Volume 9, Issue 5, May – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

Advanced Modelling of Soil Organic Carbon Content in Coal Mining Areas Using Integrated Spectral Analysis: A Dengcao Coal Mine Case Study

Gill Ammara[1*], Xiaojun NIE[2], Chang-hua LIU[3]
[1*]School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China
[2]School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China
[3]School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China

Corresponding Author: Gill Ammara[1*]

Abstract:- Effective modelling and integrated spectral analysis approaches can advance modelling precision. To develop an integrated spectral forecast modelling of soil organic carbon (SOC), this research investigated a mining coal in Dengcao Coal Mine Area, Zhengzhou. The study utilizes the Lasso and Ranger algorithms were utilized in spectral band analysis. Four primary models employed during this process include Artificial Neural Network (ANN), Support Vector Machine, Random Forest (RF), and Partial Least Squares Regression (PLSR). The ideal model was chosen. The results showed that, in contrast to when band collection was based on Lasso algorithm modelling, model precision was higher when it was based on the Ranger algorithm. ANN model had an ideal goodness acceptance, and the modelling developed by RF showed the steadiest modelling consequences. Based on the results, a distinct method is proposed in this study for band assortment at the earlier stage of integrated spectral modelling of SOC. The Ranger method can be used to check the spectral particles, and RF or ANN can be chosen to develop the prediction modelling based on different statistics sets, which is appropriate to create the prediction modelling of SOC content in Dengcao Coal Mine Area.

This research avails a position for the integrated spectral of Analysis for Advanced Modelling of Soil Organic Carbon Content in Coal Sources alongside a theoretical foundation for innovating portable device for the integrated spectral assessment of SOC content in coal mining habitats. This study might be significant for the changing modelling and monitoring of SOC in mining and environmental areas.

***Keywords:-** Near Infrared and Visible Spectroscopy; Principal Component Analysis; Three-Dimensional Slice Map; Optimal Band Combination Algorithm; Random Forest.*

I. INTRODUCTION

- *Overview of Soil Organic Carbon (SOC) Content in Coal Mining Areas*

Soil organic carbon (SOC) has an important influence on soil quality, environmental protection and carbon dioxide cycling globally. SOC helps a significant part in global ecosystem functioning working as the main source of energy for microorganisms controlling ecosystem productivity and soil structure (Wang, et al., 2020). SOC also helps in maintaining environment, soil health and plant growth as it improves soil structure, nutrient availability, and water retention, supporting strong plant productivity and growth. Linear regression approach employs models such as Partial Least Squares Regression (PLSR), principal component regression, and Multiple Linear Regression (MLR) (Wang, et al., 2022). Subsequently, machine learning techniques are popular in assessing SOC content spectral retrieval which include approaches such as Boosted Regression Tree, Random Forest (RF), Artificial Neural Network and Support Vector Machine (SVM).

In 2020, China produced coal tonnes estimated to be around 3.90×10^9 which is about 50.5% of coal produced in the world (Lu, et al., 2020). Coal mining activities have destroyed land resources, ecological environment and crop production inducing severe problem to the SOC pool. Because SOC has bigger contribution towards decreasing carbon emissions, in the terrestrial ecosystem, China as well as other countries that produce coal ought to make studies which quantify the disturbing impact from coal mining activities on SOC pool to aid scientific managing of SOC pool in places where coal is mined and notify regional low carbon land use. Currently, because of rampant human mining ventures, which leads to shifts in ecological surrounding factors in mining coal subsidence places like surface destruction, land subsidence, underground water hydrology and surface runoff, vegetation destruction and soil erosion (Lu, et al., 2020). Considering Dengcao Coal Mine Area, Zhengzhou the impact on the pool levels of soil carbon remains unavoidable, making the SOC pool in the area experience robust spatial variability. Dengcao Coal Mine Area, Zhengzhou is in Henan province with along historic agricultural fields in the region. However, recently,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

coal mining activities have intensified leading to concerns for investigating SOC in the area to help determine levels of nutrients in the region.

➢ *Significance of Dengcao Coal Mine Area in Zhengzhou*

Zhengzhou is biggest city in Henan province where Dengcao Coal Mine is situated. Dengcao Coal Mine area in Zhengzhou forms one of the leading contributors of coal mining in China providing the main local GDP totalling to about 50% in Hena province (Lu, et al., 2020). As Zhengzhou grows fast among the cities in Henan, coal mining has helped in its growth as employees have risked working in the mine coal areas. Most of the livelihood depends on the income generated from the coal mines.

➢ *Problem Statement*

Climate change has been an increasing issue in the modern world especially, in the mining areas. Mining activities have contributed to the degradation of land through its effect on SOC concentration. Mining affects the level of SOC concentration thus altering the normal environment in each area. Mining activities in Dengcao Coal Mine Area, Zhengzhou have changed land productivity in the areas. For better conservation of the areas around the mining areas, it is significant to investigate the nutrients levels in the areas for better management of the environment. To conserve the mining areas for the future generation, it is ideal to take measures to control further degradation of land nutrients that support the ecosystem. In this regard, investigating regarding soil nutrients availability, is key to ensure that ecological and environmental habitats are maintained.

➢ *Research Questions*

- What are the successful components of integrated spectral modelling technique?
- What are the primary conditions required to conduct an effective modelling?

➢ *Objectives*

- To employ really time monitoring structures of the availability of SOC in Dengcao Coal Mine Area, Zhengzhou that will help in maintaining ecological and environmental balance.
- To ensure that facts are collected about SOC in the area to help conservation management towards reducing adverse environmental degradation in the region a better future restoration of the environment.
- Innovate a modelling that is portable for analysing the SOC in coal mining areas.
- To reduce cost of analysing SOC in coal mining for quick ecological measures to be implemented to conserve the habitats.

➢ *Importance of Accurate SOC Content Modelling for Environmental Management*

Precise SOC content modelling helps in the management of the environment. For instance, ecological managers use precise modelling SOC content sustain soil structure through growing of plants. Soil health is attained through growing SOC contents, when exhausted, can improve carbon storage, advance agricultural activities as well as advance agricultural resilience system when weather conditions are extreme. For that reason, this study will evaluate the use of integrated spectral analysis for advanced modelling of soil organic carbon content in coal sources: a case study of the Dengcao coal mine area, Zhengzhou.

II. LITERATURE REVIEW

A. *Overview of Traditional Methods for SOC Content Analysis in Coal Sources*

Coal is the primary energy source worldwide. As industry and society progress, coal plays a crucial role in environmental pollution and production efficiency. Sorting coal is crucial in the process of coal production and mining as it helps guide resource estimation and production planning. However, in the current mining process, traditional methods of coal classification primarily rely on artificial categorization. This requires a significant number of materials and human resources, making it difficult to achieve automation. Therefore, determining the different types of coal quickly, in real-time, and accurately is an important area of research for the utilization and extraction of coal.

Coal is classified into three main groups - lignite, anthracite, and bituminous coal - based on how much it has been coalified. Different types of coal have diverse applications in different conditions. The quality and type of coal used in thermal power plants play a crucial role in the production process and the design of boilers. Therefore, in coal mining, the traditional production technique is to differentiate mining based on coal category. Two traditional methods for categorizing coal are manual categorization, which is faster but less precise, making automation difficult, and chemical analysis, which provides better accuracy. This method has a drawback due to its extended detection time and expensive nature.

➢ *Proximate Analysis of Coal*

Proximate analysis of coal shows percentages weight of moisture content, fixed carbon, volatile matter, and ash in coal. Fixed carbon is the left solid fuel after distilling volatile matter off in the furnace. In this context, an approximate heating coal value. Volatile matter has hydrogen, hydrocarbon, methane and gases like nitrogen and carbon dioxide found in coal that are incombustible. Normally, volatile matters range between 20% and 30%. Moisture content decreases heat in every kilogram of coal which is between 0.5% and 10%. After incinerating biomass, the moisture component is evaporated burning out volatile matter, the remaining content is referred the ash matter from the biomass (Ferfar, et al., 2024). The remnants matter has no value energy, and it is composed of inorganic

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

contents. When the ash content value in biomass reads past the acceptable confines, it causes challenges in thermochemical procedures such as gasification, pyrolysis, and combustion. when burning coal sample at 750°C temperatures for a duration around 7 minutes. Proximate analysis approach is slow compared with integrated spectral techniques which saves times and less costly. In this regard, the current study will help in reducing expenses which are incurred when using proximate method.

- *Ultimate Analysis*

Ultimate analysis entails that procedure which determines the amount of sulphur, hydrogen, nitrogen, and carbon in inorganic and organic coal sample in both liquid and solid form. Ultimate analysis when testing solid coal involves moisture, oxygen, ash, sulphur, hydrogen, carbon, and nitrogen by difference which are easily taken for reliable testing. This analysis brings the configuration of biomass of a dry ash-free and basis and dry base. An elemental assessor is utilized to define the elemental configuration. A small quantity of crushed biomass sample coal is combusted in a regulated atmosphere to generate gas for analysis. The carbon matter in the content changes to CO2 while similarly, H2 converts H2O, N2 changes NOx, and finally S converts to SO2. The emissions from the burning furnace are put on a highly heated copper purities that extracts any oxygen converting any N2 and NOx. The S, C, N, and H are identified through the generation of SO2, CO2, N2 and H2O gases for recognition (Rossini, et al., 2020). Because ultimate method takes long time to bring the research results as compared with integrated spectral approach will help to provide timely information thus helping ecological management in the region quick.

- *Conventional Kriging Model*

Kriging model entails an interpolation approach that makes forecasts at unsampled areas through a linear grouping of observing at adjacent sampled places. The impacts of all observations using the kriging model is reliant on many factors such as its geographical nearness to the locations unsampled, the spatial planning of all observed facts, like clustering of observed oversampled places. Another factor for consideration is the design involving spatial correlation from statistics collected. The usage of kriging techniques in coal mining is useful only when statistics has spatial correlation (Rossini, et al., 2020). Finally, kriging model in coal analysis of SOC allows data to be measured over various spatial supports which can be mixed and converted to support like upscaling and downscaling, can be done. Comparing kriging method with integrated spectral, kriging samples approach requires so much manpower in relation with spectral approach thus making the technique unreliable in taking instant data for immediate response. Integrated spectral analysis approach provides timely and precise for quick response from the ecological management to safe the situation.

- *The Chemi-Thermal Technique*

Thermal approaches evaluate the coal samples as functions of the time or temperature, and argon or nitrogen is normally utilized to remove the heater and stop coal sample oxidation. Thermal analysis alternatively can be is described as a combination of approaches through which the chemical or physical properties of coal samples, in a reactant and mixture are assessed over function of time or temperature at the same time as the coal sample is put to a regulated temperature control. The controls may use cooling or heating (dynamic), and maintaining the temperature steady (isothermal), in any order of these. The model involves multi-component methods which comprise thermogravimetry, Differential Scanning Calorimetry and Differential Thermal Analysis. These approaches get extensive usage in both research application and quality control on coal mining areas. Thermal analysis involves too many processes that eventually take more time compared to integrated spectral approach. The current integrated spectral method will help reduce used to assess the levels of SOC in the area to allow for faster interventions in saving the ecological conditions.

B. *Introduction to Integrated Spectral Analysis as a Novel Approach*

Integrated spectral analysis model proposes a general system that involves examining the HSI using duo parallel processing ways. One pathway devoted use of spectral characteristics and the second on spatial features, followed by integrating the outcomes of both assessments. Spectral analysis offers ways of evaluating the strength of sinusoidal (periodic) mechanisms of signals at varied occurrences. The Fourier convert assumes input function for the space or time and converts it as a multifaceted frequency function that provides the phase and amplitude of an input function. It is significant to note that the spectral analysis principle is that signals can be disintegrated into a mixture of varied frequencies. Disintegrating this, one requires to do the disintegration through Fourier Transform, converting the signals from time domain into frequency domain. Subsequently, integrated spectral analysis involves performance reliant on nonparametric approaches as well as parametric approaches (Cui, et al., 2020). Nonparametric approaches are depended on differentiating time-domain statistics into segments, employing Fourier transform for all segment, computation of squared-magnitude from the transform, averaging and summing the transform.

C. *Review of Studies Applying Integrated Spectral Analysis in Similar Contexts*

Murindangabo, et al., (2023) carried out a comparable analysis to provide a summary of the methods and approaches employed to assess the variations in Soil Organic Matter (SOM). The research involved finding suitable methods to predict and explain SOM humification, decomposition, and lability for sustainable management practices. Various models and methods for quantitatively and qualitatively evaluating SOM reviews were examined to better understand the changes in organic matter modifications. This assessment aimed to discover the uses, advantages, and constraints of these models and techniques

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

and to identify new research directions in the field. Various parameters, including oxidation kinetics, carbon management index, lability, humification degree, humification ratio, and humification index, can be used to quantitatively analyse research on SOM. Likewise, qualitative evaluation of SOM studies can involve techniques such as oxidizability, electrospray Fourier transform ionization ion cyclotron resonance mass spectrometry, visual inspection, high-performance size exclusion chromatography, odour, assessment of microbial content, plant development, decay, cation exchange capacity, and organic substance identification (Shi, et al., 2024). Additionally, the study highlighted that these factors and methods provide insights into the changes and qualities of SOM, enabling a comprehensive understanding of its transformations.

Hong, et al., (2020) conducted research to estimate Soil Organic Carbon (SOC) levels in topsoil to improve soil fertility and enhance food production quality. The study sought to determine the potential of using aerial hyperspectral imaging to estimate the soil organic carbon content in the uncovered topsoil in the southeastern area of Iowa State, USA. The researchers also observed that enhancing the subtle spectral indicators related to SOC, speeding up the calibration by using a sophisticated predictive capability development system that combines two new spectral algorithms: Optimal Band Combination Algorithm and Fractional-Order Derivative (FOD) for SOC prediction. The outcome of the case relied on 49 soil samples and a scattered aerial hyperspectral imaging. The integration of ideal spectral indices from various FOD changes using the ideal band amalgamation algorithm was utilized to create SOC estimation models with Random Forest (RF). Results indicated that as the fraction increased, overlapping baseline and peak drifts were continuously identified (Hublikar, et al., 2024). However, the level of spectral power diminished at the same time. FOD observed a significant amount of spectral variability and enhanced inconsistency in relation to the original 1st and 2nd derivatives and reflectance. The precision estimates obtained using the optimal band combination algorithm were typically better than those from full-spectrum data. The highest estimation precision for SOC was achieved by the RF method using a combination of 0.75- reflectance order and the best band amalgamation algorithm, reaching a cross-validation of 0.66 R2. This research provides guidance for future studies in selecting the best FOD modification for preprocessing spectral data and in determining the spectral index through the optimal band grouping algorithm. Airborne hyperspectral imaging-based modelling can also be utilized for organizing local-scale agricultural plans by predicting topsoil SOC levels.

Chen et al. (2020) explored the traditional environmental and geological evaluations that rely solely on the extensive efforts of soil specialists and demanding laboratory investigations. Certain aspects of Laser-Induced Breakdown Spectroscopy (LIBS), such as quick and straightforward sample analysis methods like on-site and remote detection, can significantly accelerate the identification of environmental and geological resources. Additionally, combining the LIBS technique with machine learning provides a practical way to enhance the accuracy of quantitative analysis and classification of information obtained from LIBS spectra data sets (McKenna et al., 2022).

Abdulraheem et al. (2023) also state that Remote Sensing (RS) methods offer advantages over other techniques for assessing soil properties, including large-scale imaging, non-destructive qualities, quick data collection, temporal monitoring, and multispectral capabilities. This study emphasizes the different ways of detecting soil properties, the utilization of remote sensing methods, the varieties and components involved in soil measurements, and the pros and cons of measuring soil properties. The choice of methods depends on the specific requirements of the soil capabilities project, as it is crucial to consider the constraints and benefits of each approach and the goals and circumstances of the soil measurements to determine the best remote sensing method.

D. Strengths and Limitation

The current study employs up to date research that takes facts from the ground using integrated spectral approach in reviewing the importance of conserving the ecological condition of the environment in a timely manner. Subsequently, the current study will employ use monitoring sensors that reports immediately to the ecological management department for quicker action as opposed to the traditional methods that relied on the manual sampling of the particles. However, integrated spectral approach has limitation in the need to be used by only trained experts for reliable facts interpretation. Another limitation for using integrated spectral approach is its inability to used when weather conditions are foggy thus hinders real time fact collection. Based on this understanding, it is critical to fill the gap by using modern integrated spectral analysis approach to help in offering information that will guide researchers to find ways to fully maximize the technology.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

III. METHODOLOGY

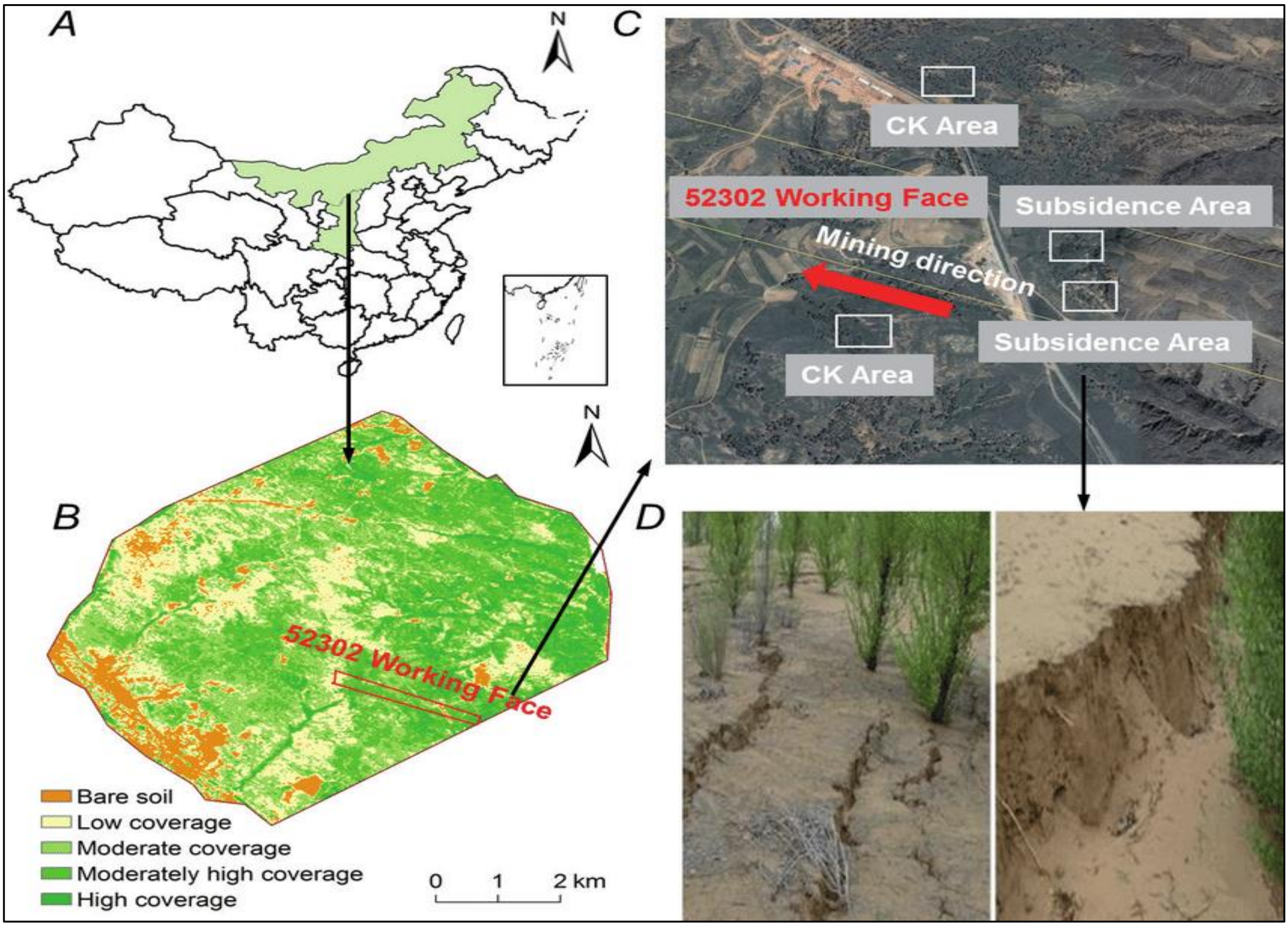

Fig 1 Showing Coal Aerial View *Study Area Dengcao Coal Mine Area, Zhengzhou*
Source: (Deng, et al., 2021)

A. Pre-processing of Spectral Data

In the planning of gathering the spectral information, the exterior light situations of the investigational Dengcao Coal Mine area were constantly varying, and due to failure of calibration not being performed regularly through the white calibration material, there was reference drift issue. Therefore, pre-processing necessity to allow spectral data for the coal samples were placed a uniform reference. First and second order view differentiation of the rates of shift of the spectra through derivatives means, which certainly weaken the impact of changes in exterior state on the testing outcomes.

Because the coal samples for the spectral experiments for this had block coal samples, it was hard for the surfaces to attain the perfect consistency of powdered coal samples; the exterior in homogeneity will create spreading as light goes through reflecting from the coal samples, bringing faults to the coal sample spectra. Standard usually transform less spectral mean from data origination and dividing the result by the coal sample deviation obtaining the normalized coal sample facts (Tao, et al., 2020). Convolutional levelling denoising is employed to eradicate arbitrary noise during operation and advance the ratio of the signal-noise.

Table 1 Sample Sizes

Data	Sample	Range g/kg	Mean g/kg	SD	Skewness	Kurtosis	CV (%)
Modelling set	80	4.99-48.92	29.56	8.1	0.312	0.69	22.89
Validation set	30	13.86-52.67	32.54	2.2	0.378	0.923	6.56

B. Selection of Spectral Indices for SOC Estimation

The narrow-band hyperspectral data contains significant spectral information without interruption. The spectral index method combines two bands to utilize hidden spectral data, enhancing sensitivity to coal particle sizes and improving understanding of the relationship between coal particles and spectrum size. Tao et al. (2020) employ four coal spectral indexes in their study, including normalized differential coal index (NDSI), bare soil index (BSI), difference coal index (DSI), and ratio coal index (RCI).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

C. Calibration and validation of the model

➢ *The PLS Technique*

The popular linear regression approaches used comprise numerous linear regressions, principal component regressions and logistic regressions. PLS, totalling the gains of Principal Component Analysis (PCA), linear regression and Canonical Correlation Analysis (CCA), can address numerous correlations between independent variable quantities and permit modelling to be done when samples numbers are less than the wavelengths (dimensions) of spectra (Li, et al., 2020).

➢ *The RF Technique*

Random forest (RF) is a method that involves training multiple decision trees to predict coal samples. It can handle complex data, and the fitting process is straightforward. Firstly, the bootstrap method randomly selects M tree samples from the training set, which are then used to generate M outcomes. Given the presence of N factors in the training set, the optimal features are selected for splitting, with each node being split by trees until the training samples of the trees fall into the same group. Secondly, every tree can grow to its full height without pruning. Finally, a multitude of decision trees are generated within a random forest to regress or classify new data. In the case of regression, the ultimate prediction result is determined by the mean predicted values of multiple decision trees.

D. Description of Data Collection Methods and Sources in Dengcao Coal Mine Area

The method is divided into three stages: training the dataset, setting up the deep learning technique, and testing the model at various locations within the Dengcao Coal Mine area. The initial step involves preparing the dataset for the Dengcao Coal Mine region by categorizing features. In this scenario, we created datasets labelled "No Coal Mines" and "Coal Mines" to build the model. The Coal Mine preparation dataset generates planning involving VGG Deep Convolutional Neural Networks (Xuesong et al., 2024). The accuracy of the model's classification is evaluated based on how the model performs on the validation set. The prepared model is then tested on concealed tiles using coal mines from various regions. The evaluation of atactic presentation on hidden tiles is determined by comparing model categorization with visually recognized coal mine categorization, using general accuracy, kappa, user's accuracy, and producer's accuracy.

E. Step-by-Step Procedure for Modelling SOC Content Using Integrated Spectral Analysis.

- Step 1: Select a particular column j from the spectral matrix, designate it as Xj in the modelling dataset, and refer to it as XK (0).
- Step two: Define the location of the column vector that was not chosen as "s," where s = {j, 1 ≤ j ≤ J, j ε {k (0), ···, k(n−1)}.
- Step three involves determining the projections of Xj onto the rest of the column vectors. Pxj equals xj minus the inner product of xJ and xK multiplied by xK divided by the inner product of xK with itself raised to the power of -1, where j belongs to the set s.
- Step 4: Identify k(n) as the argument that maximizes the magnitude of ||Pxj||, where j belongs to the set s.
- Step five: xj is not equal to Pxj for all j in the set s.
- Step 6: Here, we increase by 1. If n is still less than N, we return to step 2. This iterative process allows us to refine our analysis and ensure comprehensive results.

Finally, we assign the wavelength to each N and k (0) from the set {xk(n)= 0, ..., N − 1}. This marks the end of our preparation process. We then proceed to conduct Multiple Linear Regression Analysis (MLR) and calculate the internal Cross-Root Mean Square Error (RMSE) of the validation set. The ideal values are those that correspond to the smallest RMSE value, providing us with the most accurate and reliable results.

IV. RESULTS

➢ *Spatial Distribution of SOC Content in the Dengcao Coal Mine Area*

The MSC created a mixed pretreatment influence with four spectral index algorithms for band grouping to improve the relationship between soil particle content and spectral index. Selected foreign and domestic spectral index modelling, along with MSC spectral data, were incorporated into each modelling based on compatible mixed bands, and correlation analysis was conducted to investigate the relevance of different soil particle contents in determining soil texture in this study.

➢ *Spectral Features of Oil Organic Carbon*

Depending on the analysis of the mathematical transformation and spectral curve of soil data in coal soil, a negative correlation was observed between SOC content and spectral reflectance. The overall bands' variation, SOC content curves with R, showed a consistent trend. The R decreased as the amount of SOC increased, indicating a negative relationship. The slope of the curve appears larger for visible light (400-780 nm) compared to the slightly different curve seen for near-infrared (780-2500 nm). In the same way, the spectral curve exhibited absorption peaks at wavelengths ranging from 1300 to 1400 nm, 1750 to 1850 nm, and 2250 nm.

Table 2 SOC Distribution Probabilities

Distribution probability (g)	0.8	2	10	15	25	41	55	60	71	75
SOC	0.5	3	6	8	12	15	18	20	22	26

SOC (g)

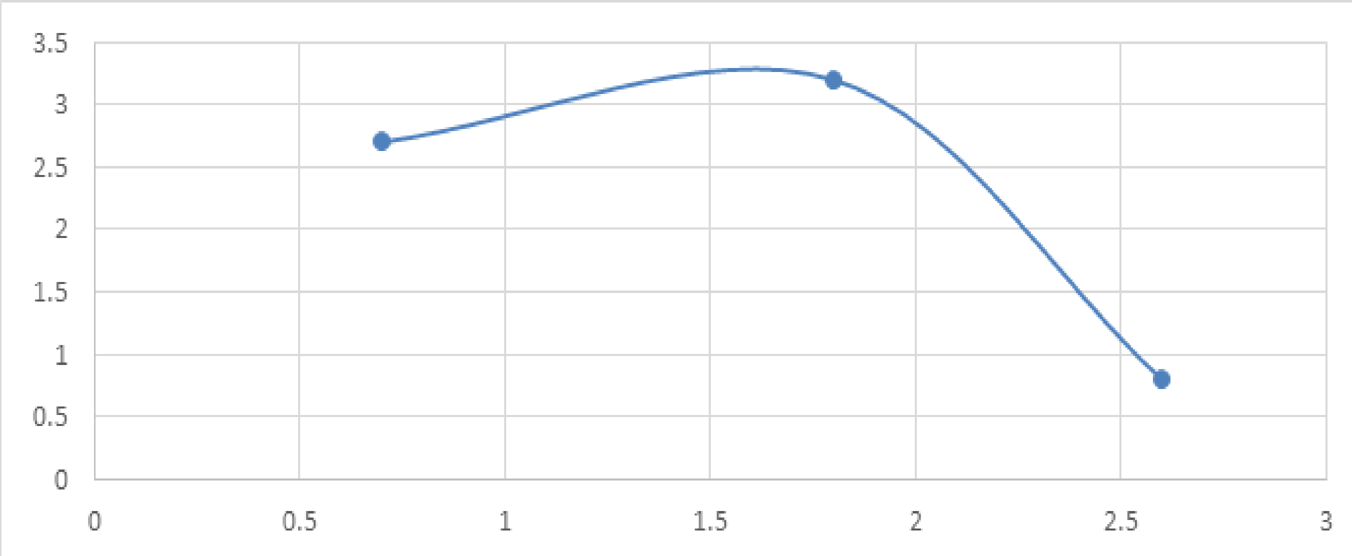

Fig 2 Soil Organic Matter Levels

Figure 2 above shows that soil organic matter (SOM) levels ranged from 16.0 to 19.5 g/kg, making up approximately half of the total distribution. The coal mine soil had SOM content ranging from 19.5 to 25.5 g/kg, accounting for approximately 47% of the total distribution, while SOM contents of over 23.0 g/kg and 13.0-16.0 g/kg made up about 40% of the total distribution. Approximately 69% of the entire distribution in the coal mine consisted of SOM contents.

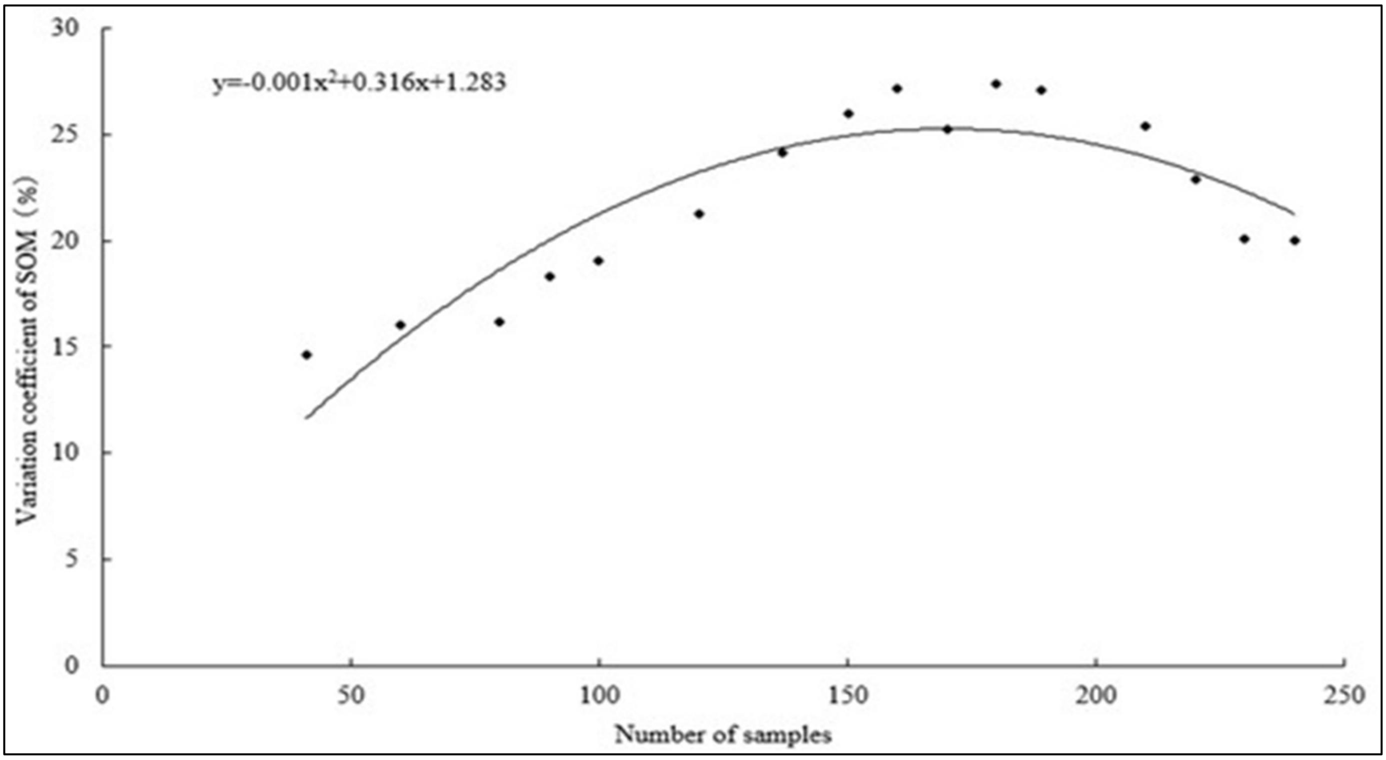

Fig 3 SOM Distribution Levels

Table 3 Sample Number and Variability Coefficients

Sample numbers	10	22	50	89	112	145	178	200	250
Variability coefficient of SOC	5	10	20	25	30	35	40	45	50

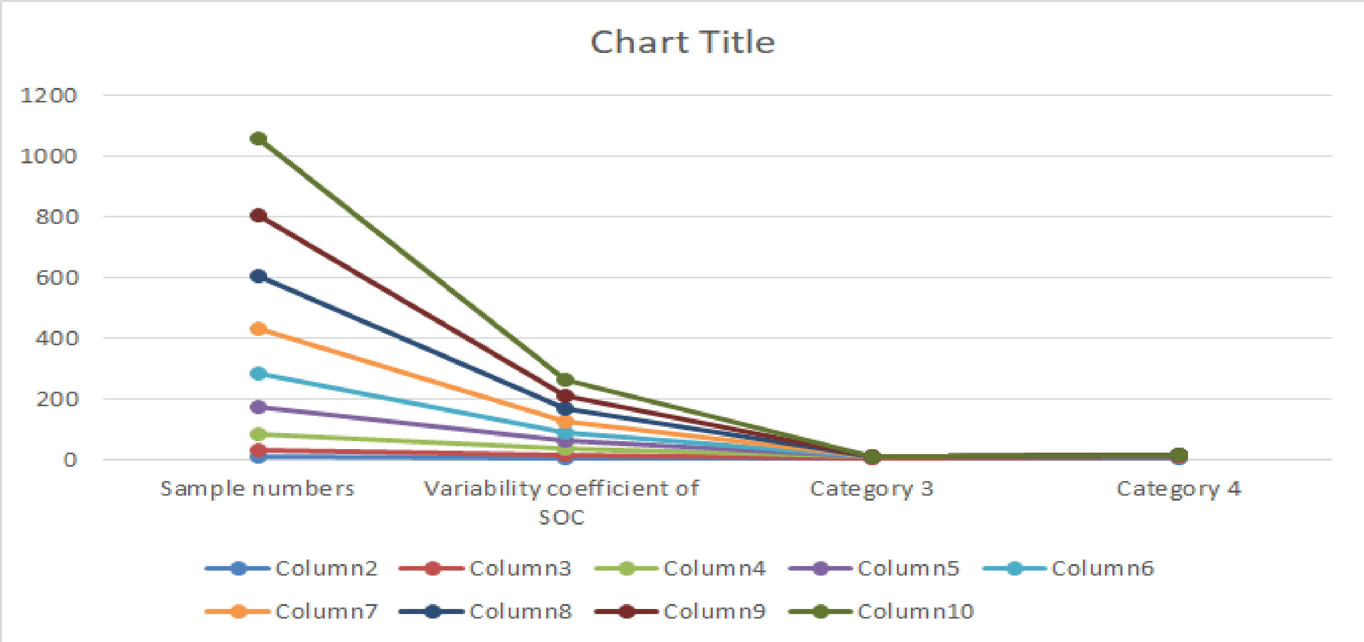

Fig 4 Association between Coefficient Variation of SOM

A research study examined the correlation among SOM samples with spatial coefficient variation, depicted in Figure 3. Among a specific set of samples, the coefficient of spatial variation of soil organic matter contents increased with the addition of soil samples, demonstrating a positive correlation between spatial variation of soil organic matter contents. The outcomes illustrate the spatial variation for SOM can be improved displayed through the three statistical units scale sets and sampling positions 150-180.

To entirely confirm the accuracy for the modelling, the model divided samples into validation and modelling sets, and 50 modelling samples and 20 validating samples selected as per the content and a soil ratio varying from higher to lower. The significant bands from DSI and BSI spectral bands and indices selected by the SPA algorithm were incorporated into the nonlinear (RF) and linear (PLS) models. The R-squared coefficient and RMSE were used to evaluate the model's accuracy. When results approach 1, the model fits more accurately. A smaller root mean square error indicates a closer match between the true value and the predicted value.

➤ *Characteristic Selection of Spectral Lines*

The research examined spectral lines of additional elements in the input data's spectral range of 190-350 nm to improve the accuracy of predictions. In the coal spectrum, the main spectral lines between 190–350 nm consist of carbon, Si, Mg, and Fe lines, along with other distinct and identifiable lines. PLSR was used to establish the linear relationship between the spectral line's intensity and the carbon levels present. Following the depiction of multiple spectral lines ranging from 190 to 350 nm, the R2 value of the standardization curve increased from 0.7892 to 0.9271.

Table 4 Spectral Lines Differences

Element	Spectral Emission Lines (nm)
C	247.86
Si	221.67, 250.69, 251.43, 251.61, 251.92, 252.41, 252.85, 288.16
Mg	279.55, 280.27, 285.21
Ca	315.89, 317.93
Al	308.22, 309.28
Fe	238.20, 239.56, 240.49, 248.81, 258.59, 259.84, 259.94, 260.71, 261.19, 273.96, 275.57, 323.46, 357.03, 358.12

➤ *Results of the Different Calibration Models for the Carbon Content under Different Calibration Sets and Prediction Sets*

Table 5 Calibration Models for the Carbon Content

Algorithm	$R^2 >$ 0.90 (%)	$R^2 >$ 0.80 (%)	R_{max}^2	R_{min}^2	RMSEC$_{avg}$	$R^2 >$ 0.90 (%)	$R^2 >$ 0.80 (%)	$R^2 >$ 0.70 (%)	R_{max}^2	R_{min}^2	RMSEPavg
100	100	0.99	0.97	0.04	16	72	90	0.95	0.60	0.24	0.1
95	100	0.99	0.85	0.10	47	85	94	0.99	0.61	0.21	0.22
97	100	0.94	0.87	0.17	14	58	80	0.97	0.45	0.27	0.18
69	100	0.95	0.83	0.17	41	90	100	0.97	0.72	0.21	0.21

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

The table shows that the R2 values in the calibration set were more stable than the prediction set values. Additionally, the R2 values for SVR, BP, and RF in the calibration set were generally higher than those in the prediction set. This could be due to overfitting resulting from excessive training on the calibration set, resulting in a too complicated model. Increasing the number of coal samples effectively decreases overfitting.

V. DISCUSSION

- *Interpretation of Results in the Context of the Dengcao Coal Mine Area*

The main objective of this study is to examine the application of integrated spectra in a coal mine. Soil contains inorganic and organic materials, particles of different sizes, and air and moisture. Other unrelated elements often influence the analysis of one element in soil particles in the spectrum. Therefore, in soil quantitative modelling, it is crucial to consider the preprocessing of spectral information to prevent issues like the expense of coal extraction (Cui et al., 2020). Research studies have demonstrated that preprocessing techniques like FDR, RL, Continuum Removal (CR), and MSC can improve modelling accuracy. This study collected samples from the field and analysed soil spectral data indoors to assess the relationship between different statistics pre-treatments and soil particle components. The results showed a clear impact of MSC preprocessing (Cai et al., 2023). The differential transformations frequently used emphasize features of the spectral absorption band. Still, when soil particle matter is inverted, the spectral curve changes caused by soil roughness are often seen as a problem (Yu et al., 2023). MSC preprocessing can enhance the correlation between soil particles and spectral elements while preserving the original spectral data characteristics.

- *Implications for Coal Mining Practices and Environmental Management*

Coal mining creates widespread degradation on natural ecosystems like forests and can destroy the land irreparably. Harm to humans, plants, and animals, happens because of environmental contamination and habitat destruction. Mine trashes are created in big quantities and should be managed. The litters are combustible and disposed to impulsive combustion. They have heavy metals proficient of leaching into local groundwater, rivers, and streams (Chow, et al., 2022). Coal washing creates similar waste issues.

- *Comparison of Integrated Spectral Analysis with Traditional Methods*

Traditional methods are expensive because they take too much time and need more workers to test. On the other hand, integrated spectral analysis takes less time and is cheap when compared with the traditional methods. Traditional methods also bring environmental damage as compared with integrated spectral analysis which ensures there is ecological balancing in sustaining the habitat as they ought to be.

- *Discussion of the Potential for Scaling Up the Methodology to Larger Areas*

Scaling use of integrated spectral analysis in the region is unavoidable to larger places in Dengcao Coal Mine area. Scaling will be accepted to all parts of the region to ensure that environmental pollution is reduced as traditional methods seem tedious and time consuming. Similarly, in large areas, the approach will help save costs as estimation of the amount of coal in the region can be predicted in determine the kind of machine to employ (Abbasi, et al., 2021). If the quantity is predicted to be low, the mining company can know the kind of machine to employ in the region thus help save resources.

VI. CONCLUSION

- *Summary of Key Findings from the Study*

Designed at forecasting coal elements using integrated spectra, this research employed four types of spectral statistics preprocessing changes and executed four types spectral index processes depended on the ideal preprocessing changes. The study employed the SPA band selection choosing nonlinear and linear models establishing a forecast model for SOC content. The researched revealed the correlation between content and original spectral data of coal was weak, however, the correlation improved after various transformations and pretreatments. MSC effectively removes noise and reduces interference from baseline changes due to spreading, enhancing spectral information in samples and spectral data. MSC also provides superior spectral coal texture data relevance.

The study has also discovered that the use of integrated spectral method for SOC analysis helps to speed in identifying nutrients in the soil guiding ecological management. In this context, the faster discovery of the nutrients provides availability allows quick intervention that help in taking measures for land restoration. Similalrly, the research finding show in Dengcao Coal Mine area in Zhengzhou has moderate SOC content signifying that the mining activities have not yet disturbed the ecological state.

- *Importance of Integrated Spectral Analysis in Advancing SOC Content Modelling*

Soil organic carbon is core component for soil organic matter and helps a main role in several soil functions. First, it stabilizes soil composition and advances aggregation, that decreases soil erosion; enhances retention and absorption of pesticides and organic pollutants through immobilization. It also helps in managing the coal mining activities as the level of carbon get low. In this context, the approach helps in maintaining balance between gases required to be emitted and used by animals and trees.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAY2382

➢ *Recommendations for Future Research And Applications In Similar Areas*

Drawing from conclusions above, this research develops recommendations oriented for effective regulation by the regional government in Henan including appropriate state agencies to discover mining is carried in agreement with law, therefore, protecting environmental rights for local communities and justifying hostile environmental influences of the development on the communities. Subsequently, constructive involvement of communities is essential to promote maintainable development in mining industry in China. The necessity to protect environmental human rights of communities the mining industry is accepted across the globe including China.

REFERENCES

[1]. Abbasi, B., Wang, X., Chow, J. C., Watson, J. G., Peik, B., Nasiri, V., ... & Elahifard, M. (2021). Review of respirable coal mine dust characterization for mass concentration, size distribution and chemical composition. *Minerals, 11*(4), 426.

[2]. Abdulraheem, M. I., Zhang, W., Li, S., Moshayedi, A. J., Farooque, A. A., & Hu, J. (2023). Advancement of remote sensing for soil measurements and applications: A comprehensive review. *Sustainability, 15*(21), 15444.

[3]. Cai, Y., Jin, Y., Wang, Z., Chen, T., Wang, Y., Kong, W., ... & Hu, H. (2023). A review of monitoring, calculation, and simulation methods for ground subsidence induced by coal mining. International Journal of Coal Science & Technology, 10(1), 32.

[4]. Chen, T., Zhang, T., & Li, H. (2020). Applications of laser-induced breakdown spectroscopy (LIBS) combined with machine learning in geochemical and environmental resources exploration. *TrAC Trends in Analytical Chemistry,* 133, 116113.

[5]. Chow, J. C., Watson, J. G., Wang, X., Abbasi, B., Reed, W. R., & Parks, D. (2022). Review of filters for air sampling and chemical analysis in mining workplaces. Minerals, 12(10), 1314.

[6]. Cui, F., Yue, Y., Zhang, Y., Zhang, Z., & Zhou, H. S. (2020). Advancing biosensors with machine learning. *ACS Sensors, 5*(11), 3346-3364.

[7]. Cui, Y., Liu, S., & Liu, Q. (2021). Navigation and positioning technology in underground coal mines and tunnels: A review. *Journal of the Southern African Institute of Mining and Metallurgy, 121*(6), 295-303.

[8]. Ferfar, M., Sakher, E., Bouras, A., Benselhoub, A., Hachemi, N., Massaoudi, M., ... & Bellucci, S. (2024). Comprehensive physicochemical characterization of Algerian coal powders for the engineering of advanced sustainable materials. *Technology Audit and Production Reserves, 1*(3 (75)), 29-36.

[9]. Hong, Y., Guo, L., Chen, S., Linderman, M., Mouazen, A. M., Yu, L., ... & Liu, Y. (2020). Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon: Effects of fractional-order derivative and optimal band combination algorithm. *Geoderma,* 365, 114228.

[10]. Hublikar, L. V., Shilar, F. A., Mathada, B. S., & Ganachari, S. V. (2024). Comprehensive investigation of green solutions for sustainable wastewater remediation: a review. *Journal of Molecular Liquids*, 124532.

[11]. Li, L., Liu, D., Cai, Y., Wang, Y., & Jia, Q. (2020). Coal structure and its implications for coalbed methane exploitation: a review. *Energy & Fuels, 35*(1), 86-110.

[12]. Lu, B., Dao, P. D., Liu, J., He, Y., & Shang, J. (2020). Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing, 12*(16), 2659.

[13]. McKenna, P. B., Lechner, A. M., Phinn, S., & Erskine, P. D. (2020). Remote sensing of mine site rehabilitation for ecological outcomes: A global systematic review. *Remote Sensing, 12*(21), 3535.

[14]. Murindangabo, Y. T., Kopecký, M., Konvalina, P., Ghorbani, M., Perná, K., Nguyen, T. G., ... & Ali, S. (2023). Quantitative approaches in assessing soil organic matter dynamics for sustainable management. *Agronomy, 13*(7), 1776.

[15]. Rossini, P. M., Di Iorio, R., Vecchio, F., Anfossi, M., Babiloni, C., Bozzali, M., ... & Dubois, B. (2020). Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts. *Clinical Neurophysiology, 131*(6), 1287-1310.

[16]. Shi, B., Meng, J., Wang, T., Li, Q., Zhang, Q., & Su, G. (2024). The main strategies for soil pollution apportionment: A review of the numerical methods. *Journal of Environmental Sciences,* 136, 95-109.

[17]. Tao, D., Yang, P., & Feng, H. (2020). Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive Reviews in Food Science and Food Safety, 19*(2), 875-894.

[18]. Wang, Y., Fu, G., Lyu, Q., Wu, Y., Jia, Q., Yang, X., & Li, X. (2022). Reform and development of coal mine safety in China: An analysis from government supervision, technical equipment, and miner education. *Resources Policy,* 77, 102777.

[19]. Wang, Y., Tian, J., Sun, Z., Wang, L., Xu, R., Li, M., & Chen, Z. (2020). A comprehensive review of battery modelling and state estimation approaches for advanced battery management systems. Renewable and Sustainable Energy Reviews, 131, 110015.

[20]. Yu, J., Li, X., Cao, S., & Liu, F. (2023). Grey fuzzy prediction model of soil organic matter content using hyper-spectral data. *Grey Systems: Theory and Application, 13*(2), 357-380.

[21]. Xuesong, H., Pu, C., Jingyan, L., Yupeng, X., Dan, L., & Xiaoli, C. (2024). Commentary on the review articles of spectroscopy technology combined with chemometrics in the last three years. *Applied Spectroscopy Reviews, 59*(4), 423-482.