Exploration of machine learning methods for prediction and assessment of soil properties for agricultural soil management: a quantitative evaluation

Sanjay Motia1 and SRN Reddy2
1USICT, GGSIPU, Dwarka, Delhi, India.
2Dept. of CSE, IGDITU, Kashmere Gate, Delhi, India.

E-mail: motia.sanjay@gmail.com

Abstract. Soil is a heterogeneous and complex natural resource that is the factual basis of almost all agriculture production activities. The soil's inherent nutrients or physiochemical properties help the researchers better understand the soil ecosystem dynamics and play a crucial role in guiding farmland decision-makers in their routine decisions. Therefore, the accurate forecasting of soil leads to improved and better soil health management (SHM). The recent advances in sensing and computational technologies have led to the expanding accessibility of farmland data either obtained distantly or proximally. The increasing availability of massive data and unreservedly accessible open-source algorithms have prompted a quickened use of machine learning (ML) procedures to investigate soil conditions. Therefore, to understand the usage of ML techniques in exploring soil properties and related applications, this paper concentrates on reviewing and analyzing ML techniques precisely to predict and assess soil properties for improved decisions on agricultural SHM. The article also explores various other vital factors like algorithms, implementation tools, and performance metrics employed in numerous soil assessment application domains and different challenges and future research directions for SHM using ML techniques. The detailed assessment concludes that the response for ML in the prediction and evaluation of soil properties for SHM is very promising for the sustainable growth of agriculture.

1. Introduction
Agriculture is an indispensable wellspring of food and livelihood for a significant portion of the global population [1]. It plays a vital role in producing both food and employment opportunities for a massive community (2). Today information-driven economy, with its emphasis on creating insightful sensor-based gadgets, data sharing framework, data-driven machines, and sensor network-driven decision support systems (DSS) [3], will play a groundbreaking activity in smart agriculture and horticulture frameworks [4]. The agriculture production systems are mainly affected by environmental and climate conditions, soil ecological characteristics, water availability, and farming practices. Out of these, the soil is the complex and heterogeneous ecosystem whose monitoring, characterization, and management, if not done timely, will lead to its degradation, loss of fertility, reduction in plant growth/yields, and economic losses. Soil degradation emerges as one of the significant challenges for sustainable global food production [5].

The technological advances in remote and proximal sensing and other related technologies like wireless sensor networks (WSNs) [6] have raised the capacity to monitor and acquire more critical spatial and temporal soil data to the manifold, so as the need to accelerate the processing and analysis of information into reliable decisions. The high availability of soil data and open-source ML algorithms on various Govt. and public web domains have raised the usage of machine learning (ML) techniques
to develop further soil analysis and management applications [7].

Due to the high availability of ML algorithms and associated applications in soil management and analysis, plenty of research papers are available in databases on the above said context. Most of the research papers mainly focused on the application of ML for a specific soil analysis problem. However, a comprehensive review of soil properties prediction and assessment considering various technical aspects like type of datasets, algorithms, performance metrics, and applications have not been specifically reviewed and presented in a single study. Therefore, the work presented in the paper concentrates on:

a) Review the existing literature on ML techniques applications on prediction and assessment of agricultural soil properties with a focus on soil nutrient and fertility management.
b) Analyzing the usage of different techniques and associated performance metric in each respective application
c) Exploration of various research issues and challenges, along with a discussion on future research directions.

The next subsection shall discuss the motivation and need for conducting the review and analysis.

1.1 Motivation
In recent times, the emergence of new-age technologies like wireless sensor networks (WSNs) and the internet of things (IOT) has enabled crop growers to adopt a data-oriented approach to collect enormous amounts of data on farm conditions (soil, plants, water) from distributed and networked sensors. Post data collection, high available storage, and efficient ML-based data analytic models utilize this data to build prediction and decision support applications for various on-farm activities like mitigation of pest and disease risk, irrigation, and fertilizer management to improve crop yields to gain economic benefits.

In the context of data analytics and its usage in building important decision-making models, the ML techniques that are one of the fastest-growing areas of artificial intelligence (AI) is playing an important role in the agriculture sector. The domain of ML involves a variety of modeling techniques and algorithms that are utilized in many agricultural applications. This paper comprehensively assesses the potential of ML techniques for soil nutrient management and fertilizer recommendation application.

1.2 Research Approach
Various steps of the adopted research approach in the presented research work are depicted in figure-1.

![Figure 1. Methodology in Research Work](image)

Stage 1 (Planning) defines the scope of study and formulate the research question, in which the information sources and keywords related to the research questions are identified and finalized. To accomplish this, six keywords, i.e., 'Machine Learning,' 'Soil Fertility,' 'Agriculture,' 'Soil Nutrients,' 'Soil Properties,' 'Fertilizer,' were identified and selected for the search of relevant articles. Step-2 covers the explicit inclusion, exclusions, and limitation criteria covering article name, publishing year, country,
journal, relevant keywords, and most importantly, relevant content. In step-3, the source of relevant studies that meet the research requirements was identified, selected, and broadly assessed based on content quality. The research content only from peer-reviewed journals and sources that fulfills the search criteria as per defined keywords are only considered for detailed review and analysis. All this is done to reduce bias errors. The relevant information on machine learning techniques, associated performance metrics, application domain, limitations, and merits were extracted, collected, and analyzed, and a detailed summary is prepared in step-5.

Post analysis/synthesis of review outcome, the future scope of research is proposed. The comprehensive investigations of various research papers studied in this research work are given in the next section. The remaining portion of the document is organized into sections, where the second section, i.e., section-2, discusses the existing related literature on ML algorithms’ applications in the soil analysis domain. Section-3 presents the applications of ML techniques in predictive analysis of agricultural soils and tools utilized for implementation of ML methods. section 4 explains the ML techniques and their quantitative usage in existing research. Section 5 explores the most widely used performance metrics in predictive analytics using ML techniques followed by the discussion on results and findings. The last section i.e. section 6 concludes the quantitative analysis.

2. Related Work
There are numerous applications in the agriculture domain, where a variety of ML techniques are used to solve different agricultural problems like crop protection, i.e., weed detection, identification of plant diseases, and integrated pest management [8], digital soil mapping [9] - [14] forecasting of crop yields and suitability [15] [16], irrigation recommendation [17], modeling of soil properties like temperature [18] [19] and, soil organic carbon (SOC) [20], development and accuracy assessment of soil grid system [21] and, assessment of soil dryness [22]. Padarian et al. [7] surveyed the usage of ML techniques in soil science applications and emphasizes that ML techniques can be applied to formulate a model that explains the distribution of soils and the reasons for spatial variation and estimation of soil properties [23] [24]. In the given study, overfitting and the interpretability of the ML models are identified as few of the research gaps for better understanding of soil conditions.

As observed from the review of existing surveys, there are various applications where ML techniques are used to solve different problems in the agriculture sector. However, there are applications where the usage of ML techniques has not been adequately reviewed and analyzed. For example, the utilization of ML specifically in prediction and assessment of soil properties, soil and plant nutrients, fertilizer management cum recommendation and, soil classification for the agricultural purpose has not been precisely surveyed, analyzed, and presented. The most widely used ML technique(s) still remains untouched and available in a single concise document. The kinds of datasets, performance metrics for assessment of widely used methods, their limitations, and future research directions have also not collectively available on the said problem domains.

Therefore, to get a detailed analysis of the said context, this review work-study the research work on the application of ML techniques that is conducted in the last five to seven-year or so. Only a few exceptional good older papers (more than ten years old) have also been considered for review. The next subsequent sections shall discuss the applications of ML Techniques for the prediction, assessment and evaluation of soil parameters for soil health management (SHM) in detail.

3. ML Techniques in Prediction and Assessment of Soil Properties: Applications and Tools
In the agricultural production system, the fertile soil plays a key role which is governed by its physical and chemical properties. The functional properties of soil are related to a its capacity to support essential ecosystem services such as inherent productivity, nutrient and water retention, and resistance to soil erosion and fertility degradation. The expanding accessibility of soil data, which is captured either distantly or proximally along with easy availability of open-source algorithms, have prompted a quickened use of ML procedures to investigate soil properties.
3.1. Categorization of ML Applications:

In reference to the context of this review work, the applications related to prediction and assessment of soil can be sub-divided into four main categories like (a) prediction of soil properties (physical and chemical properties) (PSP) [25]-[47], (b) predictive-fertilizer-recommendation (PFR) using soil properties [48]-[53], (c) soil fertility and plant nutrient assessment (SFPN) [54]-[58], (d) soil nutrient prediction (SNP)[59]-[64] and, (e) soil classification (SC)[65]-[75] (Figure-2).

The prediction of soil properties (PSP) covers estimation or forecasting of various properties of agricultural soils. Figure-3 depicts the further classification of applications of ML techniques for the prediction of soil properties. PSP caters for estimation or modeling of multiple soil parameters or properties [66][80] like soil organic carbon (SOC) [25][27][29][40],[41],[42][60][76], Bulk Density (BD) [25][26][33], Cation exchange capacity (CEC) [25][27][28][34], pH [25][27][28][44], soil texture, fractions and course fragments [25], Sand C:N Cmin Thickness [26][29], Calcium (Ca) [27][29], Phosphorous (P) [27][29][62][63], Al, H, H-Al, sum of bases, % of base saturation (V), Al saturation (m), sand, silt/clay (S/C), L. chinensis carbon [62], degree of flocculation [27][29], and Soil Organic Matter (SOM)[28][78], Potassium (K) [28] and, yield monitoring of Corn [28]. Clay content [34] and, soil salinity [43].
The ML techniques has also been applied for estimation of soil types or taxonomic classes [30] [31] [32] [65] [66] [67] [68] [69] [70] [71] [72] [73] [74] [75], soil moisture content (SMC) [37] [38] [39] [60] [82] and, soil temperature (STemp) [45] [46] [47].

Figure 4. Contribution of ML Techniques in Prediction/Assessment of Soil Properties

The graph depicted in fig-4 represents data collected from more than fifty research articles collected from useful repute journals. These papers are extensively studied and reviewed for data collection and chart preparation. Figure 4 shows that the application of soil classification utilizes ML techniques more than others. The prediction of pH, OC, P, and K is the second-highest compared to classification applications. ML’s usage in the estimation of soil temperature, moisture content, texture, and CEC is somewhat similar and near about 13.5 %, which is slightly less than the contribution of ML in the forecasting of the ratio of carbon to nitrogen. Only about 10% of papers show the contribution of ML in predicting key macro-nutrients, i.e., NPK, that are responsible for soil fertility and play an essential role in prescribing fertilizers for crop and soil specific needs.

3.2 Categorization of ML Tools:
Most of the methods have been implemented and executed using standard tools like R Software (with Classification and Regression Training (Caret), ranger, xgboost, nnet, RF, geoR, gmap, and sp packages) [25] [26][28] [3] [31] [32] [40] [41] [51][61] [63][69][71] [72][81], MATLAB [29] [36][37][47] [56] [62][67] [68] [72] [80], Weka [33][41] [52] [72] [74][77] and SPSS [39][44] or any other statistical tool[60][66] [76] [78] [79] [83], Python programming platform libraries (Keras Deep Learning and scikit-learn libraries like xGboost) [38][48][50] [53] [57][64] except a few that utilizes other tools like Microsoft Excel®2010, ARC GIS tools, land-surface model (LSM) [35], IDL (Interactive Data Language), GenStat [55], Digital Signal Processing (DSP) boards with commercial imaging libraries [70], RapidMiner [73] and other gives no information on development tools [27] [34][42][45][46] [49] [54] [58] [59][65][75][82]. The figure-5 shows the contribution of various well-known tools that helps in the analysis of agricultural soils. R software packages are the most popular tools for the implementation of ML techniques in soil analysis applications. The MATLAB and python libraries and the WEKA tool are also a good choice in plenty of applications. The usage of the SPSS tool is less as compared to the mentioned tools. The selection and choice of implementation tools largely depend on the user's skill and the features/functions present in the respective tool that meets the needs.
of the application. The next section shall discuss the ML methods/algorithms used in various soil analysis applications as given in figure-2.

Figure 5. Usage of Tool/Software in ML Implementations

4. ML Techniques in Assessment and Prediction of Soil parameters
This section shall review 57 research papers to analyze the contribution of different ML techniques in soil analysis. The extensive study helps identify the most popular and widely used methods for analyzing soil for the agriculture domain. The assessment and evaluation of soil properties categorized into various applications like soil property analysis, fertilizer recommendation, prediction of soil physio-chemical properties, and nutrient conditions. Furthermore, the quantitative analysis of research papers has been done w.r.t. usage of ML techniques. Table-1 gives the list of ML techniques and the number of papers where a particular method is used. From the analysis of the table-1, SVM, RF, TBM, ANN, and BPNN, along with their variants, are the top five methods used in 25 21,13, 12, and 10 research work papers, respectively. Also, SVM and its variants are the most preferred method in many agricultural applications due to its significant accuracy with less computation power. Also, SVM is useful for both regression and classification tasks.

RF is the second-highest supervised ML technique used for soil analysis and prediction applications. RF builds multiple decision trees and merges them to get a more accurate and stable prediction.

The tree-based models (TBM), like decision trees (DT), are instrumental in predicting and classifying the different properties and soil types. DT helps in establishing the relationship between various parameters of agricultural soils. In DT analysis, observations enter at the root node, and a iterative test is applied to differentiate the measurements into best classes, making groups cleaner. This process of testing and splitting the observations into classes continues till it reaches a terminal node. DT with fewer elements on terminal nodes yields higher accuracy, and reduced ' precision of prediction.
## Table 1. List of ML Techniques and their contribution in soil analysis

| Techniques                              | Research Papers where ML techniques used for SA                                                                 |
|-----------------------------------------|---------------------------------------------------------------------------------------------------------------|
| Random Forest (RF)                      | [25],[26],[28],[30],[31],[32],[41],[42],[47],[51],[52],[53],[57],[58],[63],[69],[71],[72],[80],[81],[82]     |
| Support Vector Machine (SVM) and variants | [28],[30],[31],[36],[37],[39],[42],[46],[52],[54],[57],[58],[59],[60],[62],[65],[66],[69],[70],[72],[73],[74],[75],[80],[81] |
| Tree Based Models (TBM)                 | [30],[31],[33],[35],[54],[55],[57],[65],[68],[69],[72],[75],[81]                                            |
| K-Nearest Neighbor (K-NN)               | [26],[30],[31],[53],[57],[69],[75],[80],[81]                                                                |
| Stochastic Gradient Boosting Model (SGBM) | [25],[28],[31],[58],[72],[80],[82]                                                                           |
| Statistical/Geo-Statistical Model (GSM) | [26],[27],[40],[45],[57],[66],[76],[79]                                                                       |
| Principle Component Analysis (PCA)      | [32],[51],[55],[61]                                                                                            |
| Extreme Learning Model (ELM)            | [29],[36],[47],[64]                                                                                            |
| Naive Bayesian                          | [58],[66],[73],[80]                                                                                            |
| k-Means                                 | [59],[61]                                                                                                     |
| Linear Discriminant Analysis (LDA)      | [30]                                                                                                          |
| Artificial Neural Network (ANN)         | [31],[34],[38],[39],[42],[46],[49],[53],[57],[65],[69],[80]                                                  |
| Backpropagation neural networks (BPNN) and Variants (Resilient/Radial) | [28],[41],[49],[47],[55],[56],[62],[73],[77],[49]                                                           |
| Multilayer Perceptron Neural Network (MLP) | [30],[41],[48],[49],[57],[78]                                                                                 |
| Single-hidden-layer neural networks (NNET) |                                                                                                               |
| Long Short-Term Memory Network (LSTM)-Class of Recurrent Neural Network (RNN) |                                                                                                               |
| Generalized Regression Neural Networks (GRNN) | [47],[54]                                                       |
| Convolutional Deep Neural Network (CDNN) | [42],[49]                                                                                                     |
| Generalized additive models (GAM)       | [26]                                                                                                          |
| Generalized linear regression (GLM)     | [26],[80]                                                                                                     |
| Boosted Regression Tree (BRT)/Extreme Gradient Boosting (xGBrading)/Gradient Boosting and its Variant | [26],[40],[42],[50],[80]                                                                                   |
Ordinary/Partial Least Squares Regression (PLSR) or Multiple Linear Stepwise Regression (MLSR)  
Regression Tree - Cubist (CU)  
Linear Regression Model (LRM) and RR/SR Variant  
Multinomial Logistic Regression (MLogR)  
Bagged Classification Tree (BCT)  
Nearest Shrunken Centroid (NSC)  
Support Vector Regression (SVR) and Variants  
Random Forest Regression (RFR)  
Multivariate/Multiple/Factorial Linear Regression (MLR)  
Least Absolute Shrinkage and Selection Operator (LASSO)  
Principal Component Regression (PCR)  
Elastic Net Regression (ENR)  
Gaussian Process (GP)  
Classification/Regression Trees (CART) Variants  
Deep Learning Regression (DNNR)  
Other Methods

[27],[44],[51],[57],[60],[66],[77],[80],[83]  
[28],[35],[60]  
[28],[33],[50],[51],[57],[66],[80],[81]  
[25],[30],[31],[57],[69],[80]  
[30]  
[30],[31],[69]  
[41],[43],[66],[80]  
[43],[80]  
[34],[37],[43],[54],[55]  
[50],[51],[80],[81]  
[51],[60],[80]  
[51]  
[53]  
[69]  
[39],[80]  
[26],[37],[67],[71],

ANN is the next highest contributor in soil analysis, precisely estimating soil properties and soil nutrients prediction. ANNs are widely used to approximate complex systems challenging to model using conventional mathematical modeling techniques. There is no specific sure-shot method for selecting proper ANNs structure and training algorithm as trial and error is the only method to get the best solution. ANNs have a high prediction capability as compared to their counterparts. BPNN method targets determining the learning rate parameter and training cycle adaptively to get the best value in stock data training to achieve high prediction accuracy.
Figure 6. ML Methods in Prediction of Soil parameters for SHM

Figure 6 and figure 7 depict the ML technique distribution graphically across the number of research papers written to predict soil parameters and soil physiochemical properties, respectively, for soil health management (SHM). As per figure 6, SVM is the top contributor for predicting soil parameters, followed by RF with 25 and 20 papers. TBM, ANN, BPNN, PLSR, MLSR are the next highest contributor to various prediction models. The prediction of physiochemical soil properties is explicitly considered, then RF and ANN come out to equally excellent and top contributors, followed by SVM usage in 9 and 8 numbers of research papers. TM or TBM and MLR are the next highest contributors with 04 numbers of articles w.r.t. estimation of soil's physiochemical soil properties.

Figure 7. ML Techniques in Prediction of Soil Physio-chemical Properties

For fertilizer prediction applications, RF and ridge regression (RR) is the highest contributor in with 06 numbers of research work followed by RR and LASSO used in 04 number of research papers related to fertilizer prescription applications. The other methods, like KNN, SVM, Neural network variants, regression-based methods, are equally participating in fertilizer recommendation applications.
Figure 8. ML Methods used in Fertilizer Recommendation

Figure 9. ML Techniques in Prediction of Soil Nutrients

In prediction of soil nutrients specifically (figure-9), SVM is the highest contributor in with 05 numbers of research work followed by BPNN, PCA, TM or TBM, K-Means and RF used in 02 number of research papers respectively. The other methods, like KNN, SVM, Neural network variants, regression-based methods, are equally participating in fertilizer recommendation applications.

The most commonly used ML models are divided into seven groups: 1) Kernel-based ML models (Gaussian process with linear kernel, Gaussian process with radial basis function kernel, SVM with linear kernel, and SVM with radial basis function kernel); 2) TBM (RF, conditional inference RF, generalized boosted models and cubist); 3) regression models (generalized linear model with stepwise feature selection, multivariate adaptive regression spline and bagged multivariate adaptive regression spline); 4) NN models (ANN); 5) instance-based model (K-NN); 6) Mathematical models and 7) Miscellaneous.
Figure 10. ML based models in Predictive Analysis of Agricultural Soils

The pictorial representation as given in the figure 10 reveals that the collective contribution of regression-based techniques like PLSR, MLSR, Cubist, MLR, GLM, BRT, XP boosting, MlogR, and their variants is higher than any other method used for predictive analysis of agricultural soils discussed in the present study with 52 contributions. The other methods like TBM and NNET follows regression-based methods with 45 and 34 grants in the research work. Therefore, regression models and tree-based models are the most preferred soil analysis methods, i.e., prediction of soil properties, prescription of fertilizers, estimation of soil nutrients, and classification of soils.

5. Performance Metrics

Figure 11 depicts the tools for measuring the performance of the ML based prediction and classification techniques applied in soil analysis applications.

Figure 11. Performance metrics used for various application

The tools have been widely used in various research works listed in the literature. The top six tools are RMSE, R2, MAE, Accuracy and correlation which have been used in 27, 21, 11, 11 and 6 related
works, respectively. Both RMSE and, $R^2$ are the top contenders for usage in maximum soil analysis and prediction applications.

6. Results and Discussion
This research work given in this paper categories the soil analysis domain in different sub-categories like prediction of soil properties, physio-chemical properties, nutrients, fertilizer recommendation etc. In addition, the review work also divides the types of ML techniques in predictive analysis of agricultural soils. The results of analysis reveal that regression-based methods are the most popular one in predictive analysis of agricultural soils. Tree based models and Neural network-based techniques are the next most preferred choices for soil analysis applications such as estimation, classification and fertilizer prescription. The other key findings or results of the review work are as under.

a) SVM and RF are the top ML Methods in prediction of soil parameters for SHM.
b) RF, RR and LASSO are the most widely used techniques in prediction of physio-chemical properties of agricultural soils.
c) RF and RR are the best fitted solution for fertilizer recommendation application.
d) SVM and BPNN are preferred methods for estimation of soil nutrients.
e) Overall regression-based models are the top choice for ML based analysis in soil health management applications
f) RMSE and $R^2$ are the top most tools for performance evaluation of ML models utilized in soil analysis.

7. Conclusion
The literature discussed in the research work has provided a complete understanding of the soil's physio-chemical properties. Also, it gives deep insight into the usage of ML techniques in the analysis and estimation of various physical, chemical, and nutrient properties of the agricultural soil. The work findings also reveal that the ML techniques are the most promising methods for predicting the soil properties in current scenario. In short, there are different techniques for different applications related to predictive analysis of soil. SVM and RF are the top most ML methods in predicting soil parameters as a whole. The estimation of soil physio-chemical properties and essential nutrients, when considered RF individually, RR and LASSO and SVM BPNN are the most preferred methods, respectively. RF and RR prove to be the best fit solution for fertilizer recommendation application.

Furthermore, the RMSE and R-square are the most popular metrics for performance evaluation of predictive algorithms. In summary, finding the best-suited ML technique is a complex task as it requires extensive study and application-specific knowledge. This research will guide the researchers in the soil analysis domain to learn about application-oriented ML techniques that are popular and most widely used in existing research.

References
[1] Arjun KM. 2013 Indian agriculture-status, importance and role in Indian economy. International Journal of Agriculture and Food Science Technology. 4(4), pp.343-346.
[2] FAO. 2017 The future of food and agriculture – Trends and challenges. Rome.
[3] Motia S, Reddy SRN. Conceptual Framework of a Prototype Data Driven Decision Support System for Farmland Health Assessment using Wireless Sensor Network. In 2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON) pp. 215-222. IEEE.
[4] Ip RH, Ang LM, Seng KP, Broster JC, Pratley JE. 2018 Big data and machine learning for crop protection. Computers and Electronics in Agriculture 151: 376-83.
[5] Singh R., and Singh, G. S. 2017 Traditional agriculture: a climate-smart approach for sustainable food production. Energy, Ecology and Environment, 2(5), 296-316.
[6] Georgieva T., Paskova N., Gaazi B., Todorov G., and Daskalov P. 2016 Design of wireless sensor network for monitoring of soil quality parameters. Agriculture and Agricultural Science Procedia, 10, 431-437.
[7] Padarian J., Minasny B., and McBratney A. B. 2020 Machine learning and soil sciences: A review aided by machine learning tools. Soil, 6(1), 35-52.
[8] Behmann J., Mahlein A. K., Rumpf, T., Römer, C., and Pfümmer, L. 2015 A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. Precision Agriculture, 16(3), 239-260.

[9] Meier M., Souza, E. D., Francelino M. R., Fernandes Filho E. I., and Schaefer, C. E. G. R. 2018 Digital soil mapping using machine learning algorithms in a tropical mountainous area. Revista Brasileira de Ciência do Solo, 42.

[10] Heung B., Ho H. C., Zhang J., Knudby A., Bulmer C. E., and Schmidt M. G. 2016 An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma, 265, 62-77.

[11] Sharififar A., Sarmandian F., Malone, B. P., and Minasny B. 2019 Addressing the issue of digital mapping of soil classes with imbalanced class observations. Geoderma, 350, 84-92.

[12] Taghizadeh-Mehrjardi R., Minasny B., Toomanian N., Zeraatpisheh M., Amirian-Chakan A., and Triantafillis J. 2019 Digital Mapping of Soil Classes Using Ensemble of Models in Isfahan Region, Iran. Soil Systems, 3(2), 37.

[13] Zare E., Huang J., and Triantafillis J. 2016 Identifying soil landscape units at the district scale by numerically clustering remote and proximal sensed data. Computers and Electronics in Agriculture, 127, 510-520.

[14] Rial M., Cortizas A. M., Taboada T., and Rodríguez-Lado L. 2017 Soil organic carbon stocks in Santa Cruz Island, Galapagos, under different climate change scenarios. Catena, 156, 74-81.

[15] González Sánchez A., Frausto Solís J., and Ojeda Bustamante W. 2014 Predictive ability of machine learning methods for massive crop yield prediction.

[16] Elavarasan D., Vincent D. R., Sharma V., Zomaya A. Y., and Srinivasan K. 2018. Forecasting yield by integrating agrarian factors and machine learning models: A survey. Computers and Electronics in Agriculture, 155, 257-282.

[17] Goldstein A., Fink L., Meitin A., Bohadana, S., Lutenberg O., and Ravid G. 2018 Applying Machine learning on sensor data for irrigation recommendations: revealing the agronomist’s tacit knowledge. Precision agriculture, 19(3), 421-444.

[18] Bilgili, M. 2010 Prediction of soil temperature using regression and artificial neural network models. Meteorology and atmospheric physics, 110(1-2), 59-70.

[19] Kim S., and Singh V. P. 2014 Modeling daily soil temperature using data-driven models and spatial distribution. Theoretical and applied climatology, 118(3), 465-479.

[20] Gomes L. C., Faria R. M., de Souza E., Veloso G. V., Schaefer C. E. G., and Fernandes Filho E. I. 2019 Modelling and mapping soil organic carbon stocks in Brazil. Geoderma, 340, 337-350.

[21] Heng T, Mendes de Jesus J, Heuvelink GB, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, Shangguan W, Wright MN, Geng X, Bauer-Marschallinger B, Guevara MA. 2017 SoilGrids250m: Global gridded soil information based on machine learning. PLoS one. Feb 16;12(2):e0169748.

[22] Coopsnersmith E. J., Minsker, B. S., Wenzel C. E., and Gilmore B. J. 2014 Machine learning assessments of soil drying for agricultural planning. Computers and electronics in agriculture, 104, 93-104.

[23] Minasny B., and Hartemink A. E. 2011 Predicting soil properties in the tropics. Earth-Science Reviews, 106(1-2), 52-62.

[24] Zhu A. X., Moore A., and Burt J. E. 2006 Prediction of soil properties using fuzzy membership. In 2nd Global Workshop on Digital Soil Mapping, Rio de Janeiro, Brazil (Vol. 4).

[25] Leenaars JG, Wheeler I, Wright MN, Batjes NH, Bauer-Marschallinger B, Blagotić A, Mantel S, Heuvelink G, Mendes de Jesus J, Guevara MA, MacMillan RA. SoilGrids250m: Global gridded soil information based on machine learning.

[26] Beguin J., Fuglstad G. A., Mansuy N., and Paré D. 2017 Predicting soil properties in the Canadian boreal forest with limited data: Comparison of spatial and non-spatial statistical approaches. Geoderma, 306, 195-205.

[27] Pinheiro É. F., Ceddia M. B., Clingensmith C. M., Grunwald S., and Vasques G. M. 2017 Prediction of soil physical and chemical properties by visible and near-infrared diffuse reflectance spectroscopy in the central Amazon. Remote Sensing, 9(4), 293.

[28] Khanal S., Fulton, J., Klopfenstein A., Douridas N., and Shearer, S. 2018 Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. Computers and electronics in agriculture, 153, 213-225.

[29] Masri D., Woon W. L. and Aung, Z. 2015 Soil property prediction: An extreme learning machine approach. In International Conference on Neural Information Processing (pp. 18-27). Springer, Cham.

[30] Brungard C. W., Boettinger J. L., Duniway M. C., Wills S. A., and Edwards Jr. T. C. 2015 Machine learning for predicting soil classes in three semi-arid landscapes. Geoderma, 239, 68-83.

[31] Houkpnatin K. O., Schmidt K., Stumpf, F., Forkuor G., Behrens, T., Scholten, T., ... and Welp, G. 2018 Predicting reference soil groups using legacy data: A data pruning and Random Forest approach for tropical environment (Dano catchment, Burkina Faso). Scientific reports, 8(1), 1-16.

[32] Machado, D. F. T., Silva, S. H. G., Curi, N., and Menezes, M. D. D. 2019 Soil type spatial prediction from Random Forest: different training datasets, transferability, accuracy and uncertainty assessment. Scientia Agricola, 76(3), 243-254.

[33] Bondi, G., Creamer, R., Ferrari, A., Fenton, O., and Wall, D. 2018 Using machine learning to predict soil bulk density distribution of visual parameters: tools for in-field and post-field evaluation. Geoderma, 318, 137-147.

[34] Sarmadian, F., Azimi, S., Keshavarzi, A., and Ahmadi A. 2013 Neural computing model for prediction of soil cation exchange capacity: A data mining approach. Intern. J. Agron. Plant Prod, 4(7), 1706-1712.

[35] Myers W., Linden S., and Wiener G. 2009 A data mining approach to soil temperature and moisture prediction. National Center for Atmospheric Research, Boulder, CO.

[36] Liu Y., Mei L., and Ooi S. K. 2014 Prediction of soil moisture based on extreme learning machine for an apple orchard. In 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (pp. 400-404). IEEE.

[37] Liu Y., Mei L., and Ooi S. K. 2014 Prediction of soil moisture based on extreme learning machine for an apple orchard.
In 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (pp. 400-404). IEEE.

Adeyemi, O., Grove, I., Peets, S., Domun, Y., and Norton, T. 2018 Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling. Sensors, 18(10), 3408.

Cai Y., Zheng W., Zhang, X., Zhangzhong, L., and Xue, X. 2019 Research on soil moisture prediction model based on deep learning. PloS one, 14(4), e0214508.

Martin, M. P., Orton, T. G., Lacarce, E., Meersmans, J., Saby, N. P. A., Paroisissen, J. B., ... and Arroaya, D. 2014 Evaluation of modelling approaches for predicting the spatial distribution of soil organic carbon stocks at the national scale. Geoderma, 223, 97-107.

Were, K., Bui, D. T., Dick, O. B., and Singh, B. R. 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afrormontane landscape. Ecological Indicators, 52, 394-403.

Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A., and Scholten, T. 2020 Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran. Remote Sensing, 12(14), 2234.

Wu, W., Zucca, C., Muhammed, A. S., Al-Shafie, W. M., Fadhil Al-Quraishi, A. M., Nangia, V., ... and Liu, G. 2018 Soil salinity prediction and mapping by machine learning regression in Central Mesopotamia, Iraq. Land Degradation and Development, 29(11), 4005-4014.

Zhang, Y., Bui, S., Shen, H., and Wang, Z. 2018 Estimating temporal changes in soil pH in the black soil region of Northeast China using remote sensing. Computers and Electronics in Agriculture, 154, 204-212.

Ahmad, M. F., and Rasul, G. 2008 Prediction of soil temperature by air temperature; a case study for Faisalabad. Pakistan Journal of Meteorology, 11(5), 19-27.

Kim, S. Y., Park, K. S., and Ryu, K. H. 2018 Outside Temperature Prediction Based on Artificial Neural Network for Estimating the Heating Load in Greenhouse. KIPS Transactions on Software and Data Engineering, 7(4), 129-134.

Feng, Y., Cui, N., Hao, W., Gao, L., and Gong, D. 2019 Estimation of soil temperature from meteorological data using different machine learning models. Geoderma, 338, 67-77.

Moreno, R. H., and Garcia, O. 2018 Model of neural networks for fertilizer recommendation and amendments in pasture crops. In 2018 ICAI Workshops (ICAIW) (pp. 1-5). IEEE.

Nabavi-Pelesaraei, A., Rafiee, S., Mohtasebi, S. S., Hosseinzadeh-Bandbafha, H., and Chau, K. W. 2018 Integration of artificial intelligence methods and life cycle assessment to predict energy output and environmental impacts of paddy production. Science of the total environment, 631, 1279-1294.

Qin, Z., Myers, D. B., Ransom, C. J., Kitchen, N. R., Liang, S. Z., Camberato, J. I., ... and Laboski, C. A. 2018. Application of machine learning methodologies for predicting corn economic optimal nitrogen rate. Agronomy Journal, 110(6), 2596-2607.

Ransom, C. J., Kitchen, N. R., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernández, F. G., ... and Sawyer, J. E. 2019. Statistical and machine learning methods for evaluating soil and weather into corn nitrogen recommendations. Computers and Electronics in Agriculture, 164, 104872.

Bondre, D. A. Prediction of crop yield and fertilizer recommendation using machine learning algorithms.

Coulibali, Z., Cambouris, A. N., and Parent, S. É. 2020 Site-specific machine learning predictive fertilizer models for potato crops in Eastern Canada. bioRxiv.

Li, H., Leng, W., Zhou, Y., Chen, F., Xiu, Z., and Yang, D. 2014 Evaluation models for soil nutrient based on support vector machine and artificial neural networks. The Scientific World Journal, 2014.

Neuhaus, A., Armstrong, L., Leng, J., Diepeveen, D., and Anderson, G. 2014 Integrating soil and plant tissue tests and using an artificial intelligence method for data modelling is likely to improve decisions for in-season nitrogen management.

Kole, S. 2014 Machine learning for soil fertility and plant nutrient management using back propagation neural networks.

Shivnath Ghosh, Santanu Koley 2014 Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks International Journal on Recent and Innovation Trends in Computing and Communication, 2(2), 292-297.

Barbedo, J. G. A. 2019 Detection of nutrition deficiencies in plants using proximal images and machine learning: A review. Computers and Electronics in Agriculture, 162, 482-492.

Keerthan Kumar, T. G., Shubha, C., and Sushma, S. A. Random Forest Algorithm for Soil Fertility Prediction and Grading Using Machine Learning.

Khan, S. A., Satyanarayana, V., and Venuopal, B. An Approach to Predict Soil Nutrients and Efficient Irrigation for Agriculture with Spatial Data Mining.

Morelos, A., Pantazi, X. E., Moshou, D., Alexandridis, T., Whetten, R., Tziotziou, G., ... and Mouazen, A. M. 2016. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosystems Engineering, 152, 104-116.

Heng, J., Leezaara, J. G., Shepherd, K. D., Walsh, M. G., Heuvelink, G. B., Mamo, T., ... and Wheeler, I. 2017 Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. Nutrient Cycling in Agroecosystems, 109(1), 77-102.

Li, Y., Liang, S., Zhao, Y., Li, W., and Wang, Y. 2017. Machine learning for the prediction of L. chinensis carbon, nitrogen and phosphorus contents and understanding of mechanisms underlying grassland degradation. Journal of environmental management, 192, 116-123.

Mohapatra, A. G., Keswani, B., and Lenka, S. K. 2017 Soil NPK prediction using location and crop specific random
forest classification technique in precision agriculture. International Journal of Advanced Research in Computer Science, 8(7).

[65] Suchithra, M. S., and Pai, M. L. 2020 Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters. Information Processing in Agriculture, 7(1), 72-82.

[66] Bhattacharya, B., and Solomatine, D. P. 2006 Machine learning in soil classification. Neural networks, 19(2), 186-195.

[67] Kovačević, M., Bajat, B., and Gajić, B. 2010 Soil type classification and estimation of soil properties using support vector machines. Geoderma, 154(3-4), 340-347.

[68] Sunangala, K., and Nithya, G. 2012 Comparative study on bio-inspired approach for soil classification. International Journal of Computer Applications, 38(4), 32-37.

[69] Shahini Shamsabadi, M., Esfandiarpour-Borujeni, I., Shirani, H., and Salehi, M. H. 2019 Application of soil properties, auxiliary parameters, and their combination for prediction of soil classes using decision tree model. Desert, 24(1), 153-169.

[70] Heung, B., Ho, H. C., Zhang, J., Kudrby, A., Bulmer, C. E., and Schmidt, M. G. 2016 An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma, 265, 62-77.

[71] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochitis, D. 2018 Machine learning in agriculture: A review. Sensors, 18(8), 2674.

[72] Gambill, D. R., Wall, W. A., Fulton, A. J., and Howard, H. R. 2016 Predicting USCS soil classification from soil property variables using Random Forest. Journal of Terramechanics, 65, 85-92.

[73] Sirsat, M. S., Cernadas, E., Fernández-Delgado, M., and Khan, R. 2017 Classification of agricultural soil parameters in India. Computers and electronics in agriculture, 135, 269-279.

[74] Harlanto, P. A., Adj, T. B., and Setiawan, N. A. 2017. Comparison of machine learning algorithms for soil type classification. In 2017 3rd International Conference on Science and Technology-Computer (ICST) (pp. 7-10). IEEE.

[75] Barman, U., and Choudhury, R. D. 2020 Soil texture classification using multi class support vector machine. Information Processing in Agriculture, 7(2), 318-332.

[76] Rahman, S. A. Z., Mitra, K. C., and Islam, S. M. 2018 Soil classification using machine learning methods and crop suggestion based on soil series. In 2018 21st International Conference of Computer and Information Technology (ICICT) (pp. 1-4). IEEE.

[77] Abdolmohamad, S., Ildoromi, A., and Heshmati, M. 2017 Factors Affecting SOC and NPK in the Rangeland, Forest and Agriculture; Case Study Halashi Catchment, Kermanshah, Iran. Journal of Geoscience and Environment Protection, 5(13), 18.

[78] Qi, H., Paz-Kagan, T., Karmel, A., Jin, X., and Li, S. 2018 Evaluating calibration methods for predicting soil available nutrients using hyperspectral VNIR data. Soil and Tillage Research, 175, 267-275.

[79] Ayoubi, S., Shahri, A. P., Karchegani, P. M., and Sahrawat, K. L. 2011 Application of artificial neural network (ANN) to predict soil organic matter using remote sensing data in two ecosystems. Biomass and remote sensing of biomass, 181-196.

[80] Bhunia, G. S., Shit, P. K., and Chattopadhyay, R. 2018 Assessment of spatial variability of soil properties using geostatistical approach of lateritic soil (West Bengal, India). Annals Of Agrarian Science, 16(4), 436-443.

[81] Sirsat, M. S., Cernadas, E., Fernández-Delgado, M., and Barro, S. 2018 Automatic prediction of village-wise soil fertility for several nutrients in India using a wide range of regression methods. Computers and electronics in agriculture, 154, 120-133.

[82] Osco, L. P., Ramos, A. P. M., Faia Pinheiro, M. M., Moriya, É. A. S., Imai, N. N., Estrabis, N., ... and Li, J. 2020 A Machine Learning Framework to Predict Nutrient Content in Valencia-Orange Leaf Hyperspectral Measurements. Remote Sensing, 12(6), 906.

[83] Wang, B., Hipsey, M. R., and Oldham, C. 2020 ML-SWAN-v1: a hybrid machine learning framework for the prediction of daily surface water nutrient concentrations. Geoscientific Model Development Discussions, 2020, 1-29.

[84] Zhu, H., Sun, R., Xu, Z., Lv, C., and Bi, R. 2020 Prediction of Soil Nutrients Based on Topographic Factors and Remote Sensing Index in a Coal Mining Area, China. Sustainability, 12(4), 1626.