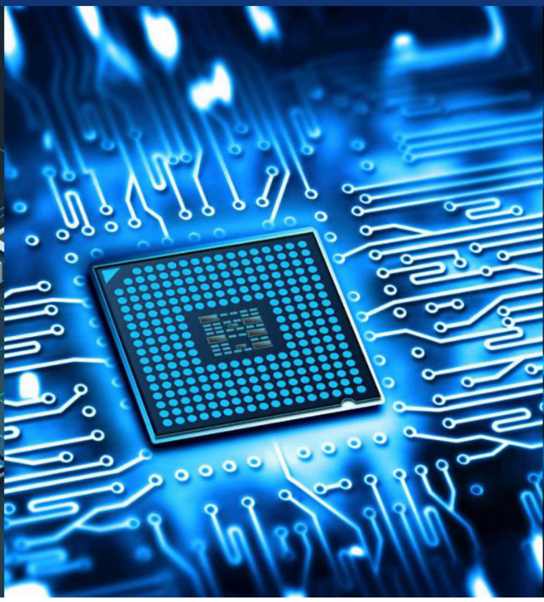
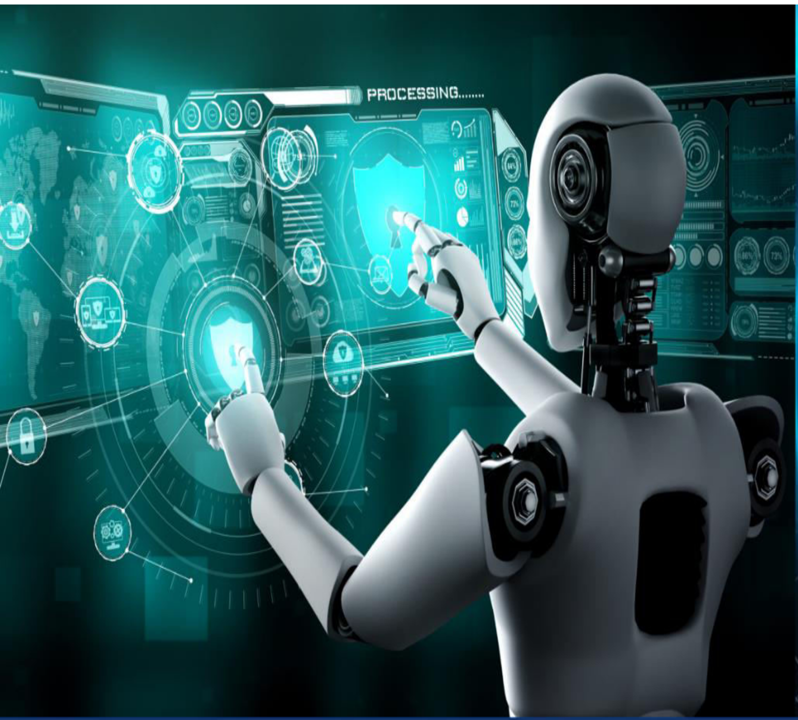
ISSN(O): 2320-9801

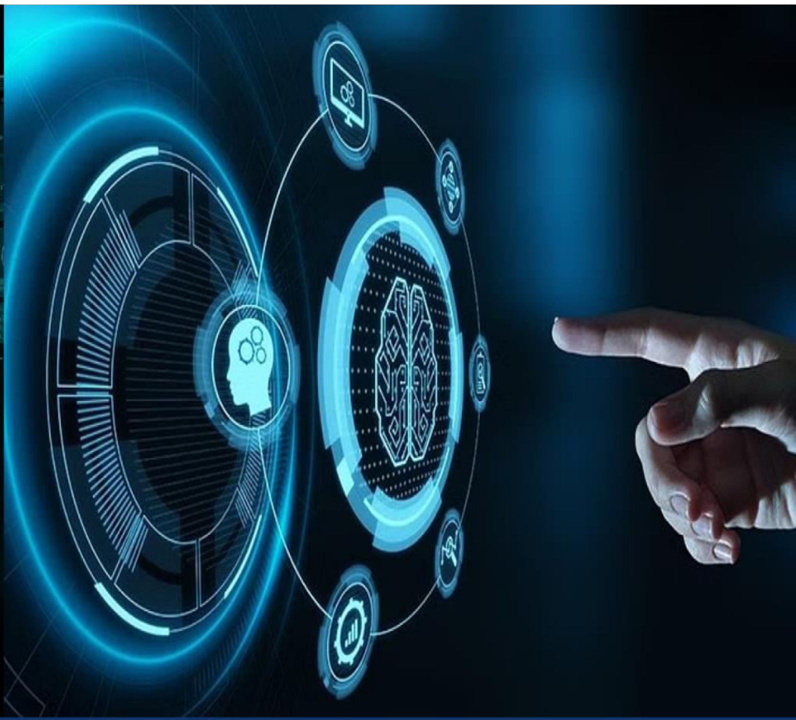
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 3, March 2025

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438 www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Drives Crop Production with Soil Carbon Sequestration using Curse of Dimensionality for Moisture Retrieval

S.Pavithra, Dr.K.Venkatasalam, Mrs.R.Charumathi M.E

Assistant Professor, Department of IT, KSR College of Engineering, India

Professor, Department of Computer Science & Engineering, Mahendra Engineering College (Autonomous), Mallasamudram, Namakkal, India

Assistant Professor, Er.Perumal Manimekalai College of Engineering, Hosur, India

ABSTRACT: The correlation decreases a bit when only meteorological data are used as input (Approach II). There are significant decreases in correlation values in Approach III, and these results are not acceptable. This application did not use meteorological variables at a future time step due to difficulties associated with their accurate estimation. If one could obtain a reasonable estimate of these variables, it is expected that results would improve significantly. The SVM results are also compared with a commonly used learning tool, the Artificial Neural Network (ANN) to validate the capabilities of SVM. It has been shown that the SVM performed better than ANN in all cases. The application presented above shows prospects for the use of statistical learning theory to predict highly complex processes that are difficult to understand and simulate using physics-based approaches, soil moisture being one of them. The SVM approach also has a strong mathematical basis and can be used to address other hydrologic phenomena.

KEYWORD: fractal dimension; feature extraction; gradient-boosting regression model; Soil spectroscopy

I. INTRODUCTION

Quantitative assessment of soil properties using visible near-infrared shortwave infrared (Vis-NIR-SWIR) spectroscopy has been demonstrated as a fast and non-destructive method. Over the past 30 years, numerous soil physical and chemical properties, such as soil texture, soil organic carbon (SOC), cationic exchange capacity (CEC), total nitrogen (N) and exchangeable potassium (K), have been investigated using the spectroscopic approach based on various multivariate statistics and machine learning approaches, and outcomes were applied in soil contamination, soil degradation, environmental monitoring and precision agriculture. As one of the attractive advantages, soil spectra can be recorded at points or by imaging from different platforms. The technique is mainly used in the laboratory, where soil samples are prepared and measured under controlled conditions, and it can be considered as an alternative to traditional analytical techniques. Portable Vis-NIR-SWIR spectrometers allow measurements operated directly in situ. Although the estimation accuracy is lower when compared to results achieved in the laboratory due to uncontrollable environmental factors in the field, in situ proximal sensing improves the efficiency of soil data collection by avoiding tedious sampling and preparation procedures. Sensors can also operate from high above, termed as air- or spaceborne imaging spectroscopy. However, there are still some limitations with respect to the application of imaging spectroscopy to the field of soil analysis, especially when vegetation is present. They have already shown the potential to map and quantify soil properties. With upcoming spaceborne sensors, like the Environmental Mapping and Analysis Program (EnMAP) from Germany and the Hyperspectral Infrared Imager (HyspIRI) from the USA, imaging spectroscopy provides the opportunity to map soil properties at regional and global scales at comparatively low costs.

Feature extraction has been proved to be successful in imaging-spectroscopy classification. The high-dimensional spectral data can be projected to a lower dimensional space with feature extraction methods, without actually losing significant information. Reduced features may increase the separation between spectrally similar classes and the classification model can perform well with a reduced number of features. In soil spectroscopy, a common approach is principal component analysis (PCA). In, PCA was used to reduce the Vis-NIR-SWIR data with more than 2000 wavelengths to a few components, the first component of which accounting for the largest variance. Also, soil information contents of the spectra consisted of PCA components, and a predictive spatial model was developed across Australia. Effective information can also be extracted with wavelet analysis. It can substantially reduce the factors

© 2025 IJIRCCE | Volume 13, Issue 3, March 2025| DOI: 10.15680/IJIRCCE.2025.1303082

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

outside the parameters to the spectrum directly or indirectly. PCA and local linear embedding (LLE) have, in a comparative way, been exploited for soil spectral distance and similarity in projected space. LLE is a nonlinear dimensionality reduction method. It can identify the underlying structure of a manifold, while PCA maps faraway data points to nearby points in the plane. The results indicate that the distances computed in the raw space have comparatively lower performance than the ones computed in low reduced spaces. Methods using PCA and LLE with Mahalanobis distance outperformed other approaches. It can be seen that an effective feature extraction method has the potential to explore the intrinsic structure of spectra, and does not only reduce the data redundancy but also improves estimation accuracy.

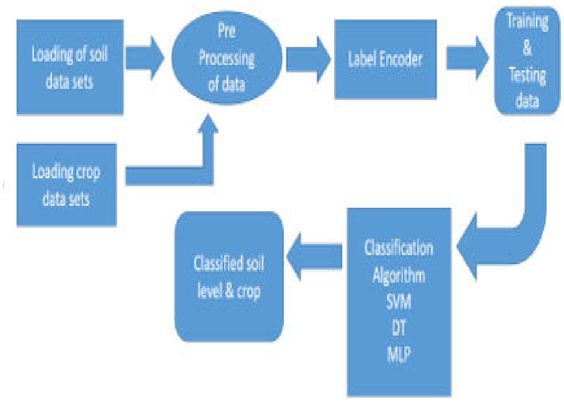


Fig 1: Work Flow Dimensionlity For Moisture Retrieval

Knowing how to effectively extract features from the spectra is crucial for a successful soil-spectral quantitative model. Studies focused on feature extraction from soil Vis-NIR-SWIR spectra are still limited. In this paper, we adopt a novel approach of fractal features based on fractal geometry using variation estimators with the different power indices 0.5, 1.0 and 2.0, which can be termed as rodogram, madogram and variogram, respectively. The concept of fractal dimension was introduced by to reduce the dimensionality of imaging spectroscopy data. Kriti Mukherjee proposed a method to generate multiple fractal-based features from imaging spectroscopy data and then further compared the performance of fractal-based dimensionality reduction using Sevcik's, power spectrum and variogram methods with conventional methods like PCA, minimum noise fraction (MNF), independent component analysis (ICA) and decision boundary feature extraction (DBFE) methods. They concluded that the classification accuracy is similar but the computational complexity is reduced. The aims of the present study are to explore fractal-based feature extraction from soil spectra and to examine its performance on the estimation of SOC, N and pH contents with soil Vis-NIR-SWIR diffuse reflectance spectra. Features generated by the fractal method were compared to PCA-transformed components, and then these two kinds of features were combined to quantify soil properties using a gradient-boosting regression method. The proposed method is further compared with partial least squares (PLS) regression, which is a frequently adopted method for the quantification of soil properties.

II. BACKGROUND OF WORK

The soil water storage capacity is critical for soil management as it drives crop production, soil carbon sequestration, and soil quality and health. It depends on soil textural class, depth, land-use and soil management practices; therefore, the complexity strongly limits its estimation on a large scale with conventional-process-based approaches. In this paper, a machine learning approach is proposed to build the profile of the soil water storage capacity. A neural network is designed to estimate the soil moisture from the meteorology data input. By taking the soil moisture as a proxy in the modelling, the training captures those impact factors of soil water storage capacity and their nonlinear interaction implicitly without knowing the underlying soil hydrologic processes. An internal vector of the proposed neural network assimilates the soil moisture response to meteorological conditions and is regulated as the profile of the soil water storage capacity. The proposed approach is data-driven. Since the low-cost soil moisture sensors have made soil moisture monitoring simple and the meteorology data are easy to obtain, the proposed approach enables a convenient way of estimating soil water storage capacity in a high sampling resolution and at a large scale. Moreover, an average root mean squared deviation at 0.0307m3/m3 can be achieved in the soil moisture estimation; hence, the trained model can be deployed as an alternative to the expensive sensor networks for continuous soil moisture monitoring. The proposed approach innovatively represents the soil water storage capacity as a vector profile rather than a single value indicator. Compared with the single value indicator, which is common in hydrology, a multidimensional vector can encode more information and thus has a more powerful representation. This can be seen in the anomaly detection demonstrated in the paper, where subtle differences in soil water storage capacity among the sensor sites can be captured even though these sensors are installed on the same grassland. Another merit of vector representation is that advanced numeric methods can be applied to soil analysis. This paper demonstrates such an advantage by clustering sensor sites into groups with the unsupervised K-means clustering on the profile vectors which encapsulate soil characteristics and land properties of each sensor site implicitly.

Soil moisture represents the water content of the soil, which is strongly affected by the storage and movement of water in the soil. Several indicators have been proposed to infer the ability of holding water in soil such as saturated water content and field capacity. However, these indicators are static measurements of the amount of water in the soil at a specific time. They do not take into account the variability in soil moisture and the changes in soil properties or climatic conditions over time. The same weaknesses are also shared in a soil water characteristic curve (SWCC), which represents a single snapshot of the soil's water-holding capacity at a given point in time. Water storage capacity of soil, on the other hand, is not limited to a specific point in time. It describes the amount of water that a soil can hold under various moisture levels over a range of time periods. It takes soil dynamics into account as well as environmental factors, such as precipitation, evapotranspiration, etc.; thus, the modelling of water storage capacity becomes very complicated and difficult. For example, the space between soil particles can be filled with water as well as air, the physicochemical interactions between soil and water can alter the density of soil water, and the relationship between soil moisture and runoff responses can be nonlinear and is attributed to many factors such as topography, soil properties, vegetation, etc. Many methods have been proposed to model the water storage capacity of soil from various perspectives, such as pore geometry, soil physical properties, initial wetness conditions, soil texture and organic matter, hydrological soil properties, etc. However, it is impossible to take all impact factors explicitly into account in a model.

Recently, the data-driven approach, which infers soil information directly from the data without considering the underlying physical processes, has become popular. Following this trend, in this paper, a neural network approach is proposed to build a profile of soil water storage capacity, without knowing the principle of water conservation or the governing processes such as infiltration or evapotranspiration, etc., a priori, but learning them entirely from the data supplied. The proposed neural network is based on LSTM, a type of recurrent neural network capable of capturing highly nonlinear relationships and handling long-term dependencies in sequential data. The neural network takes the meteorology data as predictor variables and the in situ soil moisture as target variables. Seven months of in situ soil moisture data from 10 capacitance-based sensors deployed on 10 experimental sites, together with corresponding meteorology data, are collected to build the models. The cell state vectors in the built LSTM models are then extracted out as the profiles of the soil water storage capacity for the 10 sensor sites. Comparing to single value indicators, a multidimensional vector has the ability to encode the soil responses to various impact factors over time and thus is a more powerful representation. The profile vector encapsulates soil properties and dynamics implicitly, and thus provides a convenient tool for further soil analysis with numerical methods, which will be demonstrated in this paper for anomaly detection and categorization.

III. METHODS

The proposed model outputs a sequence of predicted soil moisture values. These values are compared to in situ soil moisture measurements which are used as the ground truth in a mean squared error (MSE) loss function during the model training. The training is optimized with the stochastic gradient descent method under L2 regularization. The initial states $h0$ and $C0$ in LSTM are generally set to zeros for each training sequence in every training epoch. However, our neural network is designed to learn the long-term mechanism of interaction between the meteorology data and soil moisture response implicitly, where the cell state C is modelled as the profile of the water storage capacity in the soil and it is nonlinearly affected by many factors, such as vegetation, soil properties, land surface topography, etc. Therefore, in the training, the cell state C starts from a vector with all zeros but keeps updating with every training

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

sequence for all the training epochs, such that the cell state can be continuously regulated by many hydrological factors implicitly through the training data. Algorithm 1 shows how the training updates the cell state of our model.
To train a based model, 10% data are sampled from each dataset without overlapping each other and are pooled together to form the training set as well as the validation and testing sets. After removing a few corrupted sequences, the pooled dataset contains 805 sequences for training, 258 sequences in validation, and 368 sequences for testing.

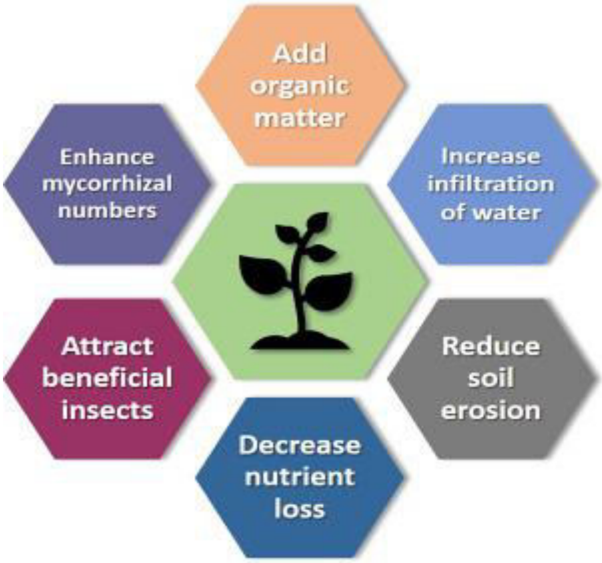


Fig 2: Soil health in agricultural ecosystems

The soil moisture response following rainfall events is unique for each location. This could confuse the learning during the base model training. However, all 10 sensors are installed on grassland. The in situ soil moisture readings from different sites should reveal some common responses of grassland to precipitation, condensation, and evaporation. The training, therefore, can still converge to a certain level, and some common characteristics of grassland would be encoded in the trained model.

The value of every point on a plot line is a Euclidean distance between two profile vectors which are from the same model training but 10 epochs apart. Along with the training, it can be seen that for all the sensor sites, the profile differences become small and the vectors become stable, even though there is still a bit of oscillation near the end of each model training. The cell state vector in our model is trained without explicit knowledge of the hydrological processes but is continuously regulated by the data with hydrological information embedded. When the cell state vector becomes stable in the training, it is deemed that the behaviour of the hydrological system has been deduced from the data and captured in the cell state vector. Our method estimates soil moisture, an indicator of the quantity of water existing in soil, from readily observed meteorology data; we, therefore, believe that the cell state vector of the model has learned the water storage capacity of the soil from the training and the vector can be used to characterize the soil in the numerical analysis, as demonstrated in the rest of this section.

IV. RESULT ANALYSIS

Soil texture has been found to play a crucial role in ecosystem health, agricultural production, and sustainable farmland management. Among the diverse soil properties, the texture plays a pivotal role in decision-making for the planning and management of agricultural land. The conventional approaches with agriculture sensors and statistical analysis were found to be non-robust, time-consuming, non-instantaneous, and expensive. However, with advanced AI processing tools and ML applications, new avenues for texture prediction and revolutionized soil management practices have been opened.

Conventional soil texture analysis is performed by sieving, sedimentation, and other hydrometric laboratory methods. Later, the results from these experiments are statistically analyzed, and conclusions are drawn manually. The complexity of this analysis can be presumed from the variable soil textures and environmental attributes that affect it. This creates heaps of data that cannot be translated into a single conclusion for correct decision-making. Therefore, all of these manual dealings require skilled professionals, a significant amount of time, and specialized instruments. However, AI tools are a promising set of alternatives for these limitations that otherwise confine soil management.

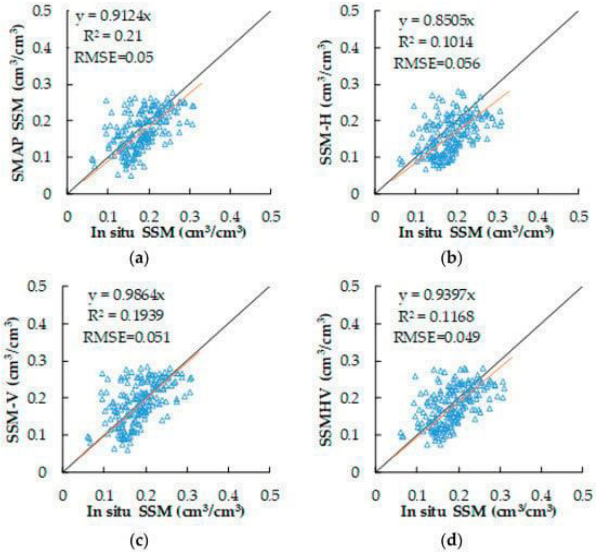


Fig 3: Result Analysis Soil Moisture

AI techniques that include machine learning (ML) and deep learning (DL) are potentially remarkable for accurate and efficient soil texture predictions. The inputs utilized by these algorithms are compositional, spectral, and geographical data sets that can be in non-numerical form. AI processing of these data sets mainly reduces the cost, time, and labor involved compared with conventional laboratory protocols. The complexity of relationships among the data sets is quickly learned and applied using the ML and DL algorithms. This is not the only scale available with this technology; cloud systems and mobile applications are another step forward. The wider scalability of AI enables farmers and land managers to access and process land management operations with ease.

V. CONCLUSIONS

Data acquisition with Vis-NIR-SWIR spectroscopy is relatively easy, and a wide range of soil properties can be analysed within a comparatively short time with relatively little effort for sample preparation. Soil spectroscopy has recently been identified as a method that has the potential to rapidly estimate soil properties. Many soil-spectral libraries are already built at regional, continental or even global scales. Various multivariate statistics methods have been successfully adopted to explore the relationship between soil spectra and soil physical/chemical properties. However, few studies are focused on feature extraction from measured soil spectra, which is also crucial to correlating spectra with soil properties.

REFERENCES

1. Viscarra Rossel, R.A.; Behrens, T.; Ben-Dor, E.; Brown, D.J.; Demattê, J.A.M.; Shepherd, K.D.; Shi, Z.; Stenberg, B.; Stevens, A.; Adamchuk, V.; et al. A global spectral library to characterize the world's soil. *Earth Sci. Rev.* **2016**, *155*, 198–230. [**Google Scholar**] [**CrossRef**]
2. Stevens, A.; Nocita, M.; Tóth, G.; Montanarella, L.; van Wesemael, B. Prediction of soil organic carbon at the European scale by visible and near infraRed reflectance spectroscopy. *PLoS ONE* **2013**, *8*. [**Google Scholar**] [**CrossRef**] [**PubMed**]

3. Chabrillat, S.; Ben-Dor, E.; Viscarra Rossel, R.A.; Demattê, J.A.M. Quantitative soil spectroscopy. *Appl. Environ. Soil Sci.* **2013**, *2013*, 616578. [**Google Scholar**] [**CrossRef**]
4. Stenberg, B.; Viscarra Rossel, R.A.; Mouazen, A.M.; Wetterlind, J. Visible and near infrared spectroscopy in soil science. *Adv. Agron.* **2010**, *107*, 163–215. [**Google Scholar**]
5. Nocita, M.; Stevens, A.; van Wesemael, B.; Aitkenhead, M.; Bachmann, M.; Barthès, B.; Ben-Dor, E.; Brown, D.J.; Clairotte, M.; Csorba, A.; et al. Soil spectroscopy: An alternative to wet chemistry for soil monitoring. *Adv. Agron.* **2015**, *132*, 139–159. [**Google Scholar**]
6. Ben-Dor, E.; Taylor, R.G.; Hill, J.; Demattê, J.A.M.; Whiting, M.L.; Chabrillat, S.; Sommer, S. Imaging spectrometry for soil applications. *Adv. Agron.* **2008**, *97*, 321–392. [**Google Scholar**]
7. Rossel, R.A.V.; Behrens, T. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **2010**, *158*, 46–54. [**Google Scholar**] [**CrossRef**]
8. Ramirez-Lopez, L.; Behrens, T.; Schmidt, K.; Stevens, A.; Demattê, J.A.M.; Scholten, T. The spectrum-based learner: A new local approach for modeling soil Vis-NIR spectra of complex datasets. *Geoderma* **2013**, *195*, 268–279. [**Google Scholar**] [**CrossRef**]
9. Soriano-Disla, J.M.; Janik, L.J.; Viscarra Rossel, R.A.; MacDonald, L.M.; McLaughlin, M.J. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Appl. Spectrosc. Rev.* **2014**, *49*, 139–186. [**Google Scholar**] [**CrossRef**]
10. Epema, G.F.; Kooistra, L.; Wanders, J. Spectroscopy for the assessment of soil properties in reconstructed river floodplains. In Proceedings of the 3rd EARSeL Workshop on Imaging Spectroscopy, Herrsching, Germany, 13–16 May 2003; pp. 13–16.
11. Udelhoven, T.; Emmerling, C.; Jarmer, T. Quantitative analysis of soil chemical properties with diffuse refectance spectrometry and partial least-square regression: A feasibility study. *Plant Soil* **2003**, *251*, 319–329. [**Google Scholar**] [**CrossRef**]
12. McBratney, A.B.; Minasny, B.; Viscarra Rossel, R.A. Spectral soil analysis and inference systems: A powerful combination for solving the soil data crisis. *Geoderma* **2006**, *136*, 272–278. [**Google Scholar**] [**CrossRef**]
13. Shepherd, K.D.; Walsh, M.G. Infrared spectroscopy—Enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. *J. Near Infrared Spectrosc.* **2007**, *15*, 1–19. [**Google Scholar**] [**CrossRef**]
14. Tóth, G.; Hermann, T.; Da Silva, M.R.; Montanarella, L. Heavy metals in agricultural soils of the European Union with implications for food safety. *Environ. Int.* **2016**, *88*, 299–309. [**Google Scholar**] [**CrossRef**] [**PubMed**]
15. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75. [**Google Scholar**] [**CrossRef**]
16. Ji, W.; Li, S.; Chen, S.; Shi, Z.; Viscarra Rossel, R.A.; Mouazen, A.M. Prediction of soil attributes using the Chinese soil spectral library and standardized spectra recorded at field conditions. *Soil Tillage Res.* **2016**, *155*, 492–500. [**Google Scholar**] [**CrossRef**]