COMPARATIVE ANALYSIS OF DIFFERENT ON-THE-GO SOIL SENSOR SYSTEMS

H. Moulay, B. Arnall, S. Phillips
Oklahoma State University, Stillwater, OK
Haitam.moulay@okstate.edu (405) 612-6315

ABSTRACT

Soil is an essential natural resource that requires crucial attention due to its significant role in crop yield. Understanding the spatial variability of soil chemical and physical attributes optimizes the profitability of nutrient management for crop development. Soil mapping systems with various types of proximal soil sensors provide crop growers with an excellent opportunity to access soil heterogeneity at a higher spatial resolution and in an efficient and less invasive manner.

Research indicates that it is possible to acquire data related to soil pH, electrical conductivity, organic matter content, soil moisture, and other factors in a cost-effective manner (Adamchuk 2007; Lund 2011; Heege 2013; Huang 2018). Different types of ongoing soil sensors can indirectly assess the range of soil characteristics that are typically measured by traditional soil testing methods currently available in the market.

The utilization of traditional soil sampling methods and subsequent laboratory analysis is both a time-consuming and costly process. Soil conductivity sensors, ion-selective sensors, and soil reflectance based on Passive gamma-ray spectroscopy sensors are commonly used in precision farming applications. These sensors are capable of scanning soils with high spatial resolution and can provide insights into variations in soil parameters. (Srinivasan, 2006).

- ***Non-contact conductivity sensors*** typically utilize electromagnetic induction to detect the electrical conductivity of soil. A main field is generated by a transmitter coil, which in turn generates a secondary field in the soil. The heterogeneity of the soil can be quantified by means of the receiver coil and subsequently translated into various soil parameters through the application of mathematical models. This sensor comprises an array of transmitter and receiver coil configurations (R1 – R4) that are specifically designed to detect changes in the upper layers of soil. By analyzing these changes, we can gain insights into various parameters such as soil compaction, water content variation, and soil type. (Grisso, Alley, Holshouser, and Thomason 2009).

- ***Proximal gamma-ray spectrometer sensors*** possess the ability to examine soils with a high spatial resolution and explain inconsistencies in soil characteristics. It has long been recognized that information on the composition of minerals or soil can be obtained from gamma radiation. As early as the 1930s, gamma detectors were developed and employed for the purpose of mineral (uranium) exploration. With the use of advanced technologies, it has been possible to separate the observed radiation into several components, including naturally occurring radioactive elements such as potassium (40K), thorium (232Th), and uranium (238U). A number of studies have investigated the associations between radionuclide concentrations and soil mineral characteristics (de Meijer, 1998).

- ***Direct contact conductivity sensors*** are one of the methods used to estimate Soil EC, Organic Matter, pH, and other parameters. This technique involves creating a small amount of electricity and transferring it into the soil using a pair of rolling electrodes colter disks. Another set of disks is used to quantify the decrease in voltage, which is directly related to the electrical conductivity of the soil at a specific position (Sudduth et al. 2005).

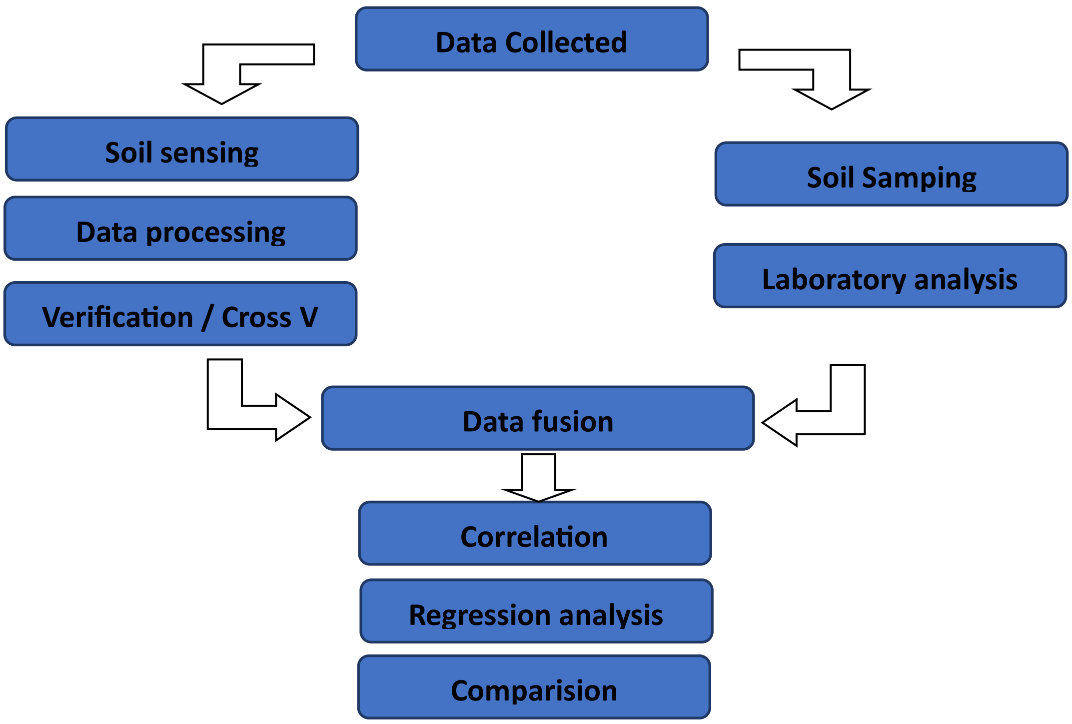
Although many sophisticated soil mapping systems that can be used to detect specific soil properties, one single system capable of responding to all soil properties does not yet exist (Adamchuk et al. 2011; Mahmood et al. 2012)

Multiple sensor systems are currently accessible in the market, and there is a continuous endeavor to create new prototypes. In this study, our main objective is to assess the effectiveness of three commercially available soil sensing systems, specifically the Veris system MSP3, SoilOptix Technologies system, and Geoprospectors system TSM. We aimed to determine their accuracy in estimating different soil parameters such as organic matter content, pH level, cation exchange capacity (CEC), soil texture, and nutrient elements.

In addition, the study utilizes machine learning to investigate the relationship between laboratory analysis and various soil sensors, focusing more on Organic matter content because this parameter can be measured by the three soil sensors.

Workflow

The steps illustrated in Figure below were followed to evaluate the predictability of the soil parameter using the three soil sensors compared to the laboratory



Study Area

The study data was conducted across seven fields namely N40_NE, N40_SW, N40_NW, LCB_HFE, LCB_HFS, LCB_North, and LCB_Canola , in three distinct geographical locations within Stillwater, Oklahoma, each of these fields has an area ranging from 1 to 5 acres, and each field also has distinct environmental conditions and soil characteristics.

The selection of these different locations was intentionally made to represent the wide range of agricultural landscapes that are common in the region.

The study recognized the importance of geographic differences in soil characteristics by conducting experiments in several places.

Soil Sample Collection and Treatment

To collect the data, we adopted a deliberate sampling strategy to evaluate the accuracy of the sensors used. Following the Three soil mapping systems Topsoilmapper (TSM), Veris MSP3, and SoilOptix were used to collect dense geo-referenced sensor measurements. Grid sampling was performed using selected 40-foot sections to ensure a representative composite soil sample. With each section, one composing sample was collected by mixing a minimum of 15 cores at a fixed depth of 15 cm with a total of 456 Samples across all the locations. These soil samples were submitted to the SWAFL lab at Oklahoma State University, Stillwater, for chemical and textural analysis to obtain soil property content that will be used to evaluate and calibrate the soil sensors.

This sampling density aimed to capture the spatial variability of soil properties accurately.

To spatially consolidate the collected data, ArcGIS Pro software was employed, facilitating the creation of a fishnet grid. This grid served as a framework for spatially averaging the dense geo-referenced points while accounting for variations in sensor speed during data collection. This step ensured a comprehensive and standardized spatial representation for subsequent analyses.

Prediction Method and Mapping

Python was employed during this study to create heat maps that visually represented the relationships between the laboratory results and the parameters obtained from the sensors. By utilizing Python's tools, such as Matplotlib and Seaborn, these heatmaps provided an easy-to-understand and detailed representation of the observed correlations.

The process of identifying robust correlations involved the utilization of statistical analytic techniques, which facilitated the identification of significant patterns from the large dataset. The study utilized statistical techniques such as Pearson's correlation coefficient and Spearman's rank correlation to determine the magnitude and direction of correlations between laboratory results and parameters collected by sensors.

The predictive modeling step involved the application of complex algorithms, including the Random Forest model. This model, developed using tools like scikit-learn in Python, aimed to utilize the identified correlations for predictive purposes. Techniques for improving and validating the model have been used to assure the accuracy and precision of the predictive results. The regression coefficient (R2) is used as a quality indicator.

The initial phase of my analysis began by creating correlation heatmaps to explore the relationships between the raw data collected from the three soil sensors and the laboratory results. These heat maps provided a comprehensive overview of the variables' interdependencies. My primary focus in the first step was on understanding the correlation between the sensor readings and the percentage of organic carbon in the soil. Following this exploratory phase, I delved into the application of machine learning techniques.

As the second step, I employed the RandomForestRegressor model to predict the percentage of organic carbon based on the sensor data. This involved training the model on the existing dataset and subsequently extracting the feature importance list to identify the most influential variables. with this information, I proceeded to the final step by re-running the model, this time incorporating only the key features identified in the feature importance list.

By optimizing the model with the best parameters derived from the feature importance analysis, my aim was to enhance the predictive performance and increase the coefficient of determination (R-squared) for a more accurate representation of the relationship between soil sensor data and organic carbon content.