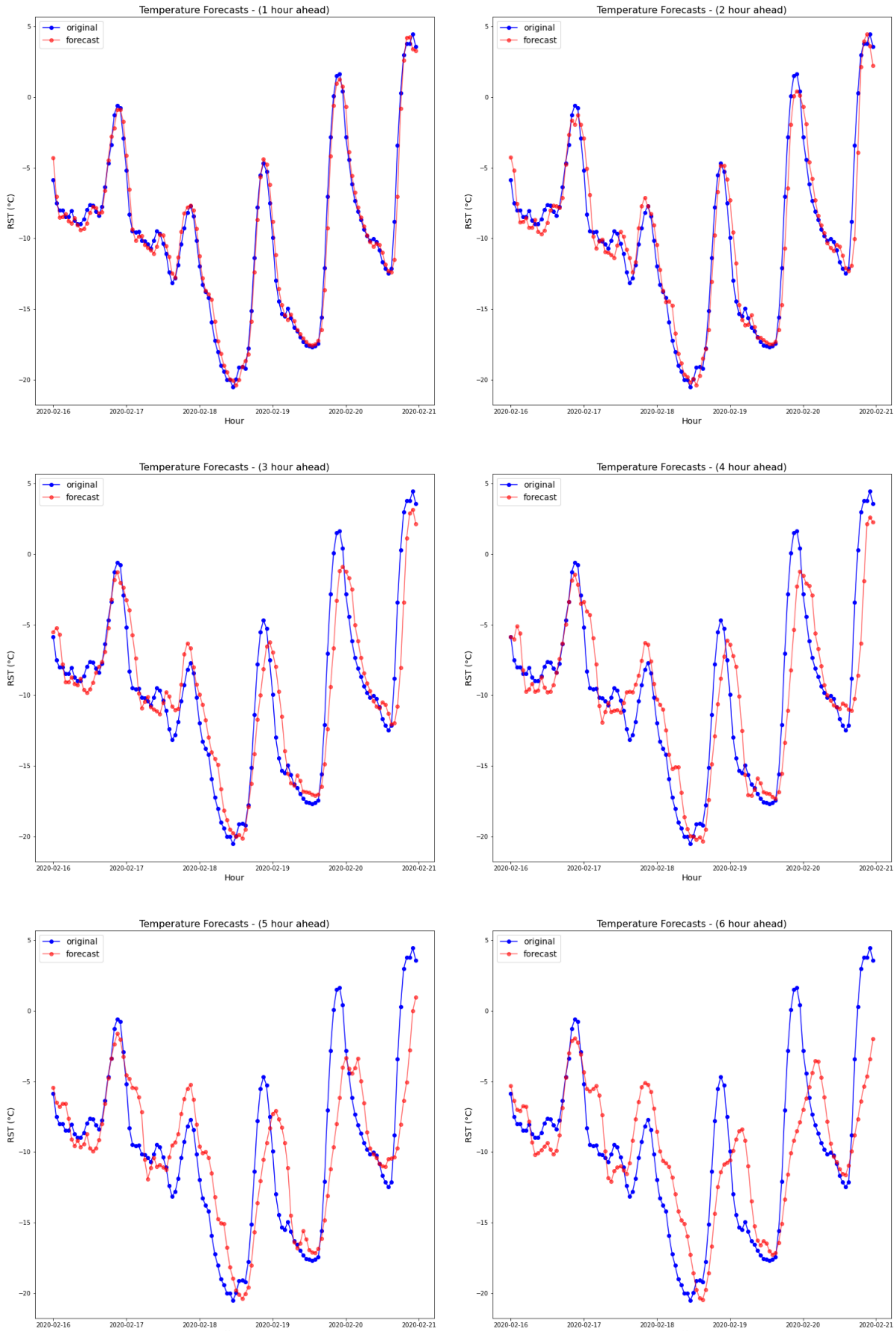


Figure B-9 Forecasted results for station E (LSTM)

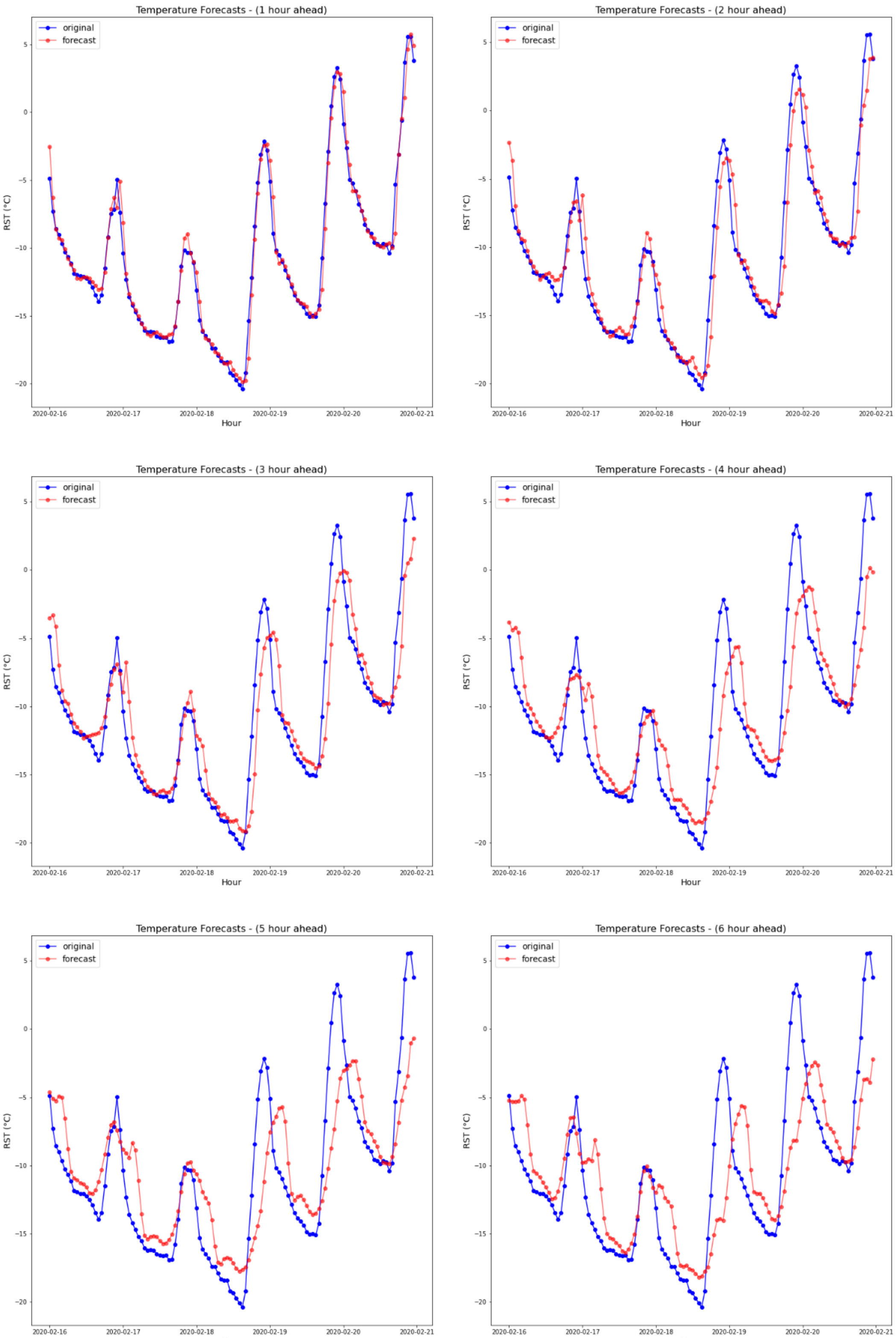


Figure B-10 Forecasted results for station F (LSTM)