Representation Learning to Improve the Performance of Graph Neural Networks in Predicting Material Properties

***Dipannoy Das Gupta, Qi Zhang, and Christopher Sutton**
Graduate Student, University of South Carolina, Assistant Professor, University of South Carolina, and Assistant Professor, University of South Carolina

Keywords: SSL, Pre-training, Fine-tuning, Contrastive loss

Advances in ML are making a large impact in materials science, chemical engineering, and computational chemistry. For example, a very promising line of research involves using ML to predict quantum mechanical (QM) properties of atomistic systems (e.g., energy, bandgap, density) for a set of reference calculations using only the structure as input. The accuracy of an ML model is significantly influenced by the selection of input features or representation. The current state of the art methods for representing materials are Graph Neural Networks (GNNs), which adeptly incorporate structural information. However, one of the constraints of GNNs is the limited availability of data labeled with QM properties for model training which can hinder the ML model accuracy.

One way to overcome the issue of labeled materials dataset is representation learning via self-supervised learning (SSL), which aims to learn from inputs without relying on explicitly labeled data. In this way, the model learns intrinsic information from the input data which has been shown to improve the ML model's performance in a subsequent application of the model to predict the properties of materials. However, the effectiveness of models trained via SSL hinges on their ability to generate robust representations.

SSL is a mechanism where an ML model is trained to learn the representations without relying on external labels and annotations. SSL works with its own objective functions focusing on learning the input patterns/representations. Examples of these objective functions include Barlow Twins Loss and Contrastive learning which aim to learn the pattern by calculating the similarity and dissimilarity between the inputs.

In this work, we used both Barlow Twins and Nt-Xent loss functions for the SSL; however, the Nt-Xent has made the model learn better which is evident from the downstream prediction tasks. Our perturbation process for SSL aims at creating a positive pair of inputs from a single structure by masking atoms and their corresponding edges. NT-Xent measures the similarity of two latent vectors extracted from a positive data pair and this principle aligns well with our perturbation strategy. Alternatively, the Barlow Twins loss assesses the cross-correlation matrix between two distorted embeddings of an input, aiming to transform the matrix into an identity matrix during the learning process.

I will present our results for the pre-trained model derived using SSL on 213,200 double perovskites (AA'BB'X6) samples. The pre-trained model was used in subsequent fine-tuning tasks afterwards. The fine-tuned model has demonstrated improved results over the prediction of different calculated properties, such as the DFT formation energy, ΔHf, and the PBE and HSE06 band gaps compared to directly trained models.

In this work, initially, we have applied Crystal twins on our 213,200 unlabeled samples. Crystal twins is a SSL framework built on CGCNN[2,4]. The pre-trained model derived after SSL is used to fine-tune another CGCNN model with 1000 labeled samples to perform regression tasks. We have reported an MAE 0f 89 meV/atom for the formation energy during prediction over 4760 left-out samples after applying a fine-tuned model.

To extract more information from atomic local environments, we developed an SSL architecture keeping SchNet on top of the MLP, SchNet is a continuous filter convolutional neural network where message passing is performed across three modules-atom centric layers, interaction modules, and filter generation networks[1]. Using the Barlow Twins loss for pre-training, the fine-tuned model after the SSL step trained using only 1000 samples leads to an MAE of 71 meV/atom for ΔHf for a test set of 4760 left-out samples. In comparison, the NT-Xent loss function leads to an MAE of 59 meV/atom for the same left-out set. Without the pre-trained model, the SchNet model shows almost the similar performance (73 meV/atom) as the Barlow Twins loss. Therefore, the SSL scheme is advantageous. We have also tested our same fine-tuned model over 946 samples from different distribution. The NT-Xent fine-tuned model has demonstrated MAE of 93 meV/atom while the Barlow Twins fine-tuned model had an MAE of 148 meV/atom. A large MAE of 239 meV/atom has been found without fine-tuning the model. Apart from the formation energy, we have also observed some improvements in the prediction of different band gaps using SSL.

In summary, our SSL framework uses cheap and affordable structures generated via bond valence method. This is an excellent solution for bypassing expensive DFT-level structures. Moreover, the labeling is another challenging task while the unlabeled structures ensure good pre-training prior to the fine-tuning. Therefore, the scarcity of labeled data is not hindering the performance. Our next goal is to incorporate more structural information during SSL. In this objective, we are working on combining bond graph and line graph. We believe this will provide more atomistic details to the models which will eventually generate finer representation for the downstream tasks.

Comparison of Bayesian and Neural Networks in Predicting Customer Retention

***Eli Hellmig, Dr. Ivan Dungan**
Undergraduate Student, Francis Marion University, Assistant Professor, Francis Marion University

Keywords: Bayesian Network, Neural Network

In a consulting project with a local software company, we compared the effectiveness of neural networks and Bayesian networks as means for predicting customer retention. Neural networks and Bayesian networks can be considered to be at opposite ends of the explainability spectrum when it comes to machine learning models. Neural networks are often said to operate as a black box, offering no real insight as to how the model chooses what variables it considers in its hidden layer(s), while Bayesian networks are intuitive and rely solely on probability methods to fit the model and predict outcomes. We examined these differences regarding customer retention and other considerations including how each model handles increasingly complex data sets, missing values in data, and how computationally efficient each model is.

Conditional Density Estimation for CMV Crash Risk Analysis and Uncertainty Quantification in Work Zones

Debbie Aisiana Indah, Judith Mwakalonge, Gurcan Comert*, and Saidi Siuhi
Graduate Student, South Carolina State University, Professor, South Carolina State University, Associate Professor, Benedict College, Professor, South Carolina State University

Keywords: Conditional density estimation, crash risk prediction, work zones, CMV Crashes

Commercial Motor Vehicle (CMV) crashes in work zones pose significant risks to public safety, leading to critical injuries, fatalities, and causing extensive disruptions. Given the high stakes, accurately predicting, and managing these risks is paramount. Traditional risk prediction methods, primarily based on regression models, offer point estimates that fall short of capturing the full spectrum of uncertainty inherent in CMV crash outcomes. To bridge this gap, our research introduces an advanced approach utilizing machine learning (ML) for initial risk probability prediction, followed by Conditional Density Estimation (CDE) to quantify the uncertainty surrounding these predictions comprehensively. Analyzing five years of CMV crash data from South Carolina, our approach not only forecasts potential crash outcomes but also delivers a quantified assessment of the uncertainties surrounding these predictions. This allows for the development of more informed, adaptable, and resilient safety strategies, tailored to address the specific challenges of CMV crashes in work zones. By providing a detailed quantification of both risk probabilities and their associated uncertainties, our research contributes a novel analytical framework to the field of traffic safety analysis. This framework aims to enhance road safety and inform infrastructure planning, paving the way for targeted safety interventions that can effectively mitigate the impact of CMV crashes in high-risk scenarios.

Evaluating Factors Influencing School Travel Mode Choice in the United States Using Explainable Artificial Intelligence

***Hannah Musau, Judith Mwakalonge, Gurcan Comert, and Saidi Siuhi**
Graduate Student, South Carolina State University, Associate Professor, South Carolina State University, Associate Professor, Benedict College, Assistant Professor, South Carolina State University

Keywords: COVID-19, mode choice, School travel, XGBoost, XAI

The Coronavirus Disease 2019 (COVID-19) pandemic significantly altered our daily lives and travel habits, yet there is limited awareness of its effects on school transportation. This study explores the dynamic factors influencing school travel mode choice in the United States across three distinct periods: pre-COVID-19 (Fall 2019), during COVID-19 (Spring 2021), and post-COVID-19 (Fall 2021). A survey methodology was used to collect data on the parents' preferred school trip mode choice over three timeframes. To achieve the study's objective extreme gradient boosting (XGBoost) was utilized to analyze a range of independent variables including locale, household size, vehicle ownership, parent's level of education, income, and employment status, distance from school, school stage and number of children enrolled at the respective school stages. Furthermore, the SHapley Additive exPlanations (SHAP) method was used to better comprehend the XGBoost model outcomes. The explainable artificial intelligence (XAI) function was utilized to rank the significance of input features based on their average values.

Results show that there were temporal variations in the importance of features before and after the COVID-19 pandemic. While distance and income remained key features in determining the mode choice, the impact of education rose while that of vehicle ownership decreased post-pandemic. These findings can be utilized to improve understanding and modeling of parents' school mode choice preferences. They can also make a significant contribution to the post-pandemic school transportation policies and in similar situations that limit public transportation activities.

Using Inertial Measurement Units to Estimate Center of Pressure in Stroke Patients During Balance Challenges

***Isabelle Museck and Jesse Dean**
Doctoral Student, Medical University of South Carolina; Associate Professor, Medical University of South Carolina

Presentation keywords: Balance, inertial measurement units, stroke

Introduction: Over 795,000 people in the United States have a stroke each year. Stroke is one of the leading causes of serious long-term disability and up to 50% of survivors experience reduced mobility. Even after receiving physical therapy, many people with chronic stroke (PwCS) exhibit deficits in their standing balance. These deficits are associated with a decreased quality of life, poorer functional mobility, and an increased risk of falls. Those with chronic stroke are particularly vulnerable to mediolateral losses of balance, as evidenced by the high proportion of sideways falls in this population. While training reactive responses to lateral losses of balance would likely be of value, an additional approach would be to prevent these losses of balance from occurring. Foundational experiments have shown that non-invasive stimulation can be used to generate artificial somatosensory feedback that elicits predictable balance responses. However, this type of sensory augmentation requires real-time measurements of the user's center of pressure (CoP), which would not be available in either a typical rehabilitation or real-world setting. Wearable sensors such as inertial measurement units (IMUs) may allow this barrier to be addressed, as these devices offer a low-cost, portable way to track motor symptoms and biomechanical parameters. The purpose of this study was to quantify the potential of IMU data to estimate real-time CoP motion while mediolateral balance is challenged in PwCS.

Methodology: We collected data from 10 PwCS. Participants wore 7 IMU sensors (sacrum, bilateral thigh, shank, and foot) and were instructed to stand still and relaxed on a force plate that translates mediolaterally in varying, unpredictable patterns. This perturbation requires participants to adjust their CoP to avoid losses of balance. Multiple 30-second trials were conducted as the challenge of the balance task was increased by scaling the platform translation velocity. Trials ended when the rms velocity exceeded 20 mm/s, an indicator of an increased risk of falls. Data collected from the mediolateral changes in CoP was compared with IMU accelerometer, magnetometer, and gyroscope data collected to quantify potential linear relationships. Identification of the correlations between IMU inputs and corresponding differences in CoP will open the door for the next steps that involve going beyond simple linear correlations and using machine learning models to make more accurate predictions.

Results: IMU and CoP data collected from 10 PwCS was analyzed to quantify possible correlations. The results demonstrated that the Medial-Lateral (ML) IMU acceleration data was most strongly correlated with CoP measurements. Across participants, the ML accelerometer measurements recorded from the trunk, left thigh, and right thigh yielded high average correlation coefficients of 0.61±0.23, 0.65±0.25, and 0.70±0.19 (mean±s.d.), respectively. These results are consistent with a recent study demonstrating that a single IMU placed on the sacrum can estimate CoP displacement with reasonable accuracy. Moreover, ML accelerometer measurements from the left and right shanks demonstrated moderate correlation with CoP displacement measurements with average coefficients of 0.48±0.25 and 0.50±0.24, respectively. The IMU data collected from the feet showed very weak correlations which is to be expected as the feet tend to not sway very much during data collection. CoP velocity and acceleration were also calculated from CoP displacement and correlated with the IMU measurements; however, only weak correlations were observed.

Conclusions: The presence of moderate to strong correlations between IMU and CoP measurements shows that IMUs placed on the trunk, shanks, and/or thighs have the potential to provide real-time information that can be used to estimate CoP displacement. However, these correlation values vary across patients and trials, likely limiting the effectiveness of a sensory augmentation device that relies on these simple estimates. Future work will apply machine learning models to personalize and enhance CoP estimations from IMU data and integrate this into a system that can provide portable therapeutic stimulation when patient instability is detected.

Making Use of Quantitative Data in Bayesian Network Models by Clustering with K-Means and Entropy

***Jacob Reeder**
Undergraduate Student, Francis Marion University **Keywords - Bayesian Networks, Clustering, Entropy**

Abstract

We developed a Bayesian Network (BN) Model for a local technology company with the primary purpose of predicting customer satisfaction based on customer data. Besides reviewing BNs and our network, the objective that we discuss here was to enhance the accuracy of the model for better analysis. The primary strategy for improving the accuracy of our model was by clustering the nodes of the network with numerical values to better convey the available information. Instead of the usual approaches to clustering, the clustering was done through combining the k-means clustering algorithm with the notion of entropy from information theory. After performing clustering on customer data we were able to successfully increase the accuracy of the model by roughly 3%.

Automating irrigation and nitrogen fertigation based on real-time data from a wireless sensor network

***Javad Alavi, Jose Oscar Payero, and Tom O. Owino**
Doctoral student, Clemson University, Associate Professor, Clemson University, and Associate Professor, Clemson University

Keywords: Automation, Center pivot, Fertigation, Sensors, Soil moisture

Introduction

The efficient application of water and nitrogen (N) fertilizers is vital for the long-term economic, and environmental sustainability of agricultural production. Insufficient applications may lead to reduced crop yields, whereas overfertilization might increase production costs and result in adverse ecological impacts. Therefore, maintaining a balance in N applications is crucial to achieving both economic benefits and environmental stewardship and even minor enhancements in NUE values would lead to substantial economic savings. Fertigation, which is defined as the application of fertilizers through irrigation systems using water, has become a widespread practice in modern agriculture [1]. Compared to traditional fertilization methods, fertigation offers a uniform distribution of fertilizers to all plants and improved nutrient absorption [2].

Precision Agriculture (PA) has gained considerable prominence in modern agricultural technology. PA addresses the need to manage variations both between and within fields, providing a conceptual framework for agricultural operations [3]. As a fundamental element of climate-smart agriculture, soil moisture, and Normalized Difference Vegetation Index (NDVI) sensors could be employed to automate and optimize the application of water and N-fertilizer. Recently, some studies have been conducted for the automation of fertigation in large-scale, open-field, and greenhouse production. Innovative algorithms for intelligent N applications have been recently introduced [4] and the application of sensors for the automation of fertigation coupled with algorithms was reported by [5]. Optical sensors and Clemson University algorithms were successfully applied for the automation of the N application [6]. However, there is still a lack of investigations on the automation of combined control of site-specific water and N applications called the ‘Intelligent Water and Nutrient Placement - (IWNP)’ in large-scale irrigation systems. The objective of this investigation was to develop an automated system retrofitted on an existing center pivot for combined control of site-specific water and N applications.

Methodology

A water and N automation system for a center pivot irrigation system was developed using a wireless sensor network driven by Arduino microcontrollers at the Edisto Research and Education Center (EREC) in Blackville, South Carolina. The initial setup and testing of the system were conducted during a field experiment with cotton in 2023. The experiment consisted of a split-plot design to evaluate five irrigation treatments and three N treatments using five replications. Soil moisture sensors were installed at three depths in the irrigation plots. The supplied water to the center pivot was separately regulated by solenoid valves installed in five irrigation zones along the length of the center pivot. Also, five NDVI sensors were installed along the length of the center pivot to estimate crop N status in different areas of the field as the center pivot rotated. Five fertilizer injection pumps were used to regulate the injection of liquid N fertilizer to different zones along the length of the center pivot. The soil moisture and NDVI data were collected and transmitted to a central receiver using a wireless sensor network driven by Arduino microcontrollers using LoRa radio communication. Based on the sensor data, the central receiver was used to control the irrigation solenoid valves and the fertilizer injection pump to regulate the application of water and N to a cotton crop when and where needed.

Results

All the equipment and sensor setups for the field experiment were constructed and initially tested in 2023, showing promising results. Field experiments to improve and fine-tune the automation system will continue during the 2024 cotton growing season.

Conclusion

The IWNP was successfully applied using smart sensing and model-based decision support systems for the first time in South Carolina to minimize water and optimize N fertilizer. The automated process demonstrated that the system enables sophisticated intelligent irrigation and fertigation and showed potential for making make agriculture more efficient, productive, and sustainable.

Efficacy of Statistical, Long Short-Term Memory (LSTM), and Quantum LSTM in Cyber-attack detection for Connected Vehicles

[*]Jean Michel Tine, Sefatun-Noor Puspa, Gurcan Comert, Balaji Iyangar, Abuhdima Esmail, and Mashrur Chowdhury
Ph.D. Student, Clemson University, Ph.D. Student, Clemson University, Associate Professor, Benedict College, Professor, Clemson University

Keywords: Cyber-Attack, False Information Attacks, Bayesian Online Change Point Detection, Long Short-Term Memory, Quantum Long Short-Term Memory

In the context of connected vehicles (CVs), external connectivity to other CVs and related infrastructure creates cyber-attack vulnerability. These cyber-attacks can compromise safe vehicle operations and disrupt transportation infrastructure. This study focuses on misinformation or false information attacks involving false data propagation within a CV's in-vehicle network, which poses significant risks and can compromise safety applications, such as forward collision warning and merging support. These attacks are so dynamic that they are hard to detect in real-time and near real-time. This research focuses on developing and evaluating models to detect disinformation attacks in a connected vehicle (CV) setting. Its objective is to enhance cyber security by effectively detecting BSM data anomalies, which will help secure V2X communications and ensure the safety and reliability of transportation systems. The advantage of Bayesian Online Change Point Detection (BOCPD) is its ability to quickly detect changes in data sequences, an essential feature for detecting real-time cyberattacks in CVs. This study compared BOCPD with long short-term memory (LSTM) and quantum LSTM (QLSTM) regarding their ability to recognize false information attacks.

In this research, we used the ability of the BOCPD to address temporal time-series data, which is a typical feature of log data. We also leverage his ability to detect multiple change points, enabling us to identify complex multi-stage cyber-attacks. Moreover, we used the LSTM, a deep learning model well known for time series data due to its capacity to address time dependencies by forgetting and remembering past information. We developed a quantum-enhanced extension of conventional long-short-term memory (LSTM) networks. Our contribution related to the QLSTM- based attack detection lies in replacing the traditional neural network components in LSTM cells with Variational Quantum Circuits (VQCs), a hybrid quantum-classical computational model that utilizes tunable parameters within quantum circuits optimized through classical optimization algorithms, to solve specific problems by learning from data. This combination of quantum computing principles and recurrent neural networks yields a more advanced framework, Quantum LSTM.

Our evaluation strategy involves using a Simulation of Urban Mobility (SUMO) to create a CV environment and generate a Basic Safety Message (BSM) from each CV. This data, including critical information like vehicle position, speed, and heading, is broadcast ten times per second. We assume that an attacker could intercept the information transmitted by the vehicle Onboard Unit (OBU) to the Roadside Unit (RSU) and inject false speed to generate an anomaly in the CV safety application.

This study revealed that QLSTM (99.7%) outperformed BOCPD (99.5%) and LSTM (99.5%) in detecting false speed attacks in a CV environment. In addition, QLSTM requires less training data, thus contributing to a more efficient approach to detecting misinformation or false information attacks.

PerovNet: A Generalizable Machine-Learned Interatomic Potential for Halide Perovskites

John T. Barber*, Nima Karimitari, Zachary J. L. Bare, Dipannoy Das Gupta, and Christopher Sutton
Doctoral Student, Postdoctoral Fellow, Postdoctoral Fellow, Doctoral Student, Assistant Professor

Keywords: Machine Learning, High-Throughput Screening, Perovskites

Halide inorganic perovskites (with the formula ABX_3) have received significant attention because of their potential as photovoltaic (PV) materials due to their ease of synthesis and optimal band gap energies. Despite the widespread interest in these materials, issues remain with stability, motivating the need to design alternative halide perovskites. Quantum-mechanics based electronic structure calculations (such as density functional theory, DFT) have become vital in understanding a wide variety of properties in these materials, which allow for the design of materials computationally before significant investments in both time and cost of synthesis.

However, one complication in the screening of halide perovskites, is that growing evidence in the literature indicates local symmetry breaking occurs from various distortions (tilting, rotations) of the corner-sharing octahedra, as well as correlated atomic disorder. These static and dynamic distortions change the atomic positions and introduce microscopic motifs (e.g., dipolar), which determine the macroscopic properties (e.g., band gaps, stability). Although these local distortions and the resulting changes in the electronic structure are saliant features that determine properties and functionality of halide perovskites for PV applications, most high-throughput calculations do not account for these structural complexities.

Using a set of DFT reference calculations, a machine learning interatomic potential (MLIP) can be trained with ab initio accuracy but at a much lower cost. These MLIPs are trained using a set of energy, forces, and virials calculated with DFT. In this work, we train a generalizable machine learning interatomic potential (MLIP) to account for these local symmetry-breaking motifs by learning the potential energy surface. The MLIP is trained within an equivariant message-passing neural network Multi Atomic Cluster Expansion (MACE) [1], which has been shown to perform well for a variety of tasks in previous work [2]. Using MACE, we train a generalized MLIP to a dataset of 341 compositions out of 195,968 halide perovskites. For an unseen dataset of 4,400, a mean average error of 45 meV/atom was obtained indicating that the model can be used to bypass DFT directly for optimizing the structures. Out of the set of 195,968 halide perovskites, we used an initial ML model from our previous work [3] to generate and predict both the stability (using the DFT decomposition energy, ΔH_d) and band gap energy (E_g) to select 10,382 halide perovskites based on the criteria $E_g < 3$ eV and $\Delta H_d < 0.3$ eV/atom using an initial template structure. For this subset, an MLIP was used to model the presence of local symmetry- breaking modes for each composition. From this analysis, we find that ca. 90% of the 10,382 compositions contain significant local symmetry-breaking modes. Current work is being performed to relate these effects to changes in the electronic structure.

Comparing Satellite-Derived SMAP Soil Moisture Data to In-Situ Measurements from a Network of Measuring Sites

***José Payero and Joe Bible**
Assistant Professor, Clemson University; Assistant Professor, Clemson University

KEYWORDS: soil moisture, remote sensing, SMAP, satellite

INTRODUCTION: NASA launched the SMAP mission in January 2015, which includes a constellation of satellites that collects active and passive remote sensing information using radar and radiometer measurements, which are transformed to estimate soil moisture (Liu et al., 2017). SMAP satellites are instrumented with an L-band radiometer (1.41 GHz) to measure brightness temperature and a rotating reflector radar (1.26 GHz non-imaging SAR) to measure radar backscatter (Burgin et al., 2017; Kim et al., 2018). The radiometer and radar sensors provided SMAP with passive and active remote sensing capabilities. In July 2015, the radar malfunctioned and is no longer operating. Therefore, SMAP lost its active remote sensing capability, but the SMAP radiometer continues to make passive microwave measurements (Burgin et al., 2017; Das et al., 2019). The daily SMAP soil moisture information is available for download from NASA as images for a defined area of interest or as daily time series for a given location. This free soil moisture information could be of much value for agriculture, especially for irrigation water management, assuming that it can accurately represent local conditions. However, there are still questions about how accurately the SMAP information can represent local conditions, which is the case in South Carolina, where a comparison of SMAP data with on-site measurements is lacking. Therefore, the overall goal of this study was to evaluate the potential use of SMAP data in agriculture. The specific objectives of this study were to: (1) evaluate the accuracy of SMAP soil moisture data compared to local ground soil moisture measurements and (2) develop statistical modeling approaches to correct the SMAP soil moisture information to predict the ground measurements.

METHODS: In 2016, a network of eighteen ground measurement stations was deployed in Barnwell county, South Carolina, to collect hourly and daily soil moisture data at two soil depths (5 cm and 30 cm). Concurrent time series of the daily SMAP surface soil moisture data from 2016 to 2018 for the locations corresponding to each ground measurement site were downloaded from NASA to be used for comparison and statistical modeling. For each measurement site, three sets of statistical models were used to estimate the measured (sensor) soil moisture from the satellite (SMAP) data.

RESULTS: A comparison of the daily SMAP and measured surface (5 cm) soil moisture data showed that, generally, the time series of the two datasets tended to follow the same trend, but SMAP consistently overestimated soil moisture compared to the measured values. Linear regression analysis resulted in a relatively poor agreement, which varied widely by measurement site, with R^2 among sites ranging between 0.20 and 0.63. Statistical modeling combining observed prior soil moisture with SMAP estimates could explain between 68% and 91% of current soil moisture, with an average increase in percent soil moisture explained of 38% compared to using the SMAP data alone to estimate the measured surface soil moisture.

CONCLUSIONS: Our results showed that using the SMAP data without local calibration could result in significant errors in estimated surface soil moisture. However, statistical modeling utilizing some local measurements could considerably improve the estimation. These results indicate opportunities for future work to extend the nowcasting of surface soil moisture for larger areas based on calibrated SMAP soil moisture readings. Conducting local calibration will require installing multiple sensors within each site (a shortcoming of the current investigation was the lack of data to estimate within plot variation in soil moisture). A clearer picture of the variation in a smaller area (within-plot variation) will enable using a relatively sparse soil moisture sensor network to nowcast over a larger geographical region.