Environmental Impacts on The Performance of Pavement Foundation Layers – Phase I

Principal Investigator:

Bora Cetin, Ph.D.

Co-Principal Investigator:

Tuncer Edil, Ph.D.

Kristen Cetin, Ph.D.

Research Team:

Debrudra Mitra

Department of Civil and Environmental Engineering

Michigan State University

May 20, 2020

PROBLEM STATEMENT

IMPACTS OF FREEZE-THAW CYCLES UNDER ROADS

- *Water in soil freezes and **expands***
- *During spring-thaw, melted water and infiltrated water trapped above the zone of frozen subgrade **– strength loss** under heavy loading*
- *Seasonal Load Restrictions – applied to **avoid/reduce damages***
- ***Prediction** of Freeze-Thaw Cycles – Monitoring systems & Computational Models*

INSTRUMENTATION

- Instrumented with an array of:
  - Soil Moisture
  - Temperature
- Weather Station to measure climate data
  - On site

OBJECTIVES

Develop a Data Driven Model to Predict the Frozen Soil Depths & Freeze-Thaw Durations

- **Inputs:**
  - Climate data (precipitation, relative humidity, percent sunshine, temperature, & wind speed)
  - Layer thicknesses
  - Material type
- **Output**
  - Number of freeze-thaw cycles at specific depths
  - Duration of freezing and thawing
  - Frost depth

4

Overview of Research Plan

- **Task 1** – Initial Memorandum on Expected Research Benefits and Potential Implementation Steps
- **Task 2** – Field Data Collection
- **Task 3** – Modelling Analyses
- **Task 4** – Final Report

TASK 2 – FIELD DATA COLLECTION

List of data that will be collected:

- Climate Data
  - Air temperature
  - Percent sunshine
  - Precipitation
  - Wind speed
  - Relative humidity
- Soil Data
  - Material data
  - Temperature
  - Water content

Task 3 – Modelling Analyses

Modeling Objectives:

Develop a **tool** that can be used to assess/predict the freeze-thaw behavior of roadways

- Simple
- Stand-alone
- For any location (where soil profile and weather data are available)

Output needed:

- number of freeze thaw cycles at certain depth
- frost depth isotherms over time

Modeling Approaches

Two types of modeling approaches to consider:

Physics-based modeling ("white box")

Data-driven modeling ("black box")

What is the appropriate approach to consider?

Different approaches towards modeling:

Physics-Based Modeling

based on physical principles and relationships between variables; described with a set of mathematical equations with variables that have physical meaning

Inputs: Many input (or assumptions) required; some may or may not be known

Pros: better at extrapolation, limited historical data required

Cons: significant knowledge of all physical properties and interactions; slower (higher computational intensity)

Data-Driven Modeling

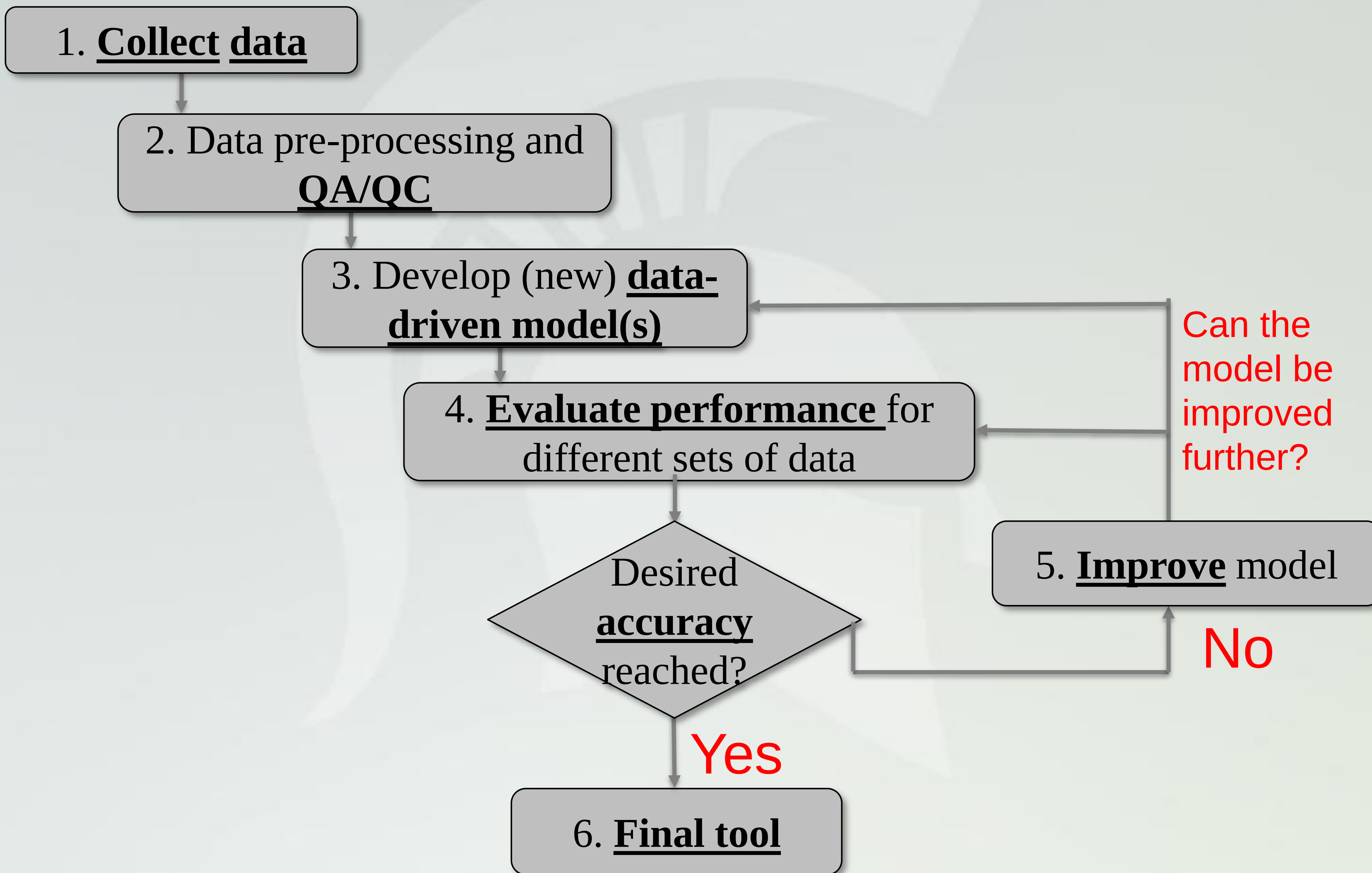
Statistical or machine learning based; uses historical data to develop a quantifiable relationship between inputs and outputs

Inputs: whatever data is available (*and ultimately found to be significant*)

Pros: lower computational intensity; no knowledge of physical properties or interactions required

Cons: worst (typically) at extrapolation outside of bounds of original data; needs larger training dataset to create and validate

Tool Development Process: *Workflow*



Step 1. Collect data: *Data Needs*

Most important requirements for data-driven modeling are:

- large(r) input datasets, which will be split into:
  - In-sample (to create the model)
  - out-of-sample (to validate the model)
- diversity of conditions (e.g. hot/cold, wet/dry, etc..)

Data needed (ideally):

- Weather data (close or near to site)
- Soil profiles/characteristics (thermal/moisture)
- Historical temperature at different depths
- A range of sites/locations of data collection

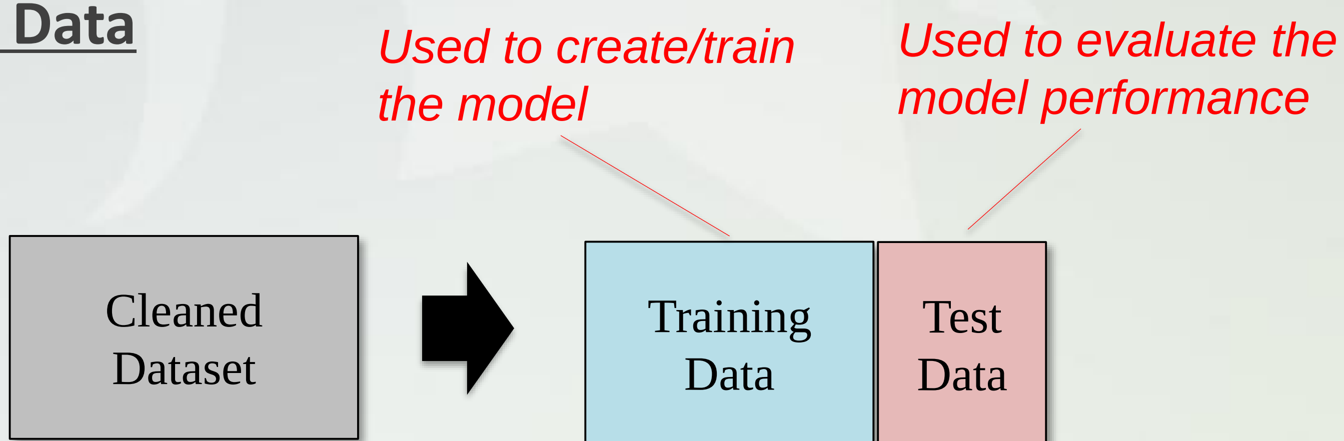
Step 2. Data Pre-Processing:

QA/QC: Types & Handling of Missing Data:

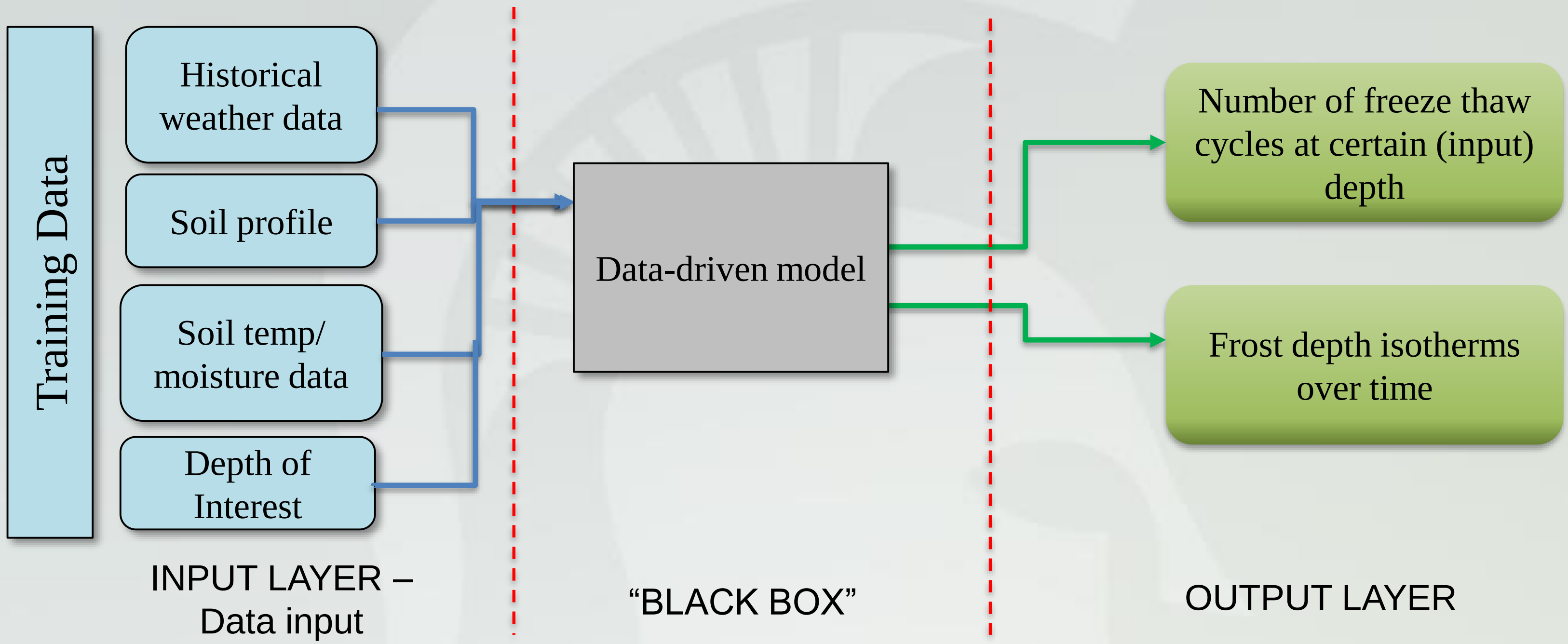
1) Short spans (less than 10 hrs)
   → Impute data (fill it in) based on trends in surrounding data
   → forward fill method

2) Long spans (more than 10 hrs) in this dataset
   → Remove the time periods with missing data

Division of Data

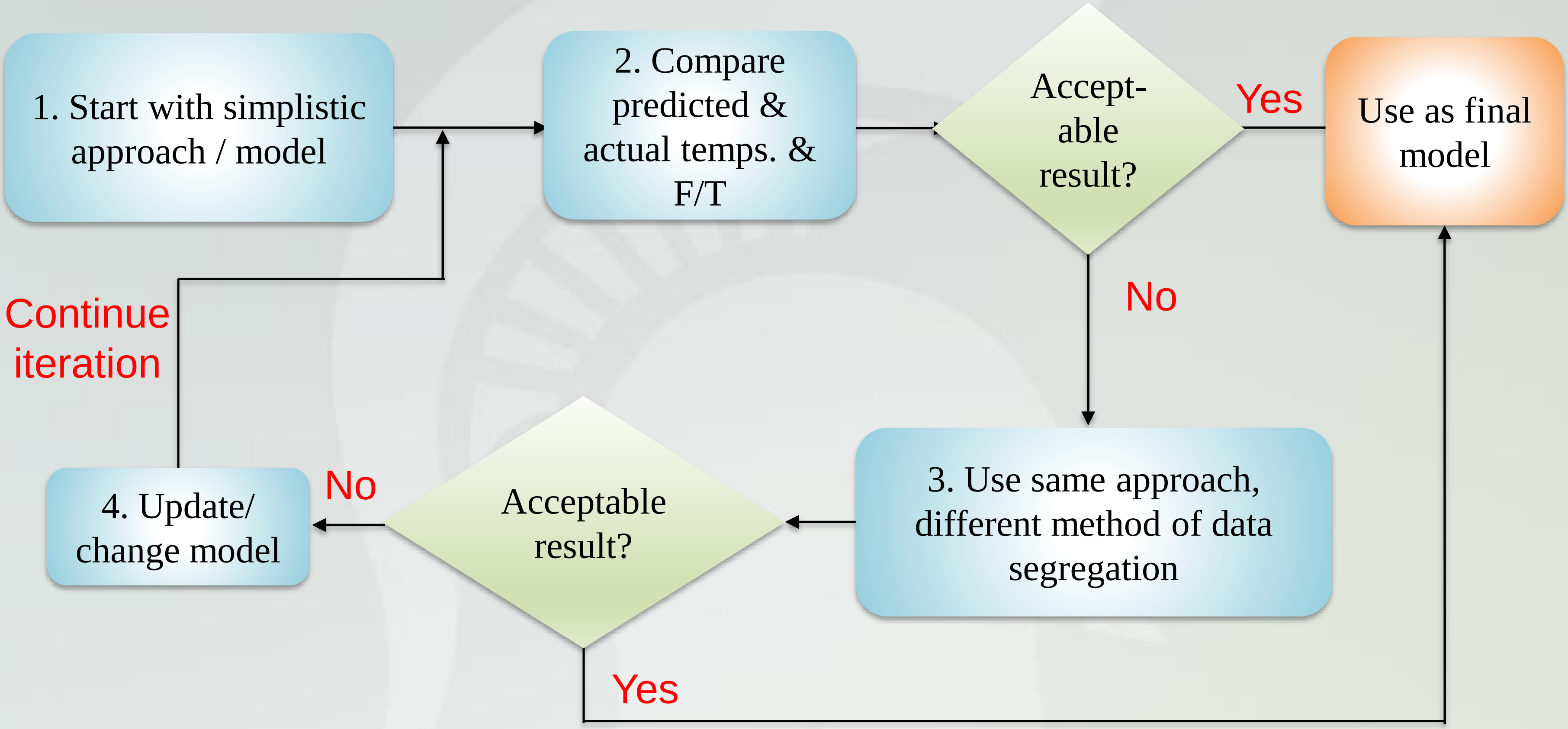


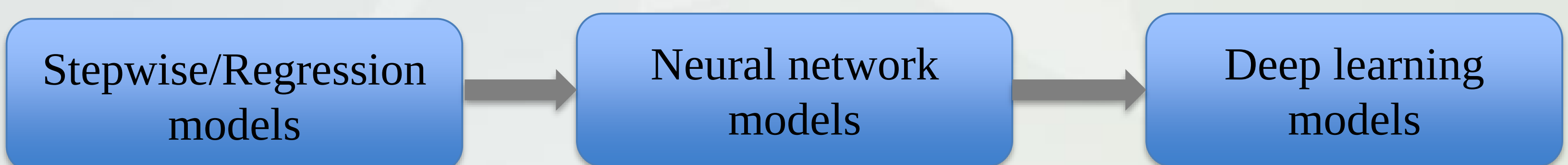
Step 3. Develop data-driven models: *Process*



Layout of model development process

Step 3-6. Refine Model: *Progressive Improvement*





<u>**Example**</u> ***(other models are considered)*** sequence from simple to complex modeling to determine relative improvement in performance

Model Selection: *(a) geotech literature review*

Previous literature on data-driven models: *Most to date have attempted to predict average daily or monthly soil temperatures, NOT hourly data, or freeze-thaw /isotherm information*

- Regression [2,5]
- Artificial Neural Networks [3-5]
- Neuro-fuzzy inference system (ANFIS) [1, 6]
- Multilayer perceptron (MLP) [6]
- Generalize regression, radial basis, and MLP neural network [7]
- Support Vector Machine (SVM) [8]

Model selection: *(b) general literature review*

Literature on modeling multi-variate time series data

Our approach: Simple → complex

Order of Evaluation / Presentation Discussion

- Regression
  - Linear & non-linear
  - Stepwise

(1,3)

- Vector autoregressive (VAR)
  - multivariate time series analysis
- Vector error correction model (VECM)
  - can be useful when there are cointegrated variables

(2)

- ANN, MLP, SVM, ANFIS (also in prev. slide)

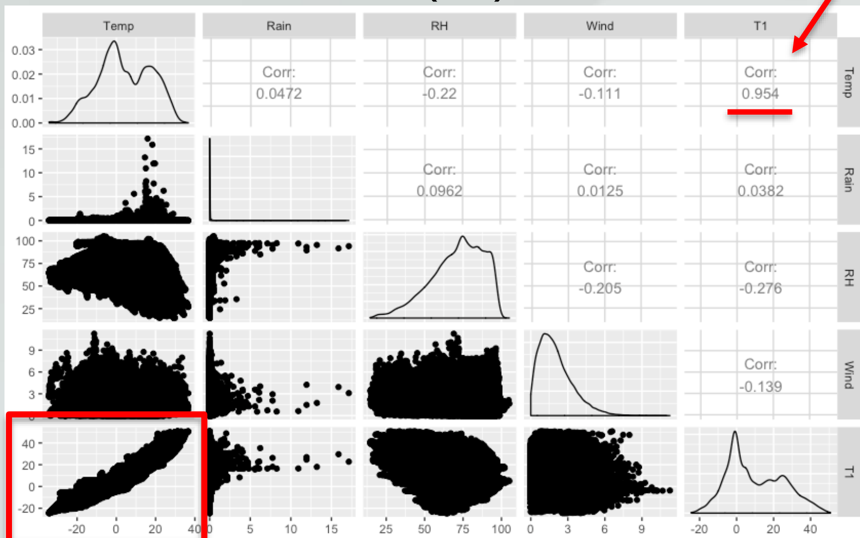
(4)

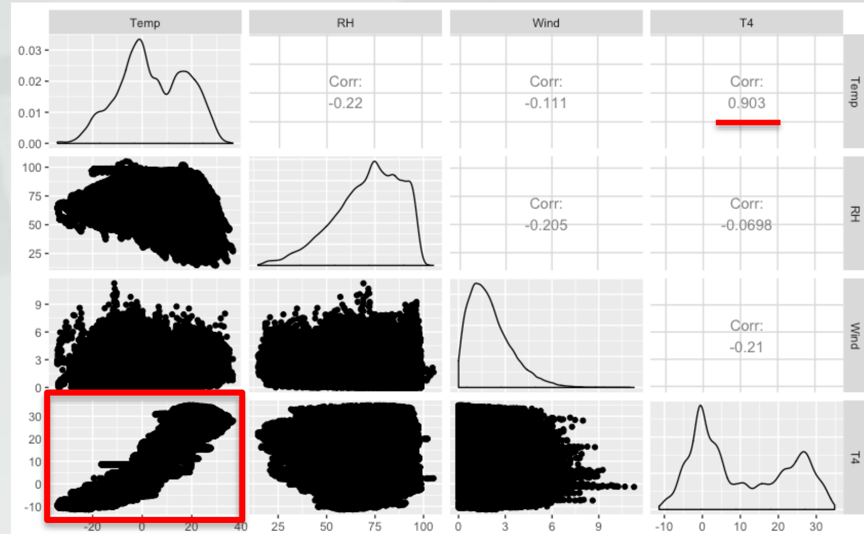
- Many others…

Soil temperature correlation with climate parameters

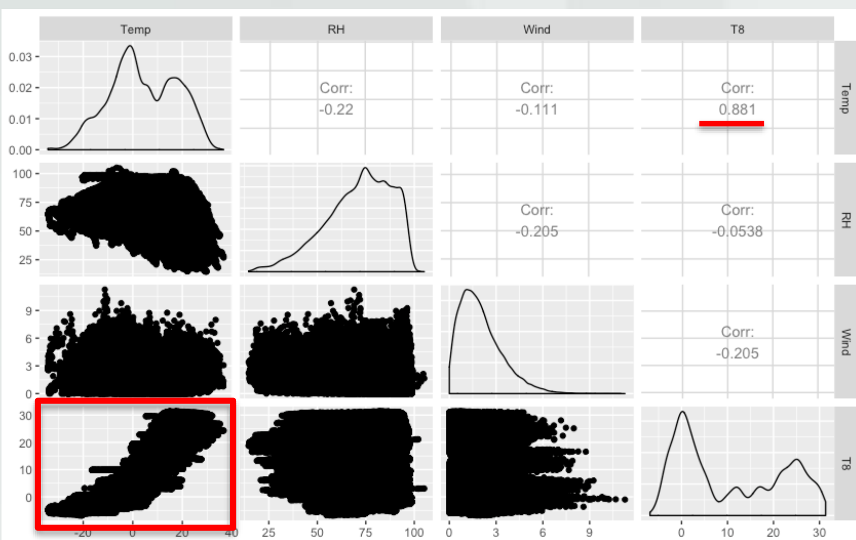
Closest to surface (T1)

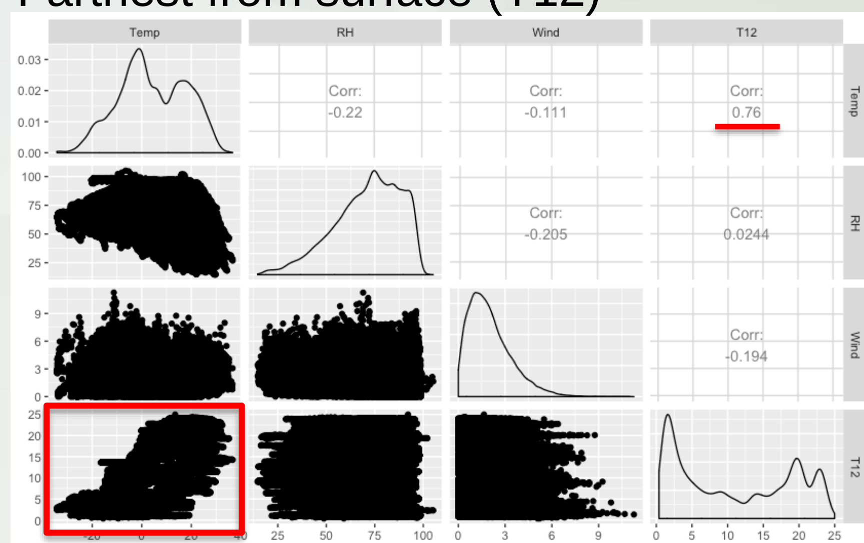
Temperature is strongest predictor



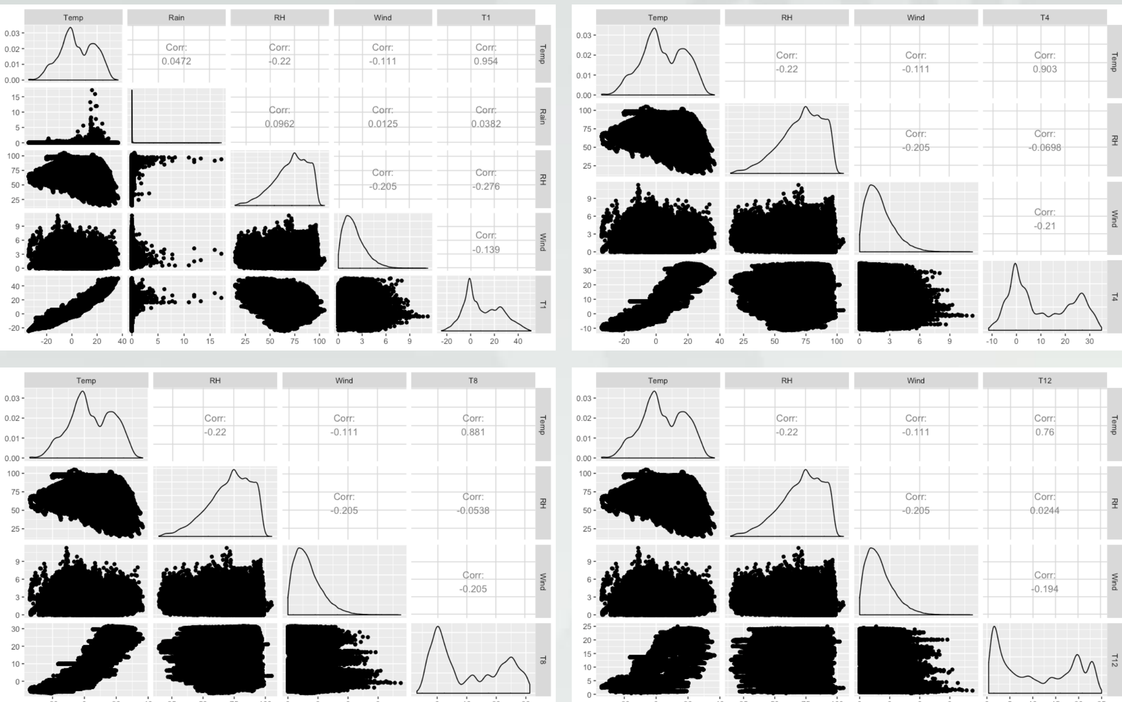


Farthest from surface (T12)





Soil temperature correlation with climate parameters



- Soil temperature is significantly corelated with <u>air temperature</u>
- Correlation coefficient reduces with the depth of soil
- Wind is negatively correlated with soil temperature
- RH is very weakly correlated with soil temperatures

(1) Regression Models: *Methods*

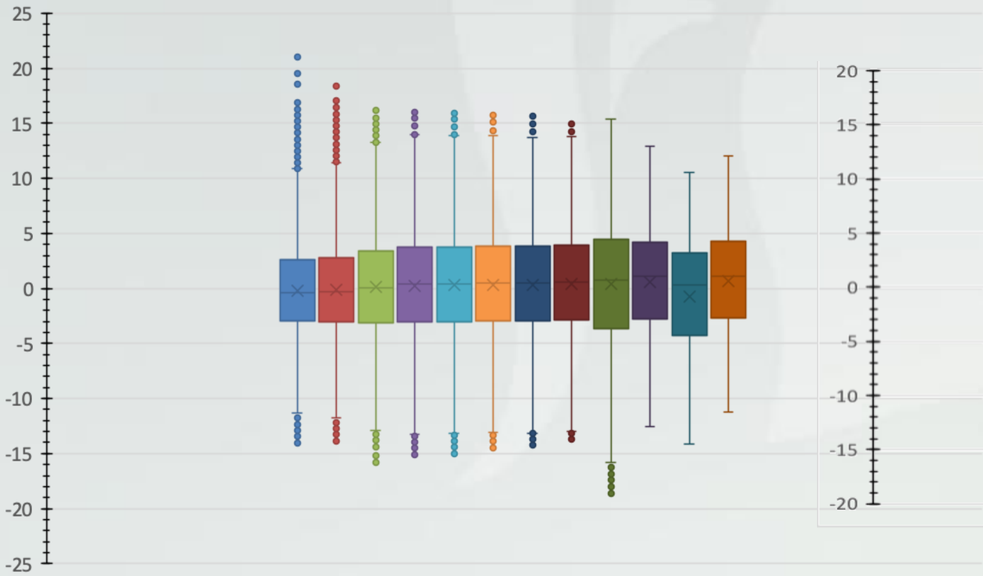
- Initially, a simple model has been selected, and then sequentially proceed towards the complex models.
- (a) Linear regression model (all variables)
- (b) Stepwise regression to evaluate the significant input variables.

Soil temperature	Regression coefficients				Regression intercept
	Air Temp	Rain	RH	Wind	
TC_1	1.04	0.19	-0.07	-0.59	12.13
TC_2	1.02	0.18	-0.05	-0.69	10.51
TC_3	0.92	0.02	0.05	-0.86	4.49
TC_4	0.84	0.02	0.08	-0.77	2.42
TC_5	0.83	0.03	0.09	-0.75	2.38
TC_6	0.81	0.06	0.09	-0.72	2.37
TC_7	0.80	0.07	0.09	-0.71	2.41
TC_8	0.76	0.12	0.09	-0.66	2.59
TC_9	0.66	0.14	0.04	-0.41	4.93
TC_10	0.60	0.11	0.09	-0.54	2.88
TC_11	0.39	0.08	0.10	-0.40	5.49
TC_12	0.47	0.04	0.09	-0.41	3.44

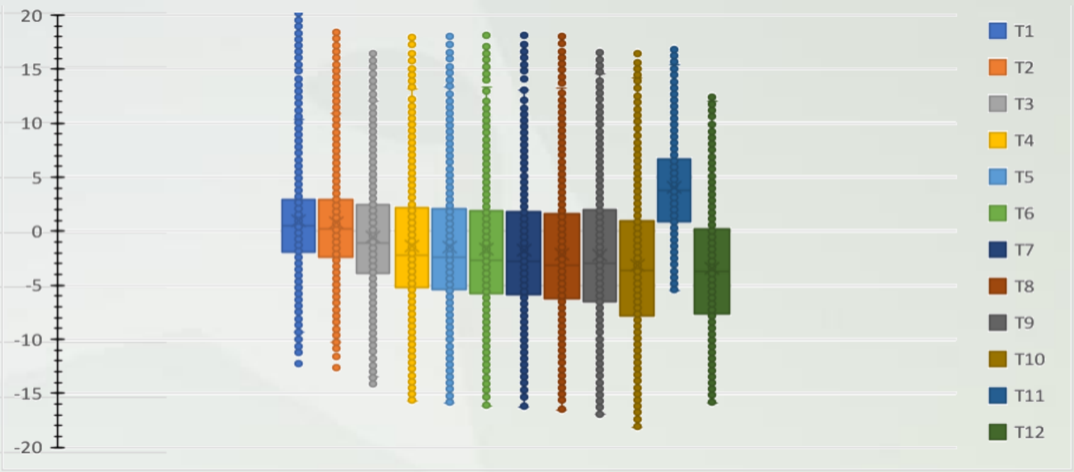
(1) Regression Models: *Data division*

- **Training Data:** first 50,000 datapoints
- **Testing Data:** remaining 9,522 datapoints

The error for all temperature values are shown below for both datasets (note all weather variables used as predictors)



Training Data



Test Data (not used to develop the model)

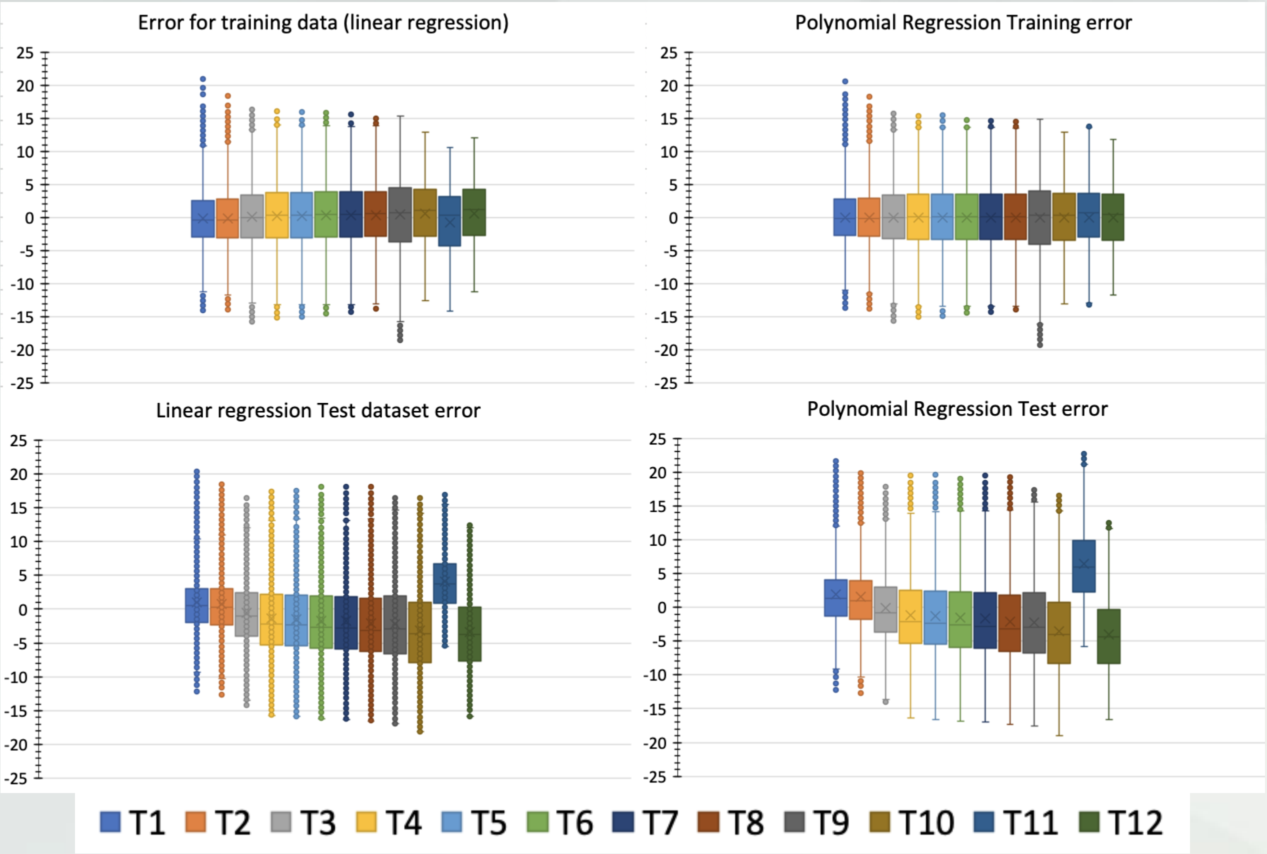
(1) Regression Models: *Stepwise*

All weather data input were considered; only those variables found to have *significant* influence are provided below, in order of most to least; **Air temperature** is most important

Temperature node	Significant inputs
TC_1	Air temperature, Relative humidity, Wind speed, Precipitation
TC_2	Air temperature, Relative humidity, Wind speed, Precipitation
TC_3	Air temperature, Relative humidity, Wind speed
TC_4	Air temperature, Relative humidity, Wind speed
TC_5	Air temperature, Relative humidity, Wind speed
TC_6	Air temperature, Relative humidity, Wind speed
TC_7	Air temperature, Relative humidity, Wind speed
TC_8	Air temperature, Relative humidity, Wind speed
TC_9	Air temperature, Relative humidity, Wind speed
TC_10	Air temperature, Relative humidity, Wind speed
TC_11	Air temperature, Relative humidity, Wind speed
TC_12	Air temperature, Relative humidity, Wind speed

(1) Regression Models: *Performance summary*

(Using weather variables only as predictors)



- Linear regression and polynomial regression models are used as the starting point
- Simplistic model
- Polynomial regression performs better compared to linear regression
- Overall, there is some amount of error in temperature prediction that can likely be improved

(1 → 3) Regression Models: *Additional considerations*

Soil temperature pattern varying depending on several parameters:

- Seasonal patterns
- Daily patterns
- Depth
- Soil characteristics

Next we tried (2) several non-regression methods, then returned to (3) an improved regression method

(2) Vector Models: *Summary*

(a) Vector Auto Regressive (VAR)

(b) Vector Error Correction Model (VECM)

(c) Vector Auto Regressive Moving Average (VARMA)

(2) Vector Models: *Data division & details*

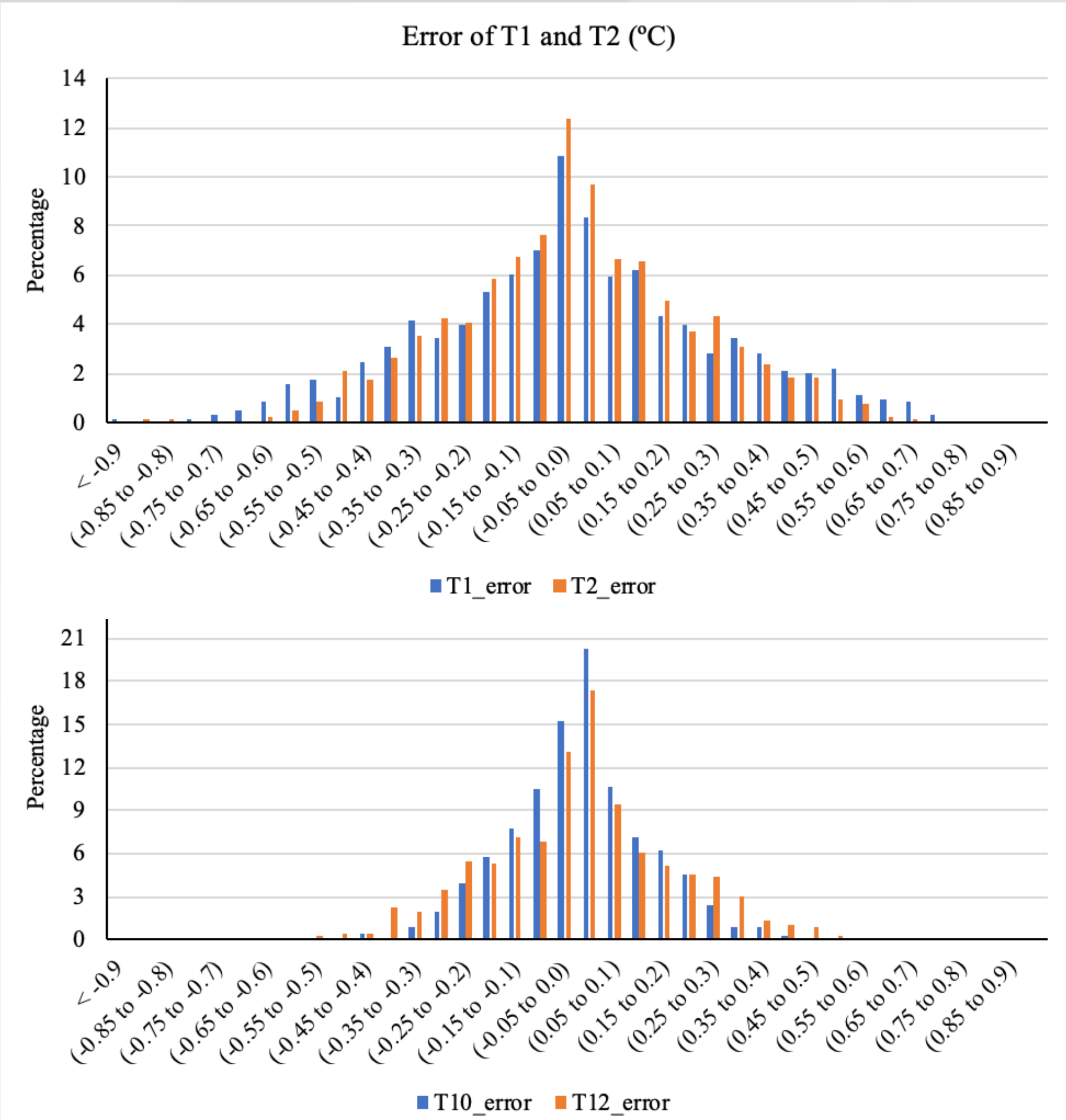
Data:

- ***Training Data***: first 45,000 datapoints *(memory limitations)*
  Aug/2017 – Nov/2018
- ***Testing Data***: remaining 9,522 datapoints

Model:

- ***Forecast length***: 10 days
- ***Maximum lag criteria:*** 24; selected based on *Bayesian information criterion (BIC)*
- First order differencing used to remove stationarity of data

(2) Vector Models: *Results Summary*



- **Unable to capture hourly** or **daily** fluctuations but **can capture seasonal variations**
- **VECM** and **VARMA** > **VAR**

Not the best approach for our needs

(1 → 3) Regression Models: *Additional considerations*

Soil temperature pattern varying depending on several parameters:

- Seasonal/Daily patterns
- Depth
- Soil characteristics

3 new variables considered:

- Day of year (1-365);
- Timestep (1 – 4 step/hour X 24 hours) per day
- Hours (1 – 24 hours/day X 365 days); *(i.e. 15 min = 0.25)*

(3) Regression Models: *Updated Methods*

- Forward Stepwise regression: *select the most important parameters*
  - Selected variables are: Day of Year, Timestep, Air Temperature, RH, Wind speed, Rain

- Consider:
  - Linear and Non-linear *(completed in R)*
  - Polynomial regression *(power of 2, 3, 4)*

(3) Regression Models: *Data division*

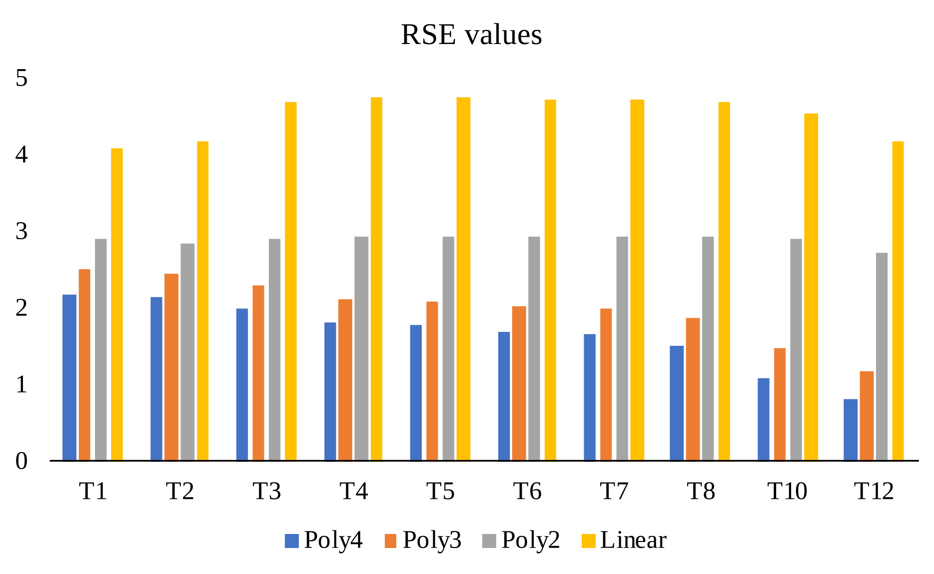
- ***Training Data***: first 35,040 datapoints *(1 year)* Aug/2017 – Aug/2018

- ***Testing Data***: remaining 10,270 datapoints *(23% of data)*

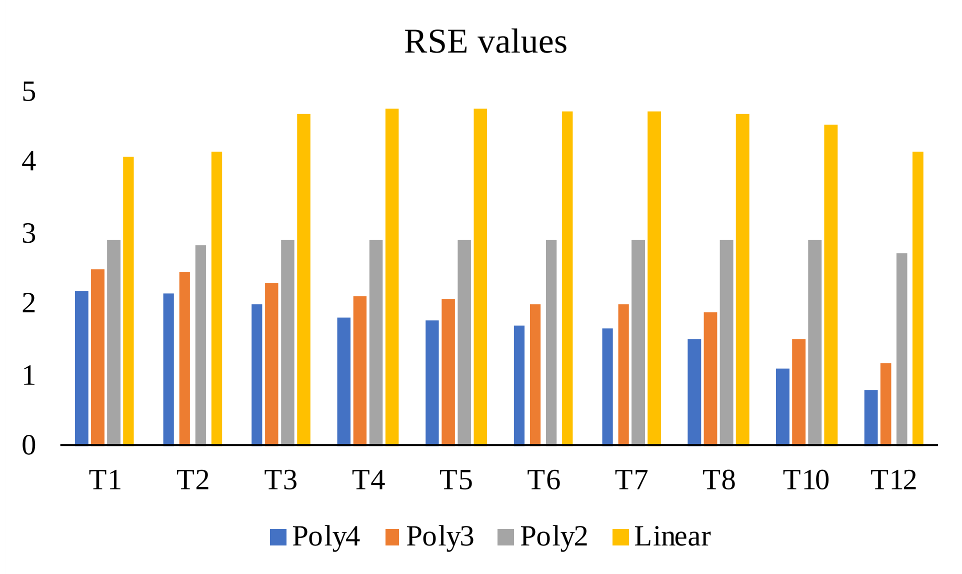
(3) Regression Models: *Results Summary*

- Non-linear regression (NLR) of 4th order performs best
- Error reduces with increasing depth
- RSE *(below)* and R-squared *(next slide)* used for evaluation

Training data



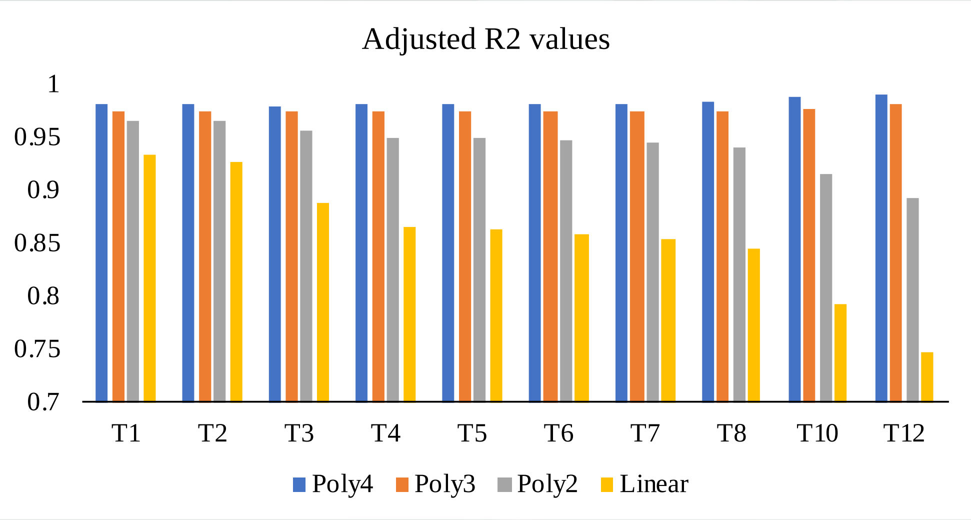
Testing data



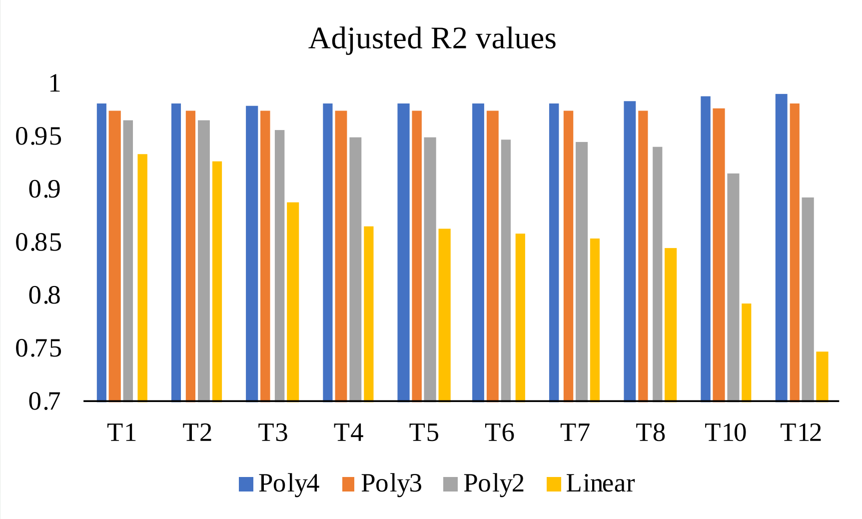
(3) Regression Models: *Results Summary*

Adjusted R^2 values are higher than 0.95 for all surfaces

Training data

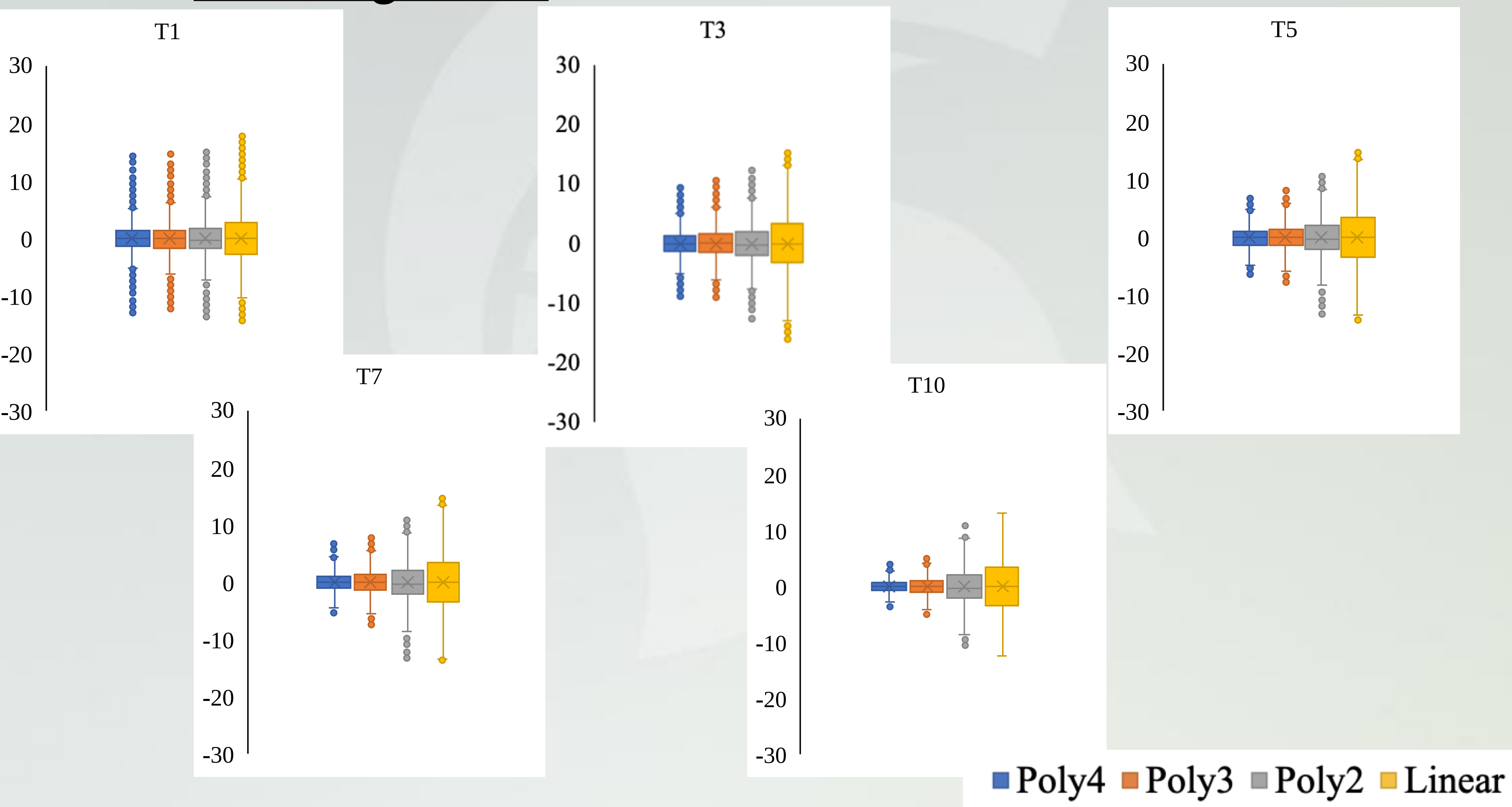


Testing data



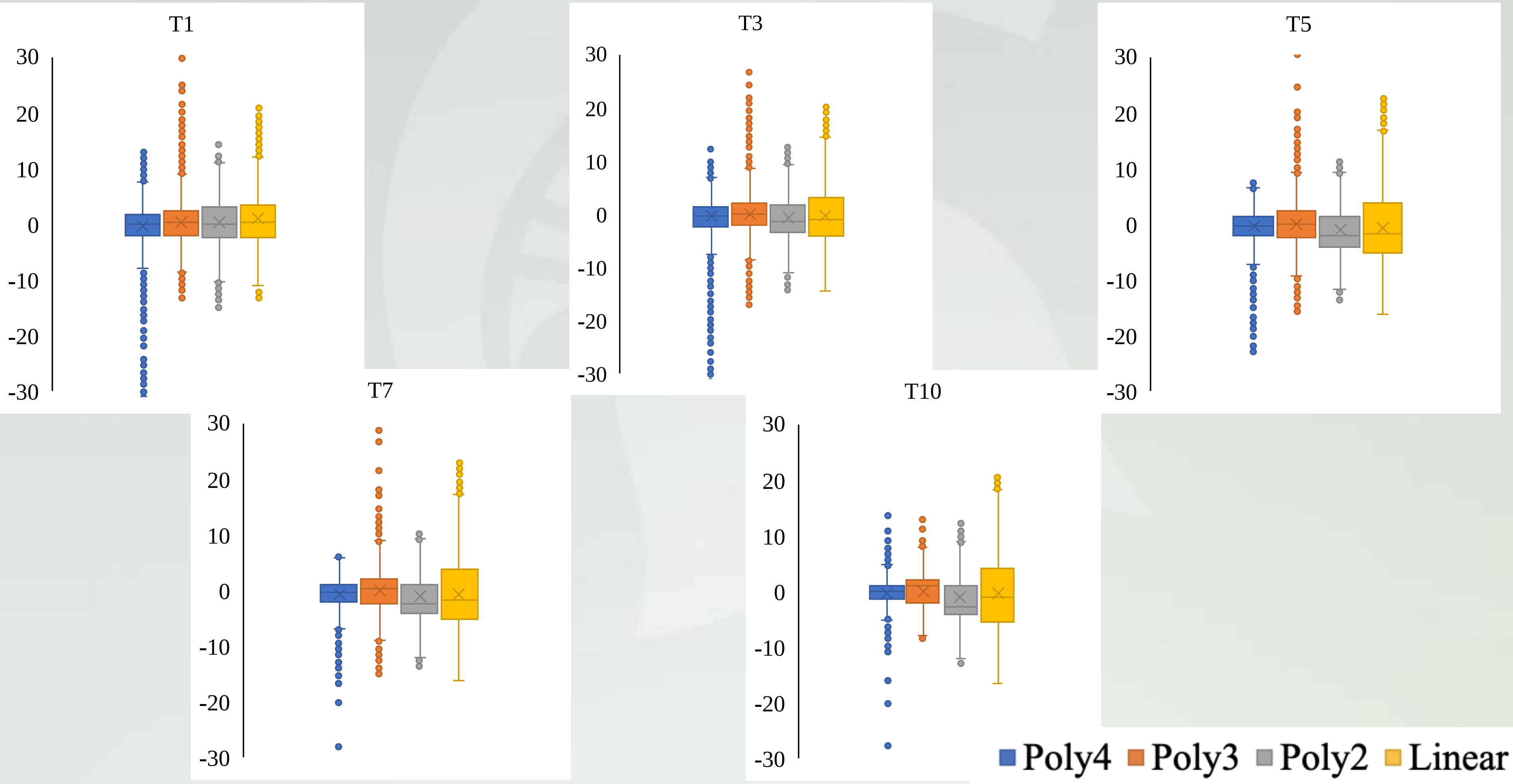
(3) Regression Models: *Error by Depth*

Training Data



(3) Regression Models: *Error by Depth*

Testing Data



(3) Regression Models: *Summary & Next Steps*

For (3):

- Polynomial regression performs better compared to other methods
- The results of (3) are better than (1) and (2)

Next Steps:

- Neural network
- Multi-layer perceptron model,
- Support Vector Machine,
- Neuro-fuzzy inference system,
- Deep learning algorithms

Calculation of Freeze-Thaw Cycles: *Questions*

By predicting temperature our ultimate goal is to predict the # of freeze-thaw (F-T) cycles

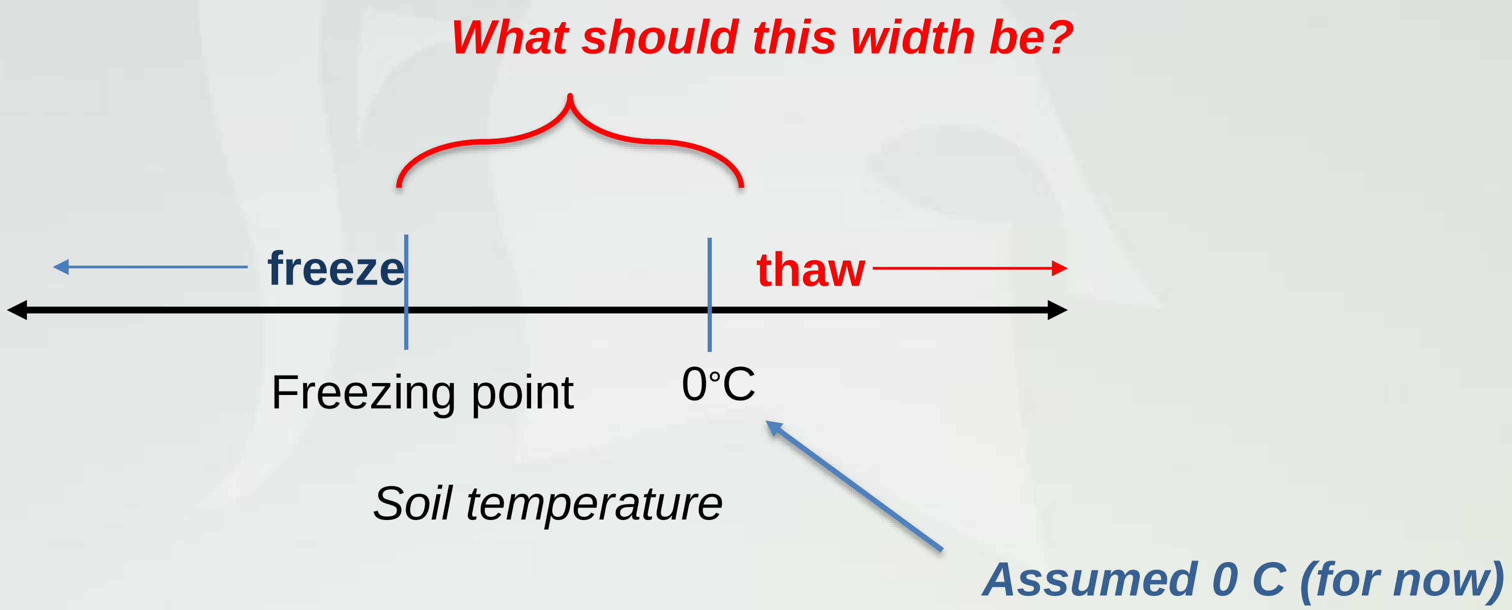
Key Questions:

(1) How do we define (calculate) a freeze-thaw cycle from soil data?

(2) How accurate are the data we are using?

Calculation of Freeze-Thaw Cycles: *Method*

- Weather and soil temperature: measured at 15 minutes time intervals
- Temperature accuracy: +/- 1 C



Calculation of Freeze-Thaw Cycles: *Method*

Method Used (*for now, focus on flexibility in code*)

1. Temperature above 0 °C => Liquid (thaw)

2. Temperature below freezing point => Solid (freeze)

3. Temperature within freezing point and 0 °C
   => phase transformation state

Calculation of Freeze-Thaw Cycles: *Method*

Freezing temperature considered (9 total):

- -0.001 °C *(i.e. no temp difference)*
- -0.1 °C
- -0.2 °C
- -0.25 °C
- -0.3 °C
- -0.4 °C
- -0.5 °C
- -0.75 °C
- -1 °C

Sensor	Depth (in)
TC_1	*3*
TC_2	*4*
TC_3	*9.5*
TC_4	*15*
TC_5	16
TC_6	18.5
TC_7	19.5
TC_8	24
TC_9	36
TC_10	48
TC_11	60
TC_12	72

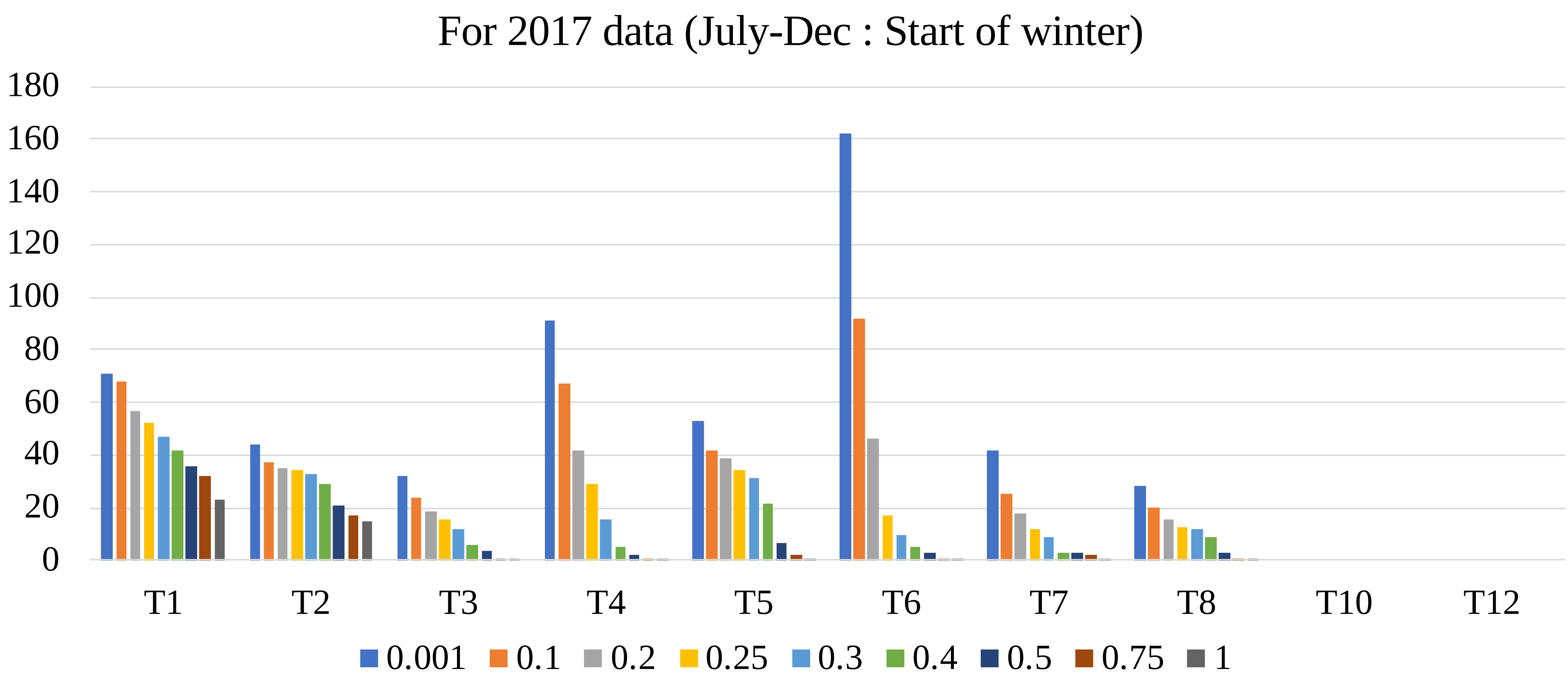
Shallow

Mid

Deep

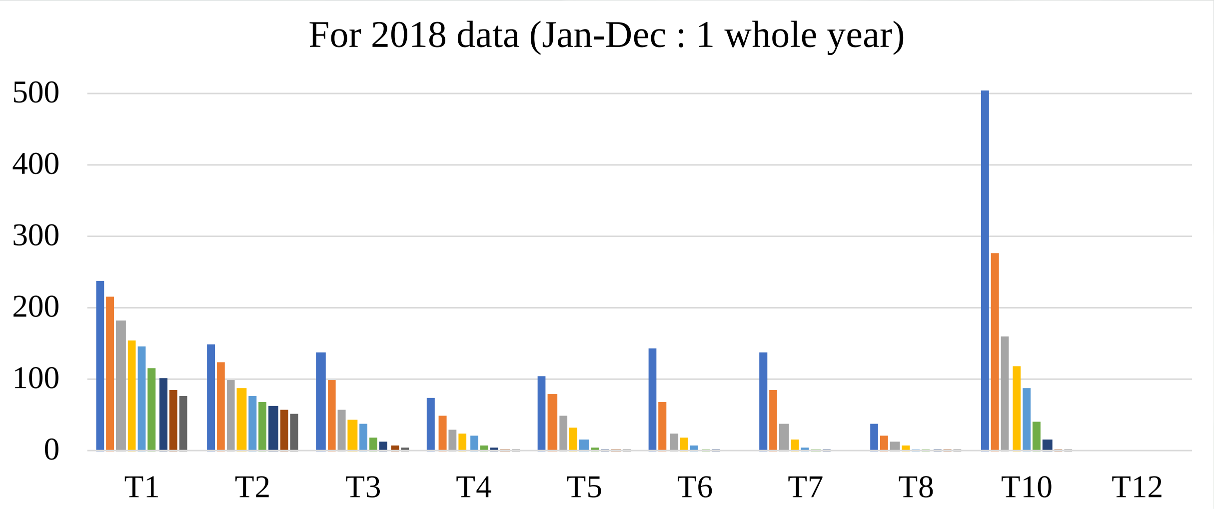
Calculation of Freeze-Thaw Cycles: *Results*

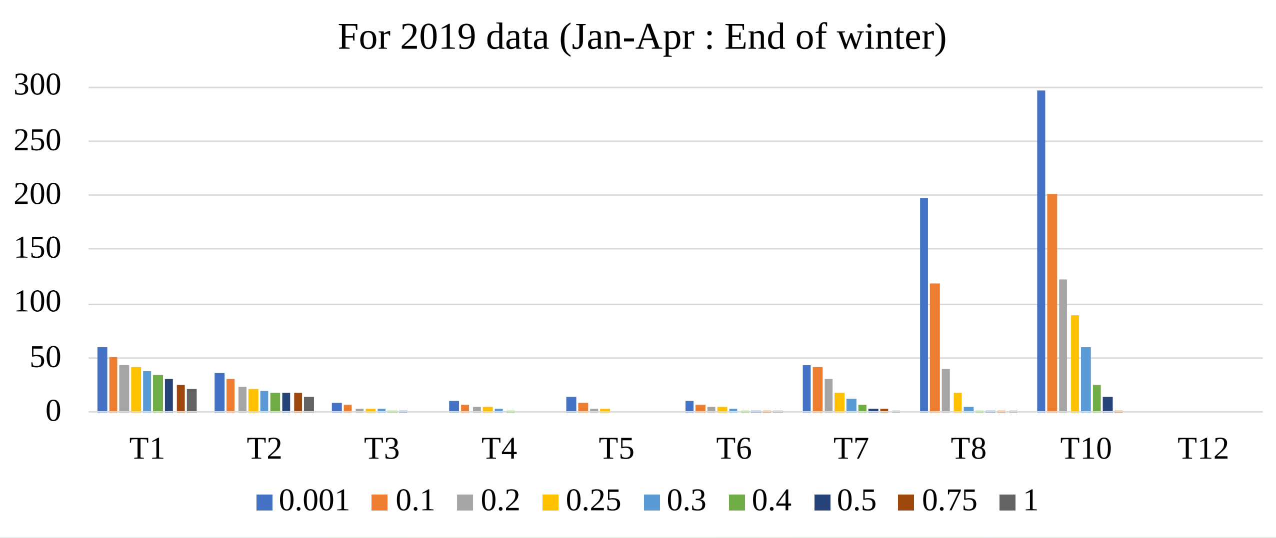
Datasets: 2017 (July-Dec), 2018 (Jan-Dec), 2019 (Jan-Apr)



Calculation of Freeze-Thaw Cycles: *Results*

Datasets: 2017 (July-Dec), 2018 (Jan-Dec), 2019 (Jan-Apr)





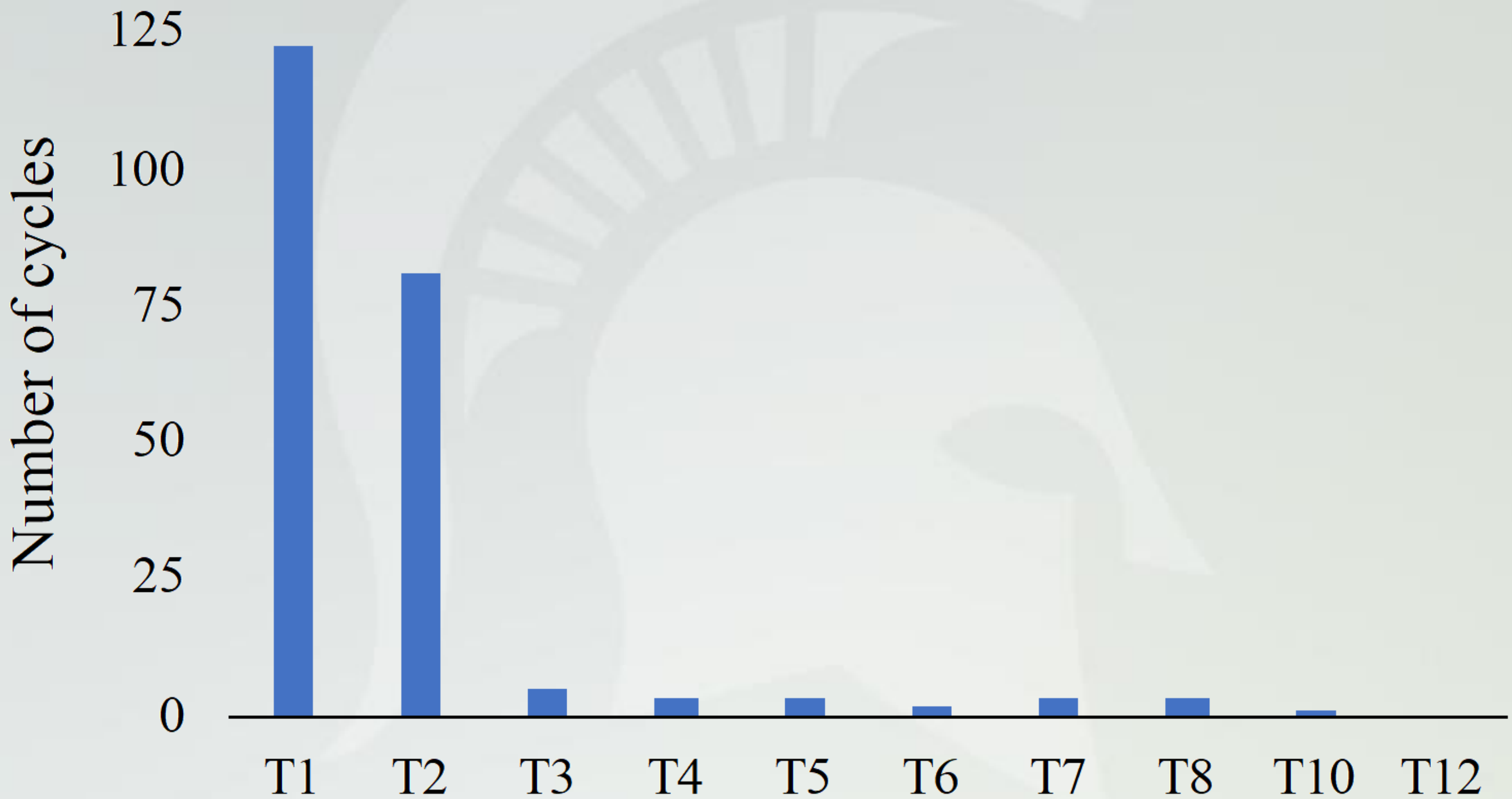
Calculation of Freeze-Thaw Cycles: *Summary*

# of F-T: dependent on freezing point temperature

- **Shallow soils (0-15 in):** more F-T cycles that deep soils;
- **Mid-level soils (16-24 in)** *(min annual temp. ~ -1 to -2°C):* # of F-T significantly influenced by F-T algorithm tolerance since more fluctuations near 0 °C range
- **Deep soils (36-72 in) :** # of F-T ~0 / generally does not go below 0 °C or change states

Calculation of Freeze-Thaw Cycles: *Summary*

# of F-T: if we choose a tolerance of 1 C



Calculation of Freeze-Thaw Cycles: *Summary*

# of F-T: if we choose a tolerance of 1 C

If we consider multiple locations (Note: sensors T3-T7 in different locations are at different depths thus cannot be easily compared)

<table>
<tr><th rowspan="2">Soil surfaces</th><th rowspan="2">Depth (in)</th><th rowspan="2"></th><th colspan="5">Sept 2017 - August 2018</th></tr>
<tr><th>cell 186</th><th>cell 188</th><th>cell 189</th><th>cell 127</th><th>cell 728</th></tr>
<tr><td>T1</td><td>3</td><td rowspan="12">2017 September to 2018 August</td><td>56</td><td>58</td><td>66</td><td>71</td><td>44</td></tr>
<tr><td>T2</td><td>4</td><td>28</td><td>30</td><td>27</td><td>64</td><td>35</td></tr>
<tr><td>T3</td><td>6.5-9.5</td><td>1</td><td>1</td><td>2</td><td>11</td><td>4</td></tr>
<tr><td>T4</td><td>9-15</td><td>1</td><td>1</td><td>1</td><td>3</td><td>1</td></tr>
<tr><td>T5</td><td>10-16</td><td>1</td><td>1</td><td>1</td><td>2</td><td>1</td></tr>
<tr><td>T6</td><td>12-18.5</td><td>2</td><td>1</td><td>1</td><td>2</td><td>1</td></tr>
<tr><td>T7</td><td>18-19.5</td><td>2</td><td>1</td><td>1</td><td>2</td><td>1</td></tr>
<tr><td>T8</td><td>24</td><td>3</td><td>1</td><td>1</td><td>2</td><td>1</td></tr>
<tr><td>T9</td><td>36</td><td>30*</td><td>1</td><td>1</td><td>2</td><td>1</td></tr>
<tr><td>T10</td><td>48</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>T11</td><td>60</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>T12</td><td>72</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
</table>

** Thermocouple error*

Calculation of Freeze-Thaw Cycles: *Summary*

# of F-T: if we choose a tolerance of 1 C

If we consider multiple locations (Note: sensors T3-T7 in different locations are at different depths thus cannot be easily compared)

Depth (in)	Sept 2018 to August 2019				
	cell 186	cell 188	cell 189	cell 127	cell 728
3	35	50	44	49	28
4	10	22	17	39	17
6.5-9.5	*3*	*2*	*3*	*7*	*4*
9-15	*2*	*1*	*1*	*3*	*4*
10-16	*2*	*2*	*1*	*4*	*5*
12-18.5	*1*	*1*	*1*	*2*	*1*
18-19.5	*1*	*1*	*1*	*1*	*1*
24	1	1	1	1	1
36	1	1	1	1	1
48	0	1	-	1	0
60	0	0	0	0	0
72	0	0	0	0	0

Review other larger tolerances above 1 C

Compare soil profiles at locations (potential impact of F-T variations)

Next Steps:

- Compare data of different locations to find the freezing temperature at different locations
- Compare actual and predicted freeze-thaw cycles obtained from the regression analysis
- Implement and test the performance of different complex models *(ANN, Multi-layer perceptron model, Support Vector Machine, Neuro-fuzzy inference system, Deep learning algorithms)*