

University of Texas at Arlington

**MavMatrix**

---

Civil Engineering Dissertations

Civil Engineering Department

---

2023

## Digital Twin Enabled Winter Operations Management Through the Integration of Artificial Intelligence, Sensory Level Data, and Publicly Available Data

Pooya Darghiasi

Follow this and additional works at: [https://mavmatrix.uta.edu/civilengineering\\_dissertations](https://mavmatrix.uta.edu/civilengineering_dissertations)



Part of the [Civil Engineering Commons](#)

---

### Recommended Citation

Darghiasi, Pooya, "Digital Twin Enabled Winter Operations Management Through the Integration of Artificial Intelligence, Sensory Level Data, and Publicly Available Data" (2023). *Civil Engineering Dissertations*. 436.

[https://mavmatrix.uta.edu/civilengineering\\_dissertations/436](https://mavmatrix.uta.edu/civilengineering_dissertations/436)

This Dissertation is brought to you for free and open access by the Civil Engineering Department at MavMatrix. It has been accepted for inclusion in Civil Engineering Dissertations by an authorized administrator of MavMatrix. For more information, please contact [leah.mccurdy@uta.edu](mailto:leah.mccurdy@uta.edu), [erica.rousseau@uta.edu](mailto:erica.rousseau@uta.edu), [vanessa.garrett@uta.edu](mailto:vanessa.garrett@uta.edu).

DIGITAL TWIN ENABLED WINTER OPERATIONS MANAGEMENT  
THROUGH THE INTEGRATION OF ARTIFICIAL INTELLIGENCE,  
SENSORY LEVEL-DATA, AND PUBLICLY AVAILABLE DATA

by:

POOYA DARGHIASI

DISSERTATION

Submitted In Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

Arlington, Texas

August 2023

**DIGITAL TWIN ENABLED WINTER OPERATIONS MANAGEMENT THROUGH  
THE INTEGRATION OF ARTIFICIAL INTELLIGENCE, SENSORY LEVEL-DATA,  
AND PUBLICLY AVAILABLE DATA**

**Committee Members:**

Dr. Mohsen Shahandashti, P.E., Advisor  
Department of Civil Engineering  
The University of Texas at Arlington

Dr. Kyeong Rok Ryu  
Department of Civil Engineering  
The University of Texas at Arlington

Dr. Nilo Tsung, P.E.  
Department of Civil Engineering  
The University of Texas at Arlington

Dr. Yuan Zhou  
Department of Industrial Engineering  
The University of Texas at Arlington

Copyright © by Pooya Darghiasi, 2023

All Rights Reserved

Dedicated to  
My Mother, Azar, and My Father, Behzad

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my academic advisor and mentor, Dr. Mohsen Shahandashti, for his endless support, invaluable guidance, and constant encouragement throughout my Ph.D. program. I am truly grateful for the valuable experiences and opportunities he provided me, which have contributed to my academic and professional growth. His guidance and expertise have played a crucial role in shaping my career path, both in the present and for the future.

I would also like to thank my supervising committee Dr. Kyeong Rok Ryu, Dr. Nilo Tsung, and Dr. Yuan Zhou for their insightful comments and advice throughout my Ph.D. program. I am also thankful to the Texas Department of Transportation (TxDOT) for their support of my research. Their assistance has played a vital role in the successful implementation of this research.

I extend my gratitude to my friends and colleagues at the University of Texas at Arlington, namely Bahram Abediniangerabi, Anil Baral, Ferika Farooghi, Abhijit Roy, Ahmad Bani Hani, Sooin Kim, Sumaya Sharveen, Sushil Bhatta, and Santosh Acharya for the great discussions and enjoyable moments we shared.

I would like to express my deepest gratitude to my parents, Azar Mehrjou and Behzad Darghiasi, as well as my brother, Alireza, and sister, Zahra. Their support, love, and encouragement have been instrumental throughout my journey.

Last but certainly not least, I am grateful to my beloved wife, Mina. Without her unwavering dedication and boundless love, this accomplishment would not have been within my reach.

August 2023

## **ABSTRACT**

Digital Twin Enabled Winter Operations Management Through the Integration of Artificial Intelligence, Sensory Level Data, and Publicly Available Data

Pooya Darghiasi

The University of Texas at Arlington, 2023

Supervising Professor: Dr. Mohsen Shahandashti

Monitoring real-time information on road conditions, especially during winter storms, is crucial to establishing winter maintenance strategies by the State Departments of Transportation (State DOTs) in the United States. States and local highway agencies allocate substantial resources each year for winter operations to improve road safety during winter storms. However, most weather-related vehicle accidents happen on snowy, slushy, or icy surfaces, leading to a significant number of fatalities and injuries annually. Traditionally, transportation agencies rely on the information provided by Road Weather Information Systems for monitoring road conditions along roadways. However, these systems are costly and only provide estimates at specific locations, resulting in distant areas being underrepresented. Additionally, the data acquisition systems integrated into snowplows, which offer real-time road condition images, comprise several components and are not extensively employed in states that encounter infrequent snowstorms. This is primarily due to the high costs associated with installation and maintenance. The main objective of this study is to develop a cost-effective system that enables the monitoring of road conditions by providing real-time road condition images and estimating road surface temperatures. The approach developed to collect and transfer real-time road conditions images proved advantageous as it eliminated the need

for complex and expensive multi-component systems, while also reducing the training requirements for users. Furthermore, a methodology was developed to leverage publicly accessible weather forecasts provided by the National Weather Service for estimating road surface temperatures on roadways (excluding bridges). This innovative approach enabled the estimation of road surface temperature without depending on expensive road weather information systems, which may not be accessible in many locations. In particular, statistical models were developed to establish relationships between road surface temperature and the weather forecasts that were publicly available in high resolution. The findings of the study indicated that linear statistical models, like multiple linear regression, could achieve an acceptable level of accuracy for estimating road surface temperature. However, it was observed that nonlinear models, such as Random Forest, could enhance accuracy by capturing the intricate and nonlinear interactions among the explanatory variables. Lastly, a data visualization platform (i.e., digital twin system) was created to display the real-time road conditions by combining the functionalities of mobile devices and capabilities of the ArcGIS application programming interface

The findings of this study emphasize the practicality of utilizing gridded weather forecasts, supplied by the National Weather Services, to estimate the temperature of road surface as well as utilizing the functionalities of mobile devices to communicate road conditions information. The proposed methodology can be integrated into a winter operation decision-making system to visualize the road conditions images and map the estimated road surface temperature on roadways without the need for road weather information systems. The estimated road surface temperatures on roadways assists highway agencies to plan winter maintenance strategies more efficiently by taking proactive measures in areas where low surface temperatures are estimated.



## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iii
ABSTRACT .....	iv
TABLE OF CONTENTS .....	vi
LIST OF TABLES .....	ix
LIST OF FIGURES .....	xii
CHAPTER 1 INTRODUCTION .....	1
CHAPTER 2 BACKGROUND .....	3
2.1. Road Condition Images.....	4
2.2. Weather Information.....	7
2.2.1. Current Weather Information .....	7
2.2.2. Graphical Forecast Maps .....	8
2.3. Road Surface Temperature .....	11
2.4. Gaps in Knowledge.....	13
2.5. Research Objective .....	14
CHAPTER 3 DEVELOPING A MULTI-PURPOSE, ALL-IN-ONE MOBILE DATA COLLECTION SYSTEM FOR WINTER OPERATIONS MANAGEMENT.....	15
3.1. Automatic Data Collection System for Road Conditions Images .....	15
3.1.1. Methodology .....	18
3.2. On-Demand Data Collection System for Road Conditions Images.....	24
3.2.1. Methodology .....	25
3.3. Results.....	28
CHAPTER 4 Estimation of Road Surface Temperature Using NOAA Gridded Forecast Weather Data .....	33
4.1. Methodology .....	34
4.1.1. Data collection .....	35

4.1.2.	Developing Statistical Models .....	40
4.1.3.	Multiple Linear Regression.....	41
4.1.4.	Leverage Gridded Forecast Weather Data to Visualize Estimated Road Surface Temperatures on a Map-based Interface.....	45
4.2.	Results and Discussions .....	48
4.3.	Summary .....	55
CHAPTER 5 Developing a Digital Twin System for Visualizing Road Conditions Information .....		57
5.1.	Introduction.....	57
5.2.	Spatial Data Collection and Processing .....	61
5.2.1.	Road Condition Images.....	61
5.2.2.	Related Weather Information .....	64
5.2.3.	Road Surface Temperatures .....	67
5.2.4.	Data Model Development for Snowplow Operations Management System .. .....	68
5.3.	Digital Twin Interface Development .....	80
5.3.1.	Use Cases .....	82
5.4.	Summary .....	86
CHAPTER 6 Improved Road Surface Temperature Prediction Using Random Forest Machine Learning Algorithm Based on Weather Forecasts .....		88
6.1.	Introduction.....	88
6.2.	Methodology .....	90
6.2.1.	Data collection .....	92
6.2.2.	Random Forest model development .....	93
6.3.	Results and Discussions .....	97
6.3.1.	Descriptive Statistics .....	97
6.3.2.	Random Forest Hyperparameter Optimization .....	99
6.3.3.	Feature Importance.....	101

6.3.4. Comparison between RF, DL, KNN, MLR, M5P Decision Tree, and SVR ..	
.....	102
6.4. Summary .....	105
CHAPTER 7 CONCLUSION AND FUTURE WORK .....	108
REFERENCES.....	111

## LIST OF TABLES

Table 2-1 Weather impacts on roads, traffic, and operational decisions (FHWA, 2020).....	3
Table 2-2 Data collection system’s components, installed on snowplow fleet, to collect road condition images .....	6
Table 2-3 State of practice for visualizing road condition information during winter storms; retrieved from state travelers’ road information websites (FHWA, 2021).....	10
Table 2-4 Examples of linear regression models to estimate the pavement temperature .....	13
Table 3-1- Technical specifications of the tablet used to test the custom Android application (Source: Samsung website).....	17
Table 3-2 Example metadata for a captured image by the custom application .....	22
Table 3-3- Use Case for the developed Android application to provide on-demand road conditions images .....	27
Table 4-1 Technical specifications of the RoadWatch® temperature sensor used to collect actual road surface temperature data (RoadWatch® manual, 2020).....	37
Table 4-2 Explanatory variables used in the statistical analysis to estimate road surface temperature .....	38
Table 4-3 Light and dark hours in North Texas during data collection months .....	43
Table 4-4 Data ranges for the statistically significant continuous attributes .....	43
Table 4-5 Data range for the statistically significant binary attribute .....	43
Table 4-6 MLR analysis results for dark group .....	50
Table 4-7- MLR analysis results for light group .....	51
Table 4-8 Summary of statistical models developed to estimate the road surface temperature ...	52

Table 4-9 Summary of the accuracy metrics for previous road surface temperature prediction models .....	53
Table 5-1 Maximum number of snowplows that the existing system can handle from the cloud storage space aspect. ....	63
Table 5-2 Maximum number of snowplows that the system can handle from the computational performance of the server— images will remain posted on the map-based interface for one hour. ....	64
Table 5-3- Conceptual data model terms description .....	69
Table 5-4 Description of the Snowplow Operations Management System’s data entities .....	71
Table 5-5 Description of widgets in the map-based Interface .....	81
Table 5-6 UC1: Login .....	83
Table 5-7 UC2: Sign up .....	84
Table 5-8 UC3: Search.....	84
Table 5-9 UC4: Return to initial map view extent.....	84
Table 5-10 UC5: Zoom-in and zoom-out of the map view .....	84
Table 5-11 UC6: Find the user location.....	85
Table 5-12 UC7: Change base map .....	85
Table 5-13 UC8: Display data entity .....	85
Table 5-14 UC9: Display the legend of the data entity .....	86
Table 5-15 UC10: Display road condition data .....	86
Table 5-16 UC11: Logout.....	86
Table 6-1 Summary of the variables used for developing the predictive models.....	92
Table 6-2- Random Forest model hyperparameters.....	95

Table 6-3 Point Biserial Correlation between the categorical variables and RST .....	97
Table 6-4 Thresholds for interpretation of Spearman correlation coefficients (Dancey and Reidy, 2007) .....	99
Table 6-5 Hyperparameters for the default and optimized Random Forest models .....	100
Table 6-6- Optimized hyperparameters of selected data-driven prediction models .....	103

## LIST OF FIGURES

Figure 2-1 U.S. states using GPS/AVL systems along with dash cameras in their snowplow fleet (Potter et al., 2016; Refai et al., 2018) .....	5
Figure 2-2- Mobile data collection system; (a) communication unit, (b) GPS antenna, (c) mobile data computer unit (Schneider et al., 2017) .....	6
Figure 2-3 A typical environmental weather station ( <i>Source: The Federal Highway Administration</i> ) .....	8
Figure 3-1 (a) Wichita Falls' snowplow, (b) Mounted tablet in the snowplow .....	16
Figure 3-2 Samsung Galaxy Tab A 8-inch tablet .....	16
Figure 3-3 Data flow from the tablet to a map-based interface .....	18
Figure 3-4- Function Modeling Methodology for the custom Android application in Java language .....	19
Figure 3-5 Uploading Image Metadata to Cloud Space.....	23
Figure 3-6 The user interface of the custom Android application .....	24
Figure 3-7 Data flow from the Android application to the ArcGIS map-based application.....	25
Figure 3-8 Function modeling methodology to develop the Android application to collect on-demand images.....	26
Figure 3-9 The user interface of the Android application to collect on-demand road conditions images .....	27
Figure 3-10 (a) Tablet mounted on the front windshield using a suction-cup mount, and (b) USB cable and power outlet used to provide power for the tablets.....	29
Figure 3-11 Examples of collected road condition images in the 2020-21 winter season (automatic) .....	30

Figure 3-12 Examples of road conditions images in the 2020-21 winter season (on-demand)....	31
Figure 4-1 Overview of the methodology from data collection to the estimation of road surface temperatures along with the locations of snowplows .....	35
Figure 4-2 Data collection: location and road functions.....	36
Figure 4-3 Temperature sensor used to collect road surface temperature data; (a) RoadWatch® sensor kit; (b) mounted sensor on vehicle’s side mirror .....	37
Figure 4-4 Total sample size required for the statistical analysis .....	39
Figure 4-5 Frequency distribution of the collected data .....	40
Figure 4-6 IDEF0 diagram for converting GRIB2 file downloaded from the NDFD to ASCII comma-separated .....	46
Figure 4-7 Overview of methodology to map the estimated road surface temperature onto the road segments.....	47
Figure 4-8 Flowchart for estimation of road surface temperature using forecast weather data from the NDFD.....	48
Figure 4-9 Pearson’s correlation coefficient between all variables in the dataset .....	49
Figure 4-10 Sample screenshot of the map-based interface showing estimated minimum road surface temperatures for the next day in TxDOT on-system roadways in North Texas (February 15, 2022). .....	54
Figure 5-1 Representation of the physical world based on the level of data integration .....	57
Figure 5-2 Digital twin technologies .....	59
Figure 5-3 Digital Twin System for Winter Operations Management .....	60
Figure 5-4- Flowchart to visualize the road condition images in the map-based interface .....	62



Figure 5-5- Road condition images data entity; displaying road condition images collected by operating snowplows (February 14, 2021, at 5:37 PM).....	62
Figure 5-6 Overview of collecting and displaying weather information that facilitates winter operations decisions. ....	65
Figure 5-7- Graphical Forecast Maps retrieved from National Weather Service: (a) Snowfall forecast by the National Weather Service (Texas, January 8, 2021, 4:15 PM), (b) Snowfall forecast by the National Weather Service (TxDOT Wichita Falls district, January 8, 2021, 4:23 PM), and (c) Precipitation forecast by the National Weather Service (TxDOT Wichita Falls district, January 8, 2021, 4:37 PM) .....	66
Figure 5-8- Examples of watches, warnings, and advisories issued by the National Weather Service (a) Advisory from the National Weather Service during a winter storm in Texas (January 8, 2021, 1:31 PM), and (b) Warning from the National Weather Service during in TxDOT Wichita Falls district (January 8, 2021, 1:50 PM) .....	66
Figure 5-9 IDEF0 diagram from downloading the data to visualizing the road surface temperatures on the map-based interface. ....	67
Figure 5-10- Road surface temperatures estimations along with snowplow locations in the Wichita Falls district.....	68
Figure 5-11- Conceptual Data Model .....	69
Figure 5-12- The Winter Operations Management System’s data entities.....	70
Figure 5-13 Sign-in webpage to access the map-based interface .....	80
Figure 5-14 Map-based interface for Winter Operations Management System .....	81
Figure 5-15 Use case diagram.....	83
Figure 6-1 Overview of the proposed methodology framework .....	91

Figure 6-2 Structure of Random Forest models (adapted from Gitconnected.com).....	93
Figure 6-3 Overview of cross-validation in hyperparameter optimization of Random Forest model .....	96
Figure 6-4 Spearman Correlation between the publicly available weather data and RST .....	98
Figure 6-5 Mean absolute error of Random Forest model based on the number of trees and maximum number of features for splitting a node.....	100
Figure 6-6 RMSE, MAE, and R-squared of optimized and default Random Forest models .....	101
Figure 6-7 The relative importance of predictive features for the Random Forest Model .....	102
Figure 6-8 Coefficients of determination for the testing datasets of developed DL, MLR, M5P decision tree, SVR, RF, and SVR.....	104
Figure 6-9 Accuracy metrics as well as the time taken to build the models.....	105

# CHAPTER 1

## INTRODUCTION

Winter road maintenance accounts for approximately 20 percent of state departments of transportation's maintenance budgets; states and local highway agencies spend an average of \$2.3 billion on winter operations every year (FHWA, 2020). Nevertheless, over 5 million vehicle crashes occur in the U.S. each year, and approximately 21% of these crashes happen in the presence of adverse weather (i.e., sleet, snow, etc.) (FHWA, 2020). According to National Highway Traffic Safety Administration (NHTSA), about 5,000 people are killed, and over 418,000 are injured in weather-related crashes each year (FHWA, 2020). These crashes also contribute to approximately \$70.7 billion in property damages, including damages to the road infrastructure (Miller and Zaloshnja, 2009). Real-time road conditions information, especially road conditions images and road weather information, provide valuable information to transportation operations managers to enhance their winter operations practices and improve the safety of the roads (Ameen et al., 2022; Shahandashti et al., 2019).

The primary objectives of this research are to: (1) develop a cost-effective approach to collect and transfer a live feed of road conditions images from mobile devices mounted on snowplows, (2) develop predictive models to estimate road surface temperatures (excluding bridges) using the publicly available weather data from the National Weather Service, and (3) and develop a digital twin system to visualize the real-time road conditions information, including the road conditions images, weather information from the National Weather Service, and estimated road surface temperatures using capabilities of ArcGIS application programming interface.

This research helps communicate adverse road conditions, improve snowplow operational decisions, and consequently decrease weather-related crashes. Real-time images of road conditions

help transportation operations managers visually monitor road conditions and make well-informed decisions during snowstorms. In addition, access to certain weather information that facilitates snowplow operations decisions, as well as information about road surface temperatures provide essential information to operations managers about possible locations of low road surface temperatures and potential ice and snow hazards on roads. This information could improve decision-making for deploying snowplows to administer anti-icing and snow-removal measures on roads during winter operations.

Chapter 2 provides a comprehensive review of the literature on the use of road conditions information to enhance winter operations practices in the U.S. Chapter 3 explains developing a multi-purpose, all-in-one data collection system for collecting and transferring road conditions images. Chapter 4 explains developing an approach to estimate road surface temperatures based on publicly available weather data using multiple linear regression models. Chapter 5 explains developing a 2D digital twin system to visualize the collected road conditions information. Chapter 6 explains developing nonlinear statistical models to improve the accuracy of the prediction models to estimate the road surface temperature by considering the nonlinear and complex interaction between the influencing variables. Chapter 7 provides a summary of this research and suggests future research directions.

## CHAPTER 2

### BACKGROUND

Winter road maintenance accounts for approximately 20 percent of State Departments of Transportation’s (State DOTs) maintenance budgets; states and local agencies spend over \$2.3 billion on winter operations annually (FHWA, 2020). Adverse weather will act through low visibility, precipitation, high winds, and extreme temperature to affect driver capabilities and vehicle performance (FHWA, 2020). Table 2-1 summarizes the impacts of various weather events on roadways, traffic flow, and operational decisions.

Table 2-1 Weather impacts on roads, traffic, and operational decisions (FHWA, 2020)

Road Weather Variables	Impacts on Roadway	Impacts on Traffic Flow	Operational Impacts
<b>Air temperature and humidity</b>	N/A	N/A	<ul style="list-style-type: none"> <li>▪ Road treatment strategy (e.g., snow and ice control)</li> <li>▪ Construction planning (e.g., paving and striping)</li> </ul>
<b>Wind speed</b>	<ul style="list-style-type: none"> <li>▪ Visibility distance (due to blowing snow and dust)</li> <li>▪ Lane obstruction (due to wind-blown snow and debris)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Traffic speed</li> <li>▪ Travel time delay</li> <li>▪ Crash risk</li> </ul>	<ul style="list-style-type: none"> <li>▪ Vehicle performance (e.g., stability)</li> <li>▪ Access control (e.g., restrict vehicle type, close road)</li> <li>▪ Evacuation decision support</li> </ul>
<b>Precipitation (type, rate, start/end times)</b>	<ul style="list-style-type: none"> <li>▪ Visibility distance</li> <li>▪ Pavement friction</li> <li>▪ Lane obstruction</li> </ul>	<ul style="list-style-type: none"> <li>▪ Roadway capacity</li> <li>▪ Traffic speed</li> <li>▪ Travel time delay</li> <li>▪ Crash risk</li> </ul>	<ul style="list-style-type: none"> <li>▪ Vehicle performance (e.g., traction)</li> <li>▪ Driver capabilities/behavior</li> <li>▪ Road treatment strategy</li> <li>▪ Traffic signal timing</li> <li>▪ Speed limit control</li> <li>▪ Evacuation decision support</li> <li>▪ Institutional coordination</li> </ul>
<b>Fog</b>	<ul style="list-style-type: none"> <li>▪ Visibility distance</li> </ul>	<ul style="list-style-type: none"> <li>▪ Traffic speed</li> <li>▪ Speed variance</li> <li>▪ Travel time delay</li> <li>▪ Crash risk</li> </ul>	<ul style="list-style-type: none"> <li>▪ Driver capabilities/behavior</li> <li>▪ Road treatment strategy</li> <li>▪ Access control</li> <li>▪ Speed limit control</li> </ul>

<b>Road Weather Variables</b>	<b>Impacts on Roadway</b>	<b>Impacts on Traffic Flow</b>	<b>Operational Impacts</b>
<b>Pavement temperature</b>	<ul style="list-style-type: none"> <li>▪ Infrastructure damage</li> </ul>	N/A	<ul style="list-style-type: none"> <li>▪ Road treatment strategy</li> </ul>
<b>Pavement condition</b>	<ul style="list-style-type: none"> <li>▪ Pavement friction</li> <li>▪ Infrastructure damage</li> </ul>	<ul style="list-style-type: none"> <li>▪ Roadway capacity</li> <li>▪ Traffic speed</li> <li>▪ Travel time delay</li> <li>▪ Crash risk</li> </ul>	<ul style="list-style-type: none"> <li>▪ Vehicle performance</li> <li>▪ Driver capabilities/behavior (e.g., route choice)</li> <li>▪ Road treatment strategy</li> <li>▪ Traffic signal timing</li> <li>▪ Speed limit control</li> </ul>
<b>Water level</b>	<ul style="list-style-type: none"> <li>▪ Lane submersion</li> </ul>	<ul style="list-style-type: none"> <li>▪ Traffic speed</li> <li>▪ Travel time delay</li> <li>▪ Crash risk</li> </ul>	<ul style="list-style-type: none"> <li>▪ Access control</li> <li>▪ Evacuation decision support</li> <li>▪ Institutional coordination</li> </ul>

Information on real-time road conditions, including road conditions images, weather information, and road surface temperature, can assist transportation operations managers in improving their winter operations practices. The following subsections synthesize the state of practice in providing real-time road conditions information in the U.S.

## 2.1. Road Condition Images

Road condition images can provide valuable information for transportation operation managers to enhance their winter operation practices (Shahandashti et al., 2019). Also, real-time road condition images could alert the traveling public about a potential safety hazard on the roads during inclement weather to help them make well-informed decisions (Hirt and Peterson, 2017). In recent years, various agencies across the U.S. have adopted the use of a GPS-based Automatic Vehicle Location (GPS/AVL) system along with cameras for their snowplow fleet to collect and transfer snowplow locations as well as road condition images during winter operations (Potter et al., 2016). Surveys have shown that integrating such a GPS/AVL system into the snowplow fleet could result

in efficiency savings from 5% to 50% due to reduced crashes on roads and more efficient fleet management (Meyer and Ahmad, 2003).

According to a survey in 2016, 26 U.S. states have reported using a GPS/AVL system for their winter operation fleet, in which eight State DOTs have integrated dash cameras into their GPS/AVL system to communicate road conditions images at predetermined time intervals (Potter et al., 2016). In 2018, Oklahoma DOT initiated a project to collect road condition images by installing tablets connected to a piece of computing equipment, along with a GPS device to collect and transfer road condition images (Refai et al., 2018). Figure 2-1 shows the states that have integrated dash cameras to the GPS/AVL system to collect and transfer road condition images during winter operations.

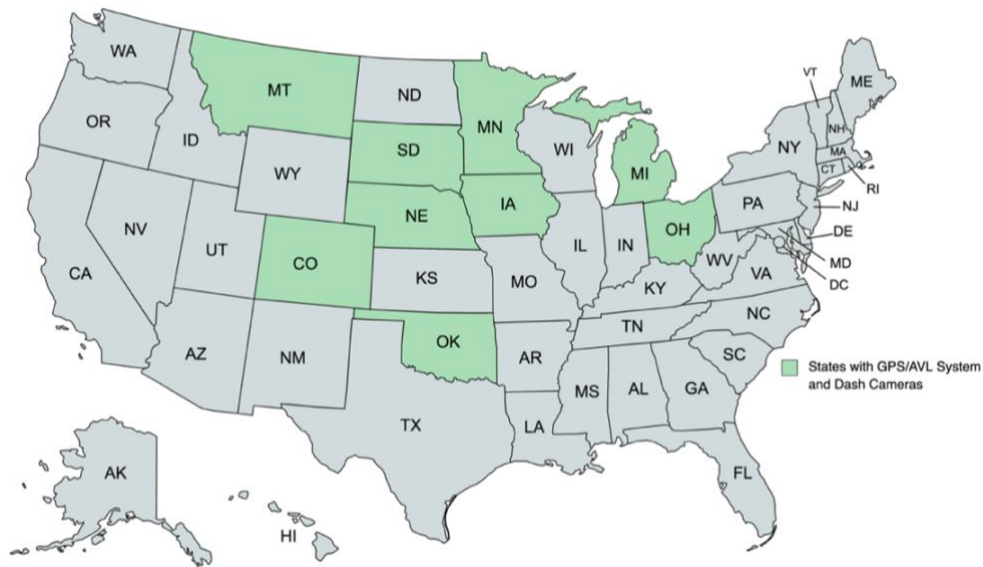


Figure 2-1 U.S. states using GPS/AVL systems along with dash cameras in their snowplow fleet (Potter et al., 2016; Refai et al., 2018)

The existing system is consisted of various detached equipment pieces, such as a GPS receiver with an antenna, a cellular modem for communication, and mobile data computer (Lee and Nelson, 2018). Figure 2-2 shows examples of equipment pieces used in the existing GPS/AVL data collection systems to communicate road condition data during snowplow operation. The existing

GPS/AVL systems could cost as much as \$25,000, with a median value of \$1,500 for operations on the snowplow fleet (Potter et al., 2016).

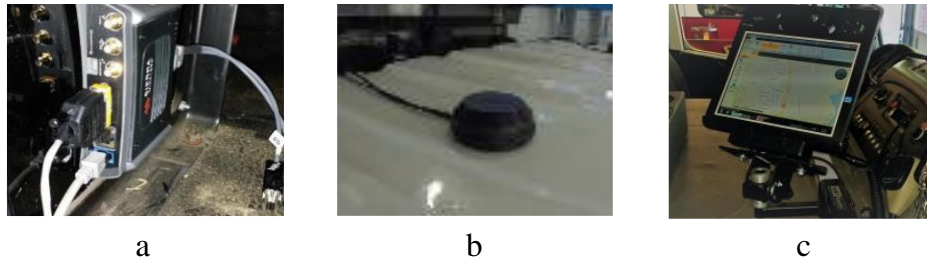


Figure 2-2- Mobile data collection system; (a) communication unit, (b) GPS antenna, (c) mobile data computer unit (Schneider et al., 2017)

A summary of the installed equipment pieces on the snowplow fleet, which gather and transfer road condition data in nine U.S. states, can be found in Table 2-2 (Potter et al., 2016; Refai et al., 2018).

Table 2-2 Data collection system’s components, installed on snowplow fleet, to collect road condition images

Department of Transportation	Data Collection System’s Components				
	GPS device	Communication modem	Computing equipment	Dash camera	Tablet/phone
Iowa	✓	✓	✓	✓	X
Colorado	✓	✓	✓	✓	X
Nebraska	✓	✓	✓	✓	X
Minnesota	✓	✓	✓	✓	X
Oklahoma	✓	✓	✓	X	✓
Michigan	✓	✓	✓	✓	X
Ohio	✓	✓	✓	✓	X
South Dakota	✓	✓	✓	✓	X
Montana	✓	✓	✓	✓	X

The collected data are typically transferred via a cellular connection, radio transmission, or satellite to a database where it could be further processed and visualized. As the collected data are from multiple sources with different data structures, highway agencies mainly use GIS-based tools to integrate all the collected data into a single platform to share spatial information through interactive maps.



## **2.2. Weather Information**

Access to weather information from observing systems and forecast providers provides essential information to winter operations managers about possible locations of low road surface temperatures and potential ice and snow hazards on roads. The following subsections provide information on the current and forecast weather information provided by the national weather service.

### **2.2.1. Current Weather Information**

The current weather information is usually collected from fixed weather stations such as Meteorological Aerodrome Reports (METAR) or Road Weather Information Systems (RWIS). Typical METAR data contain information about several weather variables for each land-based Environmental Sensor Station location (The Federal Aviation Administration, 2017):

- Air Temperature
- Dew Point
- Wind Speed and Direction
- Relative Humidity
- Horizontal Visibility
- Weather Condition
- Precipitation
- Sky Condition (Cloud Cover and Heights)
- Barometric Pressure

Environmental Sensor Stations are utilized across the U.S. to provide weather data for the public and government agencies. An Environmental Sensor Station contains various types of instrumentation, such as temperature sensors, wind sensors, and barometric pressure sensors, to collect meteorological data from the field. The Environmental Sensor Stations are mainly

administered by external agencies such as the National Weather Service, the Federal Aviation Administration, the US Geological Survey, the Department of Transportation, the Forest Service, and the Environmental Protection Agency. Figure 2-3 shows a typical environmental sensor station currently used for winter operations practices in the U.S.

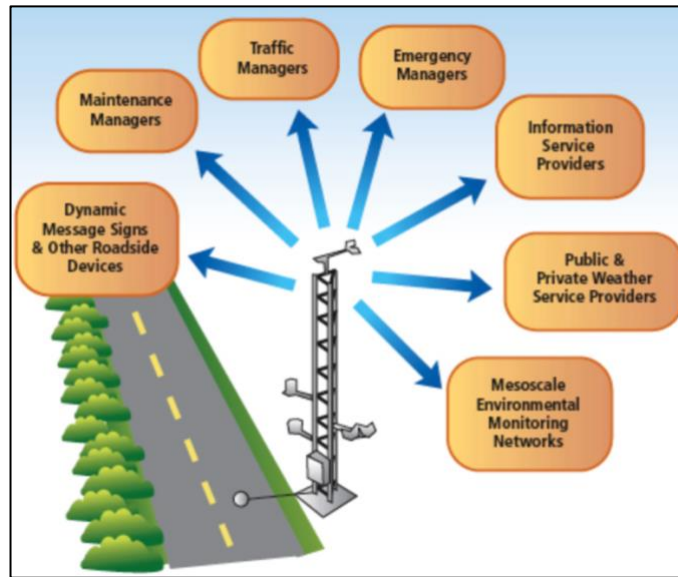


Figure 2-3 A typical environmental weather station (Source: *The Federal Highway Administration*)

### 2.2.2. Graphical Forecast Maps

The National Weather Service (NWS) is an agency of the United States federal government that provides graphical forecast maps to visualize weather forecasts, warnings of hazardous weather, and other weather-related products for organizations and the public to enhance protection, safety, and general information. The NWS's National Digital Forecast Database (NDFD) uses data from regional NWS Weather Forecast Offices and the National Centers for Environmental Prediction. Below are some examples of forecast weather data provided by the NDFD:

- Precipitation forecast
- Snowfall forecast

- Dew point forecast
- Ice accumulation forecast
- Wind speed forecast
- Relative humidity forecast

The graphical forecast maps include incremental and cumulative data for snowfall, precipitation, and ice accumulation in 6-hour intervals. The incremental and cumulative data facilitate determining the intensity and average of forecasted weather data, such as snowfall over a specific period of time.

#### **1.1.1. Warnings, Watches, and Advisories**

The National Weather Service uses weather data, radar, and satellite analysis to alert the public and related agencies to a potentially threatening weather event. Warnings, watches, and advisories are mainly issued for events such as winter storms, blizzards, ice storms, frost, and wind chill during wintertime. The issued warnings, watches, and advisories by National Weather Service can help transportation managers have their resources ready for winter operations ahead of time and deploy them to locations where the watches, warnings, and advisories are issued for. Below is a description of the terminologies used by the National Weather Service for warnings, watches, and advisories (NOAA, 2021).

- Warning: A warning is issued by National Weather Service when hazardous weather or hydrologic event is likely or imminent to occur. A warning means weather conditions pose a threat to life or property. People in the path of the storm need to take protective action.
- Watch: A watch is issued by National Weather Service when the risk of hazardous weather or hydrologic event has increased significantly, but its occurrence, location, or timing is

still uncertain. It is intended to provide enough lead time so that the public or officials will be able to set preparation plans ahead of the adverse weather condition.

- Advisory: An advisory is issued by National Weather Service when hazardous weather or hydrologic event is likely or imminent to occur. Advisories are for less severe conditions than warnings that cause significant inconvenience and, if caution is not exercised, could lead to situations that may threaten life or property.

Weather information from national services has been used by multiple U.S. highway agencies for winter operations decision-making in recent years. Examples of the road condition’s geospatial weather data used in multiple state DOTs are summarized in Table 2-3.

Table 2-3 State of practice for visualizing road condition information during winter storms; retrieved from state travelers’ road information websites (FHWA, 2021)

Department of Transportation	Related weather data entities					
	Fixed Weather Stations	Weather radar and Warnings (NWS)	Snowfall Forecast (NWS)	Ice Forecast (NWS)	Wind Forecast (NWS)	Precipitation Forecast (NWS)
Iowa	✓	✓	X	X	X	X
Colorado	✓	X	X	X	X	X
Arizona	✓	✓	X	X	X	X
Pennsylvania	✓	✓	X	X	X	X
Minnesota	✓	✓	X	X	X	X
Alaska	✓	✓	X	X	X	X
New York	✓	✓	X	X	X	X
North Dakota	✓	✓	X	X	X	X
Ohio	✓	X	X	X	X	X
Delaware	✓	✓	✓	X	X	X
Oregon	✓	✓	✓	X	X	X
Montana	✓	✓	✓	✓	✓	✓

### **2.3. Road Surface Temperature**

To ensure the safety of roads during the winter season, transportation agencies treat the roads by spreading anti-icing salts (Ameen et al., 2022; Shahandashti et al., 2019; Kimura et al., 2006). While excessive salting might result in increased cost and environmental damages (Yang et al., 2012), delayed road treatment or insufficient salting may threaten the safety of roads due to icy conditions (Hoffmann et al., 2012). Information about estimated road surface temperatures is crucial to determine the locations as well as the appropriate timing for treating the roads (Yun et al., 2018; Yang et al., 2012). In practice, transportation agencies use the data acquired from Road Weather Information Systems (RWIS) to know about road conditions information such as road surface temperatures along roads (Darghiasi et al., 2023c; Darghiasi et al., 2022; Sato et al., 2004). In recent years, several prediction models have been developed to predict road surface temperatures using data provided by RWIS stations. However, these models differ significantly in model assumptions, application areas, and solution methods. There have been two main categories of RST prediction models described in the literature: (1) numerical models and (2) statistical models.

Numerical models were developed based on the surface energy balance method. These models numerically solve the energy transfer equation through road pavement materials to calculate the temperature profile through pavement depths (Feng and Feng, 2012). Crevier and Delage (2001) developed a numerical model, METRo, to forecast 24-h road temperature using meteorological data (e.g., air temperature, humidity, wind, solar radiation) and RWIS observations (e.g., past pavement temperature profile). Feng and Feng (2012) established a numerical method based on the energy conservation method to predict the 24-h road surface temperature using solar short-wave radiation, long-wave radiation, and heat fluxes. Qin (2016) also explored developing a

numerical model to predict the maximum and minimum temperature in a concrete slab using air temperature, relative humidity, wind speed, and solar radiation in summer. Despite the benefits of numerical methods in providing an accurate solution, they are more computationally expensive than statistical models (Barnett, 1997). They also require detailed information about the thermal properties of pavement material (e.g., density, conductivity, and specific heat capacity) as well as the pavement configuration (i.e., pavement layer thicknesses) (Qin, 2016). Moreover, since the physical processes describing the interaction of local road features are too complex to be correctly measured, it sometimes results in significant errors in simulations (Yin et al., 2019).

On the other hand, statistical models establish simplified relationships between the available influencing parameters and target values rather than solving complex partial differential equations (Zamanian et al., 2022; Shahandashti et al., 2021; Chen et al., 2019). Since the statistical models can estimate road surface temperature without numerical calculations, they have been widely used in practice (Chen et al., 2019). Among the statistical models, regression models have been widely used to estimate road pavement temperatures based on various explanatory variables for different applications (Zamanian et al., 2024; Kršmanc et al., 2013). Islam et al. (2015) developed sets of statistical regression models to estimate the pavement temperature at various depths in asphalt pavements using air temperature and solar radiation data collected from a road weather station in New Mexico, USA. In another study, Asefzadeh et al. (2017) used stepwise regression analysis to determine the pavement temperature at various depths in hot mix asphalt based on air temperature and solar radiation data collected from a road weather station in Alberta, Canada. Hosseini et al. (2017) also explored estimating the surface temperature of parking lots and sidewalks where road weather stations are not available. They developed regression models based on data collected from one parking lot at the University of Waterloo, Canada, and estimated the pavement surface

temperature using standard meteorological variables (e.g., air temperature, precipitation, previous hour air temperature, solar radiation, etc.). Table 2-4 shows examples of statistical models developed to estimate pavement temperature for different applications.

Table 2-4 Examples of linear regression models to estimate the pavement temperature

Objective	Explanatory variables	Data collection	Researcher(s)
Determine the pavement surface temperature in cold climates	Air temperature, Relative humidity, wind speed during rainfall, and snowfall	Data were collected based on the meteorological data of icy pavement in Jinhua from December 2010 to January 2016	Qiu Xin et al., 2018
Determine road surface temperature along a test road concerning weather data and pavement depth	Air temperature, solar radiation, depth of pavement	Data were collected from several sensors along a 500 m road segment in Edmonton, Alberta, Canada, from Sep 2014 to Sep 2014	Asefzadeh et al., 2017
Determine the pavement surface temperature in a low-speed low, traffic parking lot	Pavement surface temperature, air temperature, previous hour air temperature, sky condition, wind speed	Tests were conducted in parking lot C at the University of Waterloo, Ontario, Canada, in the winter season of 2012-2013	Hosseini et al., 2015
Determine pavement temperature for concrete and asphalt pavement	Ambient temperature	Data were collected from 32 environmental sensor stations in Utah during 2009	Guthrie et al., 2014
Determine the pavement surface temperature in cold climates	Air temperature, dew point, lag-dependent variable (the surface temperature at earlier times)	Data were collected every 10 minutes from nine stations in Ottawa for the winter season of 2001-2002.	Sherif et al., 2011
Determine pavement temperature at different depths considering weather data	Air temperature, solar radiation, thermal history	Data were collected from two test sections on an hourly basis in Wisconsin at different pavement depths.	Boscher et al., 1998

## 2.4. Gaps in Knowledge

Despite the value of existing studies in providing road conditions information to improve winter operations management, there are still significant gaps in knowledge related to the approach used to collect road conditions information for improving winter operations practices. The following gaps were identified from the literature.

- (1) The existing approach for the acquiring information about road surface temperatures relies on fixed sensors along roads; these fixed sensors (e.g., road weather information systems) may not be available in many locations due to their high cost and maintenance requirements.
- (2) Existing camera-based GPS/AVL systems consist of different pieces of equipment, which makes them challenging and costly to use— especially for states such as Texas, which experience less frequent snowstorms.
- (3) The automatic data collection of the road condition images at predetermined intervals may result in missing specific road issues if it is not at the time when the system captures data.
- (4) There is a lack of a user-friendly interactive platform to visualize the real-time road conditions information which is helpful in managing winter operations, including road surface temperatures, road conditions images, and related weather forecasts.

## **2.5. Research Objective**

The objectives of this research are to:

- (1) Develop an approach to estimate road surface temperatures using publicly available gridded weather forecasts from the National Weather Service.
- (2) Develop an all-in-one multipurpose data collection system to collect and transfer road conditions images— both automatically and on-demand— using the capabilities of mobile devices.
- (3) Develop a 2D digital twin interface to visualize the collected real-time road conditions images, weather information from the national weather service, and road surface temperatures using capabilities of ArcGIS Application Programming Interface,



## **CHAPTER 3**

### **DEVELOPING A MULTI-PURPOSE, ALL-IN-ONE MOBILE DATA COLLECTION SYSTEM FOR WINTER OPERATIONS MANAGEMENT**

Mobile devices (e.g., tablets and smartphones) have become increasingly popular because of their low prices, increasing functionality, and computing power. These devices have made it possible to collect and process data for various purposes without the need for complicated systems, which are costly and require operator training. In recent years, several State Departments of Transportation (State DOTs) have equipped snowplow fleets with different systems to communicate road conditions images to assist winter operations supervisors in making operational decisions during winter storms. Mainly, the existing systems consist of different pieces of equipment and automatically transmit the road condition images. The currently available data collection systems are not widely used in States which do not experience frequent snowstorms because of the high installation cost and maintenance requirements. In addition, on-demand data collection systems are not widely used by operators in field because of the system's complexity and lack of training. This study proposes a cost-effective and easy-to-use mobile-based system to collect road conditions images from field— both automatically and on-demand— using the capabilities of mobile devices.

#### **3.1. Automatic Data Collection System for Road Conditions Images**

This section explains how mobile devices (i.e., tablets or smartphones) could be turned into snowplow operations management devices to provide real-time images of road conditions. For this purpose, A custom Android application is developed to collect and transfer geotagged images as a snowplow moves over five (5) mph; these images are further processed to be visualized on a map-based interface. To test the performance of the system, a pilot test was set up in the TxDOT Wichita

Falls district during the 2020-21 and 2021-22 winter seasons. Figure 3-1 shows a TxDOT snowplow in the Wichita Falls district and a programmed tablet mounted on the snowplow’s windshield to collect real-time road conditions images.



Figure 3-1 (a) Wichita Falls’ snowplow, (b) Mounted tablet in the snowplow

The developed application was tested and debugged using a Samsung Galaxy tablet, shown in Figure 3-2. The technical specifications of the tablet used in this study are summarized in Table 3-1.



Figure 3-2 Samsung Galaxy Tab A 8-inch tablet

Table 3-1- Technical specifications of the tablet used to test the custom Android application  
(Source: Samsung website)

	Element	Description
<b>Processor</b>	CPU Speed	2GHz
	CPU Type	Quad-Core
<b>Display</b>	Size (Main Display)	8.0" (203.1mm)
	Resolution (Main Display)	1280 x 800 (WXGA)
	Technology (Main Display)	TFT
	Color Depth (Main Display)	16M
<b>Camera</b>	Main Camera - Resolution	8.0 MP
	Main Camera - Auto Focus	Yes
	Front Camera - Resolution	2.0 MP
	Main Camera - Flash	No
	Video Recording Resolution	FHD (1920 x 1080) @30fps
<b>Memory</b>	RAM Size	2 GB
	ROM Size	32 GB
	Available Memory	21.3 GB
	External Memory Support	MicroSD (Up to 512GB)
<b>Network / Bearer</b>	2G GSM	GSM850, GSM900, DCS1800, PCS1900
	3G UMTS	B1(2100), B2(1900), B4(AWS), B5(850), B8(900)
	4G FDD LTE	B1(2100), B2(1900), B3(1800), B4(AWS), B5(850), B7(2600), B8(900), B12(700), B17(700), B20(800), B28(700)
	4G TDD LTE	B38(2600), B40(2300), B41(2500)
<b>Connectivity</b>	USB Version	USB 2.0
	Location Technology	GPS, Glonass, Beidou, Galileo
	Ear jack	3.5mm Stereo
	Wi-Fi	802.11 a/b/g/n 2.4+5GHz
	Wi-Fi Direct	Yes
	Bluetooth Version	Bluetooth v4.2
	NFC	No
Bluetooth Profiles	A2DP, AVRCP, DI, HID, HOGP, HSP, OPP, PAN	
<b>Operating System</b>	Android Version	Android 9.0 Pie
<b>Sensors</b>	Motion	Accelerometer
	Environment	Light
<b>Physical specification</b>	Dimension (H x W x D, inch)	8.27×4.9×0.31
	Weight (oz)	12.24
<b>Battery</b>	Battery Capacity	5100 mAh
	Removable	No
<b>Audio and Video</b>	Video Playing Format	MP4, M4V, 3GP, 3G2, WMV, ASF, AVI, FLV, MKV, WEBM
	Video Playing Resolution	FHD (1920 x 1080) @30fps
	Audio Playing Format	MP3, M4A, 3GA, AAC, OGG, OGA, WAV, WMA, AMR, AWB, FLAC, MID, MIDI, XMF, MXMF, IMY, RTTTL, RTX, OTA

### 3.1.1. Methodology

A custom Android application is developed to facilitate collecting geotagged images using mounted tablets on snowplows. The custom application is designed to collect and transfer geotagged images of road conditions at automatically predetermined time intervals (i.e., every 10 minutes) when a snowplow operates at a speed of 5 mph or more. Figure 3-3 shows the data flow from a tablet device to the map-based interface.

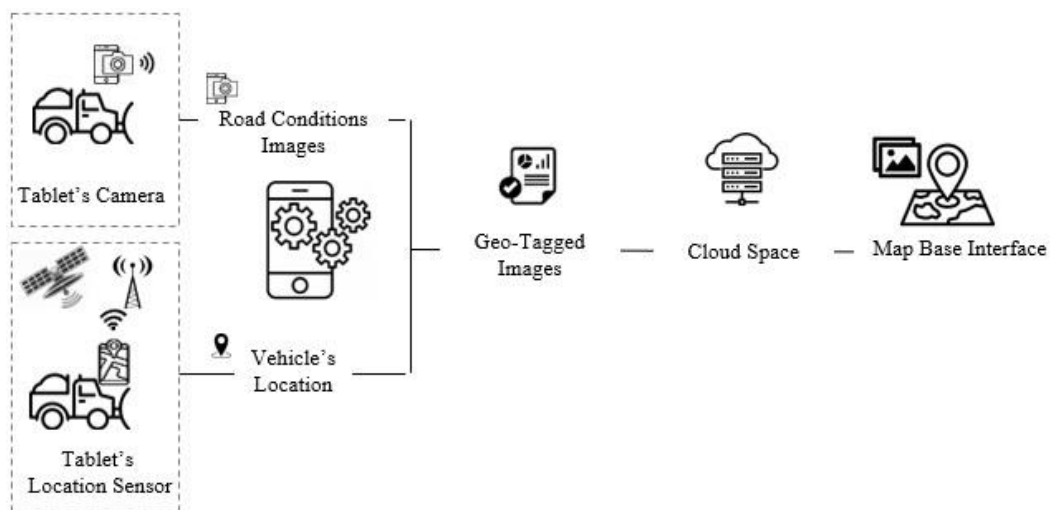


Figure 3-3 Data flow from the tablet to a map-based interface

The custom Android application is developed using Java, the official language for Android development. To develop the custom application, the different application programming interfaces (APIs), Java classes, and object methods are used to execute various tasks, such as retrieving the vehicle speed, capturing the images, determining the vehicle location, constructing metadata, and uploading the data to cloud space. Figure 3-4 shows the function modeling methodology used to develop the custom Android application, which runs on Android tablets.

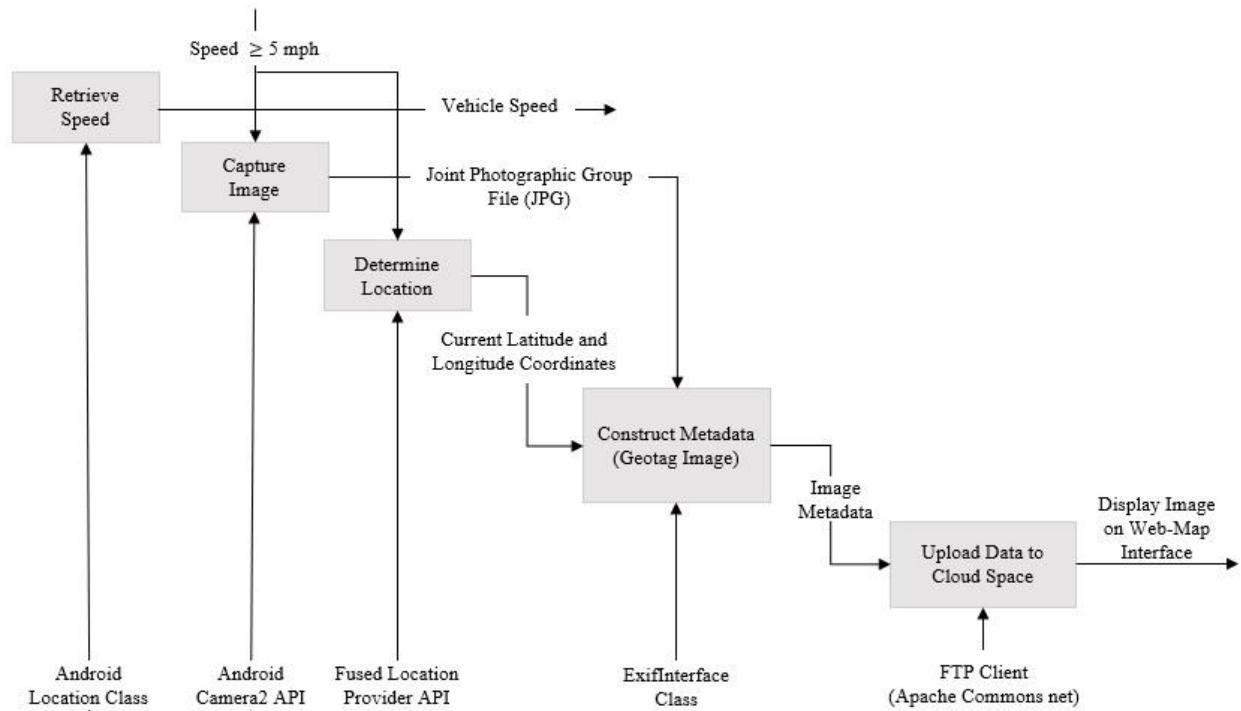


Figure 3-4- Function Modeling Methodology for the custom Android application in Java language

The following subsections provide more details about each task of the custom Android application.

### ***Vehicle Speed***

Tablets come with built-in location sensors that determine the device's location, elevation, and speed. The developed application, running on tablets, utilizes the “Location Manager” class in Android development to retrieve the device’s speed. The speed is calculated by dividing the distance the device travels by the total time it takes to travel that distance. To retrieve the speed, first, the Android “Location Manager” class obtains periodic updates of the tablet’s geographical location at specific time intervals and provides access to the device's system location services (Android developers, 2020a). Then, the “LocationListener,” another Java class, receives notifications from the LocationManager when the location of the device changes. The

LocationListener gets notified based on the specified distance intervals or the number of seconds. By having location updates, a Java location object method, “getSpeed,” returns the device's speed in metric units (Android developers, 2020a). Furthermore, appropriate conversion factor converts the metric unit (km/h) to the imperial unit (mph) within the application. The required permissions, including access to “Fine Location,” “Coarse Location,” and “Internet,” are granted in the Android manifest file to access the location services of the device to retrieve the speed.

### ***Capture Image***

The built-in camera in the tablet is used to capture the images. The “android.hardware.camera2” package uses the tablet’s camera device as a pipeline, which takes input requests for capturing a single frame in order to output one capture result as a metadata packet (Android developers, 2020b). The developed application uses the “android.hardware.camera2” package to create a “camera capture session” with a set of output surfaces within the camera device. First, the application constructs a CaptureRequest, which defines all the camera device's capture parameters to capture a single image (Android developers, 2020b). Once the request is set up, it is handed to the active capture session. After processing the request, the camera device produces a TotalCaptureResult object, which contains information about the state of the camera device at the time of capture, and the final settings used (Android developers, 2020b). Lastly, the captured image is sent to the TextureView target surface to preview (Android developers, 2020b). Appropriate permissions, including “access to the camera” and “write to external storage,” were granted in the Android manifest file to access the camera device to capture and save the images to the storage space.

### ***Vehicle Location***

Most Android devices take advantage of the signals provided by multiple sensors to determine the location. However, choosing the right combination of signals for a specific task in different conditions remains essential (Android developers, 2020c). Finding a battery-efficient solution is also critical. This custom Android application uses the “Fused Location Provider;” a Google Play Service that combines different signals (i.e., GPS and cell tower) to determine the location information of the tablet at predetermined time intervals. First, a Location Service client is created to retrieve the device location within the application. Once the Location Service client is created, the application can determine the last known location of the tablet by the “getLastLocation” Java object method. This method returns the current latitude and longitude of the device location (Android developers, 2020c). Appropriate permissions, including access to “fine location” and “coarse location,” were granted in the Android manifest file to provide access to the location services of the tablet.

### ***Construct Image Metadata***

Metadata describes data about data. Specifically, image metadata is the information embedded into an image that includes details about the image itself and information about its creation. The image metadata allows information to be transferred together with an image in a way that can be understood by software, hardware, and humans, regardless of the format. While the tablet generates default (i.e., time, camera model, focal length, etc.) metadata during the capturing process, other required metadata (i.e., location information) should be added. To integrate the required location data (longitude and latitude coordinates) into the image metadata, the Android “ExifInterface” class was used to geotag the captured images and add the location information to the image metadata. This class facilitated reading and writing Exchangeable Image File (EXIF) tags into a JPEG file or a RAW image file (Android developers, 2020d).

The standard form of GPS coordinates for the image EXIF file is in “Degree, Minutes and Seconds (DMS)” format; however, the latitude and longitude coordinates that the Fused Location Provider class provides are in decimal. Therefore, a local Java method was implemented in the application to convert the decimal coordinates to DMS format. Lastly, an ExifInterface object method known as “setAttribute” added the current latitude and longitude coordinates to the image file at the time when the image was captured. Table 3-2 shows examples of metadata created by the custom Android application for a captured image.

Table 3-2 Example metadata for a captured image by the custom application

Property	Value
Camera Make	Samsung
Camera Model	SM-T295
Exposure	1/33
Focal Length	3.8 mm
ISO Speed	176
Flash	Off
Image Width	1920
Image Height	1080
Orientation	Rotate 90 CW
Date and Time	2020:06:10 15:52:21
GPS Latitude	32.732814
GPSLatitudeRef	North
GPSLongitude	97.113237
GPSLongitudeRef	West

***Upload Image Metadata to Cloud Space***

The developed application uses the “File Transfer Protocol (FTP)” method to upload the image metadata to cloud space. FTP is a secure network protocol to transfer data between a host device



(i.e., tablet) and a remote server (i.e., cloud space). Through the FTP connection between the application and the server, the application can upload, download, or delete files from the cloud space. To establish an FTP connection, the application implements the “Apache Commons Net” library. The Apache Commons Net library contains a collection of network utilities and protocol implementations, including FTP, to set up the connection between the host device and server (Apache Software Foundation, 2020). In this application, the host device is the tablet, and the server is the University of Texas at Arlington (UTA) cloud space, which temporarily stores the images to be further processed and displayed on the map-based interface.

Every FTP client needs information about the server, including “Host Address,” “Port,” “Username,” and “Password” to establish the connection. This information was extracted from the UTA cloud account and implemented in the application development. To prevent images from backlogging in the cloud, this connection was designed to delete images older than an hour in the cloud. The appropriate permissions, including access to “network state” and “internet,” were granted in the Android manifest file to set up the FTP connection.

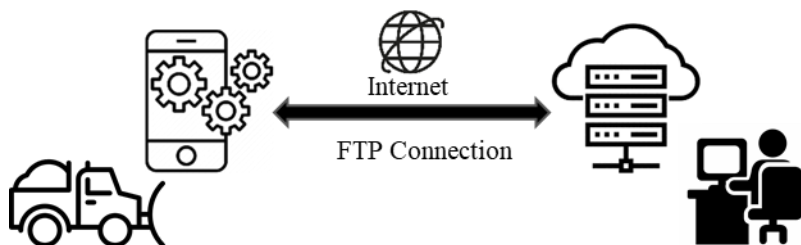


Figure 3-5 Uploading Image Metadata to Cloud Space

## *Custom Android Application's User Interface*

Collecting data using the custom application is relatively self-sufficient with minimum distraction for the snowplow operators. Figure 3-5 shows the user interface of the developed application. The user starts the data collection by pressing the “Run” button. As the button is pressed, the application begins to collect geotagged images and upload them to the cloud space if the snowplow operates at a speed of 5 mph or more. The process described in Figure 3-4 would be repeated every 10 minutes unless the user presses the “Stop” button.



Figure 3-6 The user interface of the custom Android application

### **3.2. On-Demand Data Collection System for Road Conditions Images**

The automatic data collection system collects and transfers road conditions images at predetermined time intervals (e.g., 10 minutes) without any interaction from the user—which may result in missing out specific road conditions issues if they are not at the exact time when the application records data. Since some of the particular road conditions issues (e.g., abandoned traffic control devices, damaged traffic signs, dead animals on the road, etc.) may not be at the time intervals when the system automatically collects road conditions images, it is of interest to

develop a feature in the system that facilitates providing on-demand images of road conditions upon request from users when needed for analysis in real-time.

This section explains the methodology of developing an on-demand road conditions system that enables the operations and maintenance staff to collect road conditions images when needed for analysis in real time. The collected images will be further processed and visualized on an ArcGIS map-based interface. These new functionalities enable maintenance and operations staff to collect on-demand images of road conditions.

### 3.2.1. Methodology

This Section explains how a custom Android application was developed to facilitate capturing and transferring road conditions images, associated with users' comments, upon user request. Figure 3-7 shows the data flow from the developed Android application, running on a mobile device, to a map-based interface.

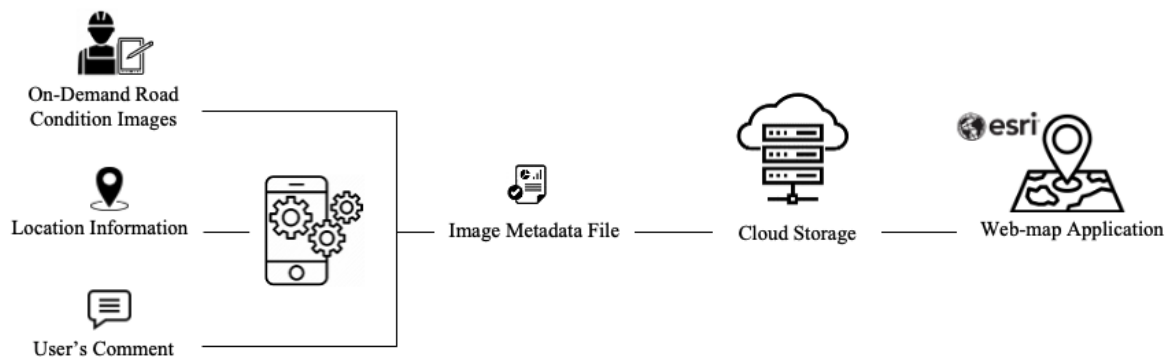


Figure 3-7 Data flow from the Android application to the ArcGIS map-based application

To develop the Android application, the capabilities of Java Application Programming Interfaces (APIs) are utilized to execute various tasks, such as capturing images, determining locations, adding users' comments, constructing metadata files, and uploading the file to cloud

space. Figure 3-8 illustrates the function modeling methodology used to develop the application, which runs on Android tablets.

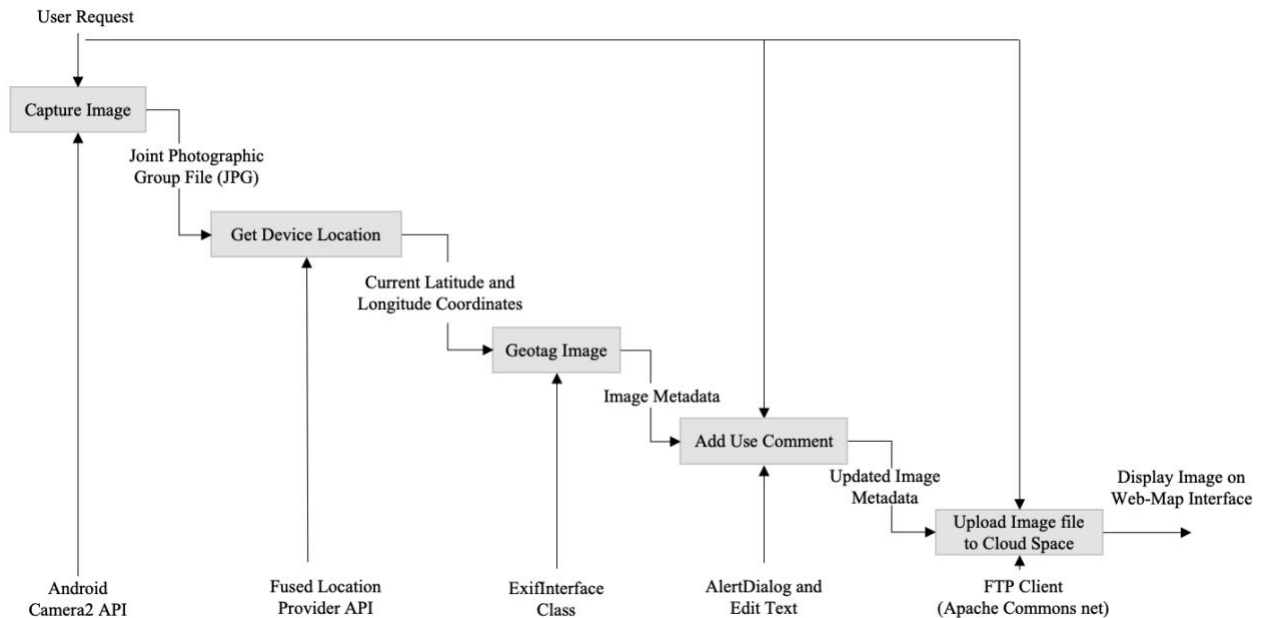


Figure 3-8 Function modeling methodology to develop the Android application to collect on-demand images

The image capturing, getting device location, geotagging images, and uploading the images to cloud space are the same as described in the previous Section. For adding the user comment, the application allows the user to comment on the captured image using the "AlertDialog" feature in Android development. The user can also skip this step and upload the image without any comments. If the user opts to add a comment about the image, an "EditText" pop-up window accepts the comment from the user. "EditText" is a user interface element in Android development that allows entering and modifying texts and numbers to the application as input (Android developers, 2020e). By having all the required information about the image (e.g., device's location, time, and user's comment), the application combines all the data into one image metadata file to be transferred to the cloud space to be further processed and visualized on the map-based interface.

### ***The User Interface of the On-demand Image Collection Application***

Data collection using the developed application is simple for the field staff. The user could capture an image by pressing the "Capture" button at the bottom of the tablet screen. When the button is pressed, the application captures the image, and a preview of the captured image will be displayed to be confirmed. Upon confirming the image preview, the user will be asked to add a comment to the image using a pop-up EditText window. The user can also skip adding a comment to the image. Lastly, by pressing the "Done" button, the image metadata file is constructed and uploaded to the cloud space, where it can be further processed and visualized on the map-based interface. Figure 3-9 illustrates the user interface of the developed Android application.



Figure 3-9 The user interface of the Android application to collect on-demand road conditions images

**Use Cases**

A use case is a set of possible sequences of interactions between a user and a system and indicates the system's action in response to a user's action. Table 3-3 presents the use case for the developed Android application to collect road conditions images upon users' demand.

Table 3-3- Use Case for the developed Android application to provide on-demand road conditions images

<b>Actor: User</b>	<b>System: On-Demand Image Collection Application</b>
--------------------	---

	0. The tablet's main screen displays a list of the installed applications on the tablet.
1. The user presses the "on-demand image collection application" icon available on the main screen of the tablet	2. The tablet launches the application and displays the camera view for the user.
3. The user presses the "Capture" button.	4. The application captures the image and displays a preview of the image for the user's confirmation.
5. The user presses the "Done" button	6. The image metadata is constructed.
7. The user presses the "Skip" button	8. The image metadata is sent to the cloud space without any comments from the user
9. The user types the comment in the text bar and presses "Done."	10. The user's comment will be added to the constructed image metadata, and the updated file will be uploaded to the cloud space.

**3.3. Results**

With the help of TxDOT, a pilot test was set up for the developed system in the TxDOT Wichita Falls district as a proof of concept and provided road conditions images to the district's transportation managers during the 2020-21 and 2021-22 winter seasons. The tablets were mounted on the front windshield of the snowplows using suction-cup mounts in a location with no distraction for the drivers. The tablets were powered using a USB power outlet in the snowplows. Figure 3-10 shows a mounted tablet on a TxDOT snowplow and the snowplow's power outlet, which powered the tablets during the operations time using a USB cable.



a



b

Figure 3-10 (a) Tablet mounted on the front windshield using a suction-cup mount, and (b) USB cable and power outlet used to provide power for the tablets

Each tablet used a data plan to transfer data to the developed snowplow operations management system for displaying the road condition images on the map-based ArcGIS interface. Figure 3-11 and Figure 3-12 show examples of collected road condition images, both automatic and on-demand, during the pilot test in the 2020-21 and 2021-22 winter seasons in the Wichita Falls district.



a



b



c



d



e



f

Figure 3-11 Examples of collected road condition images in the 2020-21 winter season  
(automatic)



(a) Wichita County, 1/1/2021, 7:36 AM, (b) Young County, 1/1/2021, 8:53 AM,  
(c) Wichita County, 1/1/2021, 7:06 AM, (d) Wilbarger County, 1/1/2021, 1:38 PM,  
(e) Wilbarger County. 12/31/2020, 3:57 PM, (f) Wichita County, 1/1/2021, 7:46 AM,



(a)



(b)



(c)



(d)

Figure 3-12 Examples of road conditions images in the 2020-21 winter season (on-demand)

(a) Throckmorton County, 11/05/2021, 1:53 PM, (b) Throckmorton County, 11/05/2021, 1:51 PM  
(c) Wichita County, 11/05/2021, 12:52 PM, (d) Wichita County, 11/05/2021, 12:47 PM

The visualized road condition information, including the geotagged images and related weather information, helps transportation operation managers to actively monitor the real-time road

condition and make well-informed decisions during winter operations. The new data collection system is expected to facilitate equipping the snowplow fleet with a mobile multi-purpose data collection system to communicate adverse road conditions during winter storms, improve snowplow operational decisions, and consequently decrease weather-related crashes on roadways.

## **CHAPTER 4**

### **Estimation of Road Surface Temperature Using NOAA Gridded Forecast Weather Data**

Monitoring road surface temperature is crucial to establishing winter maintenance strategies by the State Departments of Transportation (State DOTs) in the United States. Traditionally, transportation agencies rely on the information provided by Road Weather Information Systems for road surface temperatures along roadways. However, these systems are costly and only provide estimates at specific locations, resulting in distant areas being underrepresented. In recent years, some interpolation techniques have been considered to address this gap by estimating the road surface temperatures between the RWIS stations (Wu et al., 2022). Nevertheless, these techniques are only valid when the RWIS data are available. This study aims to estimate the road surface temperatures using forecast weather data which are available at high spatial resolution in the National Weather Service Database maintained by the National Oceanic and Atmospheric Administration (NOAA). To this end, road surface temperature data were collected from roadways using a vehicle-mounted infrared temperature sensor. Furthermore, the associated forecast weather parameters from the National Weather Service database were used to develop relationships between the publicly available weather forecast data and the actual road surface temperatures using multiple linear regression. Two estimation models were developed for dark and light groups and the gridded forecast weather data from the national weather service database were leveraged to estimate road surface temperatures along roadways using a GIS data integration approach. The results showed that the ambient temperature, relative humidity, wind speed, average temperature of the previous day, and road surface conditions (wet/dry) are statistically significant in estimating the road surface temperatures using gridded forecast weather data. The performance of the models

was validated, and satisfactory accuracy metrics (i.e., mean absolute error) of approximately 1 °C and 2 °C were achieved for the dark and light groups, respectively. The proposed method was implemented in the TxDOT Wichita Falls district as a part of the Snowplow Operations Management System to provide information about the estimated road surface temperatures to transportation managers for the 2021-22 winter season. This information facilitates establishing proactive anti-icing measures in locations where possible low surface temperatures are expected. The findings of this research contribute to a better understanding of the influence of publicly available weather forecast parameters on road surface temperatures.

#### **4.1. Methodology**

The methodology of this research is outlined in the following steps:

- Creating a dataset of actual road surface temperatures and associated weather forecasts from the NOAA’s National Digital Forecast Database.
- Developing relationships between actual road surface temperatures and associated weather forecasts, using the Multiple Linear Regression method.
- Leveraging the gridded forecast weather data from the National Digital Forecast Database to create a geospatial database in the ArcGIS tool and estimate the daily maximum and minimum road surface temperatures along road segments using the developed models.

Figure 4-1 illustrates an overview of the methodology used in this research to estimate the road surface temperatures and display them on a map-based interface.

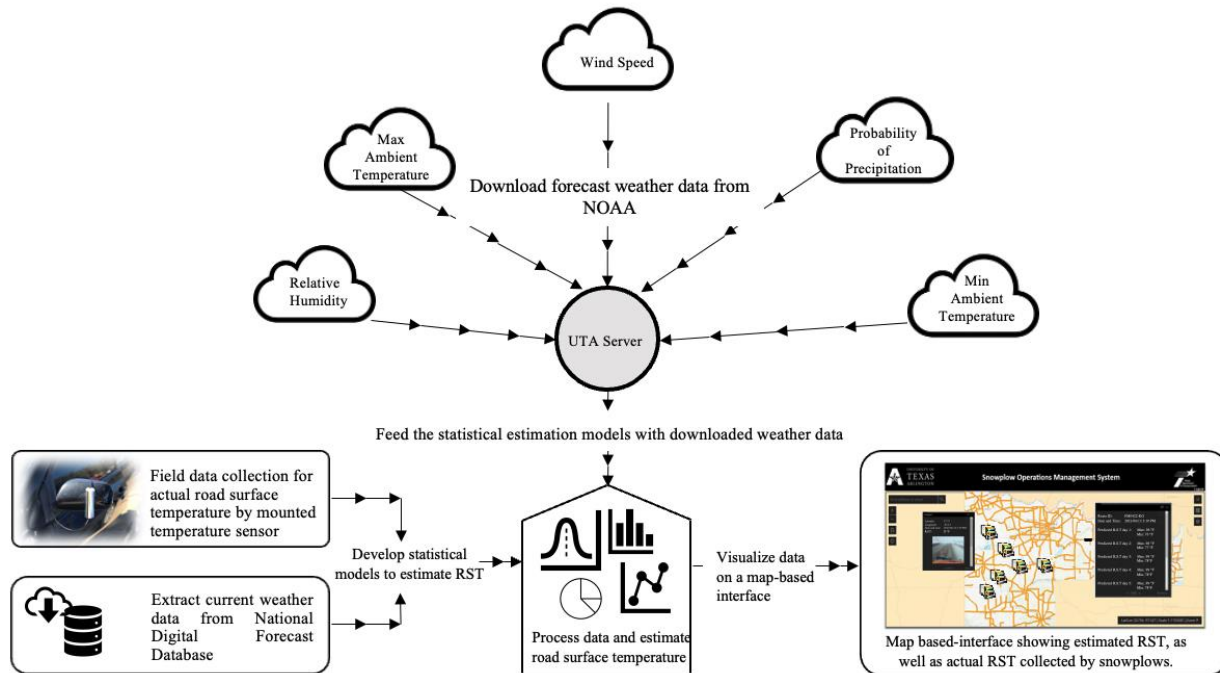


Figure 4-1 Overview of the methodology from data collection to the estimation of road surface temperatures along with the locations of snowplows

#### 4.1.1. Data collection

A vehicle-mounted temperature sensor kit— RoadWatch® temperature sensor kit— was employed on personal vehicles to randomly collect actual road surface temperature from urban collectors, urban arterials, and interstate highways within the City of Arlington, North Texas, during the winter season of 2021-22 (Figure 4-2).

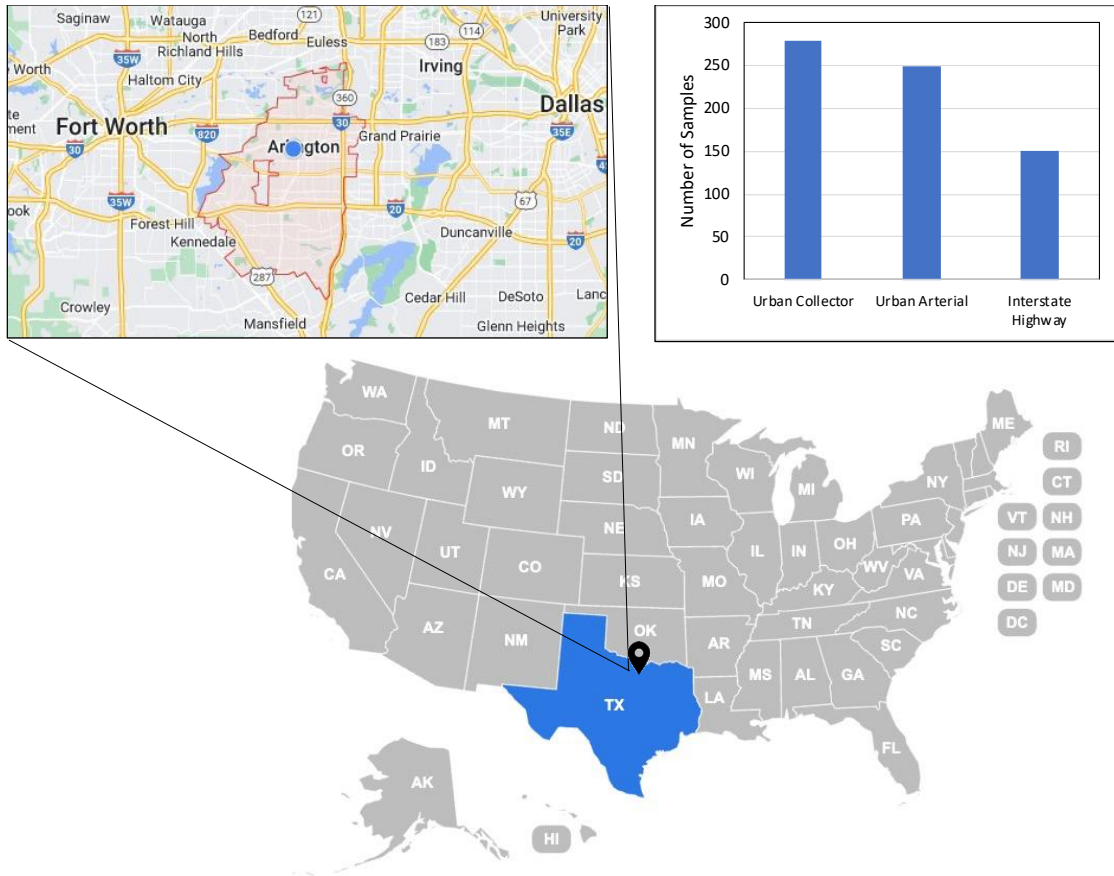


Figure 4-2 Data collection: location and road functions

The RoadWatch® sensor utilizes infrared measuring to capture the road surface temperatures. This sensor can detect a one-degree change in road surface temperature in one-tenth of a second (RoadWatch manual, 2020). According to the temperature sensor manual, the accuracy of the temperature measurements is within  $\pm 2^{\circ}\text{F}$  ( $\pm 1^{\circ}\text{C}$ ) when the ambient air temperature is between  $23^{\circ}\text{F}$  and  $41^{\circ}\text{F}$  ( $-5^{\circ}\text{C}$  to  $+5^{\circ}\text{C}$ ) (RoadWatch manual, 2020). The RoadWatch® sensor was mounted to the vehicle side mirror during data collection. Figure 4-3 shows the RoadWatch® sensor kit used to collect actual road surface temperature data. Table 4-1 summarizes the technical specifications of the RoadWatch® sensor kit that was employed to collect actual road surface temperature data.



Figure 4-3 Temperature sensor used to collect road surface temperature data; (a) RoadWatch® sensor kit; (b) mounted sensor on vehicle’s side mirror

Table 4-1 Technical specifications of the RoadWatch® temperature sensor used to collect actual road surface temperature data (RoadWatch® manual, 2020)

Properties	Description
Road surface temperature accuracy	$\pm 1$ °C (-5 °C to 5 °C ambient temperature) $\pm 3$ °C (-40 °C to -5 °C ambient temperature)
Operating voltage	12 VDC (vehicle power)
Current requirement	0.05 Amp
System operating temperature range	-40 °C to 66 °C (-40°F to +150 °F)
Sensor sample rate	Ten (10) samples per second
Sensor weight	11 oz
Vibration	Four (4) g’s two axis

Source: Data from RoadWatch®, (2020).

The associated weather data (e.g., ambient temperature, wind speed, relative humidity, and dew point) were downloaded from the National Weather Service’s grid weather database for the locations and the time that the actual road surface temperatures were measured. A summary of the explanatory variables collected from the National Weather Service’s grid weather database is shown in Figure 4-2.

Table 4-2 Explanatory variables used in the statistical analysis to estimate road surface temperature

Parameter	Abbreviation
Ambient temperature (°C)	$T_{\text{current}}$
Relative humidity (%)	Rh
Dew point temperature (°C)	$T_{\text{dewpoint}}$
Wind speed (m/s)	$W_{\text{sp}}$
Wind gust (m/s)	$W_{\text{gs}}$
Road Surface Condition (wet/dry)	RSC
6-hr precipitation (mm)	$PPT_6$
Pressure (Kpa)	P
Maximum temperature of the previous day (° C)	$T_{\text{max24}}$
Minimum temperature of the previous day (° C)	$T_{\text{min24}}$
Average temperature of the previous day (° C)	$T_{\text{avg24}}$

The collected weather data were chosen based on the frequency of their applications in the previous models for estimating pavement temperature and their availability in the National Weather Service’s grid weather database.

To ensure that the database has enough samples to demonstrate meaningful interoperations, the Cohen Statistical Power Analysis was conducted in this study. Cohen's (2013) statistical power analysis uses different methods such as a priori, compromise, criterion, posthoc, and sensitivity power analysis to calculate the sample size (Kang, 2021). This study used the priori power analysis method to calculate the sample size, as suggested by Kang (2021). The priori power analysis method calculates the sample size, which is required to detect a meaningful effect with a desired power level. For the power analysis, the probability of making Type I error (alpha) is assumed to



be 0.05 as Green (1991) suggested, and the Type II error probability (power) and effect size are assumed to be 0.8 and 0.15, respectively, as Cohen (2013) suggested. Also, the number of predictors is assumed to be the total number of collected weather parameters in the dataset. To calculate the total number of required samples, the G\*Power tool was used. Figure 4-4 shows the total required sample size, which should be included in the model to demonstrate meaningful interoperations based on assumptions in the power analysis. According to the power analysis, at least 123 samples are required for this study based on the assumptions for alpha, power, and effect size. This study meets this minimum requirement for the number of samples used in developing the statistical models. Data frequency distributions in Figure 4-5 illustrate the range and frequency of collected data.

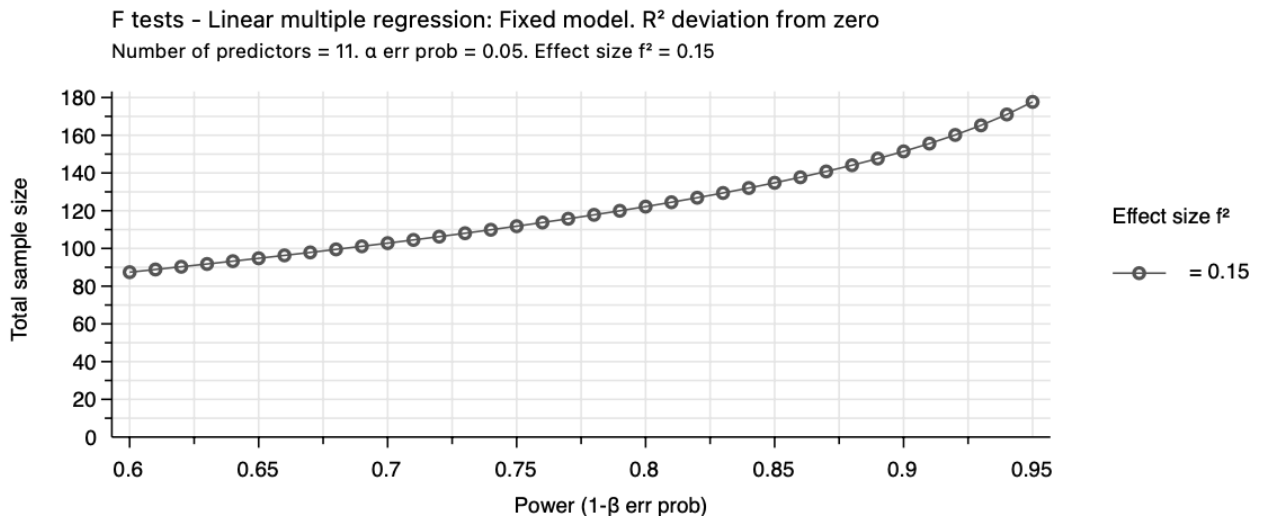


Figure 4-4 Total sample size required for the statistical analysis

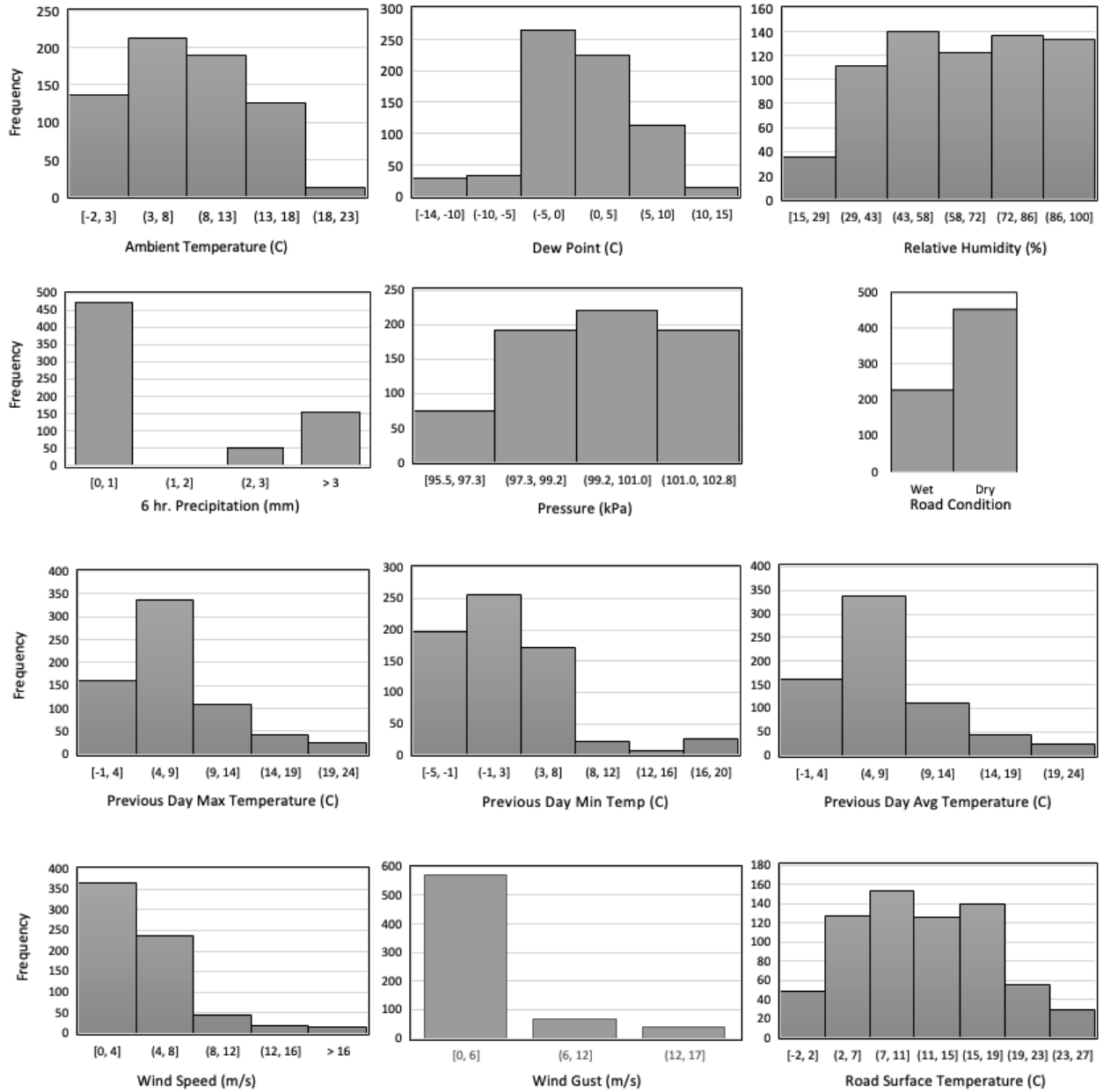


Figure 4-5 Frequency distribution of the collected data

#### 4.1.2. Developing Statistical Models

The previous statistical prediction models are either developed for applications other than winter operations (e.g., pavement mechanical performance) or require data from RWIS stations that may not be available in the NDFD gridded forecast weather data. This research focused on

developing statistical models to estimate the road surface temperature using the available gridded forecast weather data from the NDFD, with particular emphasis on the winter season. For this purpose, the statistical models were developed based on data collected during the winter season in North Texas. The following subsections elaborate on how the statistical prediction models were developed using the Multiple Linear Regression method.

#### **4.1.3. Multiple Linear Regression**

Multiple Linear Regression (MLR) is popularly used as a statistical prediction model with a linear combination of several explanatory variables (Wilks, 2011). The general formula for MLR can be described as follows (Wilks, 2011):

$$\hat{y} = \beta_0 + \sum_{i=1}^K \beta_i x_i + \varepsilon \quad (4-1)$$

where K is the number of explanatory variables,  $x_i$  is the explanatory variable,  $\beta_i$  is the coefficient of the explanatory variable,  $\beta_0$  is intercept constant,  $\varepsilon$  is the model's residual, and  $\hat{y}$  is the response variable.

Statistical analyses were performed initially on the dataset to determine which variables were statistically significant for estimating the road surface temperature. For this purpose, the M5 Regression Tree's attribute selection method was applied to the dataset. M5 Regression Tree's attribute selection method selects the best subset of variables that most improve the Akaike information criterion (AIC) (Massana et al., 2015). In M5 Regression Tree attribute selection, the feature selection starts with the full set of features. The model uses a multivariate linear model at the last node of the tree, and predictors with the smallest normalized coefficient are eliminated iteratively until the AIC cannot be further improved (Quinlan, 1992). The M5 algorithm performs

fast during training and is a promising approach in dealing with many attributes. It is also robust in handling missing values and enumerated attributes (Witten and Frank, 2002).

Since the road surface temperature fluctuates more during daytime compared to nighttime—due to the absorbance of energy from the sun— the accuracy of prediction models for nighttime data is expected to be higher than for daytime (Kwon et al., 2013). An accurate estimation of the road surface temperature, especially the minimum temperature which usually occurs overnight, facilitates the optimization of route salting strategies (Chapman and Thornes, 2006). The effect of time in pavement temperature estimation models is an important factor in improving the accuracy of models (Guthrie et al., 2014). When the data are separated into different groups based on the time of day, the fluctuations in the road surface temperature within these smaller groups are less as compared to when they are not grouped. For example, the road surface temperatures fluctuate more during daytime due to variations in the solar energy that is absorbed by the pavement compared to night when no solar energy is absorbed. Therefore, developing models to estimate the road surface temperature using these grouped data will result in higher accuracy, especially during the night when road surface temperatures are less fluctuating. As the dark times and light times vary during different months and seasons, the times suggested by Guthrie et al. (2014) were considered to separate the data into dark and light groups. As Guthrie et al. (2014) suggested, dark starts two hours after sunset and ends one hour after sunrise, and light starts one hour after sunrise and ends two hours after sunset. These cutoff points are chosen based on the corresponding times that pavements start to heat up and stop to cool down. Table 4-3 shows the dark and light hours in North Texas based on the sunrise and sunset time during the months of data collection.

Table 4-3 Light and dark hours in North Texas during data collection months

Month	Light	Dark
January	9 AM to 8 PM	8 PM to 9 AM
February	8 AM to 8 PM	8 PM to 8 AM
March	9 AM to 10 PM	10 PM to 9 AM
November	8 AM to 8 PM	8 PM to 8 AM
December	9 AM to 7 PM	7 PM to 8 AM

Data range for the continuous and binary attributes used for the model development are shown in Table 4-4 and Table 4-5, respectively. The data were collected from the city of Arlington roadways in North Texas between November 2021 and March 2022.

Table 4-4 Data ranges for the statistically significant continuous attributes

Continuous Variable	Description	Light (462 samples)			Dark (218 samples)		
		Max	Min	Standard deviation	Max	Min	Standard deviation
RST	Road Surface Temperature (°C)	27	-2	6.4	17	-1	4.2
T <sub>current</sub>	Ambient Temperature (°C)	21	-2	4.9	15	-2	4.1
Rh	Relative Humidity (%)	100	15	20	100	24	16.5
T <sub>avg24</sub>	Average Ambient Temperature of Past Day (°C)	22	0	5	22	-1	4.8
Wsd	Wind Speed (m/s)	11	0	2.3	0	26	5.6

Table 4-5 Data range for the statistically significant binary attribute

Binary Variable	Description	Light		Dark	
		Wet	Dry	Wet	Dry
RSC	Road Surface Condition	158	304	70	148

Lastly, two MLR analyses were conducted on the collected data for light and dark groups to develop the statistical models that estimate the road surface temperature. In order to evaluate the performance of the model, a 10-fold cross-validation method was used. In 10-fold cross-validation,

the data set is randomly split into ten groups for ten different runs. During each run, the model is fit to a data set consisting of nine of the original groups (90% of the dataset). The remaining group is then used for validation (10% of the dataset). This process is repeated ten times, and accuracy metrics for the out-of-sample error are calculated as the average of the accuracy metrics over the ten validation runs.

### ***Model Performance and Validation***

The performance of the models was measured using three accuracy metrics: adjusted R-Squared, mean absolute error (MAE), and root mean square error (RMSE). The adjusted R-Squared calculates how much of the target value's variance can be explained by predictors, considering the number of predictors in the model. RMSE measures the standard deviation of the residuals in the database, and MAE measures the average of the residuals in the database. Equations 4-2, 4-3, and 4-4 show the formulas for calculating MAE and RMSE.

$$\text{Adjusted } R^2 = 1 - \frac{(1-R^2)(N-1)}{(N-p-1)} \quad (4-2)$$

$$\text{MAE} = \frac{\sum_i^N |y_{\text{predicted}} - y_{\text{actual}}|}{N} \quad (4-3)$$

$$\text{RMSE} = \sqrt{\frac{\sum_i^N (y_{\text{predicted}} - y_{\text{actual}})^2}{N}} \quad (4-4)$$

Where N is the number of samples,  $R^2$  is the sample R-Squared, p is the number of predictors, and  $y_{\text{predicted}}$  and  $y_{\text{actual}}$  are the predicted and actual target values, respectively.

#### **4.1.4. Leverage Gridded Forecast Weather Data to Visualize Estimated Road Surface Temperatures on a Map-based Interface**

Although the developed statistical regression models facilitate estimating the road surface temperatures using publicly available weather data, still, estimating the road surface temperatures and mapping them onto road segments is challenging when using the live weather forecast from National Weather Service. The National Weather Service provides weather forecasts in GRIB2 format and updates them periodically via the National Digital Forecast Database (NDFD) (National Weather Service, 2019). The NDFD forecast weather data is a binary encoded format that is commonly used to store historical and forecast weather data. Since this data is originally encoded and is not compatible to be directly processed using the developed statistical models, it is of interest to develop an approach to convert these data into an appropriate format, feed them into the road surface temperature statistical models, and map the estimated road surface temperatures onto the road segment using a geospatial software. The NDFD data are based on 2.5 km grid cells that cover the Continental U.S., as well as separate grids that cover Hawaii, Alaska, the U.S. Virgin Islands, and Puerto Rico. Each grid cell has information to generate a point-specific forecast for any cell-based geographic coordinate location (National Weather Service, 2021). These forecast gridded data contain meteorological information that can be used to estimate the road surface temperature for any geographic coordinate location available in the database using appropriate correlations. As the data in the NDFD periodically updates, it is important to automatically download the data from NDFD to update the road surface temperatures. To this end, Python scripts were developed and implemented to download and store the most recent forecast from NDFD every 24 hour for the next five days. Figure 4-6 illustrates the methodology used to download and convert the forecast data into the appropriate format. The developed script in Python facilitates

connecting to the NDFD server and downloading the gridded weather forecast database for the desired parameters, such as ambient temperature, wind speed, and relative humidity, at predetermined time intervals (i.e., every 24 hours). The weather parameters are downloaded in GRIB2 format by sending HTTP requests to the national weather service database and converted to ASCII comma-separated format (CSV) using NOAA’s DeGRIB tool. The NOAA’s DeGRIB tool converts the downloaded data into an appropriate format that is compatible to be processed with the ArcGIS geospatial software.

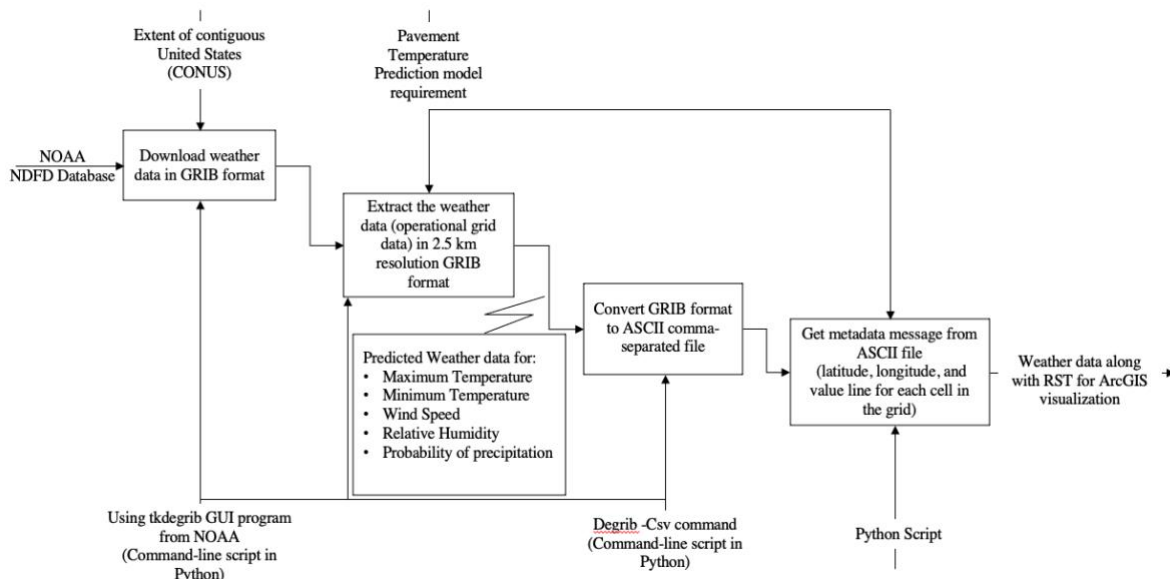


Figure 4-6 IDEF0 diagram for converting GRIB2 file downloaded from the NDFD to ASCII comma-separated

The converted data, which contains information about each weather parameter, time of the forecast, and geographical coordinates of the forecast, is imported into ArcGIS software using the ArcPy package in Python. ArcPy facilitates the conversion of data, the management of data, and the analysis of spatial data in ArcGIS software with Python (Baral et al., 2021). The geographical locations of the road segments’ centerlines can be used to obtain the associated weather parameters from the downloaded weather database in order to estimate the road surface temperatures using



developed estimation models. Figure 4-7 shows the overview of the approach for mapping the estimated road surface temperature (RST) onto road segments. The developed statistical models are used to combine the downloaded influencing weather parameters for the location of each road segment's centerline available in the road inventory database. The estimated RST for all road segments' centerlines is prepared in CSV format that contains the road segment's geographical locations as well as the associated RST estimates. This CSV file is then overwritten on the GIS shapefile of road inventory using the ArcGIS software capabilities to map the estimated RST onto the road segments. In addition, the flowchart in Figure 4-8 shows the process used in calculating road surface temperatures on the road segments.

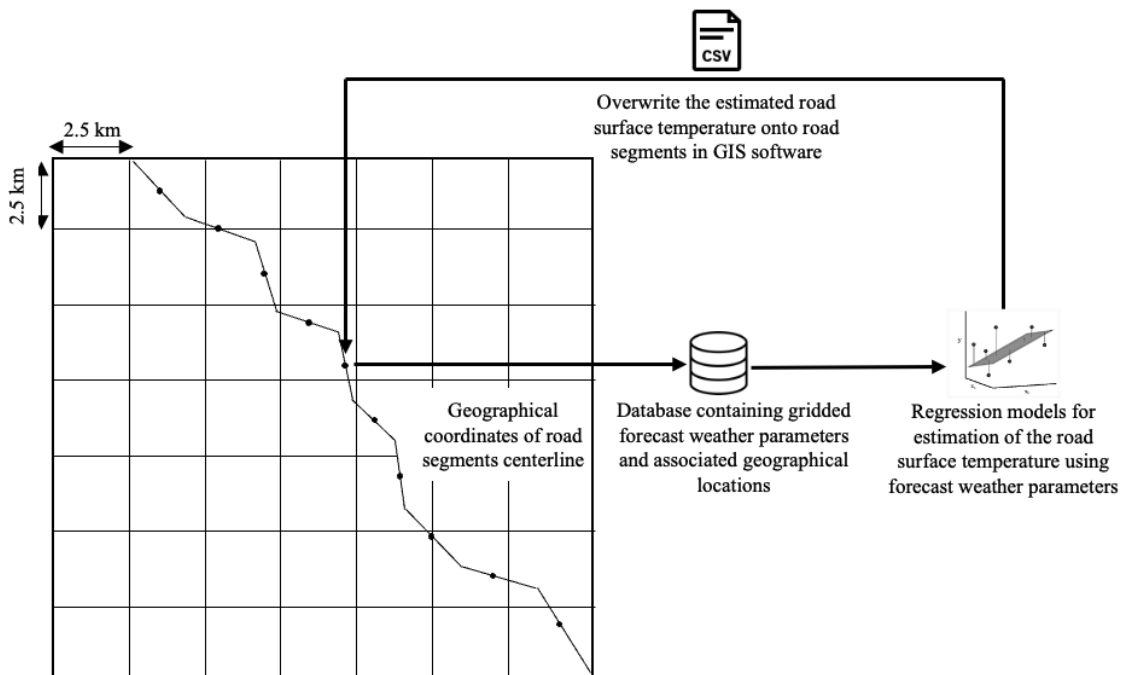


Figure 4-7 Overview of methodology to map the estimated road surface temperature onto the road segments

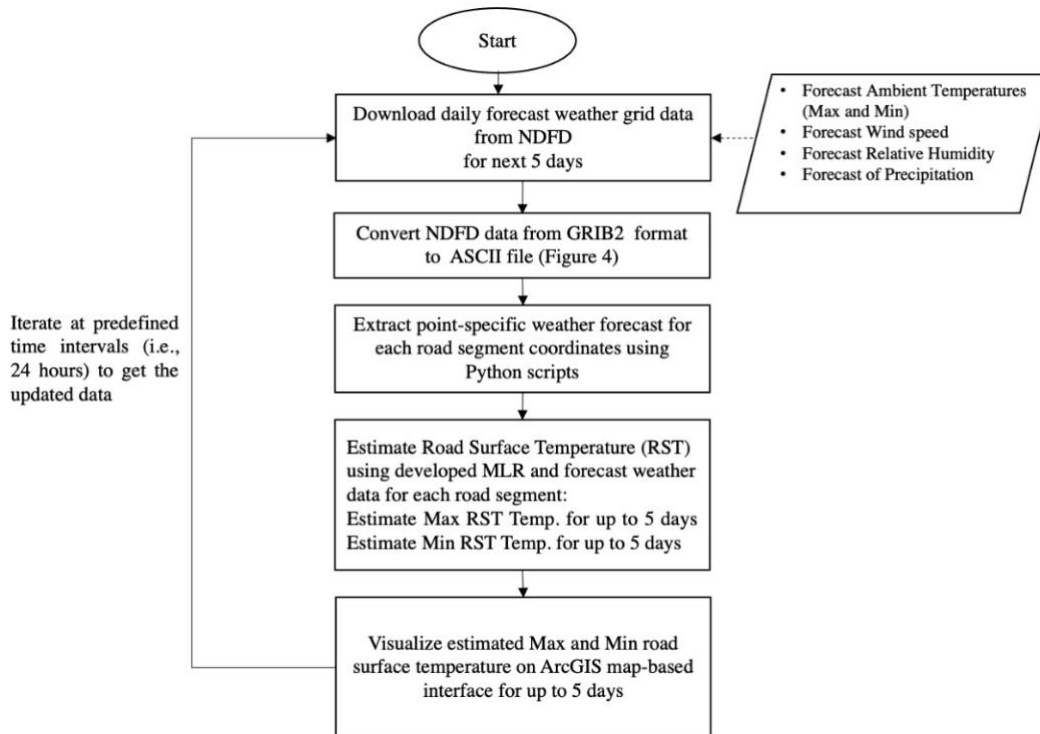


Figure 4-8 Flowchart for estimation of road surface temperature using forecast weather data from the NDFD

## 4.2. Results and Discussions

The MLR statistical models were constructed using the attributes selected by the M5 regression tree method. To avoid selecting highly correlated attributes, first, the Pearson correlation matrix for the whole dataset was calculated, and the highly correlated attributes, except one of them, were removed from the dataset.

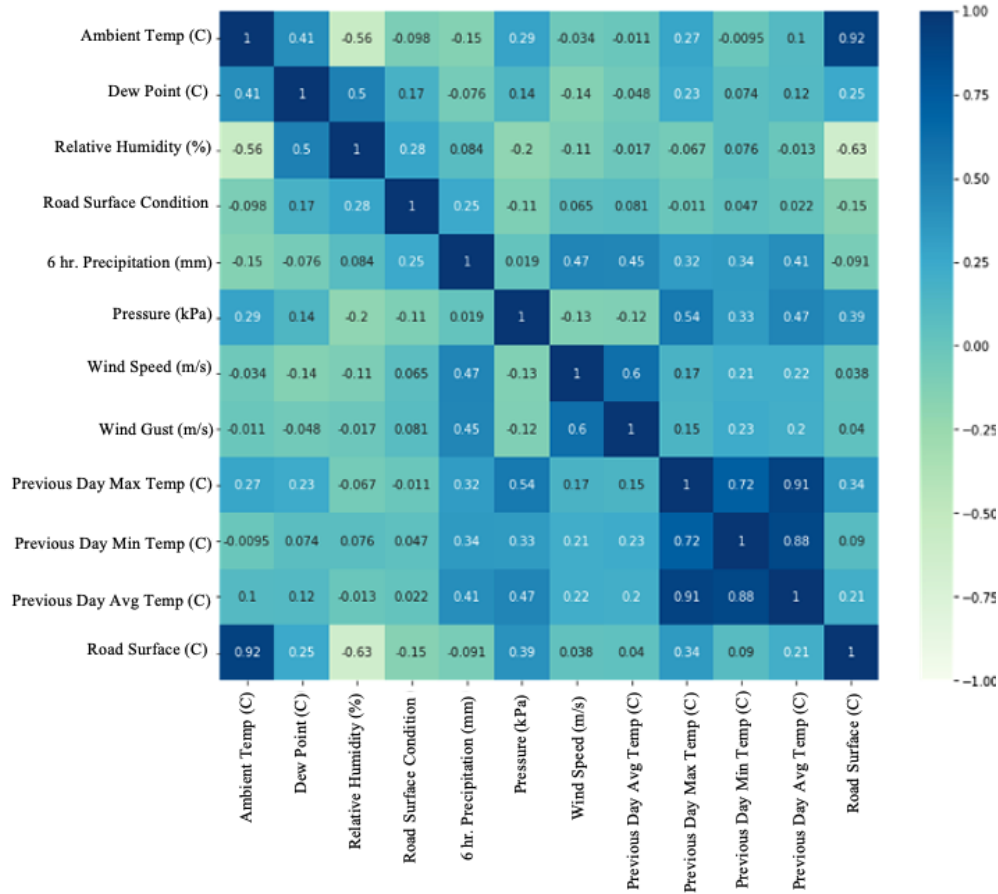


Figure 4-9 Pearson's correlation coefficient between all variables in the dataset

As shown in Figure 4-9, the maximum temperature of the previous day, the minimum temperature of the previous day, and the average temperature of the previous day are highly correlated based on the  $|R| > 0.7$  thresholds suggested by Dormann et al. (2013). Therefore, only the average temperature of the previous day was kept in the database, and the maximum and minimum temperatures of the previous day were excluded. According to the M5 regression tree method, a final selection of five attributes was chosen for the MLR analysis. These attributes included: ambient temperature, relative humidity, road surface condition, wind speed, and average ambient temperature of the previous day. The results obtained from MLR analysis for the RST estimations models for dark and light groups are summarized in Table 4-6 and Table 4-7.

Table 4-6 MLR analysis results for dark group

	Degree of freedom	Sum of Squares	Mean of Squares	F statistics	Significance F		
Regression	5	3627.33	725.467	687.536	0.0000		
Residual	211	222.64	1.055				
Total	216	3849.97					
	Coefficient	Standard Error	t-statistics	P-value	Lower 95%	Upper 95%	VIF
Intercept	0.5057	0.4285	1.1803	0.2392	-0.3389	1.3503	-
Ambient Temperature (°C)	0.9538	0.0176	54.3048	0.0000	0.9192	0.9885	1.07
Relative Humidity (%)	0.0134	0.0046	2.9014	0.0041	0.0043	0.0225	1.19
Road Surface Conditions	-0.6686	0.1614	-4.1418	0.0000	-0.9868	-0.3504	1.17
Wind Speed (m/s)	-0.0324	0.0129	-2.5056	0.0130	-0.0579	-0.0069	1.09
Average Temperature of Past Day (°C)	0.1344	0.0153	8.7958	0.0000	0.1043	0.1645	1.08

Table 4-7- MLR analysis results for light group

	Degree of freedom	Sum of Squares	Mean of Squares	F statistics	Significance F		
Regression	5	16224.06	3244.81	528.33	0.00		
Residual	456	2800.61	6.14				
Total	461	19024.67					
	Coefficient	Standard Error	t-statistics	P-value	Lower 95%	Upper 95%	VIF
Intercept	5.726	0.659	8.692	0.000	4.431	7.020	-
Ambient Temperature (°C)	1.005	0.030	34.000	0.000	0.947	1.063	1.57
Relative Humidity (%)	-0.062	0.008	-8.132	0.000	-0.077	-0.047	1.74
Road Surface Conditions	-0.374	0.260	-2.436	0.035	-0.885	0.138	1.15
Wind Speed (m/s)	0.139	0.056	2.481	0.013	0.029	0.248	1.2
Average Temperature of Past Day (°C)	0.176	0.025	6.916	0.000	0.126	0.226	1.2

According to the results, the very low significance for the F value indicates that both models have strong explanatory power for the regression analysis—the maximum acceptable Significance F at a 95% confidence level should be 0.05. The ANOVA test also indicates that all selected explanatory variables by the M5 attribute selection method are statistically significant (at the 5% significance level) for estimating the road surface temperature by having p-values less than 0.05. Among the selected attributes, ambient temperature, relative humidity, and average temperature of the previous day for the dark group are statistically meaningful at the 1% level of significance. According to the low variance inflation factors (VIF), shown in Table 4-6 and Table 4-7 for both dark and light groups, the explanatory variables in the model are not highly correlated, so there is no multicollinearity in the model—the VIF of less than 3 is considered as the multicollinearity

threshold (Imdadullah et al., 2016). The results show that the ambient temperature has the highest positive regression coefficient, and the road surface conditions has the highest negative coefficient for both light and dark groups. For example, a unit increase in the ambient temperature will result in a 0.95 and 1.01 increase in the road surface temperature for dark and light groups, respectively—assuming other independent variables are constant. Moreover, if all independent variables remain constant, changing the state of the road from wet to dry will increase road surface temperature by 0.67 °C and 0.37 °C for dark and light groups, respectively.

Furthermore, the performance of the models was evaluated using 10-fold cross-validation on the dataset used in model development. The accuracy metrics of the models were measured using R<sup>2</sup>, RMSE, and MAE. Table 4-8 summarizes the calculated accuracy metrics for both dark and light models.

Table 4-8 Summary of statistical models developed to estimate the road surface temperature

Model	Accuracy metrics using 10-fold cross validation		
	RMSE	MAE	R <sup>2</sup>
Dark	1.1	0.88	0.92
Light	2.5	1.96	0.85

The performance of the models developed in this study was compared with other pavement temperature prediction models published in the literature. The results indicate that the developed models in this study, especially the dark model, fall well within the level of acceptance in literature. By dividing the data into dark and light groups, the correlation coefficient (adjusted R<sup>2</sup>) increased from 0.86 (all data) to 0.92 (dark group), Root Mean Squared Error (RMSE) decreased from 2.3 (all data) to 1.1 (dark group) and Mean Absolute Error (MAE) decreased from 1.7 (all data) to

0.88 (dark group). This improvement in accuracy allows the dark model to fall into the ideal level of accuracy by having the mean absolute error of approximately 1°C, as suggested by Chapman and Thornes (2006). The result of the light model is also more accurate or at least consistent with previous studies in the literature, as shown in Table 4-9.

Table 4-9 Summary of the accuracy metrics for previous road surface temperature prediction models

RMSE	MAE	Researcher (s)
< 4.4 °C	< 3 °C	Nowrin and Kwon, (2022)
N/A	< 2.9 °C	Toivonen et al., (2019)
N/A	< 3 °C	Ruts and Gibson, (2013)
3.84 °C	3.1 °C	Liu et al. (2018)
N/A	< 2 °C	Yang et al. (2012)

The innovation of the models in this study lies in the use of forecast weather variables, which are available at high spatial resolutions from national weather forecasts, as the explanatory variables for the model. This will allow estimating the road surface temperatures without requiring explanatory variables from the costly fixed-station sensors.

To map the estimated road surface temperatures on roads, the grid weather data from NDFD, was leveraged along TxDOT on-system roadways in North Texas. The TxDOT on-system roadways are the routes that TxDOT maintains. These routes include interstate highways, U.S. highways, state highways, farm and ranch roads, park roads, and recreational roads (TxDOT, 2021). The TxDOT on-system route inventory is downloaded from the TxDOT open data portal as a GIS shapefile containing line data. After obtaining the on-system road inventory, HTTP requests are sent to the NDFD server to obtain a 5-days daily forecast of influencing weather data for the center points of each road segment in the route inventory database. The 5-day weather

forecast data downloaded from the NDFD includes daily forecasts on minimum overnight temperature, maximum daytime temperature, relative humidity, wind speed, and precipitation. The average temperature for the previous day is taken as the average of the maximum and minimum temperatures of the past day. The binary variable RSC is considered wet when precipitation is forecasted; otherwise, it is considered dry.

The downloaded weather data are combined using the MLR estimation models to predict the daily maximum and minimum road surface temperature for the next five days. Figure 4-10 shows the map-based interface displaying the estimated minimum road surface temperatures on TxDOT on-system roadways in North Texas for the next day.

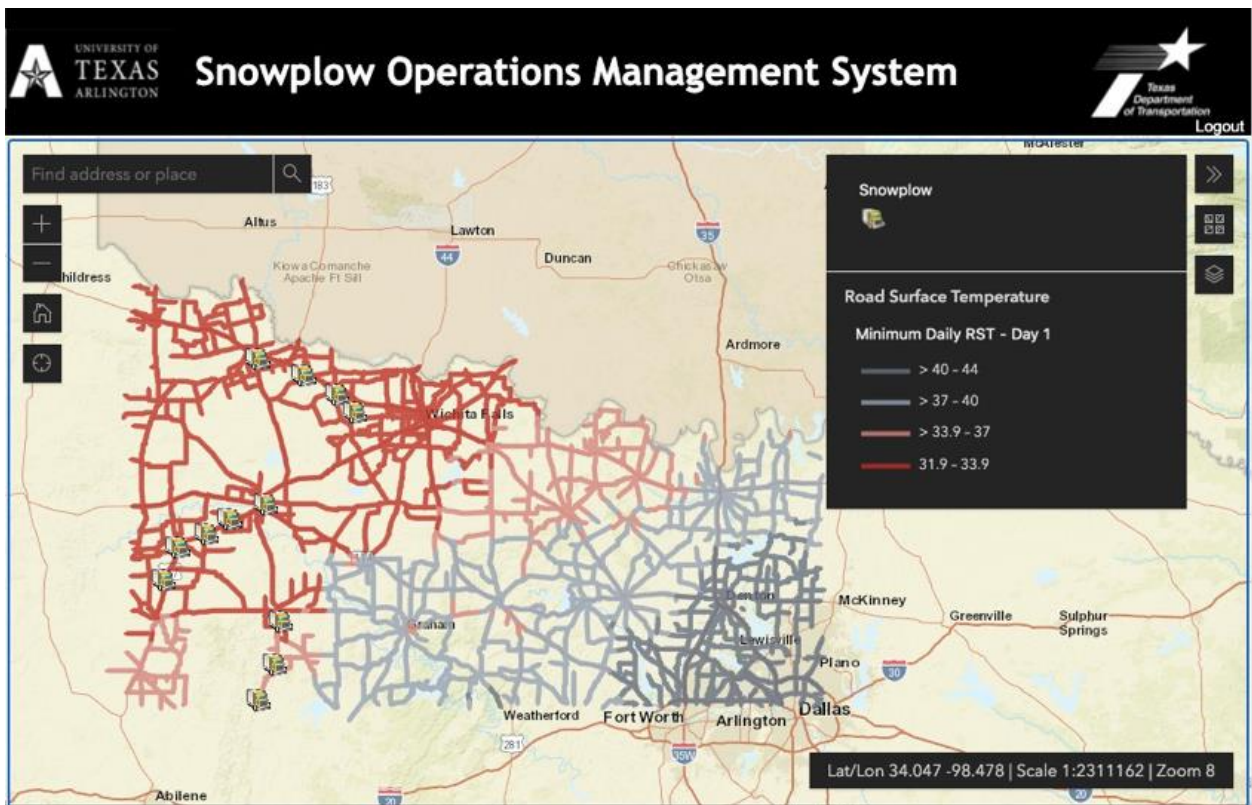


Figure 4-10 Sample screenshot of the map-based interface showing estimated minimum road surface temperatures for the next day in TxDOT on-system roadways in North Texas (February 15, 2022).



### 4.3. Summary

Monitoring road surface temperature is crucial to establishing winter maintenance strategies. The National Weather Service, administered by NOAA, provides forecasts for a variety of weather parameters, such as ambient temperature, relative humidity, wind speed, and precipitation, with an accuracy of approximately 90% for up to 5 days. However, it does not provide road surface temperature forecasts. This study aimed to develop statistical models to estimate road surface temperatures using weather forecast data available in the NOAA's National Digital Forecast Database (NDFD). Based on statistical analyses of actual road surface temperatures and associated weather data collected during the winter season in North Texas, it is concluded that ambient temperature, relative humidity, wind speed, the average temperature of the previous day, and road surface condition (wet/dry) are statistically significant factors affecting road surface temperature at 5% level of significance for both dark and light groups. At the 1% significance level, ambient temperature, relative humidity, and the average temperature of the previous day were statistically significant for estimating road surface temperature. The results also indicated that the ambient temperature has the highest positive regression coefficient, and the road surface conditions have the highest negative coefficient in both light and dark groups. It means that one unit of change in these variables will result in higher variation in the road surface temperature compared to other attributes. The MLR models developed to estimate the road surface temperature are validated by a 10-fold cross-validation method on the original dataset collected in the 2021-22 winter season. According to the results, the models have an MAE of approximately 1°C for the dark group and an MAE of approximately 2 °C for light group, which is acceptable compared to previous studies. By separating the dark group from the whole dataset, the MAE accuracy metric of the model improved by approximately 48% and fell into the ideal range of accuracy at approximately 1°C. However, the accuracy of the models totally depends on the accuracy of forecast weather

parameters provided by the National Weather Service. It should be noted that the models developed in this study are only valid within the data range of input weather parameters collected in this study. It is of interest to validate and calibrate these models for locations with weather parameters beyond the range of this study. Furthermore, the existing models only consider two periods of time for the dataset as light or dark. This period of time can be expanded to include more groups, such as early morning, late morning, early afternoon, and late afternoon, to better reflect the effect of time in the model.

Furthermore, this study leveraged forecast weather data from the NDFD to estimate the road surface temperature for each TxDOT on-system road segment using the estimation models. These findings demonstrate that the publicly available grid weather data could be used to estimate the road surface temperature along roads without requiring costly fixed-stations sensors. This method was implemented as a part of a Snowplow Operations Management System in the TxDOT Wichita Falls district during the 2021-22 winter season to provide information to winter operations supervisors for estimating road surface temperatures. Using this information, transportation operations managers can better plan their winter operations by sending snowplows to places where possible low surface temperatures are expected. Future research could evaluate if the approach in this study for the estimation of road surface temperatures can be integrated into a winter maintenance optimization framework to optimize the anti-icing material use or plowing routes. It is also of interest to expand the data collection to other regions and states in the U.S. to enhance the performance of the models and reduce overfitting.

## CHAPTER 5

### Developing a Digital Twin System for Visualizing Road Conditions Information

#### 5.1. Introduction

One of the essential components of the Winter Operations Management Systems is the inclusion of a digital twin system as an interactive data visualization platform to provide real-time road conditions for transportation operations managers. This system enables operations managers to monitor road conditions and supervise winter operations in real time without the need to be present on roads. The concept of the digital twin refers to virtual representations of physical objects throughout their lifecycle, which can be interacted with and analyzed using real-time data or simulation models (Bolton et al., 2015). The digital twin exists simultaneously with the physical object and utilizes big data technology to improve its intelligence and applicability, particularly for identifying and evaluating the actual conditions of the physical object. Digital twins can be defined based on the level of data integration as described in Figure 5-1 (Kritzinger et al. 2018).

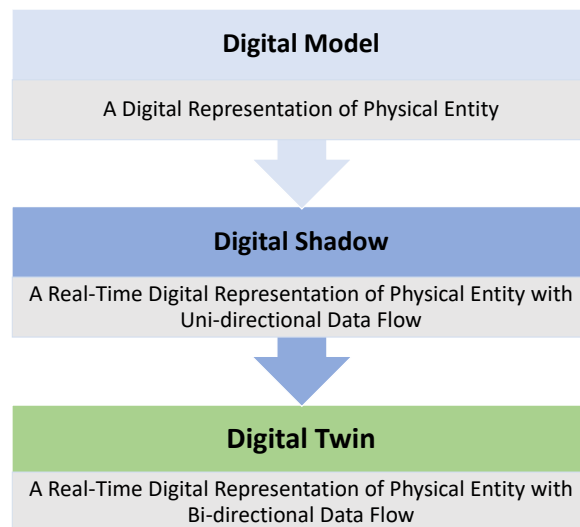


Figure 5-1 Representation of the physical world based on the level of data integration

A digital model refers to a virtual representation of a physical object or system. It is created using 2D or 3D CAD programs, allowing for the evaluation of specifications and design options in a simulated environment. The digital model aims to mimic the behavior and parameters of the real object, enabling manipulations and operations on the virtual object to produce similar results as if they were performed on the physical object. In a digital shadow, the data flow is one-way, specifically from the real object to the digital model. It involves the automatic collection and transmission of data from sensors or other sources attached to the physical object to the associated digital model. In contrast, a digital twin involves bidirectional data flow between the digital and physical objects, and it is interactive. This means that the digital twin can actively respond, adjust, and interact with the physical object in real-time based on the data it receives.

Digital Twins encompass three core elements: data acquisition, data modeling, and data visualization (Lv and Xie, 2022). Digital twins rely on some key technologies to gather real-time data, extract valuable insights, and create digital replicas of physical objects as illustrated in Figure 5-2. These technologies include the Internet of Things (IoT), Artificial Intelligence (AI), and Cloud computing. IoT serves as the foundation, using sensors to collect data from real-world objects and create digital duplicates that can be analyzed and optimized. Cloud computing provides storage and access to large volumes of data, while AI offers advanced analytics for automatic analysis and valuable predictions. Together, these technologies empower Digital Twins to bridge the physical and digital realms effectively.

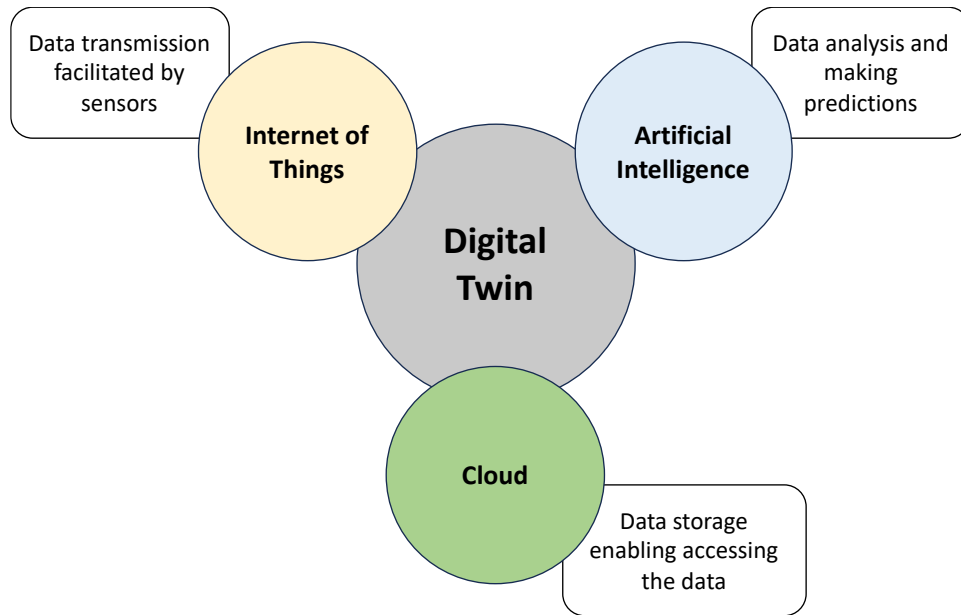


Figure 5-2 Digital twin technologies

Digital twins can be represented in both 2D and 3D formats, each offering unique advantages and applications. A 2D digital twin typically consists of graphical representations, diagrams, or schematic drawings that depict the virtual version of a physical object, system, or process. These representations are useful for capturing essential properties and analyzing specific aspects of the real-world entity, such as system schematics. The advantage of a 2D digital twin lies in its simplicity and ease of interpretation. It provides a clear and concise visual representation that is easier to understand and communicate, especially for systems or processes that do not require a full 3D representation. Furthermore, 2D digital twins can be created and analyzed with less computational resources compared to their 3D counterparts, making them more accessible and cost-effective for certain applications.

This Chapter aims to develop a digital twin system using ArcGIS Programming Interface (API) to visualize (1) live feed of road conditions images collected from mounted tablets on snowplows, (2) related weather information provided by national weather services (i.e., NOAA and NWS), (3) estimated road surface temperatures along roadways. The developed system is interactive and enables users to obtain information about the current locations of snowplows and the corresponding images of road conditions at any desired time upon their need. Figure 5-3 shows an overview of the digital twin system developed for winter operations management. The system enables operations managers to remotely monitor road conditions using computer systems, eliminating the need for physical travel and manual patrolling on roads.

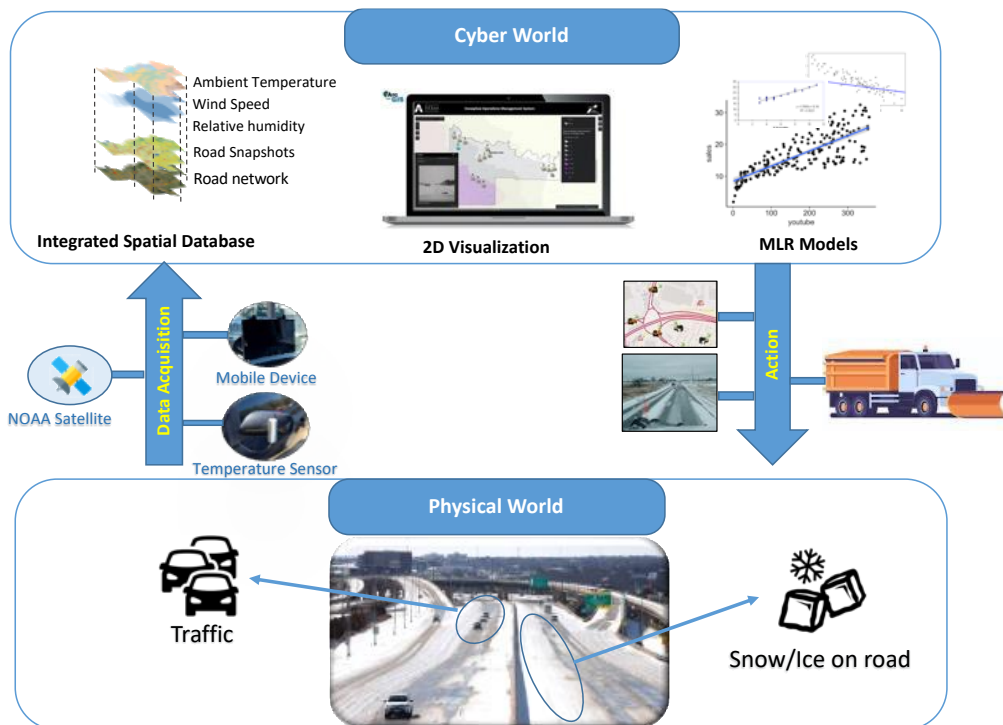


Figure 5-3 Digital Twin System for Winter Operations Management

## **5.2. Spatial Data Collection and Processing**

The digital-twin-enabled Winter Operations Management uses the collected spatial data from sensors mounted on snowplows and the national weather service to visualize real-time road conditions information using the ArcGIS application programming interface. Spatial data collection and processing for visualization, data cleaning, and data organization are explained in this Chapter.

### **5.2.1. Road Condition Images**

A geo-referenced database is developed to facilitate the real-time visualization of road condition images on a map-based interface during snowplow operations. The map-based interface uses the capabilities of the ArcGIS Application Programming Interface (API) for JavaScript and ArcGIS online. A spatial data entity representing snowplows as point features was hosted in the ArcGIS online and included in the map-based interface. The spatial data entity will be updated as the road condition images are received from operating snowplows in the field and are processed using the Exchangeable Image File (EXIF) information associated with the geotagged images. The map-based interface displays the snowplows' updated location and road condition images associated with the snowplows. The map-based application runs in a web browser and can be accessed from desktops, smartphones, and tablets, which use windows, macOS, Android, iOS, and Linux operating systems. Figure 5-4 shows the flowchart used to process the collected road condition images to visualize them in the map-based interface. An example screenshot of the map-based interface displaying the road conditions images is shown in Figure 5-5.

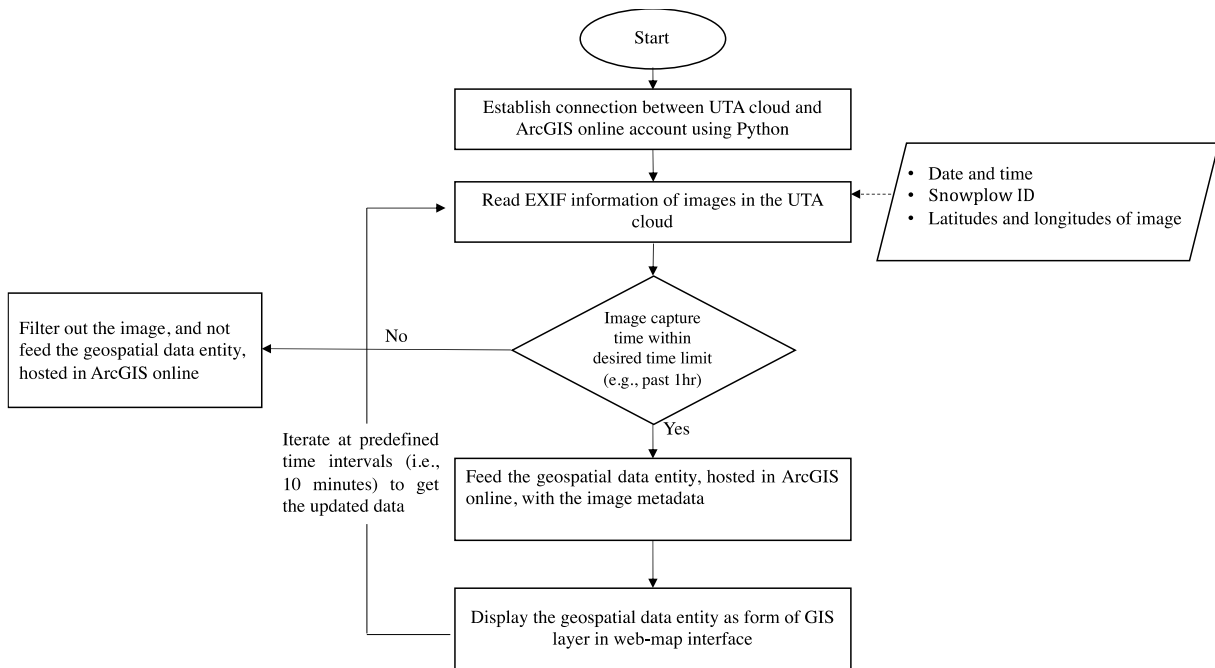


Figure 5-4- Flowchart to visualize the road condition images in the map-based interface

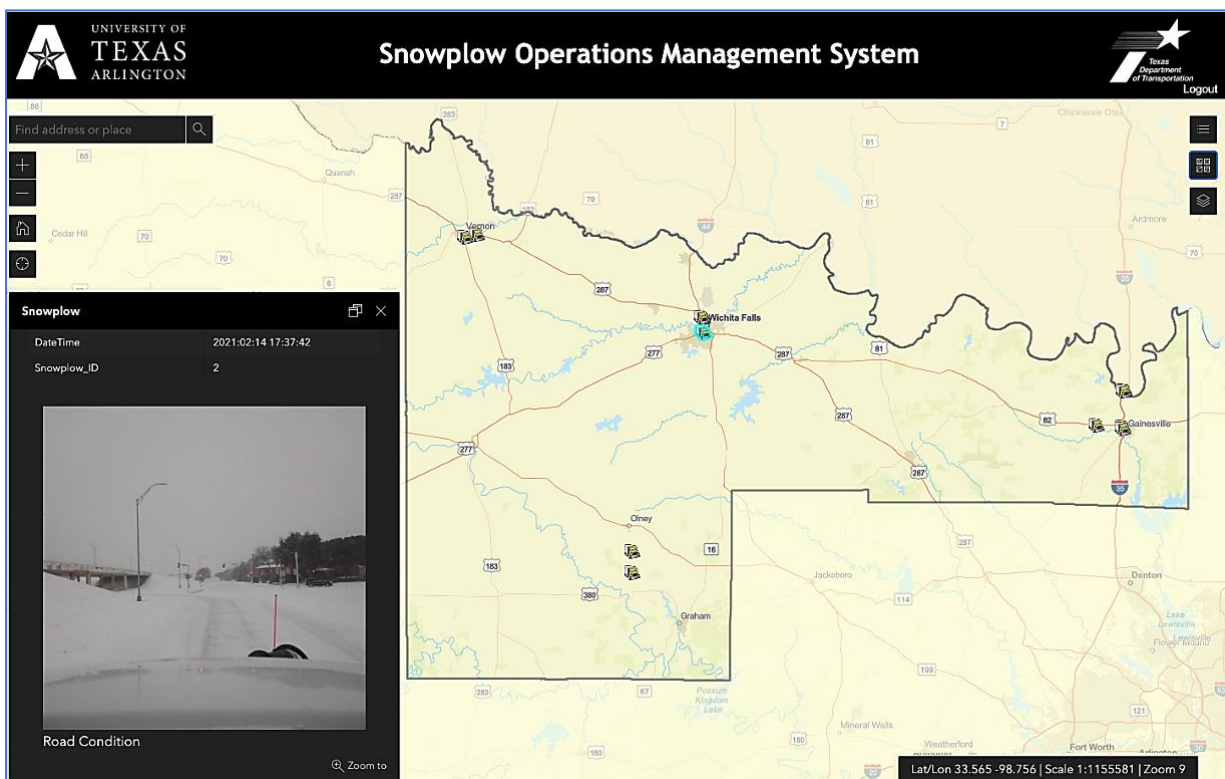


Figure 5-5- Road condition images data entity; displaying road condition images collected by operating snowplows (February 14, 2021, at 5:37 PM).



### *System Capacity*

The current system infrastructure has been developed to handle the data received from approximately ten (10) snowplows. In this study, the UTA cloud space is utilized to store the data received from snowplows, and a desktop computer (Intel® Xeon® Processor E3-1240, 3.5GHz, and 16 GB RAM) as a server to process the data. To measure the system's capacity, load testing was conducted on the system to find the maximum number of snowplows that the system, under different assumptions, can handle. The conducted load testing was evaluated from two aspects:

1. Cloud storage space
2. Computational performance of the server

In the following paragraphs, the load testing results are presented.

#### *Cloud Storage Space Aspect*

For this research, the UTA cloud storage with approximately 400 M.B. free space is used to store the data received from snowplows in field. Table 5-1 shows the load testing results under different assumptions, including the timespan that images will remain posted on the map-based interface and the time intervals between capturing images by the tablets in snowplows.

Table 5-1 Maximum number of snowplows that the existing system can handle from the cloud storage space aspect.

The time interval between capturing images by tablet	The timespan that images will remain posted on the map-based interface		
	The past one-hour	The past two-hour	The past four-hour
Every 5 minutes	30 snowplows	15 snowplows	7 snowplows
Every 10 minutes	65 snowplows	32 snowplows	15 snowplows
Every 15 minutes	100 snowplows	50 snowplows	25 snowplows

Note: Currently, the past one-hour images, captured every 10 minutes, are displayed on the map-based interface.

#### *Computational Performance of the Server Aspect*

The developed Python scripts, running on the server stationed at the University of Texas at Arlington, processes the collected data stored in the cloud space at predetermined time intervals to feed ArcGIS online map-based interface. Table 5-2 shows the results of load testing under different assumptions, including the time intervals that the server processes collected data in the cloud space and the time intervals between capturing the images by the tablets in snowplows. The timespan that the images will remain posted on the map-based interface is assumed to be one hour.

Table 5-2 Maximum number of snowplows that the system can handle from the computational performance of the server— images will remain posted on the map-based interface for one hour.

Time intervals for data processing in the server	Time intervals between capturing images		
	5 minutes	10 minutes	15 minutes
2 minutes	5 snowplows	10 snowplows	15 snowplows
5 minutes	13 snowplows	25 snowplows	38 snowplows
10 minutes	25 snowplows	50 snowplows	75 snowplows
15 minutes	38 snowplows	75 snowplows	100 snowplows

Note: Currently, the time interval between both data processing on the server and image capturing from tablets is 10 minutes.

The results of the load testing show that with the current system setting (displaying the last one-hour images received every 10 minutes from snowplows, with data processing time interval of 10 minutes), the computational performance of the server controls the system capacity and is responsive to handle the data received from approximately 50 snowplows.

### 5.2.2. Related Weather Information

In order to integrate the related weather information that facilitates winter operations into the system, Esri's map services were utilized. Esri's map services use the National Digital Forecast Database (NDFD), National Data Buoy Center, National Weather Service RSS-CAP Warnings and Advisories, and National Weather Service Boundary Overlays, all maintained by NOAA, to create data entities that contain information about current and forecast weather information. Esri's

map services download the source data and parse them using Aggregated Live Feeds methodology to return information that can be served through the ArcGIS server as a map service.

The official Esri map services were used to import the related weather information into the geo-referenced database in the Winter Operations Management System.

Based on the frequency and application of the related weather data entities used in other state DOTs, the most practical weather information for road weather information were shortlisted as follows: (1) Ambient Temperatures from NOAA (Esri 2021a), (2) Warnings, Watches, and Advisories from NWS (Esri 2021b), (3) Snowfall Forecast-Cumulative Total from NWS (Esri 2021c), and (4) Ice Forecast-Cumulative Total from NWS (Esri 2021d). Figure 5-6 shows the overview of the approach used to integrate the related weather information that facilitates winter operations practices to the map-based interface.

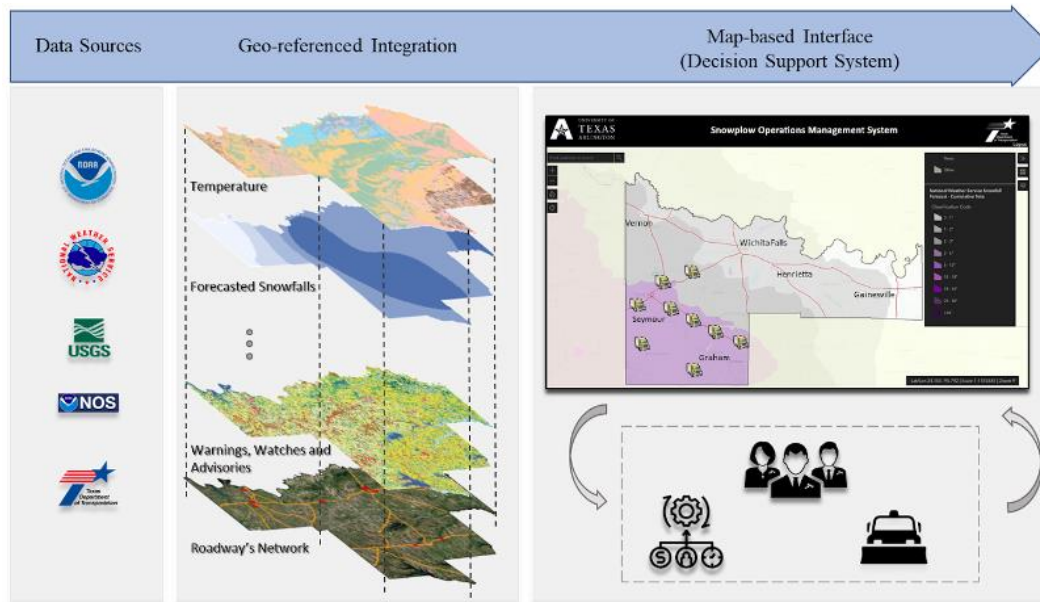


Figure 5-6 Overview of collecting and displaying weather information that facilitates winter operations decisions.

Figure 5-7 and Figure 5-8 show examples of weather-related information integrated into the map-based interface from national weather services.

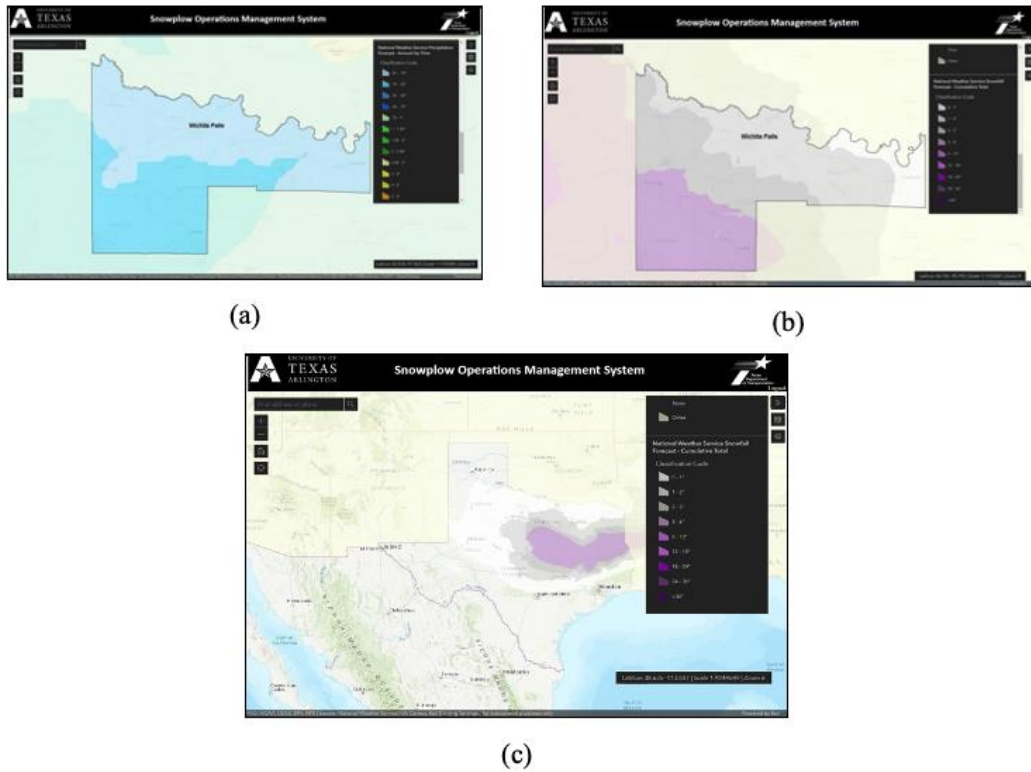


Figure 5-7- Graphical Forecast Maps retrieved from National Weather Service: (a) Snowfall forecast by the National Weather Service (Texas, January 8, 2021, 4:15 PM), (b) Snowfall forecast by the National Weather Service (TxDOT Wichita Falls district, January 8, 2021, 4:23 PM), and (c) Precipitation forecast by the National Weather Service (TxDOT Wichita Falls district, January 8, 2021, 4:37 PM)

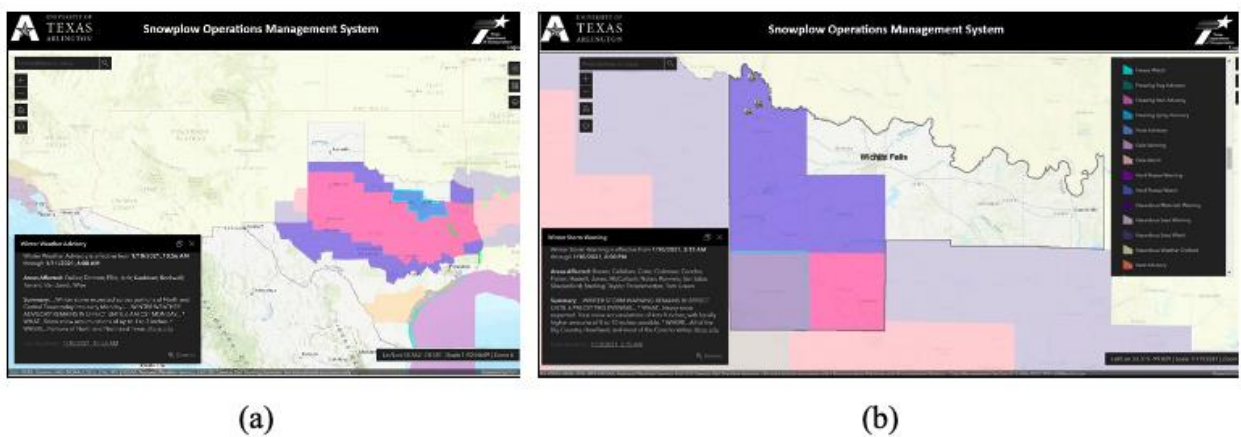


Figure 5-8- Examples of watches, warnings, and advisories issued by the National Weather Service (a) Advisory from the National Weather Service during a winter storm in Texas (January

8, 2021, 1:31 PM), and (b) Warning from the National Weather Service during in TxDOT Wichita Falls district (January 8, 2021, 1:50 PM)

### 5.2.3. Road Surface Temperatures

The road surface temperatures were estimated using the downloaded influencing weather parameters from the National Digital Forecast Database (NDFD)— These weather data are required for feeding the statistical predictive models. The data, downloaded in GRIB 2 format, are converted to ASCII comma-separated format (CSV) to be compatible with the ArcGIS data processing requirements. This research proposed an approach to download and store the data in the server stationed at UTA and process them to visualize the estimated maximum and minimum road surface temperatures on the map-based interface for up to 5 days. Figure 5-9 shows the Function Modeling illustration (IDEF0 figure) of the methodology from downloading the data to visualizing the road surface temperatures on an ArcGIS map-based interface.

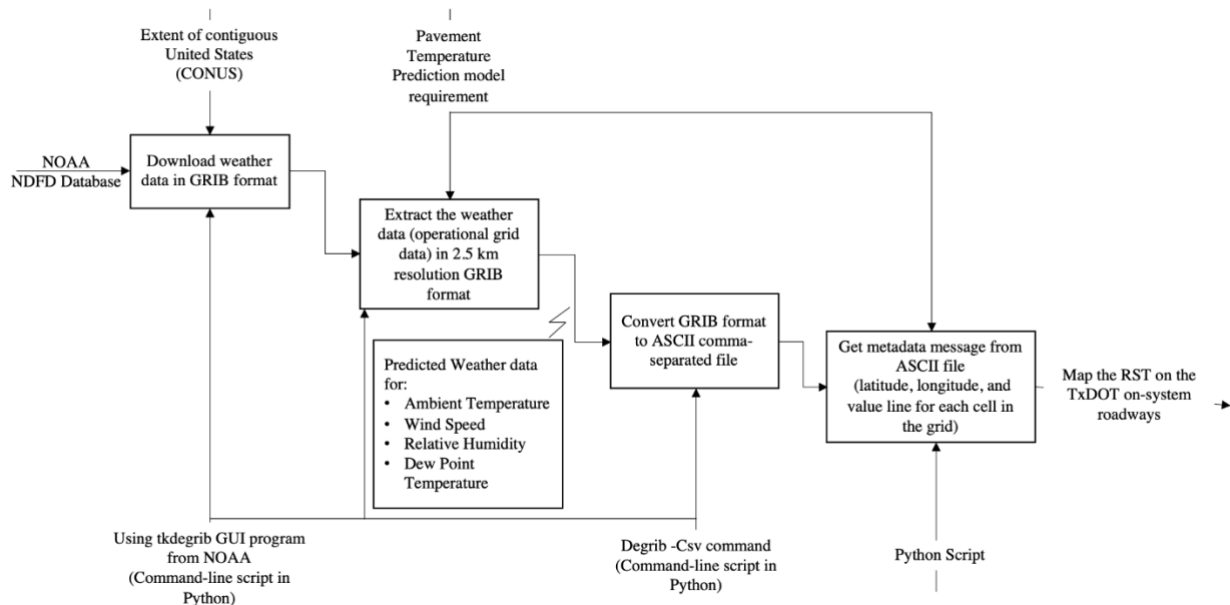


Figure 5-9 IDEF0 diagram from downloading the data to visualizing the road surface temperatures on the map-based interface.

Figure 5-10 shows the digital twin displaying the estimated road surface temperatures along with snowplow locations in the Wichita Falls district.



Figure 5-10- Road surface temperatures estimations along with snowplow locations in the Wichita Falls district

#### 5.2.4. Data Model Development for Snowplow Operations Management System

The development of a prototype data model for the Snowplow Operations Management System (SOMS) has been summarized in this section. TxDOT Data Architecture was used to create a logical data model based on the collected data. This section includes two sub-sections: (1) a conceptual data model, which represents the general idea of the SOMS, and (2) a logical data model, which provides detailed information about entities and attributes used in the SOMS.

##### 1. Conceptual Data Model

Figure 5-11 shows the conceptual data model of the system. All the data entities are imported to the system geo-referenced database and are spatially matched. Data could be imported into the geo-referenced database and exported out of the geo-referenced database. Verified users can access

and use spatial data through a map-based interface (created and designed to be used in the system) using computers and phones. An administrator is needed to manage and verify data for the users.

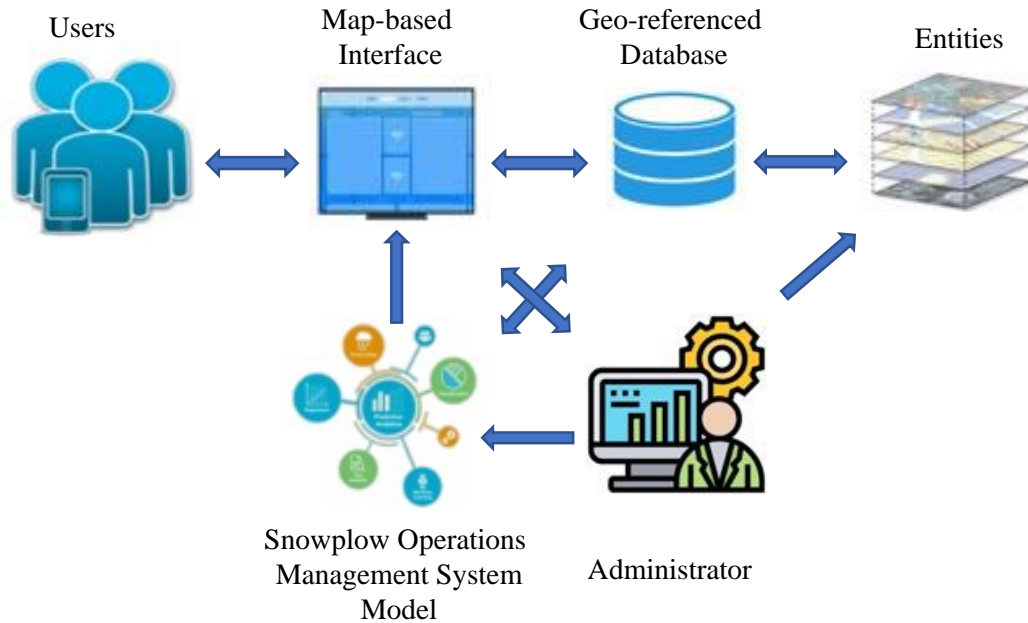


Figure 5-11- Conceptual Data Model

**Data Dictionary**

Conceptual data model terms are described in Table 5-3.

Table 5-3- Conceptual data model terms description

Term	Description
Entity	<b>Definition:</b> an entity is the detailed representation of an object of interest (e.g., road condition images).
Winter Operations Management System Geo-Reference Database	<b>Definition:</b> A Winter Operations Management System is a georeferenced database created using ESRI ArcGIS Online to store and organize GIS-based entities. Geo-referenced databases use an efficient data structure optimized for performance and storage.
Winter Operations Management System Model	<b>Definition:</b> A Winter Operations Management System model is a model that uses data such as road condition image metadata and influencing weather parameters as the inputs, processes the input data, and represents the desired output in the map-based interface.
The system map-based interface	<b>Definition:</b> The system map-based interface is a map-based interface that is based on ESRI ArcGIS online that obtains data from a geo-referenced database
User	<b>Definition:</b> a user is a TxDOT employee granted the authority to interact with the system map-based interface.

Term	Description
Administrator	<b>Definition:</b> an administrator is a TxDOT employee, or a person authorized by TxDOT who has knowledge about Winter Operations Management System and oversees managing and organizing the system as well as updating data and verifying data and users.

## 2. Logical Data Model

Collected spatial data were used to develop entities for the data model based on TxDOT Data Architecture (TxDOT, 2010). Figure 5-12 illustrates the different data entities developed for the Winter Operations Management System. The definitions of entities and their associated attributes are provided in the following subsection.

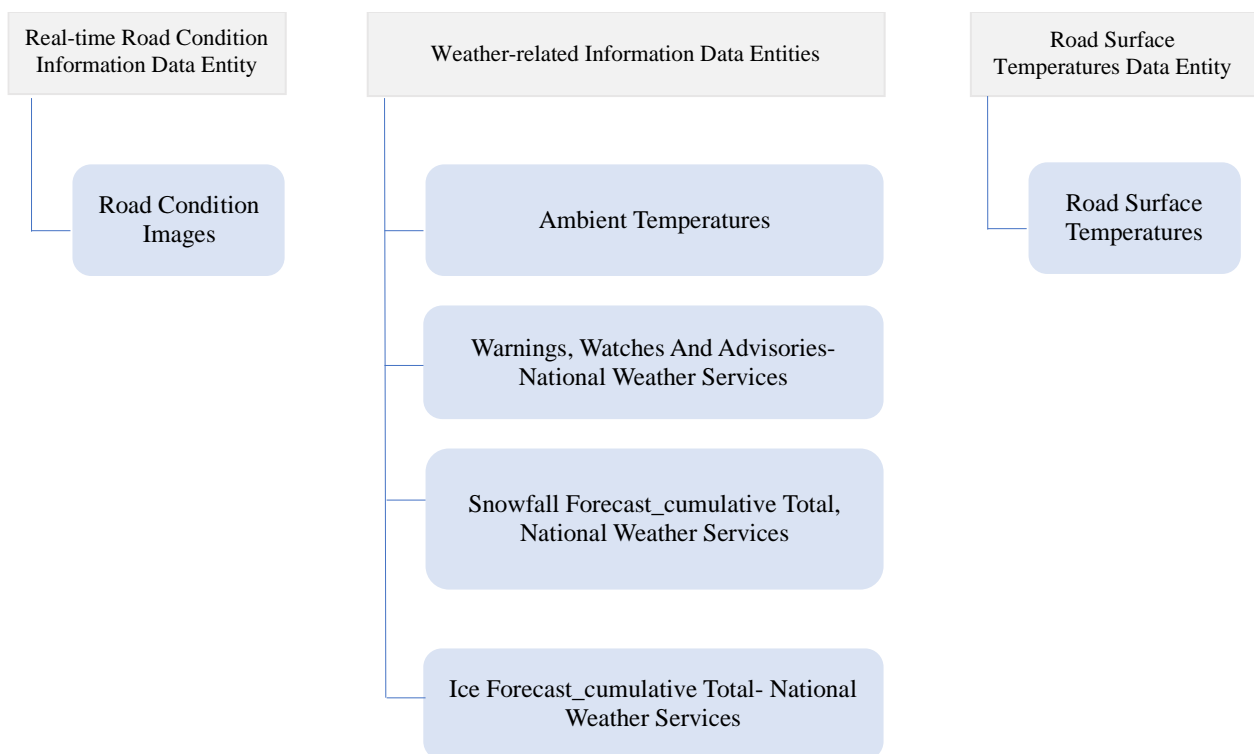


Figure 5-12- The Winter Operations Management System’s data entities

### *Data Dictionary*

The entities in the logical data model for the SOMS are described in Table 5-4.



Table 5-4 Description of the Snowplow Operations Management System’s data entities

Entity	Attribute	Description
ROAD CONDITION IMAGES	LATITUDE	<p><b>Definition:</b> LATITUDE is a float number that defines the angular distance of the image location from north/south of the Equator.</p> <p><b>Purpose:</b> LATITUDE coordinate provides an accurate locational relay from north/south of the Equator using the Global Position System (GPS)</p> <p><b>Example:</b> 37.73</p> <p><b>Valid values:</b> -90 to +90</p> <p><b>Format:</b> number</p>
ROAD CONDITION IMAGES	LONGITUDE	<p><b>Definition:</b> LONGITUDE is a float number that defines the angular distance of the image location from the East/West of the meridian at Greenwich, England.</p> <p><b>Purpose:</b> Longitude coordinate provides an accurate locational relay from East/West of the meridian at Greenwich, England, using the Global Position System (GPS)</p> <p><b>Example:</b> 96.7</p> <p><b>Valid Values:</b> -180 to +180</p> <p><b>Format:</b> number</p>
ROAD CONDITION IMAGES	DATE AND TIME	<p><b>Definition:</b> DATE AND TIME provides information about the time in which the image has been captured.</p> <p><b>Purpose:</b> DATE AND TIME identifies the exact time the image has been captured.</p> <p><b>Example:</b> 2021/02/28 1:28 PM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
ROAD CONDITION IMAGES	COMMENT	<p><b>Definition:</b> COMMENT is a combination of words that express the written remark of the user.</p> <p><b>Purpose:</b> COMMENT expresses the user's opinion or reaction to a condition in the road captured in the image.</p> <p><b>Example:</b> “Ice observed on the road”</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
AMBIENT TEMPERATURES	STATION NAME	<p><b>Definition:</b> STATION NAME is a word that defines the name of the weather station.</p>

Entity	Attribute	Description
		<p><b>Purpose:</b> STATION NAME identifies the weather station by its name.</p> <p><b>Example:</b> Wichita Falls</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
AMBIENT TEMPERATURES	DATE AND TIME	<p><b>Definition:</b> DATE AND TIME provides information about the observation time of the weather information.</p> <p><b>Purpose:</b> DATE AND TIME identifies the exact time that weather information (e.g., ambient temperature) has been observed.</p> <p><b>Example:</b> 10/3/2021, 5:51 PM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
AMBIENT TEMPERATURES	AIR TEMPERATURE (°F)	<p><b>Definition:</b> AIR TEMPERATURE is a number that defines the temperature of the air surrounding an individual, typically measured in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> AIR TEMPERATURE is to measure how cold/hot the weather is.</p> <p><b>Example:</b> 86 (°F).</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
AMBIENT TEMPERATURES	SKY CONDITION	<p><b>Definition:</b> SKY CONDITION is a combination of words describing the predominant or average sky cover based on the percent of the sky covered by opaque (not transparent) clouds.</p> <p><b>Purpose:</b> SKY CONDITION could be used as an indicator of precipitation occurrence in the near future.</p> <p><b>Example:</b> Cloudy</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
AMBIENT TEMPERATURES	WEATHER CONDITION	<p><b>Definition:</b> WEATHER CONDITION is a combination of words describing the atmosphere.</p> <p><b>Purpose:</b> WEATHER CONDITION provides the state of current atmospheric condition.</p> <p><b>Example:</b> Light Rain</p> <p><b>Valid Values:</b> N/A</p>

Entity	Attribute	Description
		<b>Format:</b> string
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	TYPE	<p><b>Definition:</b> TYPE is a combination of words that identifies the type of the issued warnings, watches, and advisories.</p> <p><b>Purpose:</b> TYPE describes the class of the issued warnings, watches, and advisories defined by National Weather Services.</p> <p><b>Example:</b> Winter Storm Watch</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	SEVERITY	<p><b>Definition:</b> SEVERITY is a combination of words that identifies the severity of the issued warnings, watches, and advisories from National Weather Services.</p> <p><b>Purpose:</b> SEVERITY describes how extreme the forecast event is expected to be.</p> <p><b>Example:</b> Mild</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	SUMMARY	<p><b>Definition:</b> SUMMARY is a combination of words giving information about the issued warnings, watches, and advisories.</p> <p><b>Purpose:</b> SUMMARY shows what should be expected during the event that the warnings, watches, or advisory is issued for.</p> <p><b>Example:</b> N/A</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	DETAILS	<p><b>Definition:</b> DETAILS is a web address (URL) to the source of the issued warning, watches, or advisories.</p> <p><b>Purpose:</b> DETAILS redirect the user to the source of the issued warnings, watches, or advisory on the web to provide more detailed information about the event.</p> <p><b>Example:</b> N/A</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>

Entity	Attribute	Description
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	UPDATED (STORED IN UTC)	<p><b>Definition:</b> UPDATED is the last time that the issued warnings, watches, or advisory has been updated, and is reported in Coordinated Universal Time UTC).</p> <p><b>Purpose:</b> UPDATED shows how recent the issued warnings, watches, or advisory is.</p> <p><b>Example:</b> 10/3/2021, 10:18 AM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	EFFECTIVE (STORED IN UTC)	<p><b>Definition:</b> EFFECTIVE is the date and time that the issued warnings, watches, or advisories take into effect.</p> <p><b>Purpose:</b> EXPIRATION shows the date and time that the issued alert is expected to take effect.</p> <p><b>Example:</b> 10/3/2021, 10:18 AM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	EXPIRATION (STORED IN UTC)	<p><b>Definition:</b> EXPIRATION is the date and time that the issued warnings, watches, or advisory expires.</p> <p><b>Purpose:</b> EXPIRATION shows the date and time that the issued alert is expected to expire.</p> <p><b>Example:</b> 10/5/2021, 10:18 AM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
WARNINGS, WATCHES, AND ADVISORIES, NATIONAL WEATHER SERVICES	AREAS AFFECTED	<p><b>Definition:</b> AREAS AFFECTED is a list of affected areas in the issued warnings, watches, or advisory.</p> <p><b>Purpose:</b> AREAS AFFECTED shows which areas are expected to be affected by the issued alert.</p> <p><b>Example:</b> Wichita Falls; Olney.</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
SNOWFALL FORECAST_CUMULATIVE TOTAL, NATIONAL WEATHER SERVICES	FROM DATE	<p><b>Definition:</b> FROM DATE is the starting date and time of the snowfall precipitation forecast.</p> <p><b>Purpose:</b> FROM DATE shows the date and time that the forecast snowfall precipitation is expected to start.</p>

Entity	Attribute	Description
		<p><b>Example:</b> 10/5/2021, 10:18 AM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
<p>SNOWFALL FORECAST_CUMULATIVE TOTAL, NATIONAL WEATHER SERVICES</p>	<p>TO DATE</p>	<p><b>Definition:</b> TO DATE is the ending date and time of the forecast snowfall precipitation.</p> <p><b>Purpose:</b> TO DATE shows the date and time that the forecast snowfall precipitation is expected to end.</p> <p><b>Example:</b> 10/8/2021, 11:28 AM</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
<p>SNOWFALL FORECAST_CUMULATIVE TOTAL, NATIONAL WEATHER SERVICES</p>	<p>DESCRIPTION</p>	<p><b>Definition:</b> DESCRIPTION is a combination of words describing the amount of total cumulative snowfall between the FROM DATE and TO DATE.</p> <p><b>Purpose:</b> DESCRIPTION provides information on how much total snow is expected to accumulate on the surface between the FROM DATE and TO DATE.</p> <p><b>Example:</b> Up to 1 inch of snowfall is expected</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
<p>ICE FORECAST_CUMULATIVE TOTAL, NATIONAL WEATHER SERVICES</p>	<p>FROM DATE</p>	<p><b>Definition:</b> FROM DATE is the starting date and time of the ice accumulation.</p> <p><b>Purpose:</b> FROM DATE shows the date and time that the forecast ice accumulation is expected to start.</p> <p><b>Example:</b> N/A</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
<p>ICE FORECAST_CUMULATIVE TOTAL, NATIONAL WEATHER SERVICES</p>	<p>TO DATE</p>	<p><b>Definition:</b> TO DATE is the ending date and time of the ice accumulation.</p> <p><b>Purpose:</b> TO DATE shows the date and time that the forecast ice accumulation is expected to end.</p> <p><b>Example:</b> N/A</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>

Entity	Attribute	Description
ICE FORECAST_CUMULATIVE TOTAL, NATIONAL WEATHER SERVICES	DESCRIPTION	<p><b>Definition:</b> DESCRIPTION is a combination of words describing the amount of total ice accumulation (inches) between the FROM DATE and TO DATE.</p> <p><b>Purpose:</b> DESCRIPTION provides information on how much total ice (inches) is expected to accumulate on the surface between the FROM DATE and TO DATE.</p> <p><b>Example:</b> Up to 1 inch of snowfall is expected</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
ROAD SURFACE TEMPERATURE	ROUTE NAME	<p><b>Definition:</b> a ROUTE NAME is a word that defines the name of the route based on the TxDOT road naming convention.</p> <p><b>Purpose:</b> ROUTE NAME identifies road by a unique designated name.</p> <p><b>Example:</b> FM-2224</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
ROAD SURFACE TEMPERATURE	MAX TEMP Day1 (°F)	<p><b>Definition:</b> MAX TEMP Day 1 is a number which is an estimation of the maximum road surface temperature for the next day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MAX TEMP Day 1 shows the estimated maximum road surface temperature for the following day.</p> <p><b>Example:</b> 67 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MIN TEMP Day1 (°F)	<p><b>Definition:</b> MIN TEMP Day 1 is a number which is an estimation of the minimum road surface temperature for the next day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MIN TEMP Day 1 shows the estimated minimum road surface temperature for the following day.</p> <p><b>Example:</b> 45 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>

Entity	Attribute	Description
ROAD SURFACE TEMPERATURE	MAX TEMP Day 2 (°F)	<p><b>Definition:</b> MAX TEMP Day 2 is a number which is an estimation of the maximum road surface temperature for the following second day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MAX TEMP Day 2 shows the estimated maximum road surface temperature for the following second day.</p> <p><b>Example:</b> 67 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MIN TEMP Day 2 (°F)	<p><b>Definition:</b> MIN TEMP Day 2 is a number which is an estimation of the minimum road surface temperature for the following second day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MIN TEMP Day 2 shows the estimated minimum road surface temperature for the following second day.</p> <p><b>Example:</b> 45 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MAX TEMP Day 3 (°F)	<p><b>Definition:</b> MAX TEMP Day 3 is a number which is an estimation of the maximum road surface temperature for the following third day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MAX TEMP Day 3 shows the estimated maximum road surface temperature for the following third day.</p> <p><b>Example:</b> 67 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MIN TEMP Day 3 (°F)	<p><b>Definition:</b> MIN TEMP Day 3 is a number which is an estimation of the minimum road surface temperature for the following third day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MIN TEMP Day 3 shows the estimated minimum road surface temperature for the following third day.</p> <p><b>Example:</b> 45 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>

Entity	Attribute	Description
ROAD SURFACE TEMPERATURE	MAX TEMP Day 4 (°F)	<p><b>Definition:</b> MAX TEMP Day 4 is a number which is an estimation of the maximum road surface temperature for the following fourth day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MAX TEMP Day 4 shows the estimated maximum road surface temperature for the following fourth day.</p> <p><b>Example:</b> 67 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MIN TEMP Day 4 (°F)	<p><b>Definition:</b> MIN TEMP Day 4 is a number which is an estimation of the minimum road surface temperature for the following fourth day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MIN TEMP Day 4 shows the estimated minimum road surface temperature for the following fourth day.</p> <p><b>Example:</b> 45 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MAX TEMP Day 5 (°F)	<p><b>Definition:</b> MAX TEMP Day 5 is a number which is an estimation of the maximum road surface temperature for the following fifth day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MAX TEMP 5 shows the estimated maximum road surface temperature for the following fifth day.</p> <p><b>Example:</b> 67 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>
ROAD SURFACE TEMPERATURE	MIN TEMP Day 5 (°F)	<p><b>Definition:</b> MIN TEMP Day 5 is a number which is an estimation of the minimum road surface temperature for the following fifth day; estimated in degrees Fahrenheit (°F).</p> <p><b>Purpose:</b> MIN TEMP Day 5 shows the estimated minimum road surface temperature for the following fifth day.</p> <p><b>Example:</b> 45 (°F)</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> number</p>



Entity	Attribute	Description
SNOWPLOW OPERATIONS MANAGEMENT SYSTEM INTERFACE USER	USER FIRST NAME	<p><b>Definition:</b> a USER FIRST NAME is a word that identifies the first name of a user who signs up for the interface</p> <p><b>Purpose:</b> USER FIRST NAME is used to identify a user.</p> <p><b>Example:</b> N/A</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
SNOWPLOW OPERATIONS MANAGEMENT SYSTEM INTERFACE USER	USER LAST NAME	<p><b>Definition:</b> a USER's LAST NAME is a word that identifies the last name of a user who signs up for the interface</p> <p><b>Purpose:</b> USER LAST NAME is used to identify a user.</p> <p><b>Example:</b> N/A</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
SNOWPLOW OPERATIONS MANAGEMENT SYSTEM INTERFACE USER	USER EMAIL	<p><b>Definition:</b> a USER EMAIL is a word that defines the email address of a user, and must be used by a user to gain access to the map-based interface</p> <p><b>Purpose:</b> USER EMAIL is used as contact information as well as sign up and sign into the map-based interface.</p> <p><b>Example:</b> John@txdot.com</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>
SNOWPLOW OPERATIONS MANAGEMENT SYSTEM INTERFACE USER	USER PASSWORD	<p><b>Definition:</b> a USER PASSWORD is a word that must be used by a user to gain access to the map-based interface</p> <p><b>Purpose:</b> USER PASSWORD is used to sign up and sign into the map-based interface.</p> <p><b>Example:</b> a-32B_74r</p> <p><b>Valid Values:</b> N/A</p> <p><b>Format:</b> string</p>

### 5.3. Digital Twin Interface Development

This section explains developing a digital twin interface for visualizing the road condition images collected from snowplows, road surface temperatures, and related weather information provided by national weather services through interactive maps hosted in ArcGIS online. This system is accessible through a password-protected web page. Figure 5-13 shows the sign-in web page to access the interface. Moreover, Figure 5-14 illustrates the developed Winter Operations Management System interface.

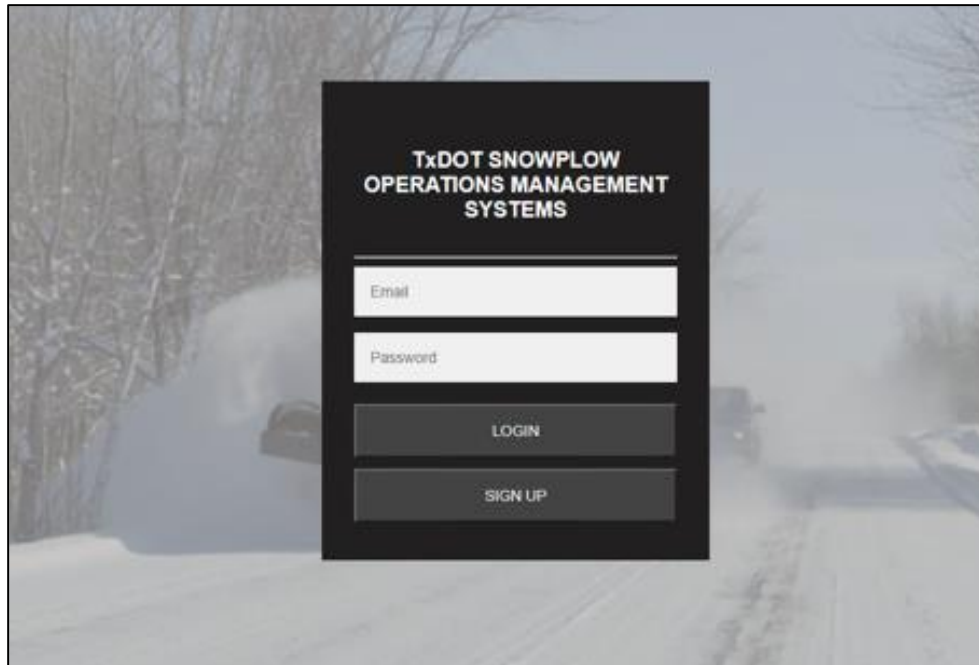


Figure 5-13 Sign-in webpage to access the map-based interface



Figure 5-14 Map-based interface for Winter Operations Management System

Descriptions of the widgets in the map-based interface are shown in Table 5-5. The entity widget helps the user control the required entity's display in the interface, i.e., the entity containing the information on road condition images can be turned on or off using the entity widget. By clicking on the displayed data on the map-based interface, the users can see information about the data (i.e., road condition images or related weather information).

The legend widget displays the legend of the entities, which are turned on in the map-based interface. The base map in the map-based interface can be changed using the base map widget. The widgets on the top-left of the interface, such as the zoom widget, home, and locator widgets, help the user navigate the interface. Table 5-5 describes the functionality of the widgets available in the map-based interface.

Table 5-5 Description of widgets in the map-based Interface

Widget Name	Description
Search	This widget helps to find a specific location in the map-based interface.

Widget Name	Description
Zoom In	This widget helps zoom in on the map view in the map-based interface.
Zoom Out	This widget helps to zoom out of the current map view in the map-based interface.
Home	This widget brings the map view to the initial view extent.
Locator	This widget helps to find the location of the user.
Base map	This widget allows the user to select the base map to be displayed in the map view of the map-based interface.
Entity	This widget displays the list of spatial data entities that can be visualized in the map-based interface.
Legend	This widget displays the legends of the spatial data entities displayed in the map-based interface.

**5.3.1. Use Cases**

A use case is a set of possible sequences of interactions between a user and a system and indicates the system's action in response to a user's action. The use case diagram is a graphical table of contents for individual use cases and defines a system boundary. Figure 5-15 represents the use case diagram for the snowplow operations management system. Use cases for the map-based application are detailed in Table 5-6 to Table 5-16.

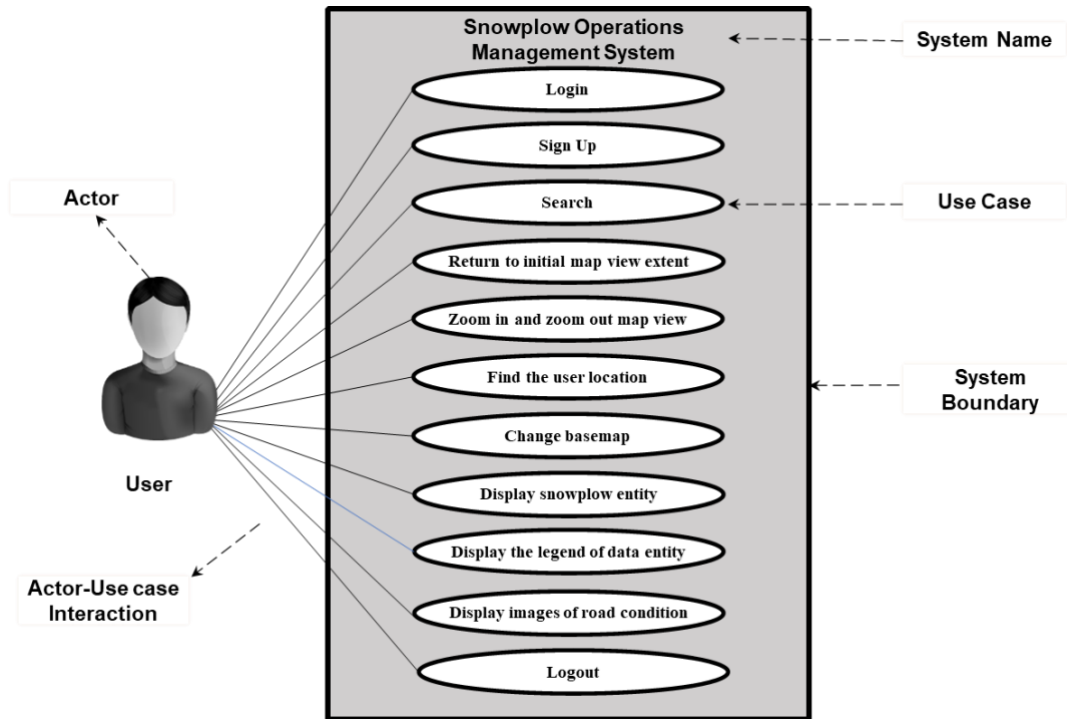


Figure 5-15 Use case diagram

Table 5-6 UC1: Login

Actor: User	System: Snowplow operations management system
	0. The browser displays a web page.
1. The user enters the web address in the address bar and press enter. URL: <a href="https://pxd4179.uta.cloud/SOMSNOV/login.php">https://pxd4179.uta.cloud/SOMSNOV/login.php</a>	2. The system displays the login page, which prompts the user to log in using a username and password.
3. The user enters the username and password, then clicks the login button.	4. The system displays <ul style="list-style-type: none"> <li>a. The map-based interface if the username and password are entered correctly.</li> <li>b. A message requesting to recheck inputs if the username or password is incorrect.</li> </ul>
5. The user sees the map-based interface, or a login error is displayed.	

Table 5-7 UC2: Sign up

Actor: User	System: Snowplow operations management system
	0. The browser displays a web page.
1. The user enters the web address in the address bar and presses enter.	2. The system displays the login page, which prompts the user to log in using a username and password, along with the option to sign up for a new account.
3. The user clicks on the signup button.	4. The system prompts the user to sign up page.
5. The user fills in the information (First Name, Last Name, Email address, Password) requested on the sign-up page and clicks the signup button to complete the process.	6. The system sends an email to the user's email address to activate the account.
7. The user opens the email and clicks the activation link to activate the account.	8. The system registers the user and displays the confirmation of registration.

Table 5-8 UC3: Search

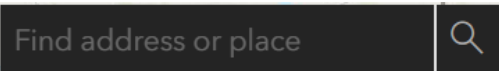
Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user enters the location on the search bar. 	2. The system displays the searched location.

Table 5-9 UC4: Return to initial map view extent


Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the home button. 	2. The system returns to the initial map view extent.

Table 5-10 UC5: Zoom-in and zoom-out of the map view

Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the zoom button.	2. The system zooms in or zooms out in the map view of the map-based interface.


	
---	--

Table 5-11 UC6: Find the user location


Actor: User	System: SRMMS
	0. The system displays the map-based interface.
1. The user clicks the locator widget. 	2. The system displays the location of the user in the map-based interface.

Table 5-12 UC7: Change base map


Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the base map widget. 	2. The system displays the available base maps from which the user can select.
3. The user clicks on the desired base map.	4. The system changes the existing base map to the base map selected by the user.
5. The user clicks on the base map widget.	6. The system closes the expanded base map widget.

Table 5-13 UC8: Display data entity


Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the entity widget. 	2. The system expands the entity widget and displays the data entities, including the road condition images, related weather information, and estimated road surface temperature entities.
3. The user clicks on the entity to turn it on and off.	4. The system displays or removes the entity from the map view.
5. The user clicks on the entity widget.	6. The system closes the expanded entity widget.

Table 5-14 UC9: Display the legend of the data entity


Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the legend widget. 	2. The system displays the legend of the entity displayed in the map-based interface.

Table 5-15 UC10: Display road condition data


Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the entity widget. 	2. The system expands the entity widget and displays the data entities, including road condition images, related weather information, and estimated road surface temperature entities.
3. The user turns on the data entities to ensure required entities are displayed on the map.	4. The system displays the snowplows in the map-based interface.
5. The user clicks on the data in the map-based interface.	6. The system displays a pop-up window containing related information about the data.
7. The user clicks on the close button of the pop-up.	8. The system closes the pop-up window.

Table 5-16 UC11: Logout

Actor: User	System: Snowplow operations management system
	0. The system displays the map-based interface.
1. The user clicks the logout button located on the top-right of the application.	2. The system exits the application.

## 5.4. Summary

As the collected data are from multiple sources with different data structures, it is important to integrate all the collected data into a system and share spatial information through a virtual model



designed to reflect the physical object. In this study, a digital twin system is developed to facilitate the real-time visualization of the spatial data during snowplow operations using ArcGIS Programming Interface. The spatial data are shared through interactive maps hosted in ArcGIS Online for real-time visualization of: (1) road condition images and snowplow location, (2) related weather information from national weather services, (3) and road surface temperatures on roadways. The digital twin system runs in a web browser and can be accessed from desktops, smartphones, and tablets, which use Windows, macOS, Android, iOS, and Linux operating systems. The developed digital twin facilitates sharing and matching spatial data through interactive maps allowing transportation operations managers to monitor road condition information graphically and facilitate decision-making on snowplow deployments during winter operations.

## CHAPTER 6

### **Improved Road Surface Temperature Prediction Using Random Forest Machine Learning Algorithm Based on Weather Forecasts**

#### **6.1. Introduction**

Insufficient and inaccurate information about road surface temperature can result in suboptimal winter maintenance decisions, leading to traffic accidents and congestion during the winter season (Chen et al., 2019). In order to enhance road safety during the winter season, states and local highway agencies allocate significant resources annually towards winter operations expenditure (FHWA, 2020). Still, most weather-related vehicle crashes occur on snowy, slushy, or icy surfaces, resulting in more than 1,300 deaths and more than 116,800 injuries annually (FHWA, 2020). Spreading anti-icing salt is one of the proactive activities that transportation agencies carry out during the winter season to ensure road safety (Ameen et al., 2022; Ameen et al., 2021; Shahandashti et al., 2019). However, spreading the appropriate amount of salt at the right locations in the road network is challenging. Oversalting can increase costs and damage the environment (Yang et al., 2012), while insufficient salting can pose a safety threat during icy weather (Hoffmann et al., 2012). To determine where and when the roads should be treated, it is helpful to have accurate information about the temperature of the road surface (Darghiasi et al., 2023a; Yang et al., 2012). Typically, transportation agencies rely on fixed sensor stations such as Road Weather Information Systems (RWIS) to obtain information about road surface temperatures (RST) (Darghiasi et al., 2022; Shahandashti et al., 2022; Sato et al., 2004). However, these systems are costly and are not available in many locations, making it challenging to obtain RST information.

Within the continental United States, the NWS provides gridded forecasts for distinct weather parameters at intervals of 2.5 kilometers through the National Digital Forecast Database (NDFD) with an accuracy of around 90% (NOAA, 2022). Darghiasi et al. (2023a) (Chapter 4) demonstrated that these gridded weather forecasts can be utilized to estimate RST on roadways using linear regression models. However, the linear regression models used in Chapter 4 may underperform due to the existence of numerous influencing parameters in the system and the nonlinearity and complexity associated with them. Lately, the use of Artificial Intelligence (AI) has evolved to a point where it offers real practical benefits in a wide range of fields (Zamanian et al., 2023; Darghiasi et al., 2023b; Darghiasi et al., 2024). In the transportation area, data analytics techniques have been used to support decision-making for different applications including infrastructure asset management (Baral et al., 2023; Baral et al., 2022; Hatamzad et al., 2022; Liu et al., 2018). For example, Qiu et al (2020) used two tree-based machine learning techniques (i.e., gradient boosting, Random Forest) and a linear regression model to estimate asphalt pavement temperature based on metrological data collected from fixed weather stations in China. Dai et al., (2023) used an ensemble deep learning model using a gated recurrent unit (GRU) network and long short-term memory (LSTM) to estimate RST based on information obtained from fixed road weather stations in China. Hatamzad et al., (2022) used three machine learning models (i.e., support vector regression, neural network, and random forest) to estimate RST based on various observations collected from three vehicle sensors and one road weather station. Moreover, Milad et al., (2021) used Markov chain Monte Carlo and Random Forest to estimate pavement temperature at different depths based on air temperature, time of day, and depth obtained from fixed field sensors. In another study, Yang et al., (2020) used K-nearest neighbor models to estimate RST based on climatological data (i.e., air humidity and temperature) obtained from a probe vehicle in South

Korea. Although the previous models have been useful in predicting pavement temperatures, these models mainly use explanatory variables obtained from actual sensors on the field (e.g., fixed road weather stations or vehicle sensors) which are not available in many locations.

This Chapter aims to demonstrate the applicability of the Random Forest algorithm to improve the accuracy of RST estimates based on gridded weather forecasts which are provided by the NWS. Random Forest models, as one of popular tree-based machine learning algorithms, have been demonstrated to be useful in estimating RST for different applications (Wang et al., 2022; Takasaki et al., 2022; Milad et al., 2021; Qiu et al., 2020). The Random Forest algorithm is highly scalable and robust against outliers; it also performs well when dealing with mixed data (e.g., continuous, and categorical) (Mantero and Ishwaran, 2020). To evaluate the prediction power of Random Forest models, the accuracy metrics of the developed models were compared with support vector machine (SVM), multiple linear regression (MLR), deep learning (DL), k-nearest neighbor (KNN), and M5P decision tree based on different accuracy metrics including mean absolute error, root mean square error, and R-squared. This Chapter demonstrates that the Random Forest model can estimate the RST more accurately than other data-driven methods (i.e., MLR, SVM, DL, KNN, and M5P decision tree) when using the gridded weather forecast from the NWS. The findings of this research contribute to the state of knowledge by providing more accurate estimates of the RST based on the gridded weather forecasts which are provided by the NWS through the National Digital Forecast Database (NDFD).

## **6.2. Methodology**

The following steps outline the methodology for this research: (1) creating a dataset consisting of actual RST, measured by a vehicle-mounted temperature sensor, and the associated weather

forecasts which are provided by the NWS, (2) developing predictive models using Random Forest methodology to estimate the RST based on the associated weather forecast data, and (3) comparing the performance the developed model with other common predictive models including multiple linear regression (MLR), deep learning (DL), support vector machine (SVM), K-nearest neighbor, and M5P decision tree. The overview of the described methodology is shown in Figure 6-1.

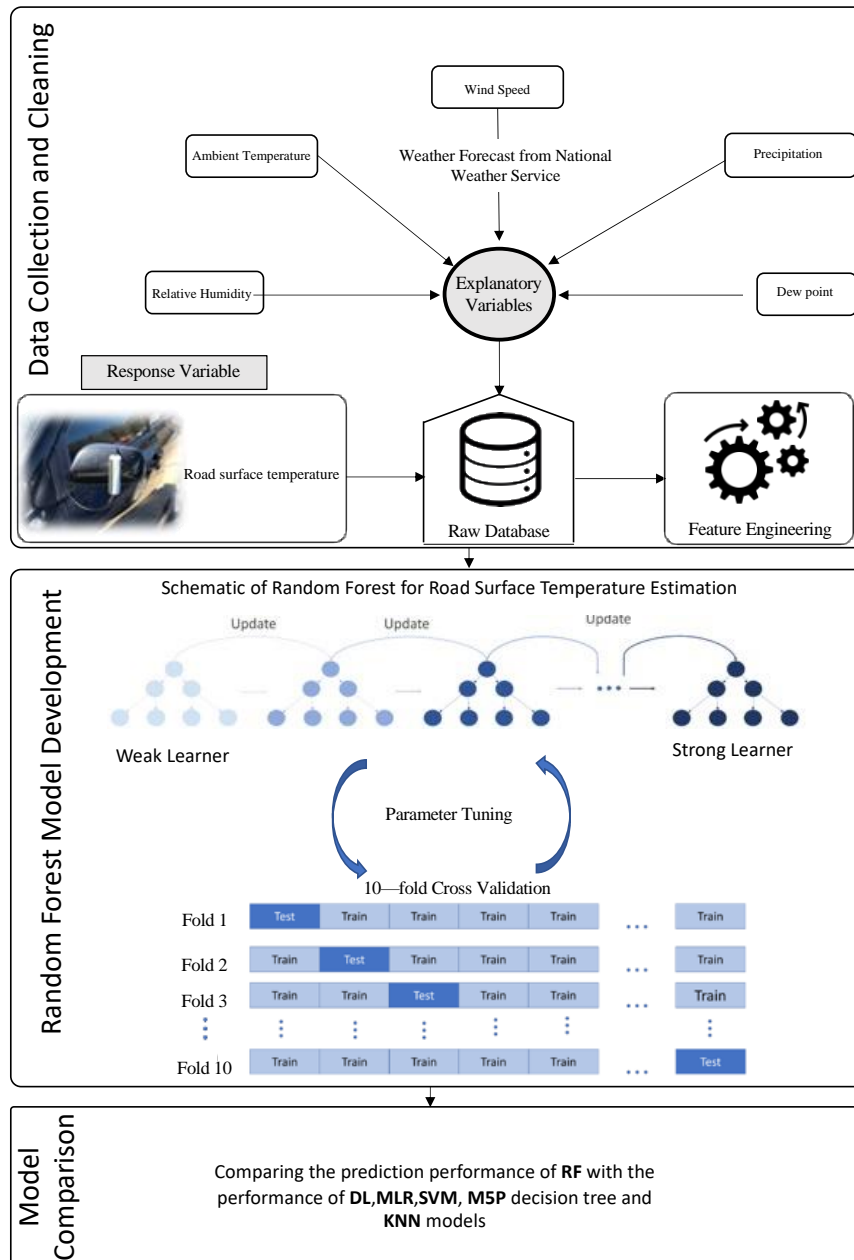


Figure 6-1 Overview of the proposed methodology framework

### 6.2.1. Data collection

For data collection, a temperature sensor kit, as described in Chapter 4 (Section 4.4.1), was utilized to randomly measure road surface temperature (RST) from various roadways in North Texas during the winter season of 2021-22. The National Digital Forecast Database (NDFD), which is maintained by the NWS, was utilized to obtain the weather forecast data that was associated with the collected RST. The corresponding weather forecasts such as air temperature, relative humidity, and wind speed were extracted from the NDFD based on the geographical coordinates and the time of the measurements for each RST sample. The dataset included eleven numerical variables and two categorical variables. The categorical variable of “road conditions” was considered “dry” when no precipitation was forecasted, and “wet” when precipitation was forecasted. The “sky conditions” categorical variable was classified into four categories as suggested by NWS based on the cloud coverage forecast (National Oceanic and Atmospheric Administration, n.d.). Table 6-1 summarizes the variables in the dataset that were used for developing the predictive models.

Table 6-1 Summary of the variables used for developing the predictive models

Type	Unit	Variable	Descriptive Statistics		
			Max	Min	Standard Deviation
Numerical	°C	Road Surface Temperature	27	-2	6.4
Numerical	n/a	Hour	24	0	4.9
Numerical	°C	Air Temperature	21	-2	5
Numerical	%	Relative Humidity	100	0	22
Numerical	°C	Average Air Temperature of Last Day	22	-1	5
Numerical	°C	Maximum Air Temperature of Last Day	28	1	6
Numerical	°C	Minimum Air Temperature of Last Day	20	-5	5

Numerical	m/s	Wind Speed	12	0	2.5	
Numerical	m/s	Wind Gust	17	0	4.2	
Numerical	°C	Dew Point	15	-14	5	
Numerical	kPa	Pressure	102.65	95.28	1.7	
Categorical	N/A	Sky Conditions	Clear (336 samples)	Partly Cloudy (80 samples)	Cloudy (124 samples)	Mostly Cloudy (140 samples)
Categorical	N/A	Road Conditions	Wet (234 samples)		Dry (446 samples)	

**6.2.2. Random Forest model development**

A Random Forest model is an ensemble learning algorithm that combines multiple decision trees to make predictions, resulting in improved accuracy and robustness for classification and regression tasks (Breiman, 2001). The final prediction of the Random Forest model is determined by aggregating the predictions of all the individual trees, typically through majority voting for classification or averaging for regression. This ensemble approach helps improve the model's accuracy, robustness, and ability to handle complex datasets. Random Forest models are known for their effectiveness in handling high-dimensional data, feature selection, and handling missing values. Figure 6-2 shows the structure of a typical Random Forest model.

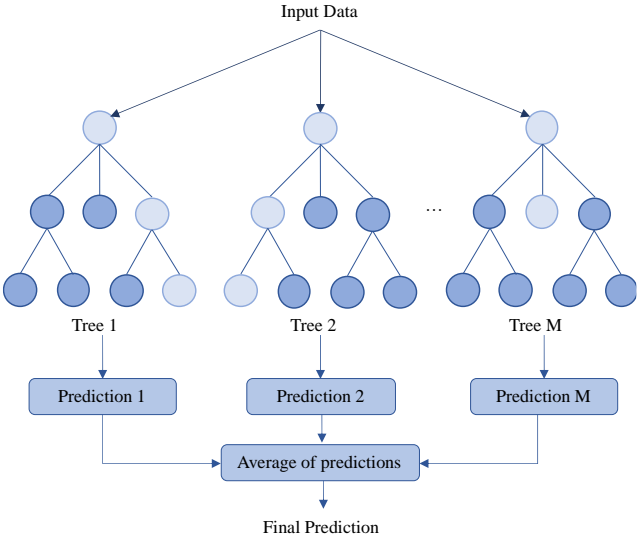


Figure 6-2 Structure of Random Forest models (adapted from Gitconnected.com)

The general mathematical form of a Random Forest model is:

$$f(X_i) = \frac{1}{M} \sum_{m=1}^M T_m(X_i; \Theta_m) \quad (6-1)$$

where  $M$  is the number of trees,  $X_i$  is a set of features in the dataset, and  $\Theta_m$  is a set of parameters that define each tree (Montgomery and Olivella, 2018). The procedure for developing a Random Forest regressor can be outlined as follows:

1. Select random  $m$  samples from the training dataset,
2. Construct a decision tree associated with these  $m$  samples,
3. Choose an arbitrary number of trees and repeat the first two steps,
4. Predict the target value using individual trees and assign the mean of all predictions as the final prediction.

The advantage of a Random Forest model over other multiple-tree models (e.g., gradient boosting machine and Bayesian additive regression) is that the Random Forest model reduces the levels of correlation between the trees, which in turn increases the model performance (Montgomery and Olivella, 2018). The Random Forest is more robust to changes in the input data as well as outliers in predictors compared to individual tree models (Breiman, 2001). The generalization error always converges by increasing the number of trees in the model. Since each tree is a completely independent random experiment, the risk of overfitting is low in Random Forest modeling (Youssef et al., 2016).

### ***Modeling Approach***

The Random Forest regression model was trained to predict the RST based on the created dataset described in the data collection section. The model hyperparameters, shown in Table 6-2,



were optimized using the random sampling method and 10-fold cross-validation. The main hyperparameters of the Random Forest model that influence the model error are the number of trees, the maximum depth of the tree, and the maximum number of features for splitting the node (Shreyas et al., 2016). In order to achieve a good balance between performance, memory consumption, and processing time, it is recommended to keep the maximum number of trees between 100 and 1000 (Callens et al., 2020). The parameter maximum depth of the tree is the number of nodes from the furthest leaf node to the tree's root. Low maximum depth values might result in underperformance of the model, while high maximum depth values may lead to overfitting (Putrada et al., 2021).

Table 6-2- Random Forest model hyperparameters

Row	Hyperparameters	Values
1	Number of trees	[100, 200, 300, 400, 500, 600, 700, 800, 900, 1000]
2	Maximum depth of the tree	[10, 20, 30, 40, 50, 60, 70, 80, 90, 100, none]
3	Maximum number of features for splitting a node	[2,4,6,8,10]

The random sampling technique is employed to identify the optimal subset of hyperparameters that yield the greatest enhancement in the model's prediction performance. Figure 6-3 shows the overview of the modeling approach used in this study. The random sampling algorithm randomly takes samples based on the statistical distribution for each hyperparameter, allowing to control the number of attempted hyperparameter combinations. To perform the 10-fold cross-validation, the dataset is randomly partitioned into ten groups for ten distinct iterations. In each iteration, the model is trained using nine groups, which constitute 90% of the dataset, while the remaining group, comprising 10% of the dataset, is employed for validation purposes. After repeating this process

ten times, the mean square error of the model was calculated using the average value for all iterations, and the best combination of hyperparameters was selected based on the hyperparameter set that resulted in the lowest error.

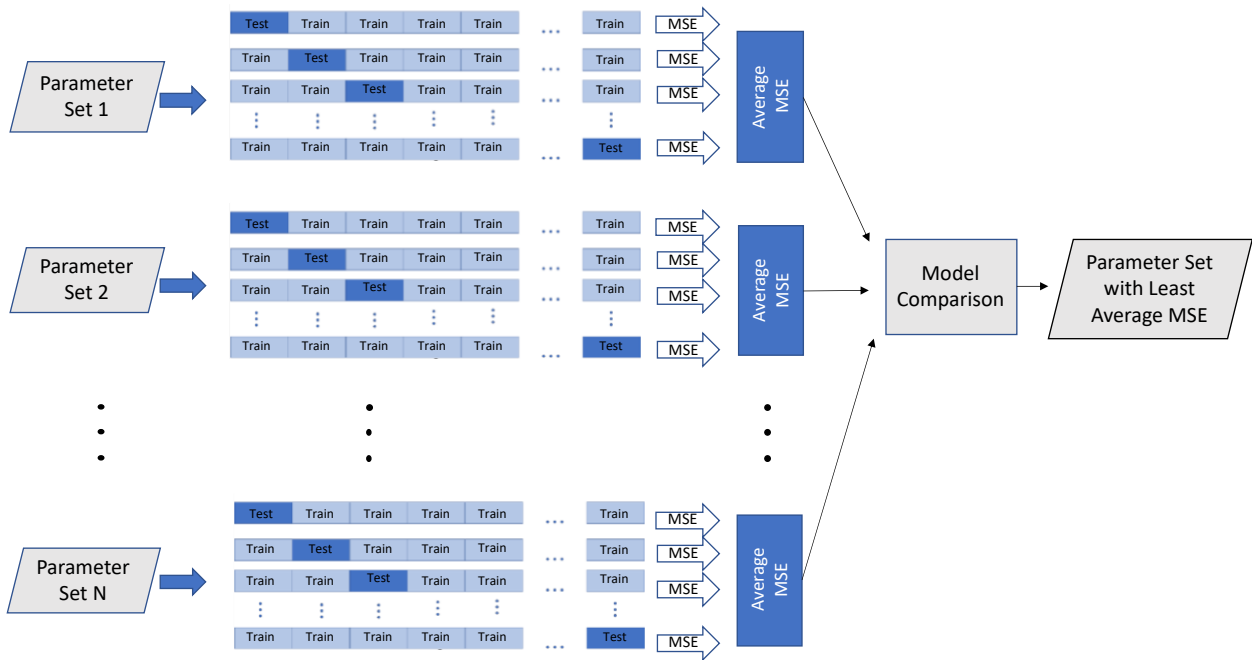


Figure 6-3 Overview of cross-validation in hyperparameter optimization of Random Forest model

### *Performance Metrics*

To evaluate the predictive performance of the models, three metrics were employed: R-squared ( $R^2$ ), mean absolute error (MAE), and root mean squared error (RMSE). The R-squared metric quantifies the proportion of the dependent variable that can be explained by the independent variables, while the Mean Absolute Error calculates the average magnitude of the residuals within the dataset. Additionally, Root Mean Square Error quantifies the standard deviation of the residuals.

## 6.3. Results and Discussions

### 6.3.1. Descriptive Statistics

This section provides descriptive statistics of the collected data described in the data collection section. The dataset includes two categorical variables: “road surface condition”, and “sky condition.” The point biserial correlation, a special case of Pearson’s correlation that reflects the relationship between categorical and continuous features, was conducted to examine the correlation between these categorical variables and the RST. The correlation results, summarized in Table 6-3, show that these two categorical features are not highly correlated with the RST. However, the “road surface condition” is slightly more correlated with the RST than the “sky condition.”

Table 6-3 Point Biserial Correlation between the categorical variables and RST

	Road Surface Temperature	
	Correlation	p-value
Road Surface Condition	0.18	1.86e-6
Sky Condition	0.13	5.7e-3

To determine the correlation between the continuous variables, the Spearman correlation analysis was conducted on the dataset. Figure 6-4 illustrates the Spearman correlation between the RST and weather forecast parameters.

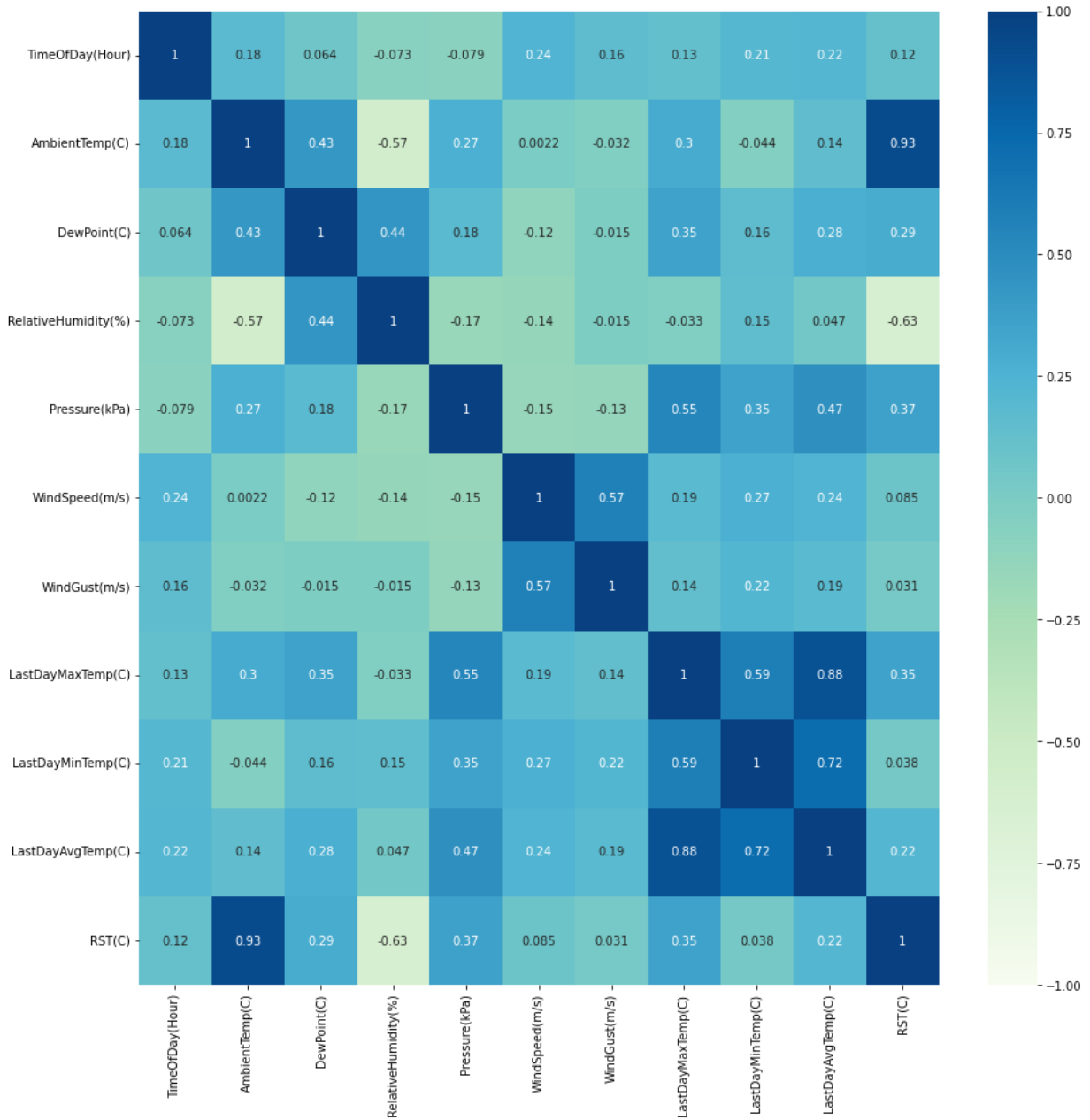


Figure 6-4 Spearman Correlation between the publicly available weather data and RST

The results indicate that the correlation between the RST and weather forecast parameters are all positive except for the relative humidity. The strongest positive correlation is between the RST and ambient temperature. The only negative correlation between the relative humidity and RST means that any increase in the relative humidity will reduce the RST. According to the thresholds shown in Table 6-4, RST is very weakly correlated with wind speed, wind gust, last day minimum

temperature, and time of day (hour); moderately correlated with dew point, last day maximum temperature, and air pressure; and weakly correlated with last day average temperature. It should be noted that the Spearman correlation only considers the monotonic relationships between the predictive features and the target variable. However, the predictive features and the target variable may be correlated non-monotonically.

Table 6-4 Thresholds for interpretation of Spearman correlation coefficients (Dancey and Reidy, 2007)

Spearman Correlation	Correlation
$\geq 0.7$	Very Strong
0.4 – 0.69	Strong
0.3 – 0.39	Moderate
0.20 – 0.29	Weak
0.01 – 0.19	Very Weak

### 6.3.2. Random Forest Hyperparameter Optimization

The Random Forest model was developed using the 10-fold cross-validation technique based on the mean absolute error loss function. The prediction performance of the Random Forest model was optimized by selecting the best set of hyperparameters that minimize the model’s error the most. Table 6-5 shows the default and optimized values of the model’s hyperparameters. The results show that the optimal number of trees was increased from 100 (default) to 400 (optimized), whereas other hyperparameters remained the same during the hyperparameter optimization. Figure 6-5 shows the mean absolute error of the model based on the number of trees and the maximum

number of features for splitting a node. The results indicate that the mean absolute error of the model varies by approximately 16% for random combinations of the number of trees and the maximum number of features for splitting a node.

Table 6-5 Hyperparameters for the default and optimized Random Forest models

Hyperparameter	Default	Optimized
Number of trees	100	400
Maximum depth of the tree	None	None
Maximum number of features for splitting a node	4	4

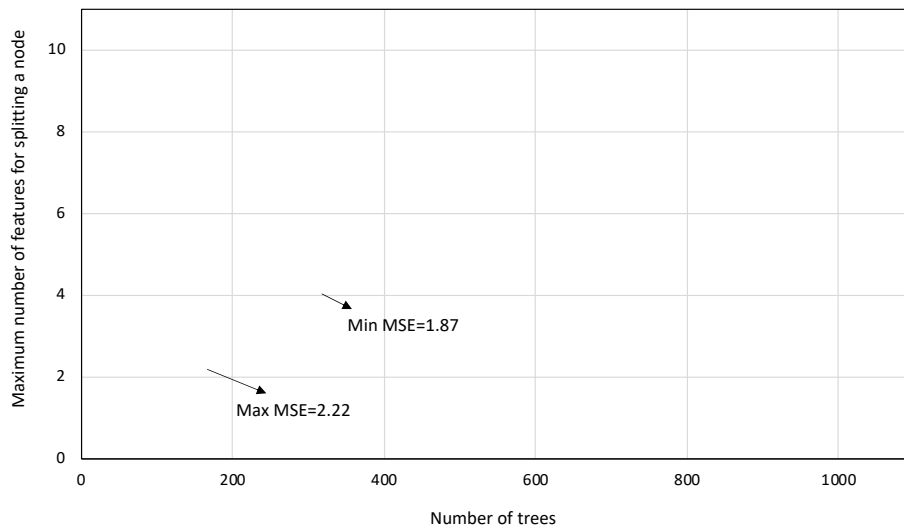


Figure 6-5 Mean absolute error of Random Forest model based on the number of trees and maximum number of features for splitting a node

Furthermore, the RMSE, MAE, and R-squared were calculated for the default and optimized models. The findings, presented in Figure 6-6, indicate that the model with optimized hyperparameters achieves approximately 2% improvement in RMSE, MAE, and R-squared.

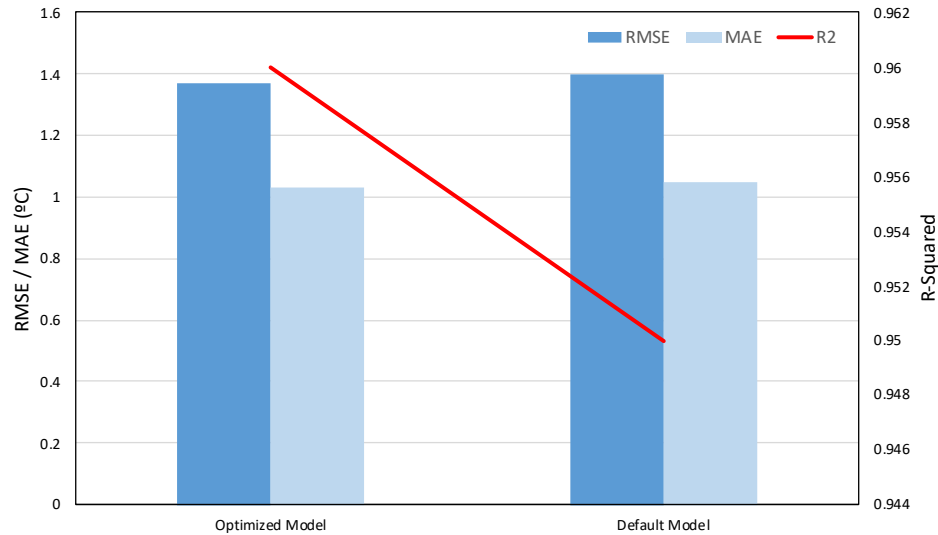


Figure 6-6 RMSE, MAE, and R-squared of optimized and default Random Forest models

### 6.3.3. Feature Importance

In addition to estimating RST estimates, Random Forest methods are capable of providing importance scores that indicate the importance of predictors in predicting the target value. By determining feature importance scores, predictors can be ranked according to their ability to predict a target variable (Abediniangerabi et al., 2021). The importance of a feature can be calculated by measuring the increase in the model's prediction error following its permutation. If shuffling the value of a feature results in an increase in the model's error, it is deemed "important". In contrast, if changing the feature's values does not affect the model error, it is considered "unimportant" (Fisher et al., 2019; Breiman, 2001). In this study, the importance of independent features on RST was evaluated using the permutation feature importance technique. The feature importance score is calculated by calculating the average variation reduction over all trees caused by selecting each feature for the internal nodes. Figure 6-7 illustrates the relative feature importance scores for predicting the RST based on the weather forecast parameters. Among the predictors, the ambient temperature, relative humidity, and time of day (hour) have a higher

influence on the RST by having relative importance scores of 0.45, 0.18, and 0.13, respectively. Although the time of day (hour) was very weakly correlated to RST using the Spearman correlation, it is among the most important factors based on the Random Forest feature importance score. The reason is that the Spearman method only considers the monotonic relationship between the variables. However, the Random Forest feature importance considers arbitrary relationships between the features including the non-monotonic relationships.

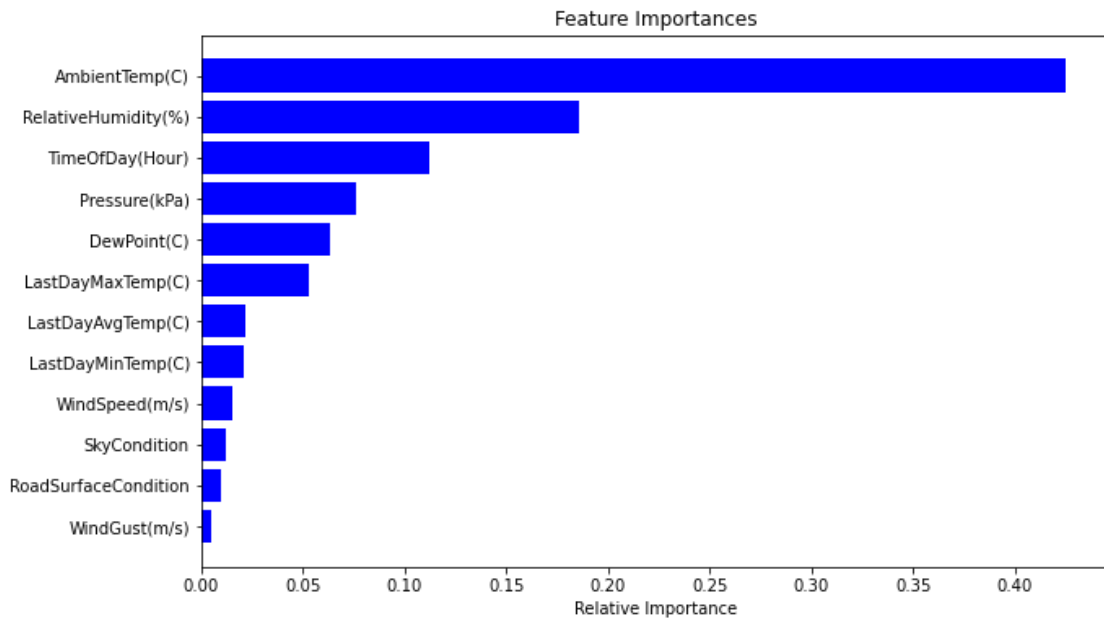


Figure 6-7 The relative importance of predictive features for the Random Forest Model

#### 6.3.4. Comparison between RF, DL, KNN, MLR, M5P Decision Tree, and SVR

The testing accuracies of the Random Forest model were compared with multiple prediction models including multiple linear regression (MLR), deep learning (DL), K-nearest neighbor (KNN), M5P decision tree, and support vector machines (SVM). The hyperparameters of each predictive model were optimized using the random search method. Table 6-6 summarizes the selected hyperparameters for each model.



Table 6-6- Optimized hyperparameters of selected data-driven prediction models

Model	Optimized Hyperparameters
K-nearest neighbor	<ul style="list-style-type: none"> <li>• Distance: Manhattan</li> <li>• Number of neighbors: 5</li> </ul>
Deep Learning	<ul style="list-style-type: none"> <li>• Momentum: 0.2</li> <li>• Number of Layers: 2 (first layer 4, second layer 5)</li> <li>• Learning rate: 0.3</li> </ul>
Support Vector Machine	<ul style="list-style-type: none"> <li>• Kernel: Polynomial (degree 4)</li> <li>• C=10</li> <li>• Gamma=0.001</li> </ul>
M5P decision tree	<ul style="list-style-type: none"> <li>• Minimum number of instances per leaf:4</li> </ul>

The deep learning model trained to predict the RST comprised two fully connected hidden layers with 4 and 5 hidden nodes, and one output layer. The network was trained through 100 iterations with a learning rate of 0.3 and momentum of 0.2. The K-nearest Neighbor model was trained using Manhattan distance and five (5) neighbors. The support vector machine was trained using a Polynomial kernel (degree of 4), C value of 10 (penalty parameter of the error), and gamma value of 0.001. Moreover, the M5P decision tree— a conventional decision tree with the addition of a linear regression function at the nodes— was trained using at least 4 instances allowed at each node. In addition, a multiple linear regression model was trained using attributes selected by the M5 Regression Tree’s attribute selection method. The attribute selection method of the M5 Regression Tree algorithm identifies the most suitable subset of features that contribute to maximizing improvements in the Akaike information criterion (AIC) (Massana et al., 2015). The selected attributes for the MLR model included time of day (hour), ambient temperature, dew point, relative humidity, air pressure, and the average temperature of the last day. To avoid multicollinearity in the model, the Spearman correlation values were used to ensure that highly

correlated attributes are not selected—a correlation coefficient of  $|R| > 0.7$  is commonly used as the multicollinearity threshold (Dormann et al., 2013).

Figure 6-8 shows the measured and predicted RST for the testing data using 10-fold cross-validation. The vertical lines for each data point indicate the relative error per prediction. The results indicate that the Random Forest model—with the coefficient of determination (R-squared) of 0.96—can explain more variability in the target value, compared to other prediction models (i.e., DL, MLR, SVM, KNN, and M5P decision tree).

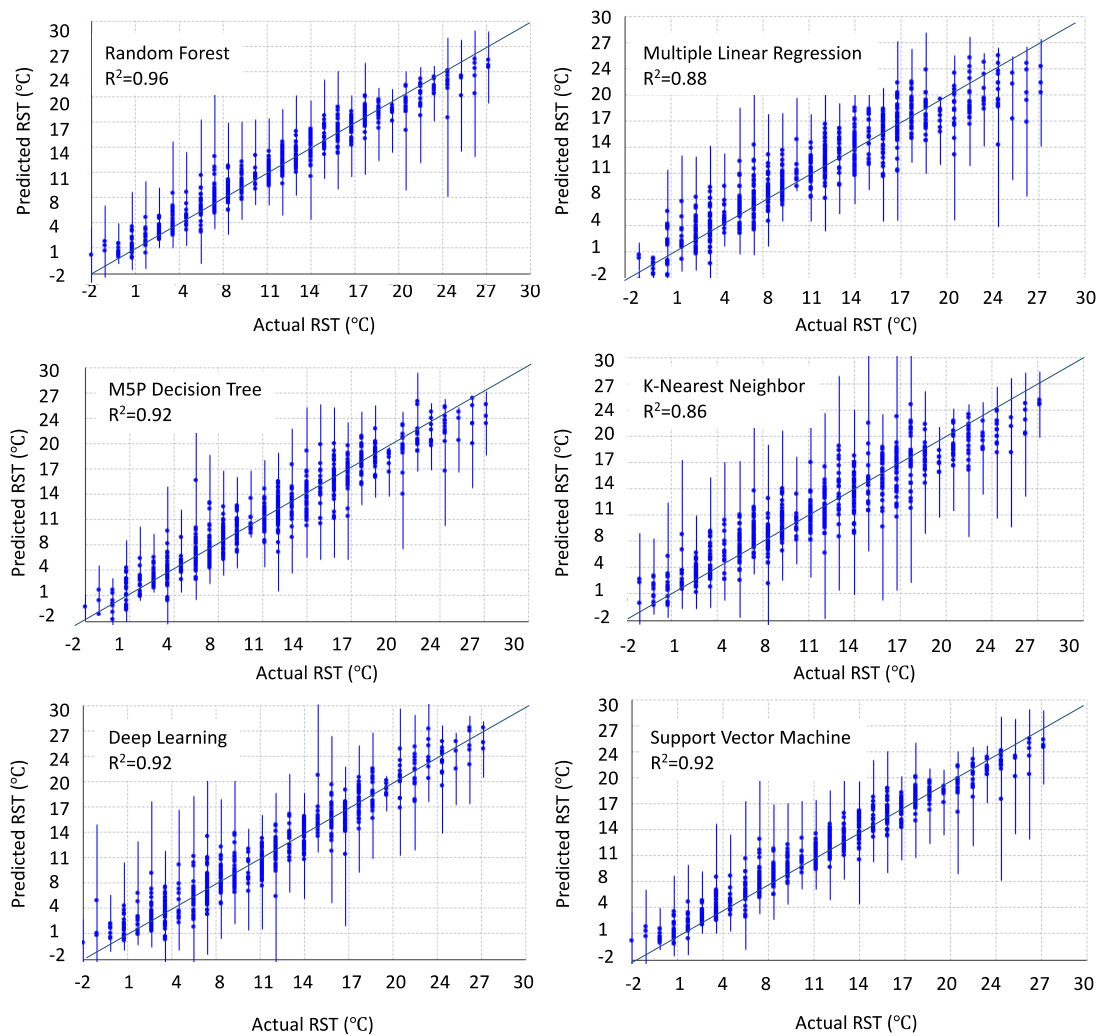


Figure 6-8 Coefficients of determination for the testing datasets of developed DL, MLR, M5P decision tree, SVR, RF, and SVR

Furthermore, a comparison of the prediction performance of the models, including root mean square error, mean absolute error, R-squared, and model construction time is shown in Figure 6-9. According to the results, the Random Forest model trained to predict the RST had the highest accuracy metrics among other models with a root mean square error of 1.38 °C and mean absolute error of 1.03 °C. Model construction for the SVR model takes about 16 seconds, which is relatively long compared to other models. This can be due to the use of a complex kernel and a high C value (penalty parameter of the error) for the model. On the other hand, the KNN model is the fastest model, taking approximately 0.01 seconds to build, followed by multiple linear regression, M5P decision tree, Random Forest, and deep learning.

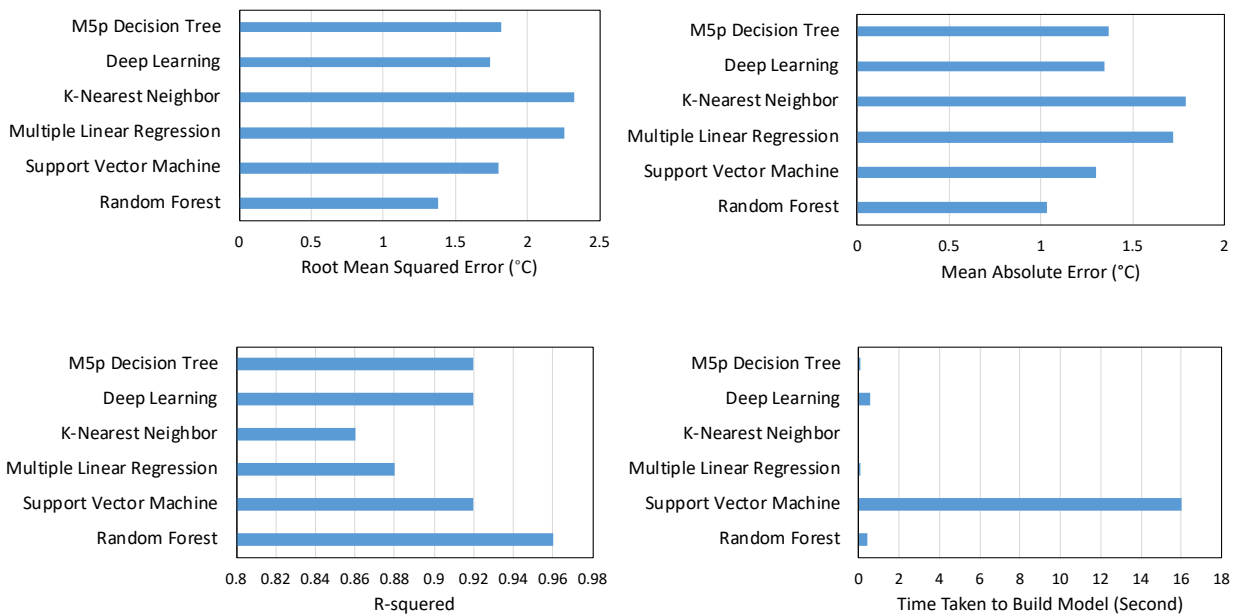


Figure 6-9 Accuracy metrics as well as the time taken to build the models

#### 6.4. Summary

Road surface temperature (RST) monitoring is essential for establishing effective winter maintenance plans to enhance road safety and prevent congestion. Road weather information systems typically provide actual and estimated road surface temperatures. As road weather

information systems (RWIS) are not available in many places, it can be challenging to obtain information about RST in areas with no access to RWIS. This study aimed to evaluate the applicability of the Random Forest model for predicting the road surface temperature based on the weather forecasts from NWS. To develop the Random Forest model, a dataset was compiled, incorporating observed RST collected via a vehicle-mounted temperature sensor, along with corresponding weather variables extracted from the NDFD. The actual RST was randomly measured from road networks in North Texas in the 2021-22 winter season. The corresponding weather data for each RST sample were obtained from the national digital forecast database (NDFD) for the time and geographical location where the actual RST was collected. The prediction performance of the Random Forest model was compared with the prediction performance of multiple predictive models including multiple linear regression, deep learning, support vector machine, support vector machine, M5P decision tree, and k-nearest neighbor. The results indicated that the Random Forest model outperforms other prediction models, indicating the advantage of the Random Forest model in estimating the road surface temperature using the weather forecast data from the NWS. The feature importance scores of the explanatory variables were also derived for the Random Forest model. According to the feature importance scores, the relative humidity, ambient temperature, and time of day had a score higher than 0.1, indicating their significant influence on the prediction of the RST.

The findings of this study emphasize the practicality of employing the Random Forest algorithm to enhance the precision of RST estimation, utilizing weather forecasts from NWS. The proposed methodology can be integrated into a winter operation decision-making system to map the estimated RST on roadways without the need for road weather information systems. The estimated

RST on roadways assists highway agencies to plan winter maintenance strategies more efficiently by taking proactive measures in areas where low surface temperatures are estimated.

The methodology presented in this study provides a framework that can be extended to include additional geospatial data including road vegetation, average daily traffic, latitude, longitude, and altitude of roads to increase the accuracy of RST estimates. Furthermore, it would be interesting to examine the performance of the prediction models during extreme weather, especially during winter storms with data ranges that go beyond those used in this study. Future research may investigate whether the proposed methodology can be incorporated into a snowplow operations optimization framework to optimize the plowing routes.

## CHAPTER 7

### CONCLUSION AND FUTURE WORK

Accurate and up-to-date information on road conditions is crucial for transportation supervisors in effectively managing winter operations during winter storms. This research developed a cost-effective system to collect and transfer real-time images of road conditions during winter operations using the capabilities of mobile devices. Furthermore, an innovative approach was developed to estimate the road surface temperature utilizing the available weather forecasts which are provided by the National Weather Service. This innovative method is expected to eliminate the necessity for relying on costly fixed sensor stations along roads, such as Road Weather Information Systems. Additionally, a 2D digital twin system was developed to facilitate the visualization of the collected data including the road conditions images, estimated road surface temperatures, and related weather forecasts.

The developed system was implemented in the TxDOT Wichita Falls district for two winter seasons. The findings revealed that the developed system can serve as a cost-effective alternative to the expensive GPS-AVL systems that are consisted of multiple pieces of equipment such as GPS devices, cellular modem communication, and mobile data computer, which may not be justifiable for regions with infrequent winter storms. The system could be easily detached when not in use, and users do not require extensive training to operate it. Moreover, the statistical prediction models demonstrated effective outcomes in estimating the road surface temperature by utilizing the weather forecast provided by the National Weather Service. The linear regression models, developed to estimate the road surface temperature for light and dark groups indicated that the ambient temperature, relative humidity, wind speed, the average temperature of the previous day, and road surface condition (wet/dry) are statistically significant factors affecting

road surface temperature at 5% level of significance for both dark and light groups. Based on the linear regression findings, the models exhibited a Mean Absolute Error (MAE) of around 1°C for the dark group and approximately 2°C for the light group. These values are considered acceptable when compared to previous studies conducted in this field.

Additionally, in order to enhance the accuracy of the road surface temperature estimation models, nonlinear statistical models, with a particular emphasis on Random Forests, were developed. The Random Forest models were employed to estimate the road surface temperature, and their performance was compared to that of other widely used models, such as deep learning, support vector machine, k-nearest neighbor, and M5P decision tree. The findings demonstrated that the Random Forest model exhibited superior performance compared to other prediction models, highlighting its advantage in accurately estimating road surface temperature by leveraging weather forecast data from the National Weather Service (NWS). According to the random forest feature importance scores, relative humidity, ambient temperature, and time of day were identified as the top three influential factors in predicting the road surface temperature, with scores higher than 0.1.

This research is expected to contribute towards enhancing the communication of adverse road conditions, thereby improving the decision-making process for snowplow operations. Real-time road condition images enable transportation operations managers to visually monitor and assess road conditions, without the need to travel on roads. Furthermore, access to information about road surface temperature along with specific weather information (e.g., snow and ice accumulation forecasts), which aids in snowplow operation decisions, provides crucial insights for operations managers regarding areas prone to low surface temperatures and potential ice and snow hazards. This information significantly improves the decision-making process for deploying snowplows to

implement anti-icing and snow removal measures during winter operations without relying on information obtained from costly fixed weather stations along roads such as road weather information systems.

### ***Future Works***

In order to expand upon the findings of this research, future work could focus on the following ideas:

- It is recommended to broaden the scope of data collection to encompass a wider range of extreme weather conditions, thus extending the validity of statistical models.
- It is recommended to assess the feasibility of integrating the developed framework, which estimates road surface temperature using weather forecasts, into a route optimization framework for snowplows. This integration has the potential to optimize plowing routes and material spreading operations.
- It is recommended to explore the inclusion of additional features, such as traffic volume, road vegetation, pavement material type (e.g., concrete or asphalt), and road class (urban or rural), in the development of estimation models. By assessing the significance and potential impact of these features, it is possible to enhance the accuracy of the estimation models.
- It is recommended to evaluate the performance of additional prediction models, including gradient boosting, extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost), for the estimation of road surface temperature. Conducting such assessments will provide insights into the effectiveness and suitability of these models in accurately predicting road surface temperature.



- It is recommended to employ machine learning image classification techniques to create an automated approach for categorizing road conditions into different categories, such as snowy, icy, and wet.

## REFERENCES

Abediniangerabi, B., Makhmalbaf, A., & Shahandashti, M. (2021). Deep learning for estimating energy savings of early-stage facade design decisions. *Energy and AI*, 5, 100077.

Ameen, W., & Shahandashti, M. (2021). Surveying the Effectiveness of Methods for Enhancing the Visibility of Snowplows in Texas. In *International Conference on Transportation and Development 2021* (pp. 309-320).

Ameen, W., Farooghi, F., Shahandashti, M., & Mattingly, S. (2022). Visibility of Winter Operations Vehicles: The State of Practice in the United States. *Journal of Cold Regions Engineering*, 36(2), 06022003.

Android developers. (2020a). Location, Retrieved June 10, 2020, from the Android Developer. Website: <https://developer.android.com/reference/android/location/Location>

Android developers. (2020b). `android.hardware.camera2`, Retrieved June 8, 2020, from the Android Developer. Website: <https://developer.android.com/reference/android/hardware/camera2/package-summary>

Android developers. (2020c). Get last known location, Retrieved June 12, 2020, from the Android Developer. Website: <https://developer.android.com/training/location/retrieve-current>

Android developers. (2020d). ExifInterface, Retrieved June 12, 2020, from the Android Developer. Website: <https://developer.android.com/reference/android/media/ExifInterface>

Android developers. (2020e). EditText, Retrieved May 10, 2021, from the Android Developer website:  
<https://developer.android.com/reference/android/widget/EditText>

Apache Software Foundation. (2020). Class FTP connection, Retrieved June 12, 2020, from the apache commons. Website:  
<https://commons.apache.org/proper/commons-net/apidocs/org/apache/commons/net/ftp/FTPClient.html>

Asefzadeh, A., Hashemian, L., & Bayat, A. (2017). Development of statistical temperature prediction models for a test road in Edmonton, Alberta, Canada. *International Journal of Pavement Research and Technology*, 10(5), 369-382.

Baral, A., Darghiasi, P., & Shahandashti, M. (2023). Risk-Averse Maintenance Decision Framework for Roadside Slopes along Highway Corridors. In *International Conference on Transportation and Development 2023* (pp. 383-392).

Baral, A., Poumand, P., Adhikari, I., Abediniangerabi, B., & Shahandashti, M. (2021). GIS-Based Data Integration Approach for Rainfall-Induced Slope Failure Susceptibility Mapping in Clayey Soils. *Natural Hazards Review*, 22(3), 04021026.

Baral, A., Sadegh Nasr, M., Darghiasi, P., Abediniangerabi, B., & Shahandashti, M. (2022). Detection and Classification of Vegetation for Roadside Vegetation Inspection and Rehabilitation Using Deep Learning Techniques. In *International Conference on Transportation and Development 2022* (pp. 143-152).

- Barnett, R. (1997). Higher education: A critical business. McGraw-Hill Education (UK).
- Bolton, R. N., McColl-Kennedy, J. R., Cheung, L., Gallan, A., Orsingher, C., Witell, L., & Zaki, M. (2018). Customer experience challenges: bringing together digital, physical and social realms. *Journal of service management*, 29(5), 776-808.
- Bosscher, P. J., Bahia, H. U., Thomas, S., & Russell, J. S. (1998). Relationship between pavement temperature and weather data: Wisconsin field study to verify super pave algorithm. *Transportation Research Record*, 1609(1), 1-11.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Callens, A., Morichon, D., Abadie, S., Delpey, M., & Liquet, B. (2020). Using Random Forest and Gradient boosting trees to improve wave forecast at a specific location. *Applied Ocean Research*, 104, 102339.
- Chapman, L., & Thornes, J. E. (2006). A geomatics-based road surface temperature prediction model. *Science of the Total Environment*, 360(1-3), 68-80.
- Chen, J., Wang, H., & Xie, P. (2019). Pavement temperature prediction: Theoretical models and critical affecting factors. *Applied Thermal Engineering*, 158, 113755.
- Cheng, L., Zhang, X., & Shen, J. (2019). Road surface condition classification using deep learning. *Journal of Visual Communication and Image Representation*, 64, 102638.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.
- Commercial Vehicle Group, (2014). *RoadWatch® manual*. Accessed November 1, 2021. <https://msfoster.com/wp-content/uploads/2014/09/Road-Watch-Bullet-Installation-Instructions-060414.pdf>

- Crevier, L. P., & Delage, Y. (2001). METRo: A new model for road-condition forecasting in Canada. *Journal of applied meteorology*, 40(11), 2026-2037.
- Dai, B., Yang, W., Ji, X., Zhu, F., Fang, R., & Zhou, L. (2023). An Ensemble Deep Learning Model for Short-Term Road Surface Temperature Prediction. *Journal of Transportation Engineering, Part B: Pavements*, 149(1), 04022067.
- Dancey, C. P., & Reidy, J. (2007). *Statistics without math's for psychology*. Pearson education.
- Darghiasi, P., Baral, A., & Shahandashti, M. (2023c). Developing a Cost-Effective Mobile-Based System for Collecting On-Demand Road Condition Images for Snowplow Operations Management. In *International Conference on Transportation and Development 2023* (pp. 127-137).
- Darghiasi, P., Baral, A., Abediniangerabi, B., & Shahandashti, M. (2022). A Multi-purpose All-in-One Mobile Data Collection System for Snowplow Operation Management. In *International Conference on Transportation and Development 2022*, forthcoming.
- Darghiasi, P., Baral, A., Abediniangerabi, B., Makhmalbaf, A. & Shahandashti, M. (2023b). Multilevel optimization of UHP-FRC sandwich panels for building façade systems. In *Artificial Intelligence in Performance-Driven Design*. Wiley.
- Darghiasi, P., Baral, A., Mattingly, S., & Shahandashti, M. (2023a). Estimation of Road Surface Temperature Using NOAA Gridded Forecast Weather Data for Snowplow Operations Management. *Journal of Cold Regions Engineering*. Advance online publication. <https://doi.org/10.1061/JCRGEI/CRENG-691>

Darghiasi, P., Zamanian, M., & Shahandashti, M. (2024). Enhancing Winter Maintenance Decision Making through Deep Learning-Based Road Surface Temperature Estimation. In Construction Research Congress 2024. Forthcoming

Dormann CF, Elith J, Bacher S. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36, 27–46.

Esri (2021a) Current Weather and Wind Station Data, retrieved in June 2021, retrieved in June 2021, available at <https://www.arcgis.com/home/item.html?id=cb1886ff0a9d4156ba4d2fadd7e8a139>

Esri (2021b) USA Weather Watches and Warnings, retrieved in June 2021, retrieved on June 221, available at <https://www.arcgis.com/home/item.html?id=a6134ae01aad44c499d12feec782b386>

Esri (2021c) Snowfall forecast for the next 72 hours across the Continental United States. retrieved on June 221, available at <https://www.arcgis.com/home/item.html?id=be1bb766bf1c44a9be97bbb7c04355ff>

Esri (2021d) National Weather Service Ice Accumulation Forecast for the USA, retrieved in June 2021, retrieved in June 2021, available at <https://www.arcgis.com/home/item.html?id=633a82d5917e4ceab13319266ce6ec3f>

Feng, T., & Feng, S. (2012). A numerical model for predicting road surface temperature in the highway. *Procedia Engineering*, 37, 137-142.

FHWA. (2020). Office of Operations Road Weather Management Program, "How do Weather Events Impact Roads?", Retrieved June 5, 2020. Website: [http://ops.fhwa.dot.gov/weather/q1\\_roadimpact.htm](http://ops.fhwa.dot.gov/weather/q1_roadimpact.htm).

- Fisher, A., Rudin, C., & Dominici, F. (2019). All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *J. Mach. Learn. Res.*, 20(177), 1-81.
- Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate behavioral research*, 26(3), 499-510.
- Gu, L., Wu, M., & Kwon, T. J. (2020). An enhanced spatial statistical method for continuous monitoring of winter road surface conditions. *Canadian Journal of Civil Engineering*, 47(10), 1154-1165.
- Gupta A, Choudhary A (2018) A framework for camera-based real-time lane and road surface marking detection and recognition. *IEEE Trans Intell Veh* 3(4):476–485
- Guthrie, W. S., Dye, J. B., & Eggett, D. L. (2014). Comparison of winter temperature profiles in asphalt and concrete pavements (No. UT-14.01). Utah. Dept. of Transportation.
- Haas, Karen, and David Hensing. Why your agency should consider asset management systems for roadway safety. No. FHWA-HRT-05-077. The United States. Federal Highway Administration. Office of Safety Research and Development, 2005.
- Hatamzad, M., Pinerez, G. C. P., & Casselgren, J. (2022). Intelligent cost-effective winter road maintenance by predicting road surface temperature using machine learning techniques. *Knowledge-Based Systems*, 247, 108682.
- Hoffmann, M., Nutz, P., & Blab, R. (2012). New findings in winter maintenance and their implementation in Austria. Standing International Road Weather Commission, 2012.

- Hong, L., Jun, L., & Yanhui, F. (2009, April). Road surface condition recognition method based on color models. In 2009 First International Workshop on Database Technology and Applications (pp. 61-63). IEEE.
- Hosseini, Faranak, et al. (2015) "Prediction of Pavement Surface Temperature Using Meteorological Data for Optimal Winter Operations in Parking Lots." Cold Regions Engineering.
- Imdadullah, M., Aslam, M., & Altaf, S. (2016). mctest: An R Package for Detection of Collinearity among Regressors. R J., 8(2), 495.
- Islam, M. R., Ahsan, S., & Tarefder, R. A. (2015). Modeling temperature profile of hot-mix asphalt in flexible pavement. International Journal of Pavement Research and Technology, 8(1), 47.
- Jia, L., Sun, L., & Huang, L. (2007). A numerical temperature prediction model for asphalt concrete pavement. JOURNAL-TONGJI UNIVERSITY, 35(8), 1039.
- Jonathan J. Rutz, Chris V. Gibson. (2013) Integration of a Road Surface Model into NWS Operations. Bulletin of the American Meteorological Society 94:10, 1495-1500.
- Kang, H. (2021). Sample size determination and power analysis using the G\* Power software. Journal of educational evaluation for health professions, 18.
- Kawai, S., Takeuchi, K., Shibata, K., & Horita, Y. (2012, November). A method to distinguish road surface conditions for car-mounted camera images at night-time. In 2012 12th International Conference on ITS Telecommunications (pp. 668-672). IEEE.
- Kimura, K., Namikawa, Y., Sone, S., & Kuwabara, M. (2006). Research on Environmental Impact of De-icing Salts. In PIARC XII INTERNATIONAL WINTER ROADS CONGRESS, TORINO-SESTRIERE, ITALY, 2006.

- Kršmanc, R., Slak, A. Š., & Demšar, J. (2013). Statistical approach for forecasting road surface temperature. *Meteorological applications*, 20(4), 439-446.
- Kwon, T. J., & Gu, L. (2017, August). Modeling 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.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Lee, M. S., & Nelson, D. (2018). Utilization of AVL/GPS Technology: Case Studies (No. CR 16-01).
- Liu, B., Yan, S., You, H., Dong, Y., Li, Y., Lang, J., & Gu, R. (2018). Road surface temperature prediction based on gradient extreme learning machine boosting. *Computers in Industry*, 99, 294-302.
- Liu, M., Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators B: Chemical*, 177, 970-980.
- Lv, Z., & Xie, S. (2022). Artificial intelligence in the digital twins: State of the art, challenges, and future research topics. *Digital Twin*, 1(12), 12.
- Mammeri, A., Ulmet, L., Petit, C., & Mokhtari, A. M. (2015). Temperature modeling in pavements: the effect of long-and short-wave radiation. *International journal of pavement engineering*, 16(3), 198-213.
- Mantero, A., & Ishwaran, H. (2021). Unsupervised Random Forests. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 14(2), 144-167.



- Massana, J., Pous, C., Burgas, L., Melendez, J., & Colomer, J. (2015). Short-term load forecasting in a non-residential building contrasting models and attributes. *Energy and Buildings*, 92, 322-330.
- Meyer, E., & Ahmed, I. (2003, August). Benefit-cost assessment of automatic vehicle location (AVL) in highway maintenance. In *Proceedings of the 2003 Mid-Continent Transportation Research Symposium*, Ames, Iowa.
- Milad, A. A., Adwan, I., Majeed, S. A., Memon, Z. A., Bilema, M., Omar, H. A., ... & Yusoff, N. I. M. (2021). Development of a Hybrid Machine Learning Model for Asphalt Pavement Temperature Prediction. *IEEE Access*, 9, 158041-158056.
- Miller, T. R., & Zaloshnja, E. (2009). On a crash course: The dangers and health costs of deficient roadways.
- Montgomery, J. M., & Olivella, S. (2018). Tree-Based Models for Political Science Data. *American Journal of Political Science*, 62(3), 729-744.
- National Oceanic and Atmospheric Administration (2021), National Weather Service, “Watch Warning Advisory Explained”, Retrieved January 21, 2021. Website: <https://www.weather.gov/sjt/WatchWarningAdvisoryExplained>
- National Oceanic and Atmospheric Administration (2022), SciJinks, “How Reliable Are Weather Forecasts.” Accessed April 15, 2022. <https://scijinks.gov/forecast-reliability/#:~:text=A%20seven%2Dday%20forecast%20can,right%20about%20half%20the%20time.>

- National Weather Service, (2019). National Digital Forecast Database and Local Database Description and Specification. National Weather Service Instruction 10-201, 19pp. retrieved from: <https://www.weather.gov/media/mdl/ndfd/pd01002001curr.pdf>
- National Weather Service, (2021). National Digital Forecast Database GIS Tutorial. retrieved from: <https://www.weather.gov/media/mdl/ndfd/pd01002001curr.pdf>
- National Weather Service. (2019). National Digital Forecast Database and Local Database Description and Specification. National Weather Service Instruction 10-201, 19pp. Accessed November 15, 2021. <https://www.weather.gov/media/mdl/ndfd/pd01002001curr.pdf>
- Nowrin, T., & Kwon, T. J. (2022). Forecasting short-term road surface temperatures considering forecasting horizon and geographical attributes—an ANN-based approach. *Cold Regions Science and Technology*, 202, 103631.
- Onstad, K. L. (2008). Methodology and process to accompany a virtual geotechnical database for the St. Louis metropolitan area.
- Potter, A. S., Gallagher, M. R., & Bayer, C. W. (2016). Synthesis on GPS/AVL equipment used for winter maintenance (No. CR 14-01). Minnesota. Dept. of Transportation. Clear Roads Pooled Fund.
- Putrada, A. G., & Perdana, D. (2021, October). Improving Thermal Camera Performance in Fever Detection during COVID-19 Protocol with Random Forest Classification. In *2021 International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS)* (pp. 1-6). IEEE.
- Qin, Y. (2016). Pavement surface maximum temperature increases linearly with solar absorption and reciprocal thermal inertial. *International Journal of Heat and Mass Transfer*, 97, 391-399.

- Qiu, X., Xu, J. X., Tao, J. Q., & Yang, Q. (2018). Asphalt Pavement Icing Condition Criterion and SVM-Based Prediction Analysis. *Journal of Highway and Transportation Research and Development (English Edition)*, 12(4), 1-9.
- Qiu, X., Xu, W. Y., Zhang, Z. H., Li, N. N., & Hong, H. J. (2020). Surface temperature prediction of asphalt pavement based on GBDT. In *IOP Conference Series: Materials Science and Engineering* (Vol. 758, No. 1, p. 012031). IOP Publishing.
- Refai, H., Stevenson A., (2018). Maintenance Activity Truck Tracking [PowerPoint Slides]. ITS Heartland Annual Meeting.
- RoadWatch manual (2020). Accessed November 1, 2021. <https://msfoster.com/wp-content/uploads/2014/09/Road-Watch-Bullet-Installation-Instructions-060414.pdf>
- Sato, N., Thornes, J. E., Maruyama, T., Sugimura, A., & Yamada, T. (2004, June). Road Surface Temperature Forecasting. In *Sixth International Symposium on Snow Removal and Ice Control Technology* (p. 414).
- Schneider, W., Crow, M., Holik, W. A., Bakula, C., Maistros, A. R., Gould, Z. T., & Lurtz, J. M. (2017). Evaluation of the GPS/AVL systems for snow and ice operation resource management (No. FHWA/OH-2017-31). Ohio. Dept. of Transportation. Office of Statewide Planning and Research.
- Shahandashti, M., Hossain, S., Baral, A., Adhikari, I., Pourmand, P., and Abedinangerabi, B. (2020). Slope Repair and Maintenance Management System, Final Report, TxDOT.
- Shahandashti, M., Hossain, S., Zamanian, M., & Akhtar, M. A. (2021). Advanced Geophysical Tools for Geotechnical Analysis.

- Shahandashti, M., Mattingly, S. P., Ameen, W., & Farooghi, F. (2019). Synthesis of Safety Applications in Winter Weather Operations.
- Shahandashti, M., Mattingly, S., Darghiasi, P., Baral, A., & Abediniangerabi, B. (2022). Snowplow Operations Management System (No. FHWA/TX-22/5-6996-01-1). Texas Department of Transportation. Research and Technology Implementation Office.
- Sherif, A., & Hassan, Y. (2004). Modeling pavement temperature for winter maintenance operations. *Canadian Journal of Civil Engineering*, 31(2), 369-378.
- Shreyas, R., Akshata, D. M., Mahanand, B. S., Shagun, B., & Abhishek, C. M. (2016, August). Predicting popularity of online articles using Random Forest regression. In 2016 Second International Conference on Cognitive Computing and Information Processing (CCIP) (pp. 1-5). IEEE.
- Takasaki, Y., Saldana, M., Ito, J., & Sano, K. (2022). Development of a method for estimating road surface conditions in winter using Random Forest. *Asian Transport Studies*, 8, 100077.
- Texas Department of Transportation (2021). Highway Designations Glossary. Accessed November 15, 2021: <https://www.txdot.gov/inside-txdot/division/transportation-planning/highway-designation/glossary.html>
- Texas Department of Transportation (2021a). Highway Designations Glossary. Accessed November 15, 2021: <https://www.txdot.gov/inside-txdot/division/transportation-planning/highway-designation/glossary.html>
- Texas Department of Transportation (July 2010). TxDOT Data Architecture (Version 4.2)
- The Federal Aviation Administration (2017), “Official Guide to Basic Flight Information and ATC Procedures, Aeronautical Information Manual,”

- Thornes, J. E., & Shao, J. (1991). Spectral analysis and sensitivity tests for a numerical road surface temperature prediction model. *Meteorological Magazine*, 120(1428), 117-124.
- Toivonen, E., Hippi, M., Korhonen, H., Laaksonen, A., Kangas, M., & Pietikäinen, J. P. (2019). The road weather model RoadSurf (v6. 60b) driven by the regional climate model HCLIM38: evaluation over Finland. *Geoscientific Model Development*, 12(8), 3481-3501.
- Wang, C., Jia, B., Zhou, J., Feng, L., & Chen, J. (2022). Retrieval of Road Surface (Bridge Deck) Temperature near 0° C Based on Random Forest Model. *Atmosphere*, 13(9), 1491.
- Wilks, D.S. (2011) *Statistical Methods in the Atmospheric Sciences*. New York, NY: Academic Press, 1 p
- Witten, I. H., & Frank, E. (2002). Data mining: practical machine learning tools and techniques with Java implementations. *Acm Sigmod Record*, 31(1), 76-77.
- Wu, M., Kwon, T. J., & Fu, L. (2022). Spatial mapping of winter road surface conditions via hybrid geostatistical techniques. *Journal of Cold Regions Engineering*, 36(4), 04022009.
- Yang, C. H., Yun, D. G., & Sung, J. G. (2012). Validation of a road surface temperature prediction model using real-time weather forecasts. *KSCE Journal of Civil Engineering*, 16(7), 1289-1294.
- Yang, C. H., Yun, D. G., Kim, J. G., Lee, G., & Kim, S. B. (2020). Machine learning approaches to estimate road surface temperature variation along road section in real-time for winter operation. *International Journal of Intelligent Transportation Systems Research*, 18(2), 343-355.
- Yin, Z., Hadzimustafic, J., Kann, A., & Wang, Y. (2019). On statistical nowcasting of road surface temperature. *Meteorological Applications*, 26(1), 1-13.

- Youssef, A. M., Pourghasemi, H. R., Pourtaghi, Z. S., & Al-Katheeri, M. M. (2016). Landslide susceptibility mapping using Random Forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. *Landslides*, 13(5), 839-856.
- Yun, D. G., Kim, S., Yang, C. H., Kim, J. G., & Park, J. H. (2018). Monitoring the road surface temperature by evaluating the climatological and ambient temperature data collected from probe vehicles using machine learning technique (No. 18-02505).
- Zamanian, M., & Shahandashti, M. (2022). Investigation of Relationship between Geotechnical Parameters and Electrical Resistivity of Sandy Soils. In *Construction Research Congress 2022* (pp. 686-695).
- Zamanian, M., Asfaw, N., Chavda, P., & Shahandashti, M. (2023). Classifying Soil Sulfate Concentration Using Electrical Resistivity Imaging and Random Forest Algorithm. In *Airfield and Highway Pavements 2023* (pp. 204-213).
- Zamanian, M., Darghiasi, P., Shahandashti, M. (2024). Empirical Study of the Correlation Between Geoelectrical and Soil-Index Properties of Clayey Soils. In *Construction Research Congress 2024*. Forthcoming