Environmental Impacts on the Performance of Pavement Foundation Layers-Phase I

MnDOT Project

Task 3 – Modeling Analyses

January 21st, 2021
Revised March 28th, 2021

Prepared by:
Bora Cetin – Principal Investigator
Kristen Cetin - Co-Principal Investigator
Tuncer B. Edil – Co-Principal Investigator
Debrudra Mitra – Graduate Research Assistant

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Literature Review

The objective of this task is to create and validate a model to predict soil temperatures and number of freeze-thaw cycles for different soil depths. To achieve the target, initially a literature review was completed to assess the current state of art in soil temperature prediction.

There are several studies that have focused on implementing data-driven methods to predict the soil temperatures. Mihalakakou et al. [1] used an Artificial Neural Network model to predict the temperature of a bare and short grass covered soil surface and compared its performance with physical models. Six years of data was used for training and 1 year of soil temperature data was used for model testing. They found that although the physical model performed better compared to the data-driven model, it is much more complex and required many more inputs. The relative error in the predicted soil surface temperature was 10% to 15%. In another study by Tang et al. [2], a linear regression model was used to predict the mean annual ground temperature, then a Feed-Forward Neural Network model was used to predict the daily mean ground temperature in Chengdu, China. Approximately 50 years of daily soil temperature data was used to create the model. Among the considered variables, ambient temperature and relative humidity were found to be able to be used to best predict the daily surface temperature. Talaee and Hosseinzadeh [3] predicted daily soil temperatures at six different soil depths using a Coactive Neuro-Fuzzy Inference System (CANFIS) in Iran [3]. 10 years of data was used to create the model where mean, maximum, and minimum air temperatures, relative humidity, hours of sunshine and solar radiation were used as variable inputs to the model.

In another study by George [4], weekly average soil temperature was predicted using air temperature, relative humidity and wind speed data using both Neural Network and Multiple Linear Regression models. In [5], monthly soil temperatures at four different depths were predicted and compared using three different models, including Multi-Layer Perceptron, Radial Basis Neural Network, and Generalized Regression Neural Network models. The air temperature was found to be the most effective variable to predict monthly soil temperature. In addition, the accuracy of the models generally reduced with an increase in depth. Kim et al. [6] modeled daily soil temperatures at two depths in Illinois using Multilayer Perceptron and Adaptive Neuro-Fuzzy Interference System using climate data as input. Another study by Bilgili [7] predicted the monthly soil temperature data with approximately 8 years of climate data in Turkey, using regression models, including Linear and Nonlinear Regression and Artificial Neural Networks. Stepwise Regression was also used to select the most important variables for analysis. In a similar study, 20 years of soil temperature data was used to predict the monthly soil temperature in Turkey using Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System and Multi-Linear Regression models [8]. Air temperature, month number, soil depth and monthly precipitation were determined to be the best combination of variables for soil temperature prediction. Daily soil surface temperatures were predicted using a combination of two different Support Vector Machine models, where one model was used to predict the annual average soil temperature and the other model was used to predict the daily ground temperature amplitude with respect to the annual average temperatures [9]. It was obtained that the combination of the two models performed much better compared to a single Support Vector Machine model in predicting the soil temperature.

In summary, the above-discussed studies used different data-driven methods, most commonly various types of Neural Network and Regression models to predict soil temperatures. However, there is no commonly accepted method nor a common set of variables which have been used to predict soil temperatures. In addition, in the majority of these studies, the shortest timestep used for temperature prediction is the daily level, rather than a more granular level. Similarly, none of these studies focus on the prediction of the number of freeze-thaw cycles based on soil temperature data.

2. Data Processing

Two different datasets are being used in this study, as discussed in the Task 2 report. For ease and clarity of discussion, we have named these **Dataset 1** and **Dataset 2**. Along with the soil surface temperatures at different depths, climate data was in Dataset 1. These data were collected across a 2-mile span of roadway in MNROAD facility at Monticello, Minnesota. Air temperature, relative humidity, wind speed, precipitation and solar radiation data were collected at a 15-minute timestep to be consistent with the collected temperature data granularity. In Dataset 2, soil surface temperature data were collected at a 1-hour time interval in three different counties in Minnesota, including Koochiching, Olmsted and Wright. For brevity, we focus on the data collected in Olmsted county in this Task, as data timespan of the three locations are similar.

The climatic variables were pre-processed to create a list of potential variables which can be used as the input parameters in the modeling process. The final list of parameters considered can be divided in two categories: time variables and climate variables. The list of all the variables is shown below:

The time variables considered include:
1. Month number (1 to 12)
2. Week number (1 to 52)
3. Day of year (1 to 365)
4. Timestep (1 to 4*24 for 15-minute timestep data)

The climatic variables considered include:
1. Air temperature (AirTemp)
2. Relative humidity (RH)
3. Rain or precipitation (Rain)
4. Windspeed (Wind)
5. Radiation (Rad)
6. Daily average air temperature (avgTemp)
7. Daily average relative humidity (avgRH)
8. Daily average precipitation (avgRain)
9. Daily average windspeed (avgWind)
10. Daily average solar radiation (avgrad)
11. Variation of the air temperature with respect to the daily average value (varTemp)
12. Variation of the relative humidity with respect to the daily average value (varRH)
13. Variation of the precipitation with respect to the daily average value (varRain)
14. Variation of the windspeed with respect to the daily average value (varWind)

15. Variation of the solar radiation with respect to the daily average value (varRad)
The variation values were calculated by subtracting the daily average values from the instantaneous value at a specific time interval, using the following equation:
Variation = Instantaneous value – Average value of a specific day

To implement the Regression model, variables should be independent of one other. To check the correlation among the variables, Stepwise Regression was used. The results of the Stepwise Regression analysis are shown in Table 1.

The correlation coefficient values are shown in Table 1, where any value close to ±1 represents high correlation and 0 represents no correlation. The cells are color coded based on the level of correlation coefficients. Yellow represents highly correlated parameters with a correlation coefficient higher than 0.7. Green cells represent moderate correlations where the absolute value of correlation coefficient value between 0.3 to 0.7. Non-colored cells represent variables with low correlation, with a correlation coefficient less than 0.3. These are used in the final Regression model.

Table 1. Correlation analysis using stepwise regression

	Week	*Month*	*Day of Year*	*Timestep*	*Air Temperature*	*Rain*	*RH*	*Wind*	*Radiation*	*Variation Air Temperature*	*Variation Rain*	*Variation RH*	*Variation Wind*	*Variation Radiation*	*Average Air Temperature*	*Average Rain*	*Average RH*	*Average Wind*	*Average Radiation*
Week	1																		
Month	1	1																	
DayofYear	1	1	1																
Timestep	0	0	0	1															
AirTemp	0.22	0.21	0.22	0.11	1														
Rain	0.02	0.02	0.02	0.01	0.04	1													
RH	0.22	0.22	0.22	-0.25	-0.24	0.09	1												
Wind	-0.09	-0.09	-0.09	0.04	-0.11	0.01	-0.23	1											
rad	0	0	0	0.04	0.46	-0.03	-0.5	0.18	1										
varTemp	0	0	0	0.4	0.27	-0.01	-0.55	0.26	0.5	1									
varRain	0	0	0	0.01	0	0.96	0.03	0.02	-0.02	-0.01	1								
varRH	0	0	0	-0.36	-0.21	0.04	0.71	-0.27	-0.51	-0.78	0.04	1							
varWind	0	0	0	0.05	0.1	0.02	-0.27	0.71	0.35	0.37	0.02	-0.38	1						
varRad	0	0	0	0.04	0.15	-0.02	-0.4	0.28	0.91	0.55	-0.02	-0.56	0.39	1					
avgTemp	0.22	0.22	0.22	0	0.96	0.04	-0.09	-0.18	0.34	0	0	0	0	0	1				
avgRain	0.09	0.09	0.09	0	0.15	0.26	0.24	0	-0.04	0	0	0	0	0	0.16	1			
avgRH	0.31	0.31	0.31	0	-0.13	0.09	0.71	-0.06	-0.2	0	0	0	0	0	-0.13	0.34	1		
avgWind	-0.12	-0.12	-0.12	0	-0.25	0	-0.06	0.71	-0.11	0	0	0	0	0	-0.26	0	-0.08	1	
avgrad	-0.01	-0.01	-0.01	0	0.77	-0.02	-0.33	-0.18	0.42	0	0	0	0	0	0.8	-0.09	-0.46	-0.25	1

As shown in Table 1, several variables cannot be considered individually in the Regression model as they are either highly or moderately correlated among each other. As an example, the time variables such as *Week* with respect to *Month*, and *Week* with respect to *Day of Year* are highly correlated. Similarly, the variation values of precipitation, relative humidity, wind speed and solar radiation are highly correlated with their respective measured data values. Similarly, the average air temperature is highly correlated with both the air temperature and average radiation values. The average radiation values are highly correlated with actual air temperature values. Thus, based on the results obtained of the stepwise regression method, 8 different variables were selected as the input variables for the Regression model. These include, *Day of Year*, *Timestep*, *Air Temperature*, *Radiation*, *Variation in Air Temperature*, *Variation in Rain*, *Variation in Relative Humidity* and *Variation in Windspeed*.

Dataset 1 and Dataset 2 (see Task 2 report) were then used to create the data-driven models. Dataset 1, with approximately 1.5-2 years of data available at each cell (16 months) from January 1, 2018 to April 16, 2019, was split into two datasets, one for use in training the model, and a second for model testing. This split was a 75%-25% split where 12 months of data (2018) was used as training data, and January 1, 2019 to the end of the dataset was used for the testing of the model. For Dataset 2, with data across multiple years, data was split in a 80%-20% division, where 80% of it was used for the training purpose and the rest of the data was used as testing data. For example, for the Olmsted location, data from September 2005 to February 2007 was used as testing data and January 2000 to September 2005 was used at training data. The reason that the datasets were split slightly differently is because for Dataset 1, training dataset was selected to include one whole year and similarly for Dataset 2, the objective was to select an entire winter season for training data. For weather data, for Dataset 1, onsite field collected data was used; for Dataset 2, the nearest available weather station data was used for each location.

3. Model Development:

As mentioned in the literature review, there are no well accepted algorithm that can be used for the prediction of the soil surface temperatures. At the initial stage of the study, different models were implemented to study which algorithm performs best. Initially, a Linear Regression model was used, followed by Regression models used for time series data forecasting, including Vector Auto Regression, Vector Auto Regression Moving Average, and Vector Error Correction Models. However, the above-mentioned models were unable to accurately capture the granular trend in soil temperature. An example result of soil temperature prediction is shown in Figure 1. This was obtained for a specific cell location and at a specific depth using Vector Auto Regression modeling methods. The results show the predicted temperature for testing dataset where x-axis of the figure shows timestep of the year. However, the results obtained for the other mentioned algorithms were also similar for all the soil surfaces. Thus, Nonlinear Regression models were implemented, and the performance of the soil temperature prediction improved significantly using this method.

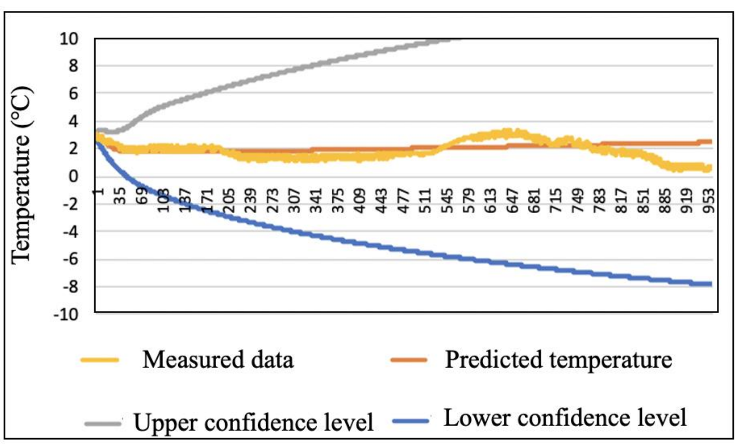


Figure 1. Temperature prediction using Vector Auto Regression Model

3.1. Temperature prediction: A Nonlinear Regression modeling method was implemented using a fourth order polynomial model. Two modeling methods were considered based on a Nonlinear Regression algorithm. Model 1 predicts temperature at each depth <u>individually</u>, using a single model. Two time variables and six climate variables were used as input. These include *Day of Year*, *Timestep*, *Air Temperature*, *Radiation*, *Variation in Air Temperature*, *Variation in Rain*, *Variation in Relative Humidity* and *Variation in Windspeed*. The equation for the Model 1 is in Appendix. Model 2 was developed using a combination of two parts - one to predict the daily average soil temperatures, and a second to predict the variation in soil temperature with respect to the daily average values. To predict the daily average soil temperatures, six variables were used (*Day of Year*, *Timestep*, *Average Air Temperature*, *Average Rain*, *Average Relative Humidity* and *Average Wind Speed*. For the second half of the model, *Day of Year*, *Timestep*, *Air Temperature*, *Radiation*, *Variation in Air Temperature*, *Variation in Rain*, *Variation in Relative Humidity* and *Variation in Wind Speed* were used as the input parameters.

The results of the soil temperature prediction are shown in Figure 2 for both initial Models, including the temperature at three depths (3-inch, 14-inch, 72-inch). These three depths were selected to be presented because they represent the top surface (largest diurnal fluctuations), intermediate depth, and deep depth (least fluctuations). The timespan of the temperature prediction shown in Figure 2 is for 4 months of data from January 2019 to April 2019, which represents the second part of winter. This time span is used because January to December 2018 was used as training data, thus the remaining 4 months was used as testing data. This divides the dataset in a 75%-25% split where 75% of the data was used for training and rest 25% was used for testing. The Models are labeled as "initial" models in Figure 2 because further adjustments were made to the modeling method to arrive at the final Models for this work.

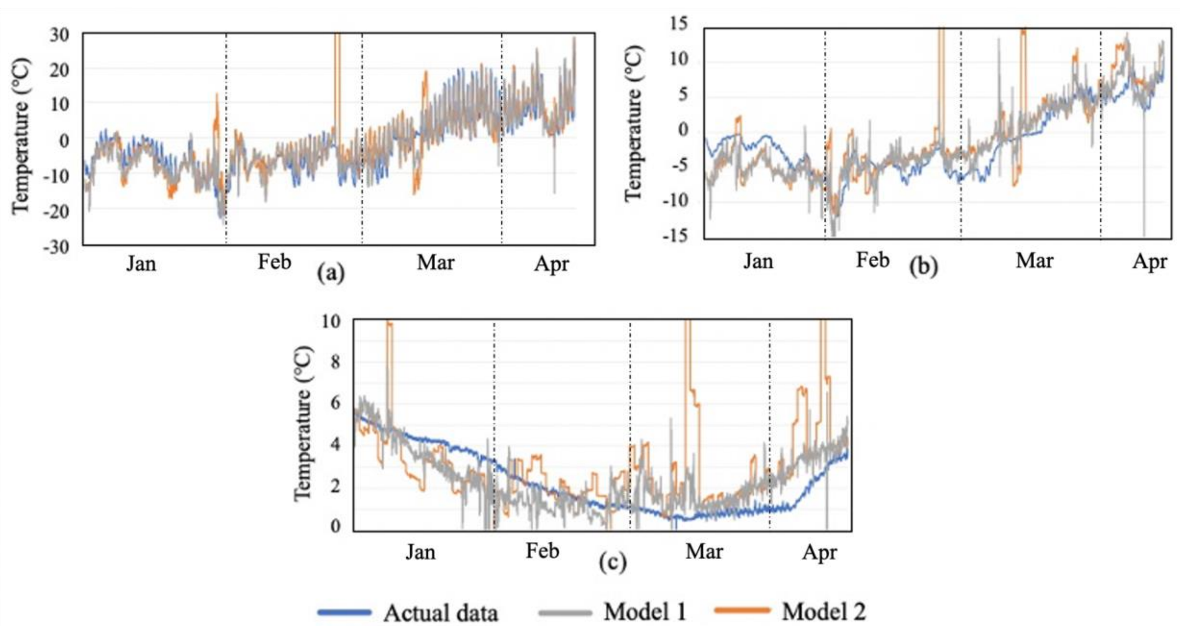


Figure 2. Comparison between the measured and predicted temperature for depths of (a) 3 inch, (b) 14 inch and (c) 72 inch below the surface for the initial Model 1 and Model 2 for the timeframe of January to April

As shown, the temperature fluctuations at the top surface are greater, whereas this fluctuation decreases with depth, as expected. Both the Models can predict the temperature at the top surface quite well. However, the performance of these initial Models reduces at deeper depths. At a depth of 14-inch, Model 2 predicts the overall trend of the temperature for most of the timespan. However, several spikes occur where the predicted temperature deviates significantly compared to the measured temperature data. Model 1 did not generate any spikes in the temperature prediction. However, the predicted temperature deviates from the measured soil temperature more than Model 2.

To reduce the spikes in the model predictions, two different filters were used. The objective of the filters was to limit the allowable variation in the model results from timestep (*n-1*) to the next timestep (*n*). In addition, the filters were also designed to limit temperature predictions which were significantly higher or lower than the measured temperature bounds (i.e. the minimum and maximum soil temperatures observed for each depth throughout the dataset). To identify the temperature bounds, the range of soil temperatures (maximum and minimum) expected for each depth across the dataset was assessed. The RMSE (root mean squared error) values for both the initial Models and the models which include the developed filters for all six locations at the MNROAD test facility and several depths are shown in Figure 3. A smaller RMSE values equates to better model performance over the timespan of evaluation.

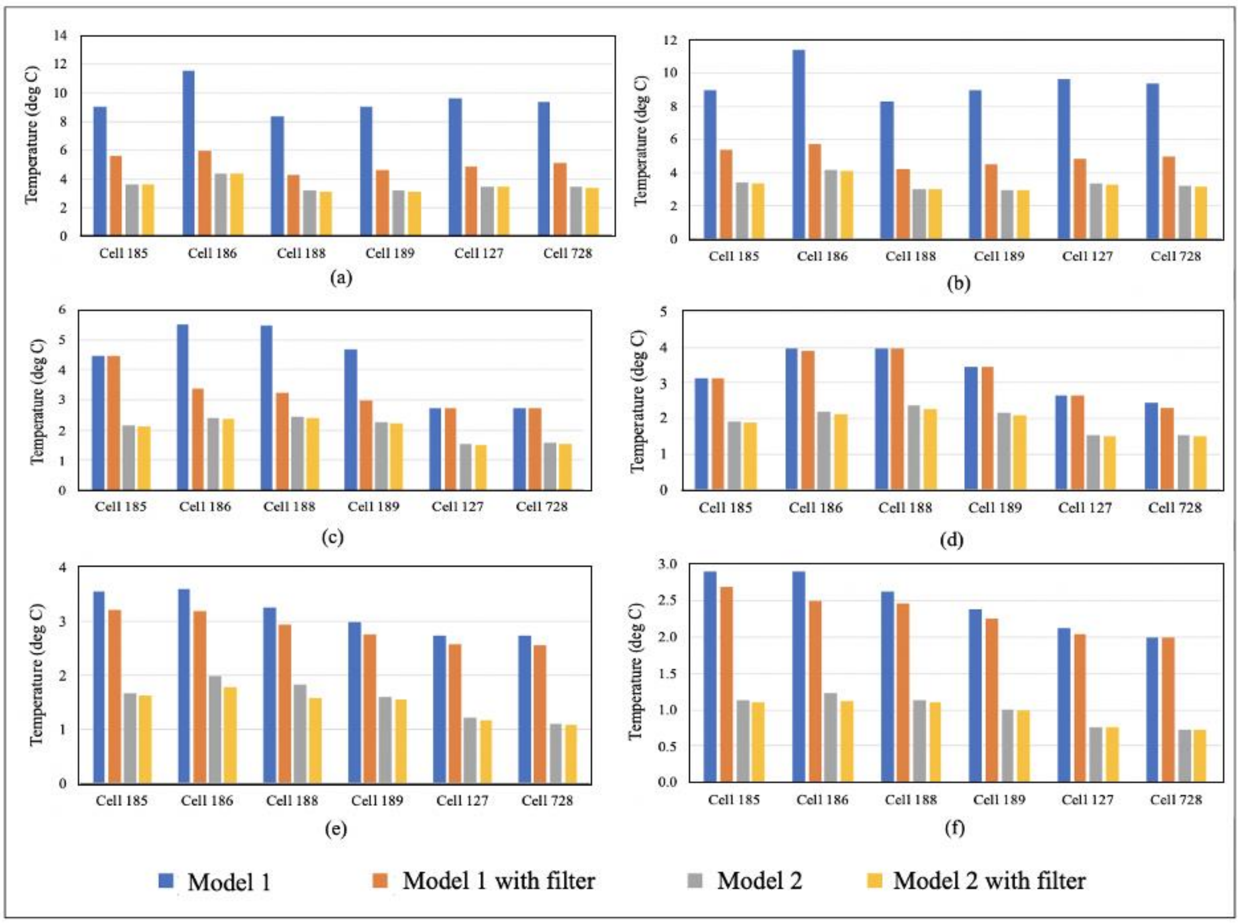


Figure 3. Comparison of RMSE values for initial Model 1 and 2, and Model 1 and 2 with filters applied, at depths of (a) 3 inches, (b) 4 inches, (c) 18.5 inches, (d) 24 inches, (e) 48 inches and (f) 72 inches for all cell locations for Dataset 1

The results show that, overall, across the models, the RMSE values for predicting the soil temperature decrease at deeper depths. In addition, in comparing the models, Model 1 generally has larger error than Model 2. The use of the filters helps to reduce this model error. For 3-inch and 4-inch depths, RMSE values for Model 1 without the filters was approximately 8 to 12°C. After implementing both filters, the RMSE values were reduced to half, or approximately 4 to 6 °C. Model 2 has a significantly smaller RMSE values of 2 to 4°C with the use of filtering. For intermediate and deeper depths, Model 2 with the filters had the smallest RMSE values among the 4 model variations evaluated for all the depths in all locations. For the deepest depths, the smallest RMSE values were approximately 1 to 2°C for 18.5- and 48-inch depths, and less than 1°C for the 72-inch depth. These results demonstrate that Model 2 with filters provides the best prediction of temperature among the modeling methods implemented.

A similar study was conducted using Dataset 2 which was the longer timescale data collected throughout Minnesota with a 1-hour timestep. The performance of both initial Model 1 and 2, and Model 1 and 2 with filters were evaluated for a time period starting from January 2000 to February 2007. This is substantially longer time period than the Dataset 1. The performance of both Model 1, Model 2 and the two models with filters are shown in Figure 4. As shown in the Figure, the RMSE values decreased at deeper depths as obtained for Dataset 1. However, the RMSE values were higher for this dataset compared to the model performance with previous set of data. The reason for these higher RMSE may be because the weather data used for model prediction was not collected on site and instead was obtained from the closest available weather station.

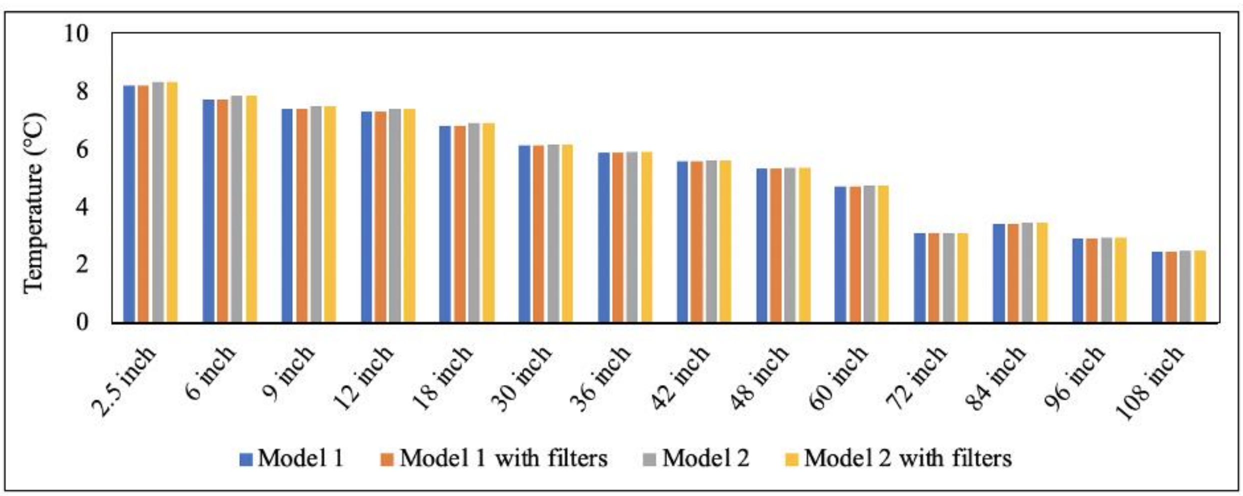


Figure 4: Comparison of RMSE values for Dataset 2

Based on the resulting performance of the models considered, the Model 2 with filters performed best in predicted temperatures at the various depths.

3.2. Prediction of number of freeze-thaw cycles: Based on substantial discussions with the project Technical Advisory Panel, a method was finalized to calculate the number of freeze-thaw cycles. To do this, 0°C was used as threshold temperature for melting and -1°C was used as the threshold temperature for freezing. Along with the temperature, a time delay is also considered in the calculation, where a 24 hour time delay is required after the -1°C threshold is passed, to ensure complete freezing. For thawing, a minimum 5 consecutive hours above 0°C was required.

Based on the above-mentioned method, the number of freeze-thaw cycles was evaluated from Dataset 1 using the four considered modeling methods, and compared with the number of cycles obtained from measured soil temperature (Figure 5). This figure also depicts the freeze-thaw cycle comparison for 4 months in the beginning of 2019 which represents the end of the winter. As discussed above, the reason this time period was used for this dataset is because for training one whole year of data was utilized, and the rest of the data was used as testing dataset.

As shown in Figure 5, the number of freeze-thaw cycles obtained from the Model 1 is similar to the number of freeze-thaw cycles obtained from the measured data. At deeper depths, the number of freeze-thaw cycles decreased, as expected. Model 2 generally slightly overpredicts the number of cycles for most soil depths and locations, as compared to Model 1. The number of freeze-thaw cycles was also calculated for the Dataset 2, as shown in Figure 6.

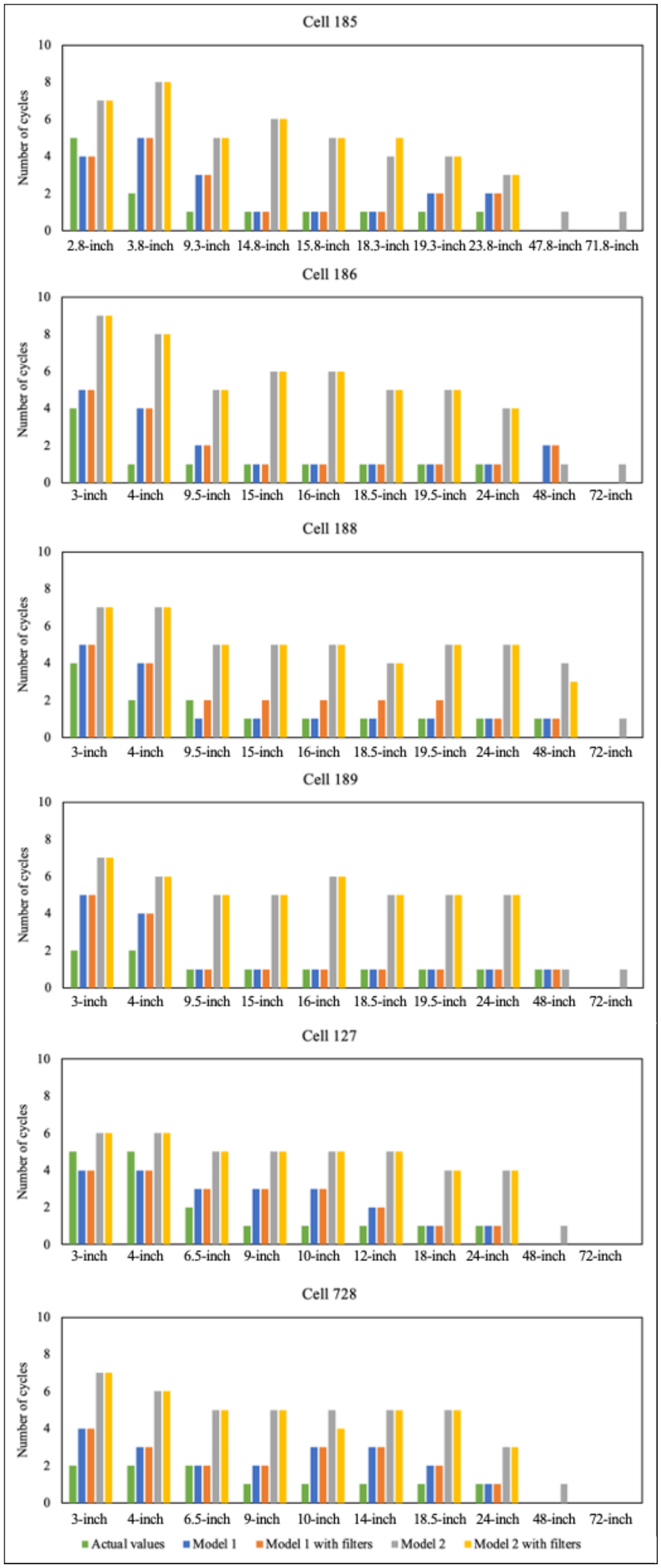


Figure 5. Comparison of the freeze-thaw cycle variations for the four Models

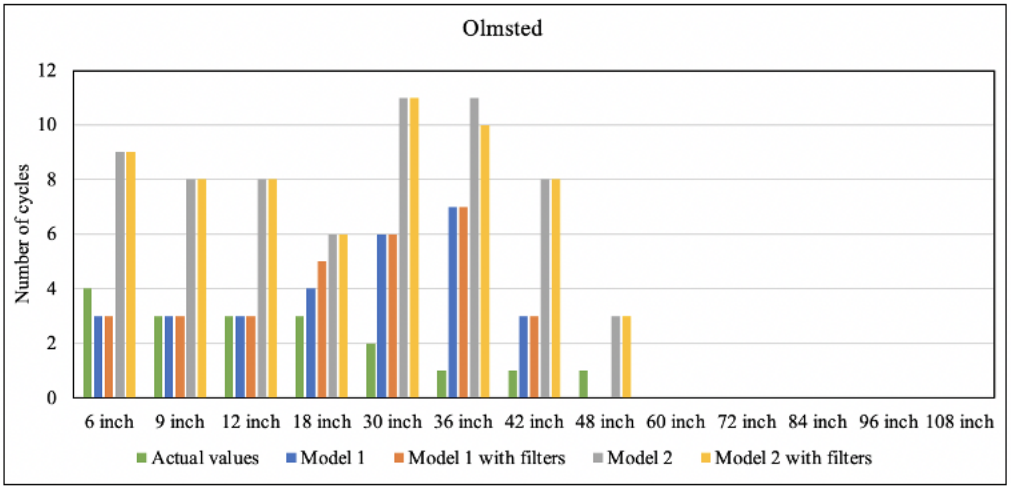


Figure 6. Number of freeze-thaw cycle comparison for Dataset 2

As shown in Figure 6, apart from the 30-inch and 36-inch depths, the number of freeze-thaw cycle predicted by the Model 1 is similar to the actual number of freeze-thaw cycles. However, for the two mentioned depths, Model 1 and 2 overpredict the number of cycles. One possible reason for this is that, as mentioned in the previous section, the location of soil temperature collection and location of weather data was not same, which may impact the level of error in the input data. This result was consistent with the results obtained with the Dataset 1. Thus, from these two sections, we can conclude that the Model 1, is better suited to evaluate the number of freeze-thaw cycles effectively.

3.3. Isotherm calculation: 0°C isotherm distribution for different depths and different time periods were calculated from the measured data. The results of cell 185 are shown in Figure 7. Rest of the figures are shown in Appendix section. The 0°C isotherm curve reaches its deepest depth during the months of December to February and then the depth of the isotherm curve reduces. A similar trend can be seen for all cell locations.

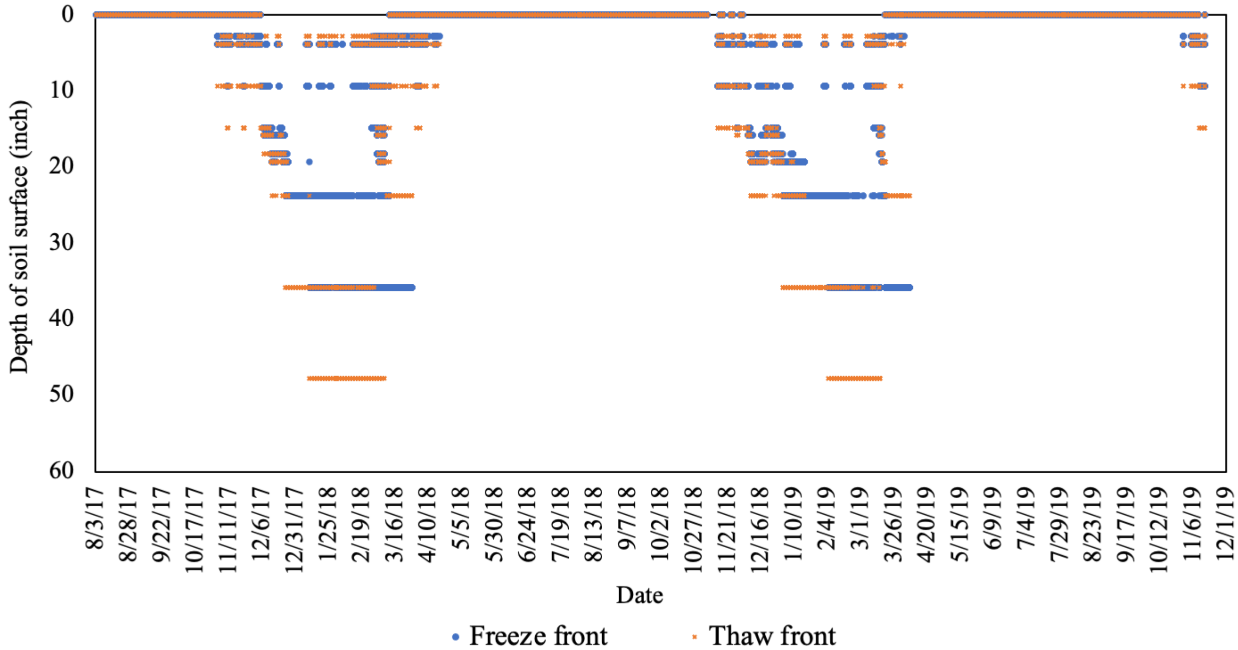


Figure 7. 0°C isotherm for cell 185 in Dataset 1

3.4. Time duration of freezing and thawing phase and their occurrence: The duration of the frozen period including the start and end dates, was calculated for each soil depth along with the freezing and thawing periods for each test location. Using Dataset 1, these were evaluated for approximately 18 months of data, from September 2017 to April 2019 and shown in Table 2. It was obtained that for some years, for shallower depths, such as 6.5 inch in both Cell 189 and 127, the soil froze intermittently, in some cases for as little as 1-day periods. As expected, for deeper depths the temperatures fluctuate less thus there are not short freeze periods at these depths.

The starting and ending time of freezing phase and their duration was compared for the predicted soil temperatures with the actual values. The performance of both the Model 1 and Model 2 were evaluated, as shown in Table 3. Model 1 performed significantly better compared to the Model 2. The performance of the Model 1 and its comparison with respect to the actual values is shown in Table 3. Note the freeze and thaw periods calculated for these cells assume the use of the shallowest soil depth that is not within the pavement foundation layers (i.e. base, subbase and subgrade layer) (e.g. cell 189 uses a soil depth of 6.5 in to calculate when the freezing and thawing period starts and ends).

Table 2. Starting and ending day of frozen soil surface with the duration for the location of (a) cell 185, (b) cell 186, (c) cell 188, (d) cell 189, (e) cell 127, (f) cell 728

Soil depth (inch)	Number of freeze-thaw cycles	Freezing start date	Freezing end date	Freeze duration (days)	Freezing period	Thawing period
9.5	2	12/7/17	2/27/18	82	Year 2017-2018: Dec 7 to Feb 27	Year 2017-2018: Feb 27 to Mar 20
		12/7/18	3/11/19	94		
14.8	2	12/25/17	3/3/18	68		
		1/2/19	3/15/19	72		
15.8	2	12/26/17	3/3/18	67		
		1/3/19	3/16/19	72		
18.3	2	12/26/17	3/4/18	68		
		1/19/19	3/18/19	58		
19.3	2	12/27/17	3/4/18	67	Year 2018-2019: Dec 7 to Mar 11	Year 2018-2019: Mar 11 to Mar 20
		1/20/19	3/18/19	57		
23.8	2	12/28/17	3/10/18	72		
		1/22/19	3/20/19	57		
47.8	0	-	-	-		
71.8	0	-	-	-		

(a)

Soil depth (inch)	Number of freeze-thaw cycles	Freezing start date	Freezing end date	Freeze duration (days)	Freezing period	Thawing period
9.5	3	12/13/17	3/1/18	78	Year 2017-2018: Dec 13 to Mar 1	Year 2017-2018: Mar 1 to Mar 12
		12/7/18	12/22/18	15		
		12/30/18	3/14/19	74		
15	2	12/26/17	3/4/18	68		
		1/1/19	3/17/19	74		
16	2	12/26/17	3/4/18	68		
		1/3/19	3/17/19	73		
18.5	2	12/27/17	3/7/18	70		
		1/21/19	3/19/19	57	Year 2018-2019: Dec 7 to Mar 14	Year 2018-2019: Mar 14 to Mar 21
19.5	2	12/28/17	3/7/18	69		
		1/21/19	3/19/19	57		
24	2	12/31/17	3/12/19	71		
		1/26/19	3/21/19	54		
48	0	-	-	-		
72	0	-	-	-		

(b)

Soil depth (inch)	Number of freeze-thaw cycles	Freezing start date	Freezing end date	Freeze duration (days)	Freezing period	Thawing period
9.5	2	12/13/17	3/1/18	78	Year 2017-2018: Dec 13 to Mar 1	Year 2017-2018: Mar 1 to Mar 15
		12/8/18	3/15/19	97		
15	2	12/26/17	3/5/18	69		
		1/2/19	3/18/19	75		
16	2	12/26/17	3/9/19	73		
		1/2/19	3/19/19	76		
18.5	2	12/27/17	3/11/18	74		
		1/3/19	3/20/19	76		
19.5	2	12/27/17	3/11/18	74	Year 2018-2019: Dec 8 to Mar 15	Year 2018-2019: Mar 15 to Apr 1
		1/3/19	3/20/19	76		
24	2	12/29/17	3/15/18	76		
		1/21/19	3/22/19	60		
48	1	3/6/19	4/1/19	26		
72	0	-	-	-		

(c)

Soil depth (inch)	Number of freeze-thaw cycles	Freezing start date	Freezing end date	Freeze duration (days)	Freezing period	Thawing period
6.5	3	11/10/17	11/11/17	1	Year 2017-2018: Nov 10 to Feb 27	Year 2017-2018: Feb 27 to Mar 21
		12/6/17	2/27/18	83		
		11/27/18	3/13/19	106		
9	2	12/7/17	2/28/18	83		
		11/28/18	3/14/19	106		
10	2	12/7/17	2/28/18	83		
		11/28/18	3/14/19	106		
12	3	12/26/17	3/4/18	68		
		12/10/18	12/19/18	9		
		1/2/19	3/15/19	73	Year 2018-2019: Nov 27 to Mar 13	Year 2018-2019: Mar 13 to Mar 22
18	2	12/31/17	3/17/18	77		
		1/22/19	3/22/19	60		
24	2	1/1/18	3/21/18	80		
		1/25/19	3/22/19	57		
48	0	-	-	-		
72	0	-	-	-		

(d)

Soil depth (inch)	Number of freeze-thaw cycles	Freezing start date	Freezing end date	Freeze duration (days)	Freezing period	Thawing period
6.5	3	11/10/17	11/11/17	1	Year 2017-2018: Nov 10 to Feb 27	Year 2017-2018: Feb 27 to Mar 21
		12/6/17	2/27/18	83		
		11/27/18	3/13/19	106		
9	2	12/7/17	2/28/18	83		
		11/28/18	3/14/19	106		
10	2	12/7/17	2/28/18	83		
		11/28/18	3/14/19	106		
12	3	12/26/17	3/4/18	68		
		12/10/18	12/19/18	9		
		1/2/19	3/15/19	73	Year 2018-2019: Nov 27 to Mar 13	Year 2018-2019: Mar 13 to Mar 22
18	2	12/31/17	3/17/18	77		
		1/22/19	3/22/19	60		
24	2	1/1/18	3/21/18	80		
		1/25/19	3/22/19	57		
48	0	-	-	-		
72	0	-	-	-		

(e)

Soil depth (inch)	Number of freeze-thaw cycles	Freezing start date	Freezing end date	Freeze duration (days)	Freezing period	Thawing period
6.5	2	12/7/17	3/3/18	86	Year 2017-2018: Dec 7 to Mar 3	Year 2017-2018: Mar 3 to Mar 21
		11/27/18	3/15/19	108		
9	3	12/8/17	3/4/18	86		
		11/28/18	12/1/18	3		
		12/7/18	3/15/19	98		
10	3	12/8/17	3/4/18	86		
		12/8/18	12/21/18	13		
		12/25/18	3/15/19	80		
14	2	12/27/17	3/17/18	80		
		1/20/19	3/19/19	58	Year 2018-2019: Nov 27 to Mar 15	Year 2018-2019: Mar 15 to Mar 23
18.5	2	12/27/17	3/17/18	80		
		1/20/19	3/21/19	60		
24	2	1/2/18	3/21/18	78		
		1/27/19	3/23/19	55		
48	0	-	-	-		
72	0	-	-	-		

(f)

Table 3. Comparison of freezing time for actual and predicted soil temperatures for Dataset 1 for the location of (a) cell 185, (b) cell 186, (c) cell 188, (d) cell 189, (e) cell 127, (f) cell 728

Cell 185

Depths	Method	Number of cycles	Freezing start day	Freezing end day	Freezing duration	Total Freezing duration
9.5 inch	Actual value	1	Jan-02	Mar-08	65	65
	Method 1	3	Jan-02	Jan-07	5	60
			Jan-09	Feb-23	45	
			Feb-26	Mar-08	10	
14.8 inch	Actual value	1	Jan-02	Mar-15	72	72
	Method 1	1	Jan-02	Mar-09	66	66
15.8 inch	Actual value	1	Jan-03	Mar-16	72	72
	Method 1	1	Jan-02	Mar-09	66	66
18.3 inch	Actual value	1	Jan-19	Mar-18	58	58
	Method 1	1	Jan-02	Mar-03	60	60
19.3 inch	Actual value	1	Jan-20	Mar-18	57	57
	Method 1	2	Jan-02	Jan-07	5	57
			Jan-10	Mar-03	52	
23.8 inch	Actual value	1	Jan-22	Mar-20	57	57
	Method 1	2	Jan-02	Jan-07	5	57
			Jan-10	Mar-03	52	

(a)

Cell 186

Depths	Method	Number of cycles	Freezing start day	Freezing end day	Freezing duration	Total Freezing duration
9.5 inch	Actual value	1	Jan-02	Mar-09	66	66
	Method 1	2	Jan-02	Feb-23	52	63
			Feb-25	Mar-08	11	
15 inch	Actual value	1	Jan-02	Mar-19	76	76
	Method 1	1	Jan-02	Mar-09	66	66
16 inch	Actual value	1	Jan-03	Mar-20	76	76
	Method 1	1	Jan-02	Mar-09	66	66
18.5 inch	Actual value	1	Jan-21	Mar-22	60	60
	Method 1	1	Jan-02	Mar-09	66	66
19.5 inch	Actual value	1	Jan-21	Mar-22	60	60
	Method 1	1	Jan-02	Mar-09	66	66
24 inch	Actual value	1	Jan-26	Mar-25	58	58
	Method 1	1	Jan-10	Mar-09	58	58

(b)

Cell 188

Depths	Method	Number of cycles	Freezing start day	Freezing end day	Freezing duration	Total Freezing duration
9.5 inch	Actual value	2	Jan-02	Jan-07	5	63
			Jan-09	Mar-08	58	
	Method 1	1	Jan-02	Mar-08	65	65
15 inch	Actual value	1	Jan-02	Mar-18	75	75
	Method 1	1	Jan-02	Mar-09	66	66
16 inch	Actual value	1	Jan-02	Mar-18	75	75
	Method 1	1	Jan-02	Mar-09	66	66
18.5 inch	Actual value	1	Jan-02	Mar-19	76	76
	Method 1	1	Jan-02	Mar-09	66	66
19.5 inch	Actual value	1	Jan-02	Mar-19	76	76
	Method 1	1	Jan-02	Mar-09	66	66
24 inch	Actual value	1	Jan-03	Mar-20	76	76
	Method 1	1	Jan-02	Mar-09	66	66
48 inch	Actual value	1	Feb-16	Mar-29	41	41
	Method 1	1	Jan-30	Mar-02	31	31

(c)

Cell 189

Depths	Method	Number of cycles	Freezing start day	Freezing end day	Freezing duration	Total Freezing duration
6.5 inch	Actual value	1	Jan-02	Mar-09	66	66
	Method 1	1	Jan-02	Mar-08	65	65
9 inch	Actual value	1	Jan-02	Mar-18	75	75
	Method 1	1	Jan-02	Mar-09	66	66
10 inch	Actual value	1	Jan-02	Mar-19	76	76
	Method 1	1	Jan-02	Mar-09	66	66
12 inch	Actual value	1	Jan-03	Mar-20	76	76
	Method 1	1	Jan-02	Mar-09	66	66
18 inch	Actual value	1	Jan-03	Mar-20	76	76
	Method 1	1	Jan-02	Mar-09	66	66
24 inch	Actual value	1	Jan-21	Mar-22	60	60
	Method 1	1	Jan-02	Mar-09	66	66
48 inch	Actual value	1	Mar-06	Apr-01	26	26
	Method 1	1	Jan-30	Feb-14	15	15

(d)

Cell 127

Depths	Method	Number of cycles	Freezing start day	Freezing end day	Freezing duration	Total Freezing duration
6.5 inch	Actual value	2	Jan-02	Jan-07	5	63
			Jan-09	Mar-08	58	
	Method 1	3	Jan-02	Jan-07	5	65
			Jan-09	Mar-08	58	
			Mar-13	Mar-15	2	
9 inch	Actual value	1	Jan-02	Mar-14	71	71
	Method 1	3	Jan-02	Jan-07	5	65
			Jan-09	Mar-08	58	
			Mar-13	Mar-15	2	
10 inch	Actual value	1	Jan-02	Mar-14	71	71
	Method 1	3	Jan-02	Jan-07	5	65
			Jan-09	Mar-08	58	
			Mar-13	Mar-15	2	
12 inch	Actual value	1	Jan-02	Mar-16	73	73
	Method 1	2	Jan-02	Mar-09	66	70
			Mar-11	Mar-15	4	
18 inch	Actual value	1	Jan-22	Mar-23	60	60
	Method 1	1	Jan-10	Mar-10	59	59
24 inch	Actual value	1	Jan-25	Mar-23	57	57
	Method 1	1	Jan-13	Mar-10	56	56

(e)

Cell 728

Depths	Method	Number of cycles	Freezing start day	Freezing end day	Freezing duration	Total Freezing duration
6.5 inch	Actual value	2	Jan-02	Jan-07	5	68
			Jan-09	Mar-13	63	
	Method 1	2	Jan-02	Jan-07	5	70
			Jan-09	Mar-15	65	
9 inch	Actual value	1	Jan-02	Mar-15	72	72
	Method 1	2	Jan-02	Jan-07	5	70
			Jan-09	Mar-15	65	
10 inch	Actual value	1	Jan-02	Mar-15	72	72
	Method 1	3	Jan-02	Jan-07	5	68
			Jan-09	Mar-09	59	
			Mar-11	Mar-15	4	
14 inch	Actual value	1	Jan-20	Mar-19	58	58
	Method 1	3	Jan-02	Jan-06	4	68
			Jan-10	Mar-10	59	
			Mar-11	Mar-15	4	
18.5 inch	Actual value	1	Jan-20	Mar-21	60	60
	Method 1	2	Jan-02	Jan-05	3	62
			Jan-10	Mar-10	59	
24 inch	Actual value	1	Jan-27	Mar-23	55	55
	Method 1	1	Jan-17	Mar-10	52	52

(f)

As shown in Table 3, Model 1 can generally predict the freezing start and end day and the duration of the freezing phase for the Dataset 1. The freezing and thawing period for all cell locations are shown in Table 4 and compared their performance using Model 1. It can be seen from Table 4 that the Model 1 can predict the freezing period efficiently for most of the cell locations. However, it underpredicts the time for the thawing period.

Table 4. Comparison of freezing and thawing period for all cell locations using Dataset 1

Cell location	Freezing period		Thawing period	
	Actual value	Predicted	Actual value	Predicted
Cell 185	Jan 2 – Mar 8	Jan 2 – Mar 3	Mar 8 – Mar 20	Mar 3- Mar 9
Cell 186	Jan 2 – Mar 9	Jan 2 – Mar 8	Mar 9 – Mar 25	Mar 8- Mar 9
Cell 188	Jan 2 – Mar 8	Jan 2 – Mar 2	Mar 8 – Mar 29	Mar 2- Mar 9
Cell 189	Jan 2 – Mar 9	Jan 2 – Feb 14	Mar 9 – Apr 1	Feb 14- Mar 9
Cell 127	Jan 2 – Mar 8	Jan 2 – Mar 10	Mar 8 – Mar 23	Mar 8- Mar 15
Cell 728	Jan 2 – Mar 15	Jan 2 – Mar 10	Mar 15 – Mar 23	Mar 10- Mar 15

4. Conclusion

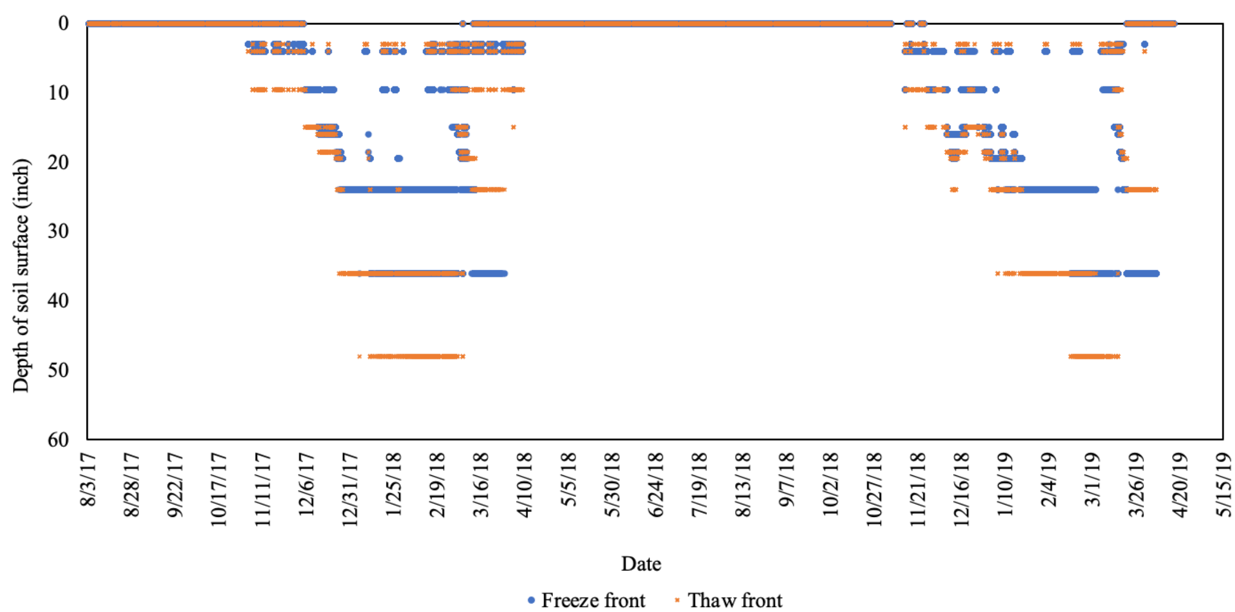
In this task, 4 different variations of data-driven model algorithms were developed and implemented to predict soil temperature and the number of freeze thaw cycles at different soil depths. Overall, Model 2 with filters performs better for predicting soil temperatures. Model 1 with filters performs better in predicting the number of freeze-thaw cycles. For calculating the freezing and thawing periods, both the models performed fairly well, but Model 1 performed better. Moving forward in future work, based on the availability of the soil temperature and climate variable data, the performance of these models can be evaluated in further detail.

5. References

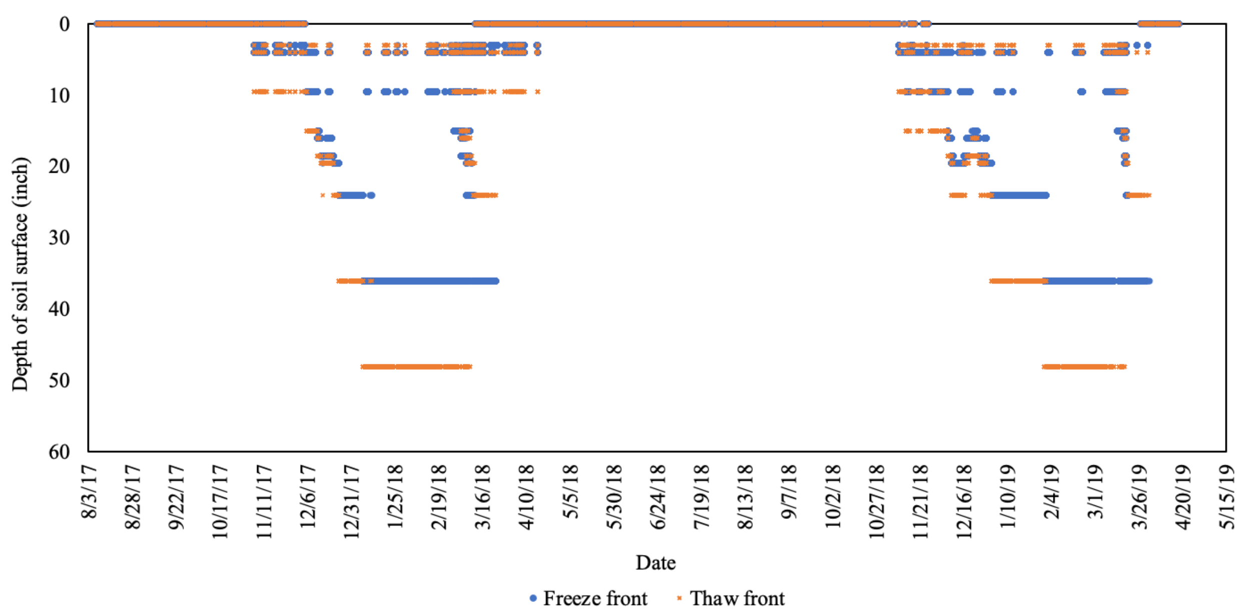
[1] Mihalakakou, G. "On estimating soil surface temperature profiles." Energy and Buildings 34, no. 3 (2002): 251-259.
[2] Tang, Wang, and Shangchang Ma. "Application of Regression and Artificial Neural Network in Ground Temperature Processing." In 2019 International Conference on Meteorology Observations (ICMO), pp. 1-4. IEEE, 2019.
[3] Talaee, P. Hosseinzadeh. "Daily soil temperature modeling using neuro-fuzzy approach." Theoretical and applied climatology 118, no. 3 (2014): 481-489.
[4] George, Raju K. "Prediction of soil temperature by using artificial neural networks algorithms." Nonlinear Analysis: Theory, Methods & Applications 47, no. 3 (2001): 1737-1748.
[5] Kisi, Ozgur, Mustafa Tombul, and Mohammad Zounemat Kermani. "Modeling soil temperatures at different depths by using three different neural computing techniques." Theoretical and applied climatology 121, no. 1-2 (2015): 377-387.
[6] Kim, Sungwon, and Vijay P. Singh. "Modeling daily soil temperature using data-driven models and spatial distribution." Theoretical and applied climatology 118, no. 3 (2014): 465-479.

[7] Bilgili, Mehmet. "Prediction of soil temperature using regression and artificial neural network models." Meteorology and atmospheric physics 110, no. 1-2 (2010): 59-70.
[8] Citakoglu, Hatice. "Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey." Theoretical and Applied Climatology 130, no. 1-2 (2017): 545-556.
[9] Xing, Lu, Liheng Li, Jiakang Gong, Chen Ren, Jiangyan Liu, and Huanxin Chen. "Daily soil temperatures predictions for various climates in United States using data-driven model." Energy 160 (2018): 430-440.

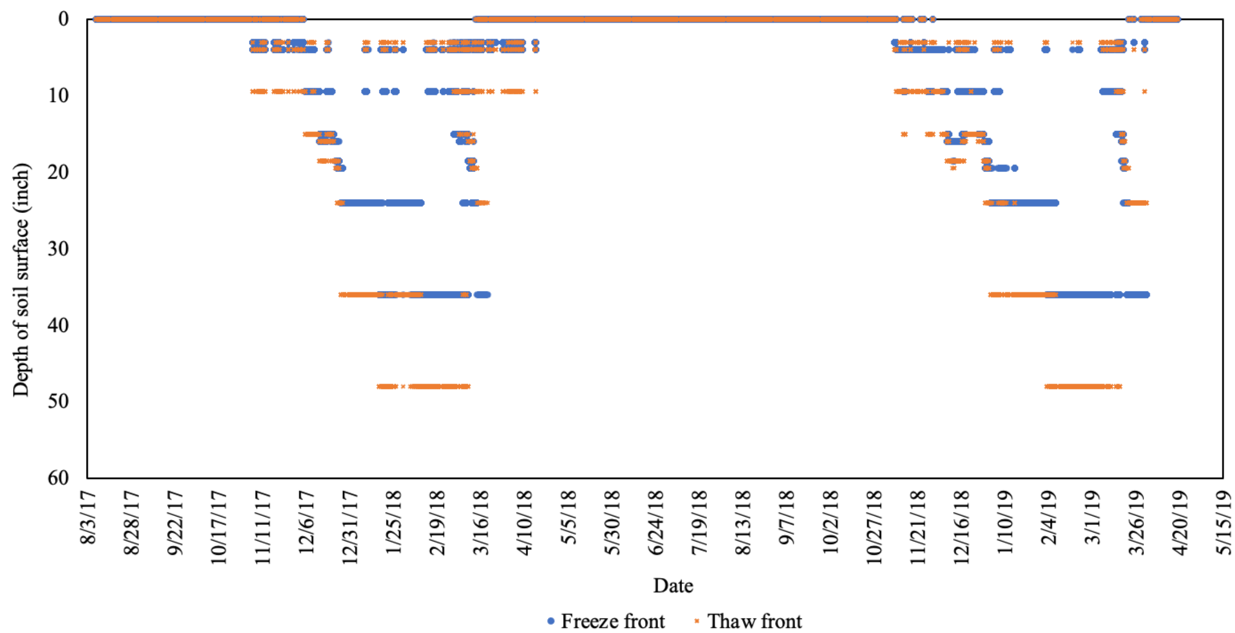
6. Appendix



(a)

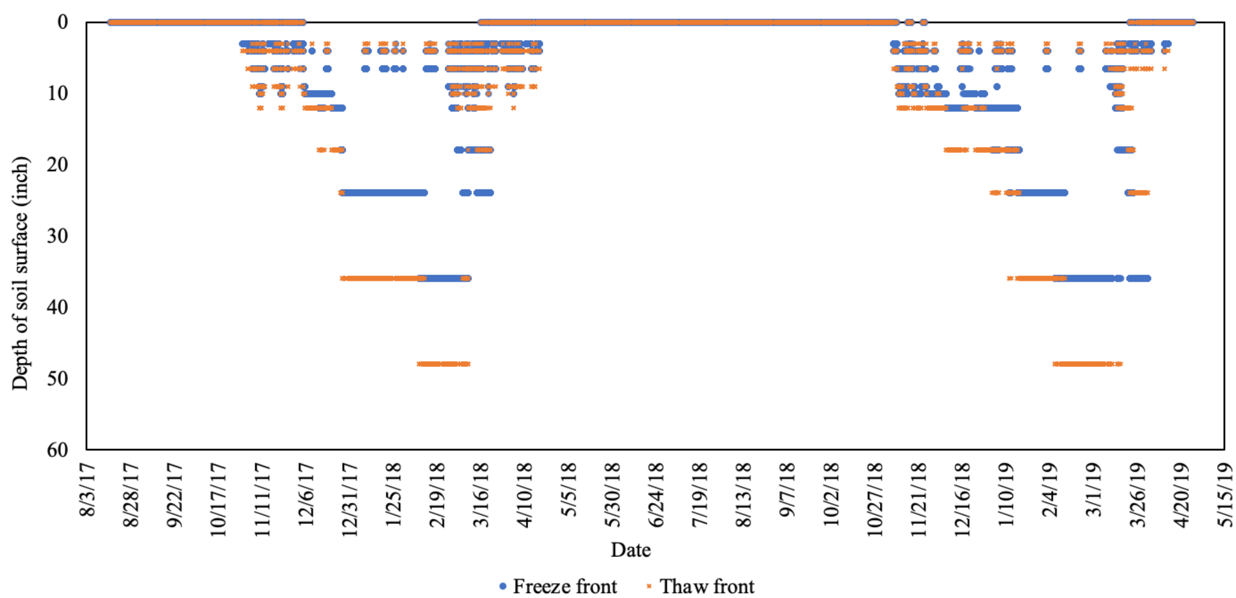


(b)



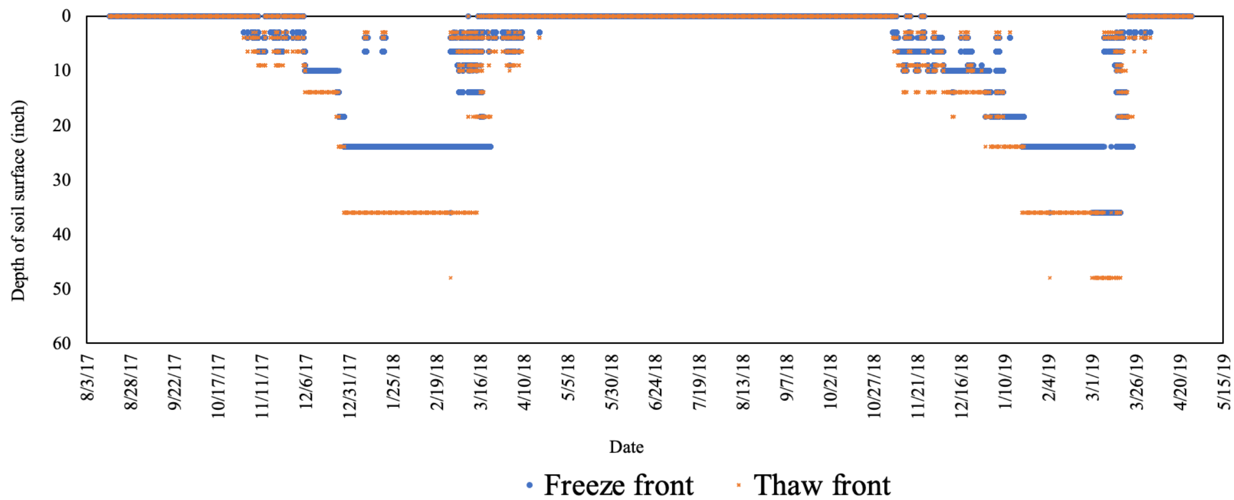
0
10
20
30
40
50
60
Depth of soil surface (inch)
8/3/17
8/28/17
9/22/17
10/17/17
11/11/17
12/6/17
12/31/17
1/25/18
2/19/18
3/16/18
4/10/18
5/5/18
5/30/18
6/24/18
7/19/18
8/13/18
9/7/18
10/2/18
10/27/18
11/21/18
12/16/18
1/10/19
2/4/19
3/1/19
3/26/19
4/20/19
5/15/19
Date
Freeze front
Thaw front

(c)



0
10
20
30
40
50
60
Depth of soil surface (inch)
8/3/17
8/28/17
9/22/17
10/17/17
11/11/17
12/6/17
12/31/17
1/25/18
2/19/18
3/16/18
4/10/18
5/5/18
5/30/18
6/24/18
7/19/18
8/13/18
9/7/18
10/2/18
10/27/18
11/21/18
12/16/18
1/10/19
2/4/19
3/1/19
3/26/19
4/20/19
5/15/19
Date
Freeze front
Thaw front

(d)



(e)

Figure. 0°C isotherm for test sites (a) cell 186, (b) cell 188, (c) cell 189, (d) cell 127 and (e) cell 728 in Dataset 1

Table A1. Significant coefficients used in regression equations for temperature prediction at different soil depths for Dataset 1, Cell 186.

Terms	3 inch	4 inch	9.5 inch	15 inch	16 inch	18.5 inch	19.5 inch	24 inch	48 inch	72 inch
Constant	-2.587	-2.418	-1.385	0.276	0.575	1.390	1.596	2.654	5.437	7.630
DayofYear	-0.204	-0.206	-0.233	-0.278	-0.285	-0.307	-0.311	-0.338	-0.362	-0.322
DayofYear^2	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.005
DayofYear^3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Timestep	-0.153	-0.133	-0.009	-0.018	-0.024	-0.040	-0.044	-0.058	-0.046	-0.027
DayofYear*Timestep	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Timestep^2	0.001	0.000	-0.004	-0.002	-0.002	-0.001	-0.001	0.000	0.000	0.000
Timestep*Rad*varAirTemp	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Rad	0.657	0.527	0.095	0.021	0.009	0.079	0.098	0.223	0.824	0.673
Timestep*Rad	-0.021	-0.012	0.012	0.018	0.019	0.018	0.018	0.016	0.008	0.008
Timestep^2*Rad	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Rad^2	0.261	0.241	0.110	0.005	0.000	-0.034	-0.038	-0.073	-0.208	-0.126
DayofYear*Rad^2	-0.005	-0.005	-0.004	-0.003	-0.003	-0.002	-0.002	-0.001	0.001	0.001
Timestep*Rad^2	0.001	0.000	-0.002	-0.003	-0.003	-0.003	-0.003	-0.002	-0.001	-0.001
Rad^3	-0.075	-0.065	-0.021	0.000	0.001	0.006	0.007	0.010	0.018	0.002
Timestep*Rad^3	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Rad^4	0.006	0.005	0.003	0.002	0.001	0.001	0.001	0.001	0.000	0.001
Rad^3* varAirTemp	-0.001	-0.001	-0.001	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000
DayofYear*varRH	0.019	-0.017	-0.126	-0.231	-0.245	-0.250	-0.253	-0.254	-0.172	-0.109
varRH	-15.700	-12.260	-3.544	5.319	6.798	7.507	7.992	9.426	10.670	8.896
Timestep*varRH	0.330	0.366	0.573	0.610	0.601	0.608	0.603	0.545	0.163	0.003
Timestep^2*varRH	0.001	0.001	-0.003	-0.005	-0.005	-0.005	-0.005	-0.005	-0.001	0.001
AirTemp*varRH	-0.316	-0.394	-0.413	-0.357	-0.350	-0.336	-0.329	-0.285	-0.098	-0.014
Timestep*AirTemp*varRH	0.010	0.011	0.010	0.007	0.007	0.006	0.006	0.005	0.002	0.001
Rad* varRH	-3.546	-4.806	-6.432	-6.166	-6.108	-5.785	-5.715	-5.378	-3.867	-3.096
Timestep*Rad* varRH	0.041	0.041	0.016	-0.020	-0.025	-0.036	-0.040	-0.052	-0.034	-0.007
Rad^2* varRH	-1.204	-1.003	-0.267	0.042	0.071	0.177	0.204	0.294	0.329	0.229
Timestep*Rad^2* varRH	-0.008	-0.008	-0.005	-0.002	-0.001	-0.001	0.000	0.000	0.000	-0.002
DayofYear*Rad^2* varRH	-0.001	-0.002	-0.003	-0.004	-0.004	-0.003	-0.003	-0.003	-0.002	-0.001
Rad^3* varRH	0.226	0.219	0.149	0.095	0.090	0.072	0.068	0.054	0.031	0.025
varAirTemp* varRH	-0.289	-0.470	-0.549	-0.424	-0.405	-0.378	-0.370	-0.324	-0.127	-0.063
Rad*varAirTemp* varRH	0.173	0.199	0.219	0.214	0.212	0.202	0.199	0.177	0.092	0.057
varRH^2	19.920	15.660	-5.840	-18.780	-20.540	-22.790	-23.420	-24.040	-21.470	-14.830
DayofYear*varRH^2	-0.124	-0.083	0.086	0.190	0.205	0.225	0.230	0.233	0.198	0.136
DayofYear^2*varRH^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Timestep*varRH^2	-0.140	-0.162	-0.165	-0.100	-0.093	-0.077	-0.070	-0.043	-0.007	0.004
Timestep^2*varRH^2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000	0.000
Rad*varRH^2	3.037	3.594	5.564	5.680	5.594	5.170	5.109	4.854	4.159	3.591
DayofYear*Rad*varRH^2	-0.023	-0.025	-0.026	-0.023	-0.023	-0.021	-0.020	-0.019	-0.014	-0.012
Timestep*Rad*varRH^2	0.037	0.039	0.026	0.021	0.021	0.020	0.020	0.020	0.016	0.011
AirTemp*varRH^2	-0.066	-0.059	0.016	0.055	0.057	0.063	0.064	0.071	0.070	0.067
Rad^2*varRH^2	0.309	0.303	0.219	0.149	0.141	0.127	0.124	0.107	0.023	-0.004
Timestep*Rad^2*varRH	-0.008	-0.008	-0.005	-0.002	-0.001	-0.001	0.000	0.000	0.000	-0.002
varRH^3	-2.375	-2.343	-1.462	-0.999	-0.937	-0.820	-0.804	-0.765	-0.539	-0.631
DayofYear*varRH^3	0.013	0.013	0.007	0.004	0.003	0.003	0.003	0.002	0.001	0.001
Timestep*varRH^3	-0.012	-0.012	-0.006	-0.004	-0.004	-0.003	-0.003	-0.003	-0.002	-0.001
Rad*varRH^3	-0.306	-0.308	-0.281	-0.251	-0.248	-0.237	-0.235	-0.229	-0.195	-0.156
varRH^4	0.031	0.034	0.044	0.048	0.048	0.048	0.048	0.047	0.039	0.031
varRH^2* varWind	-1.426	-1.407	-0.918	-0.682	-0.658	-0.613	-0.615	-0.586	-0.418	-0.375