IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (Early Access)

Soil Carbon Estimation From Hyperspectral Imagery With Wavelet Decomposition And Frame Theory

Bishal Roy, Vasit Sagan, Haireti Alifu, Jocelyn Saxton, Dorsa Ghoreishi, Nadia Shakoor

***Abstract*—Assessing soil organic carbon (SOC) stocks is crucial for understanding the carbon sequestration potential of agroecosystems and mitigating climate change. This study presents a novel method for assessing SOC and mineral content at various soil depths in sorghum crops using hyperspectral remote sensing. Conducted at Planthaven Farms, MO, the research encompassed 10 genotypes across 30 plots, yielding 180 soil samples from six depth intervals (0-150 cm) for bare soil. Chemical analyses determined the SOC and mineral levels, which were then compared with spectral data from HySpex indoor sensors. We utilized time-frequency analysis methods, including Discrete Wavelet Transformation (DWT), Continuous Wavelet Transformation (CWT), and Frame transformation along with traditional spectral transformations, specifically Fractional Derivatives and Continuum Removal. Analysis revealed the shortwave infrared (SWIR) region, particularly the 1800-2000 nm range, as exhibiting the strongest correlations with SOC content (exceeding 0.8). The visible near-infrared (VNIR) region also provided valuable insights. Models incorporating CWT achieved high accuracy (Test R2 exceeding 0.9), while Frame transformation achieved strong accuracy (Test R2 between 0.7-0.8) with fewer features. The Random Forest Regressor proved most robust, demonstrating superior accuracy and reduced overfitting compared to Support Vector Regression, Partial Least Squares Regression, and Deep Neural Network models. The models demonstrated the efficacy of hyperspectral data for SOC estimation, suggesting potential for future applications that merge this data with above-ground biomass to improve SOC mapping across larger scales. This research offers a promising spectral transformation approach for effective carbon management and sustainable agriculture in a changing climate.**

***Index Terms*—Hyperspectral remote sensing; Soil organic carbon; Wavelet decomposition; Frame theory; Signal processing**

Manuscript received 17 May 2024; revised 18 Aug 2024; accepted 06 Sep 2024. Date of publication; date of current version. This research was funded by the NSF/USDA award under Grant Number: 2020-67021-31530 and the U.S. Geological Survey under Grant/Cooperative Agreement No. G23AP00683 (GY23-GY27). (Corresponding author: Vasit Sagan) Bishal Roy, Vasit Sagan, and Haireti Alifu are with the Taylor Geospatial Institute and Department of Earth, Environmental and Geospatial Science, Saint Louis University, St. Louis, MO 63108 USA (e-mail: bishal.roy@slu.edu; vasit.sagan@slu.edu; alifu.haireti@taylorgeospatial.org)

Jocelyn Saxton and Nadia Shakoor are with the Donald Danforth Plant Science Center, St. Louis, MO 63132 USA (e-mail: jsaxton@danforthcenter.org; nshakoor@ danforthcenter.org).

Dorsa Ghoreishi is with the Department of Mathematics, Saint Louis University, St. Louis, MO 63108 USA (e-mail: dorsa.ghoreishi@slu.edu).

Digital Object Identifier

I. INTRODUCTION

AGROECOSYSTEMS, which account for one-third of the global arable land, are vital in the global carbon cycle, serving as significant carbon sinks by sequestering large amounts of organic carbon in the soil [1]. This process is critical for mitigating climate change by reducing atmospheric carbon dioxide levels [2]. Soil Organic Carbon (SOC) is a key component of soil health, enhancing water retention, nutrient availability, soil biodiversity, and erosion resistance, thereby supporting optimal crop productivity. SOC is comprised of carbon found in living, dead, and decomposing organic materials within the soil, including plant material, soil organisms, and animal remains [3]. The role of SOC in improving the physical, chemical, and biological properties of soil is well documented, making it essential for fostering optimal crop productivity [4].

Accurate Accurately quantifying SOC is inherently challenging due to soil's high spatial variability and the complexities introduced by differing sampling techniques, such as variations in depth and analysis increments [5]. Traditional field methods are labor-intensive and primarily focus on topsoil layers, often overlooking the substantial carbon stocks in deeper soils, which account for nearly half of the global SOC reserves [6].

The advent of hyperspectral imaging has significantly advanced soil analysis by capturing continuous spectra for each pixel, allowing for the identification of distinct spectral signatures associated with SOC [7]. In its developmental stages, hyperspectral imaging was utilized primarily for identifying minerals, rocks, and soils through handheld reflectance spectroscopy, with its applications expanding significantly over time [8]. Despite these advancements, the application of hyperspectral imaging in depth-sensitive SOC estimation remains limited. Recent developments in preprocessing methods, such as Savitzky-Golay derivatives and wavelet decomposition, have improved SOC prediction accuracy by enhancing the signal-to-noise ratio and extracting valuable spectral information [9]. However, these improvements have not fully addressed the complexities of SOC distribution across different soil layers.

Recent advancements in AI techniques, such as SpectralGPT, have rapidly developed and are increasingly applied in hyperspectral remote sensing for SOC estimation [10]. However, hyperspectral data often suffer from degradation and variability, necessitating robust models [11]. Integrating multi-feature fusion approaches, like combining Graph Neural Networks with CNNs, enhances hyperspectral image classification [12]. Methods like CNNs with Bandwise-Independent Convolution and Graph Attention Networks show

promise for band selection and classification [12]. While these methods offer future potential, this study utilizes traditional machine learning to identify key spectral regions for SOC detection through signal transformation methods. These foundational insights can be utilized in future studies on a larger scale, with more diverse soil types and geographical coverage, potentially integrating advanced AI models for enhanced prediction accuracy. Furthermore, research underscores the critical role of precise variable selection in enhancing the accuracy of SOC estimation models, as variable selection directly influences model performance by filtering out irrelevant data [13]. The challenges associated with transferring Vis-NIR models across different environmental contexts further emphasize the need for adapting models to the specific spectral and soil conditions of diverse study areas [14, 15].

Wavelet decomposition and frame theory offer innovative approaches to address these challenges. Wavelet decomposition excels in capturing both low and high-frequency components of hyperspectral signals, facilitating tasks like feature extraction and denoising, while frame theory provides a robust framework for processing complex, high-dimensional datasets [16]. Frame theory, with its foundation in linear algebra and signal processing, provides a comprehensive framework for analyzing complex, high-dimensional datasets, ensuring perfect reconstruction, stability, redundancy, and efficient, signal-independent linear inversion procedures [17]. The theoretical principles of frame theory, which are essential for processing and interpreting high-dimensional datasets in hyperspectral imagery analysis, are well articulated and supported by the foundational work of Kovačević & Chebira [18]. This attribute is crucial for analyzing intricate or non-stationary characteristics of hyperspectral data, thereby enhancing the precision of SOC estimations [19].

Enhanced SOC quantification methods can improve carbon management strategies, aiding climate change mitigation through accurate monitoring of carbon sequestration in agroecosystems [20]. The study's findings could influence soil health and climate change policies, emphasizing the need for advanced remote sensing technologies for environmental monitoring [21]. Additionally, the scalability of these techniques for airborne and spaceborne applications offers new opportunities for large-scale SOC quantification, critical for global carbon cycle studies. The versatility of these methods in various environmental contexts, such as water management and land use planning, highlights their broad relevance. By providing detailed SOC assessments across different soil depths, this research supports innovative soil management approaches, contributing to the resilience and sustainability of agricultural systems worldwide [22]. The objectives of the work presented in this article are to: (1) Identify spectral regions most sensitive to SOC changes and the influence of soil depth; (2) Quantify the improvement of wavelet decomposition and frame theory over traditional hyperspectral analysis methods; and (3) Develop an optimized SOC estimation model incorporating wavelet decomposition and frame theory.

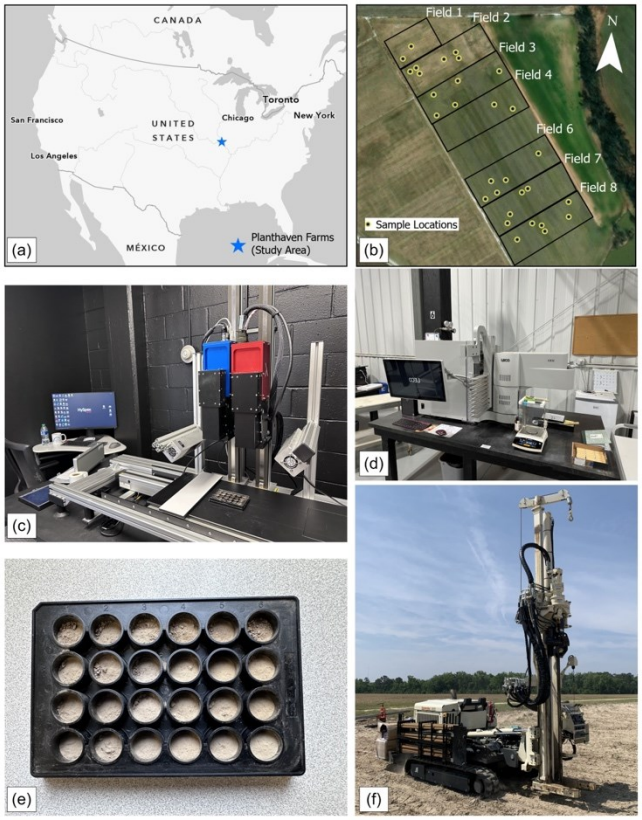

Fig. 1: Location of the study area and experimental setup. (a) Geographic location of the study area; (b) Location of the seven test fields; (c) Indoor Hyperspectral Scanner; (d) LECO 832 Analyzer; (e) Processed soil samples in tray; (f) GeoProbe machine.

II. Data

A. Field Data Collection

The field experiment was conducted at Planthaven Farms, MO, encompassing seven fields strategically positioned near the Missouri River watershed (Fig. 1a, b). These fields, totaling 2.2 hectares, were arranged in a randomized complete block design. We planted 308 sorghum plots, including 10 select genotypes replicated three times for comprehensive analysis. Plot dimensions were standardized at 20 ft x 20 ft, with 8 rows of sorghum per plot. Planting occurred in late June, and baseline soil samples were collected prior to planting. Soil samples were collected using a GeoProbe machine (Fig. 1f) to extract cylindrical cores of 150 cm depth and 7.6 cm diameter. These cores were stored at temperatures between 0-4 degrees Celsius to preserve their integrity until further processing.

B. Laboratory Analysis

Soil core processing involved several steps (Fig. 2). The PVC-encased cores were sectioned using a bandsaw into six distinct depth intervals: 0-15 cm, 15-30 cm, 30-60 cm, 60-90 cm, 90-120 cm, and 120-150 cm. Two cups of soil were reserved from each depth in a plastic bag, then dried at 40-45 degrees Celsius until completely devoid of moisture and subsequently stored for analysis. We selected Ward Laboratories Inc. (Kearney, NE) for our carbon analysis needs.

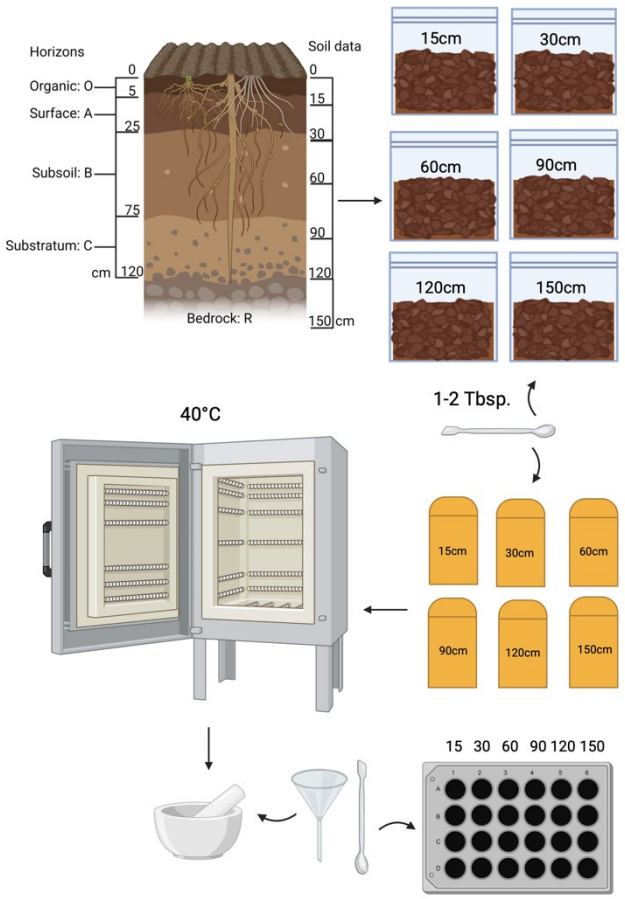

Fig. 2: Soil Sample Processing workflow.

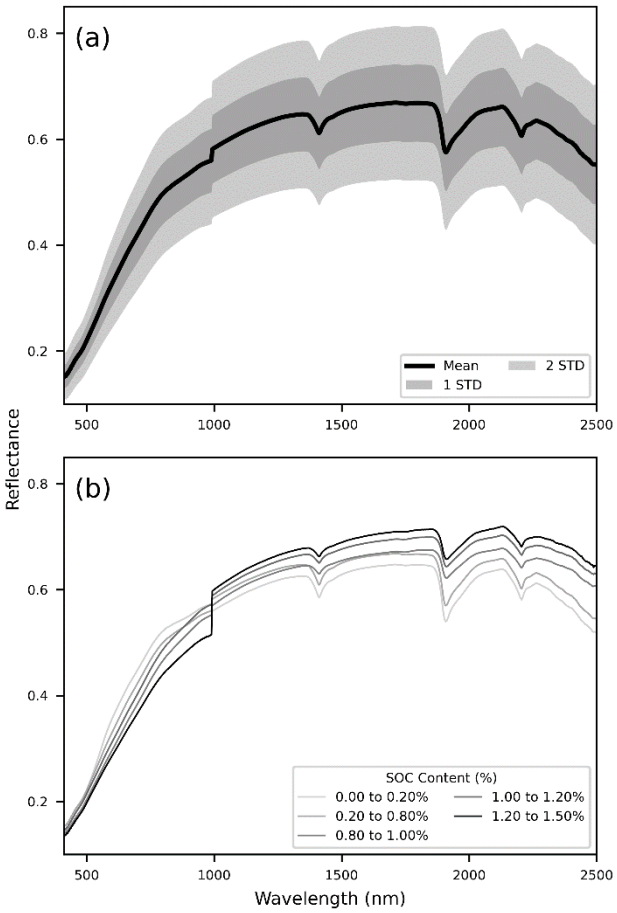

Fig. 3: Exploratory results, (a) Mean, 1 and 2 standard deviations of the sample spectra (n=180); (b) Reflectance spectra for soils with varying SOC content.

Their LECO 832 instrument (Fig. 1d), operating on a combustion-based principle, provided precise SOC quantification.

C. Hyperspectral Imagery

To complement soil carbon analysis, we utilized advanced indoor hyperspectral scanning of dried soil samples using the HySpex sensor system, which captures both visible and near-infrared (VNIR) and short-wave infrared (SWIR) regions. The VNIR sensor covers 400-1000 nm across 300 bands with 0.03 mm/pixel resolution, while the SWIR sensor spans 960-2500 nm across 362 bands at 0.15 mm/pixel. Custom close-up lenses were used to maintain a 30 cm working distance, and the setup included vertical sensor mounting with halogen lamps at 45-degree angles to minimize noise. A 50% zenith white reflectance panel corrected the raw data, and the HySpex RAD module facilitated radiometric calibration, converting raw values into absolute radiance($W/sr \cdot nm \cdot m^2$) [23]. Pre-processed samples were organized in 24-cell assay plates, with each row representing six soil depths from a plot for systematic scanning. Figures 3a and 3b illustrate the mean and variability of the samples and the spectra's relationship with SOC content, while Figure 4 shows the decline in SOC with increasing depth. Table 1 details the dataset's central tendency and spread.

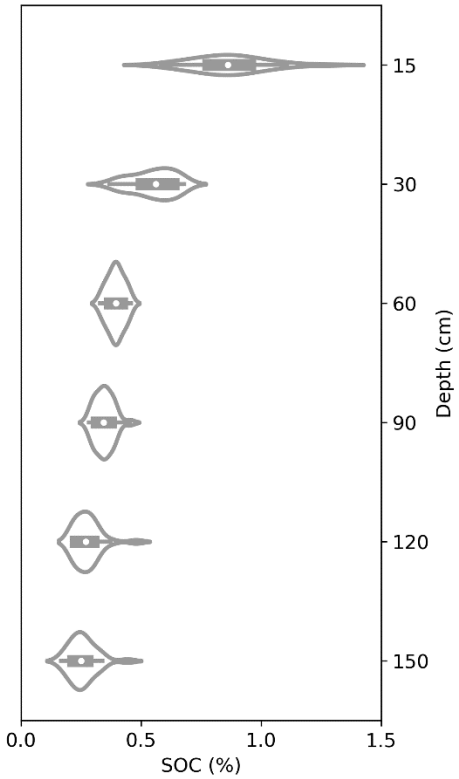

Fig. 4: Soil Organic Carbon (SOC) percentage by depth.

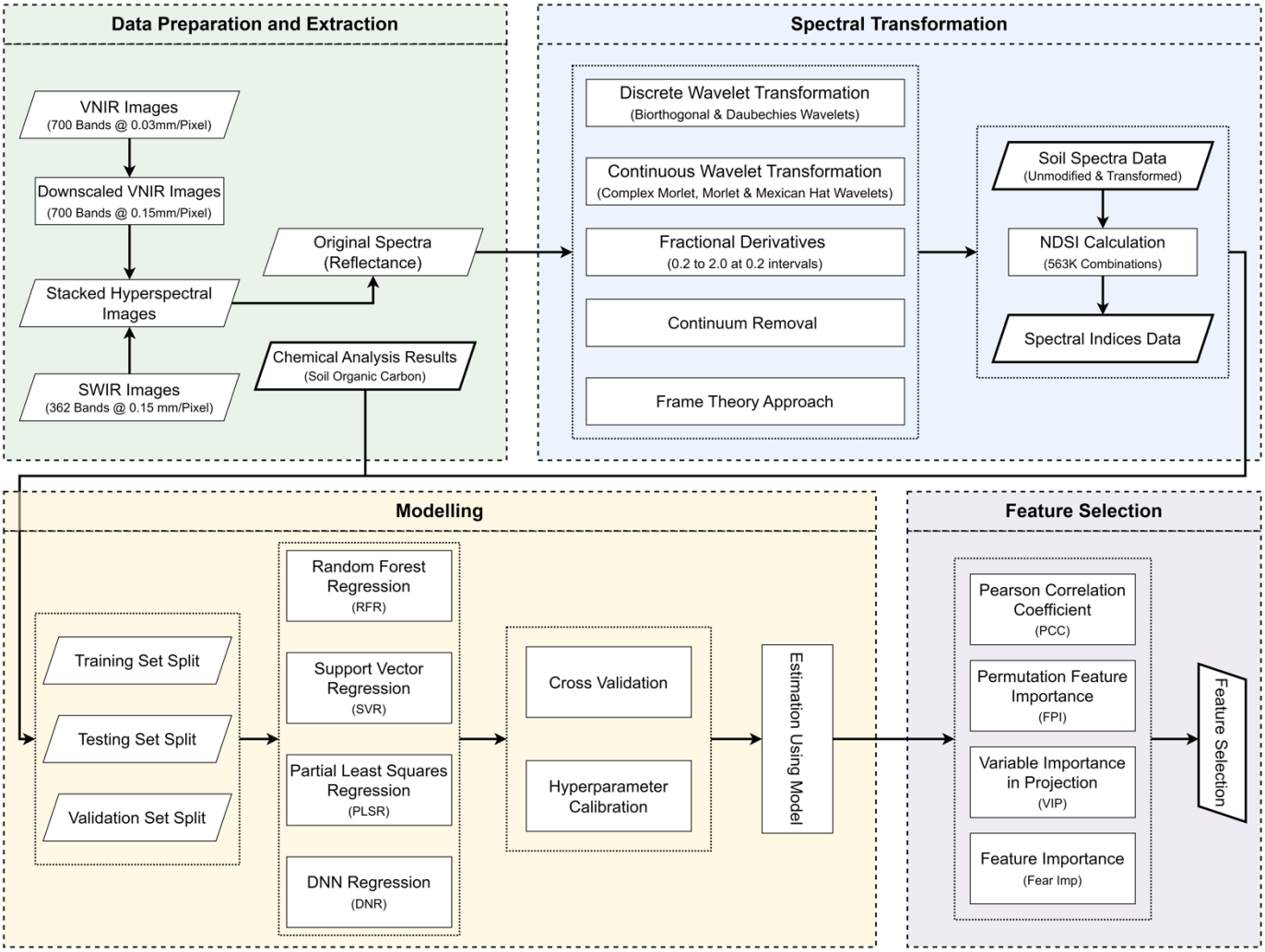

Fig. 5: Overall workflow of the data preparation to modelling pipeline.

Table 1: Descriptive statistics of SOC Samples (in %)

Depth (cm)	Sample Size	Min	Max	Mean	SD	CV (%)
0-15	30	0.582	1.273	0.866	0.152	17.53
15-30		0.368	0.680	0.551	0.089	16.16
30-60		0.330	0.458	0.394	0.034	8.786
60-90		0.283	0.453	0.346	0.038	11.14
90-120		0.214	0.480	0.275	0.055	20.32
120-150		0.166	0.443	0.257	0.057	22.32

III. Methods

A. Workflow

The workflow is visualized in Figure 5.

B. Spectral Transformation

1) Wavelet Transformation

In our study, we employed both discrete and continuous wavelet transformations to analyze the spectral data. Although wavelet transform has limitations in characterizing local time-domain features, it is effective for decomposing signals into frequency components, which is essential for analyzing spectral features related to SOC. Despite the complexities of hyperspectral data, the wavelet transform's balance of time and frequency resolution made it a suitable and practical tool for our objectives. Discrete Wavelet Transformation (DWT) decomposes signals into wavelet coefficients at various scales and positions, providing insights into the signal's frequency composition by recursively filtering the spectrum with high-pass and low-pass filters derived from the selected wavelet function [24]. This process yields detail coefficients representing high-frequency details and approximation coefficients representing the low-frequency trend within the spectrum [25]. On the other hand, Continuous Wavelet Transformation (CWT) offers a comprehensive analysis of hyperspectral data, capturing both broad and localized spectral features that indicate variations in soil organic carbon content. CWT is favored for its ability to provide a continuous spectrum of scales for feature extraction, enhancing estimation capabilities in soil properties an [26]. To address the sensitivity of wavelet packet transformation to noise, hyperspectral imaging was conducted in a controlled indoor environment using a specialized sensor, minimizing atmospheric noise and ensuring clean, accurate data. Basic denoising techniques were also applied during pre-processing to further enhance data quality. A Python package named “PyWavelets” was used to transform the spectral data [27]. We selected wavelet basis functions by testing several

common types and referencing previous studies to ensure they effectively captured the spectral characteristics of soil. In our study, we employ the following mother wavelets:

Biorthogonal Wavelet (bior 1.3): The Biorthogonal 1.3 wavelet ("bior1.3") is chosen for its precise symmetry and additional vanishing moments, which are critical for accurate feature extraction in hyperspectral data analysis [28]. Its robust edge-preserving capability ensures the integrity of nuanced spectral variations, enhancing the accuracy of SOC estimation by effectively isolating both approximation and detail coefficients. The transformation can be expressed as equations 1 and 2.

$$A[n] = \sum_k h_k \cdot x[2n-k] \quad (1)$$

$$D[n] = \sum_k g_k \cdot x[2n-k] \quad (2)$$

Where, $A[n]$ are the approximation coefficients, $D[n]$ are the detail coefficients, h_k are the low-pass decomposition filter coefficients, g_k are the high-pass decomposition filter coefficients, $x[n]$ is the input signal.

Daubechies Wavelet (db2): Daubechies wavelets, particularly "db2," are valued for their compact support and significant vanishing moments, which are essential for processing complex signals like hyperspectral data [29]. The "db2" wavelet effectively captures sharp spectral discontinuities, making it ideal for SOC estimation by decomposing signals into approximation and detail components, thereby enhancing the detection of key spectral signatures related to soil organic carbon [30]. The transformation can be expressed as equations 3 and 4.

$$A[n] = \sum_k h_k^{db2} \cdot x[2n-k] \quad (3)$$

$$D[n] = \sum_k g_k^{db2} \cdot x[2n-k] \quad (4)$$

Where, $A[n]$ are the approximation coefficients; $D[n]$ are the detail coefficients; h_k^{db2} are the low-pass decomposition filter coefficients; g_k^{db2} are the high-pass decomposition filter coefficients; $x[n]$ is the input signal.

Morlet Wavelet (morl): The Morlet wavelet, characterized by a cosine wave modulated by a Gaussian envelope [31], is especially effective for identifying spectral features that vary gradually across different frequencies [32]. This wavelet is well-suited for hyperspectral data analysis due to its ability to balance time and frequency resolution, making it ideal for detecting subtle changes in soil properties reflected in the spectral signatures. This decomposition can be expressed as equation 5.

$$\psi(t) = \exp^{-\frac{t^2}{2}} \cos(5t) \quad (5)$$

Where, $\psi(t)$ represents the wavelet function, t is the time or position parameter, exp represents the base of the natural logarithm.

Complex Morlet Wavelets (cmor 1.5-0.5): The Complex Morlet wavelet, defined by its adjustable bandwidth and center frequency, is crucial for spectral analysis in soil carbon estimation [33]. The bandwidth controls frequency resolution and feature localization, while the center frequency targets specific soil-related frequencies, optimizing spectral feature detection [34]. The center frequency parameter allows for tuning the wavelet to target frequencies associated with soil characteristics, enhancing the detection of relevant spectral features [35]. Proper tuning of these parameters balances spectral detail with spatial precision, making the Complex Morlet wavelet a powerful tool for accurately extracting and analyzing soil carbon spectral information. This decomposition can be expressed as equation 6.

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} e^{2i\pi f_c t} e^{-t^2/f_b} \quad (6)$$

Where, $\psi(t)$ represents the wavelet function; f_b denotes the bandwidth parameter.; f_c is the center frequency.; t is the time parameter.

Mexican Hat Wavelet (mexh): The Mexican Hat wavelet, or the Ricker wavelet, is a second derivative of a Gaussian function [36], producing a waveform that is adept at highlighting rapid changes in the signal [37]. Its shape makes it particularly useful for detecting inflection points and boundaries within the spectral data [38], which can signify transitions between different soil constituents or layers. This decomposition can be expressed as equation 7.

$$\psi(t) = \frac{2}{\sqrt{3}\sqrt[4]{\pi}} \exp^{-\frac{t^2}{2}} (1-t^2) \quad (7)$$

Where, $\psi(t)$ represents the wavelet function, t is a non-dimensional time parameter, exp represents the base of the natural logarithm.

We used the standard settings of PyWavelets for signal decomposition, which effectively handled energy distribution during wavelet packet transformation.

2) Fractional Derivatives

Fractional derivatives, extending differentiation to non-integer orders, offer a powerful tool for enhancing feature extraction in hyperspectral imagery analysis for soil carbon estimation [39]. By emphasizing subtle spectral changes, fractional derivatives highlight variations in soil properties across spectral bands, improving SOC estimation accuracy. The derivative order controls signal smoothness, allowing adjustment of sensitivity to specific spectral features [40]. In this study, we employ the Grunwald-Letnikov (G-L) definition due to its straightforward coefficients [41]. The G-L definition is expressed as equation 8.

$$d^{\alpha} f(x) = \lim_{h \to 0} \frac{1}{h^{\alpha}} \sum_{m=0}^{(t-a)/h} (-1)^m \frac{\Gamma(\alpha+1)}{m!\Gamma(\alpha-m+1)} f(x-mh) \quad (8)$$

Where, α is the order; h is the step size; t is the upper limit of fractional derivatives; a is the lower limit of fractional derivatives; Γ is the gamma function.

3) Continuum Removal

Continuum removal in spectral analysis normalizes the spectrum against a convex hull-based continuum curve, isolating absorption features critical for identifying soil properties linked to organic carbon [42]. By interpolating between reflectance maxima, the technique emphasizes the depth and shape of absorption features, enhancing the accuracy and comparability of spectral data across different samples by mitigating effects like lighting variations [43]. This transformation can be expressed as equation 9.

$$R_{cr}(\lambda) = \frac{R(\lambda)}{C(\lambda)} \tag{9}$$

Where, $R_{cr}(\lambda)$ is the continuum-removed reflectance at wavelength λ; $R(\lambda)$ is the original reflectance at wavelength λ; $C(\lambda)$ is the continuum curve at wavelength λ obtained by interpolating the reflectance values at the local maxima.

4) Frame Theory Applications

Frame theory offers a mathematically rich and flexible framework for signal analysis, extending the capabilities of traditional basis-oriented approaches [44]. A frame for a Hilbert space is a collection of vectors that satisfy a specific frame condition, allowing for redundancy and enhanced representational power [45]. For our study, we employ Gabor frames due to their suitability for pinpointing localized changes within signals [46]. A Gabor system in L2(ℝ) has the form $\{e^{2\pi imbx} g(x - na)\}_{m,n\in\mathbb{Z}}$ for some parameters a, b > 0 and for some function $g \in L^2(\mathbb{R})$. Using the translation operators and the modulation operators we can denote a Gabor system by $\{g_{m,n}\} = \{M_{mb} T_{na} g\}$ [47]. This construction results in frame elements that are well-localized in both the time and frequency domains. The core of a Gabor function is defined as equation 10 [48].

$$g(t) = e^{-\pi\beta t^2} \cos(2\pi\omega t) \tag{10}$$

Where, t represents a time point, ω is the frequency of the oscillation within the Gabor function, and β is the bandwidth of the Gabor function, determining its width. To further enhance the ability to capture subtle spectral nuances, we adopt an adaptive bandwidth strategy [49]. The bandwidth of each Gabor function is dynamically adjusted based on its frequency, as expressed in equation 11.

$$\beta(\omega, \omega_{min}, \omega_{max}) = 1.0 + \kappa \times ln\left(1 + \frac{\omega - \omega_{min}}{\omega_{max} - \omega_{min}}\right) \tag{11}$$

Where, ω denotes the frequency, κ is the scaling factor and β is the bandwidth parameter. Compositional variations in soil, reflective of differing SOC content, manifest as localized features across the spectrum [50]. The adaptive Gabor frame enables the extraction of these features with the precision necessary for accurate SOC estimation. Higher frequencies often necessitate a broader bandwidth to accurately represent rapid oscillations within the signal. In contrast, lower frequencies can be effectively analyzed with a narrower bandwidth [51]. Our simplified Gabor frame construction focuses on this frequency-dependent modulation as expressed in equation 12.

$$g_\omega(t) = e^{-\pi\beta(\omega)t^2} \cos(2\pi\omega t) \tag{12}$$

Where, ω denotes the frequency, $g_\omega(t)$ represents a single frame element indexed by frequency, and $\beta(\omega)$ is the bandwidth of a Gabor element, calculated based on its frequency using the adaptive bandwidth function. By tailoring the analysis to the spectral characteristics in this manner, we achieve a representation exceptionally sensitive to the subtle shifts in the soil spectra, directly correlated with SOC variations. Frame theoretic analysis was performed to extract the spectral characteristics encoded within the adaptive Gabor frame. For a fixed vector $g \in L^2(\mathbb{R})$ and $a, b > 0$, the associated Gabor frame or Weyl-Heisenberg frame G(a, b, g) is a sequence $\{g_{m,n}\}$ for which there exist two constants $0 < A < B < \infty$ such that the frame inequality is satisfied [52]. This is expressed om equation 13,

$$A\|f\|^2 \leq \sum_{m,n\in\mathbb{Z}} |\langle f, g_{m,n}\rangle|^2 \leq B\|f\|^2 \text{ for all } f \in L^2(\mathbb{R}) \tag{13}$$

An analysis operator for a Gabor frame is a linear map defined as T_g which takes f to $\{\langle f, g_{m,n}\rangle\}_{m,n}$ and a synthesis operator T_g^* which is the adjoint of the analysis operator, takes elements $\{c_{m,n}\} \in \ell^2(\mathbb{Z})$ to $\sum_{m,n\in\mathbb{Z}} c_{m,n}\, g_{m,n}$. The Gabor frame operator $S_g = T_g^* T_g$ is defined by equation 14,

$$Sg f = \sum_{m,n\in\mathbb{Z}} \langle f, g_{m,n}\rangle g_{m,n} \tag{14}$$

This process mathematically decomposes the soil spectrum f into a combination of the frame elements ϕ_i as expressed in equation 15 [53],

$$Sf = \sum_{i\in I} \langle f, \phi_i\rangle \phi_i \tag{15}$$

Where the S is the frame operator, and frame coefficient f, ϕ_i quantify how strongly the spectrum aligns with each frame element. In Python, we defined custom functions to construct individual Gabor functions and assemble the complete Gabor frame. To dynamically adjust the bandwidth of each Gabor function based on its frequency, the adaptive bandwidth strategy was incorporated. Frame analysis is performed using NumPy's [54] matrix-vector multiplication capabilities, projecting the spectral data onto the adaptive Gabor frame.

C. Modelling

In this study an automated modeling pipeline (visualized in Figure 5) was deployed to apply machine learning regression techniques across spectrally transformed datasets for the assessment of SOC. These models are: Random Forest Regressor (RFR), Support Vector Regression (SVR), Partial Least Squares Regression (PLSR), and Deep Neural Network Regression (DNN), for their application in SOC estimation from spectral data. The implementation of these models was facilitated by the Scikit-learn package in python [55]. All spectral features were included in the model inputs, and Pearson Correlation Coefficient (PCC) analysis was

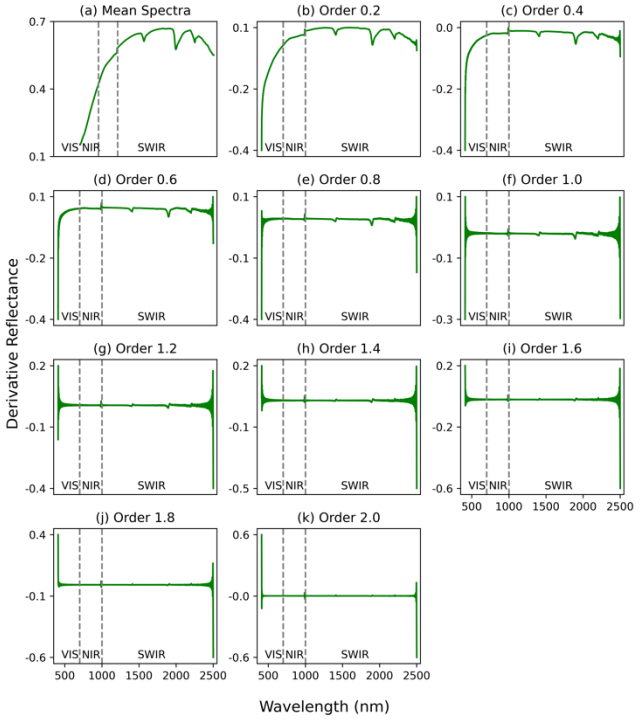

Fig. 6: Transformation of soil reflectance spectra with increasing fractional derivative orders (0.2 to 2.0 at 0.2 intervals).

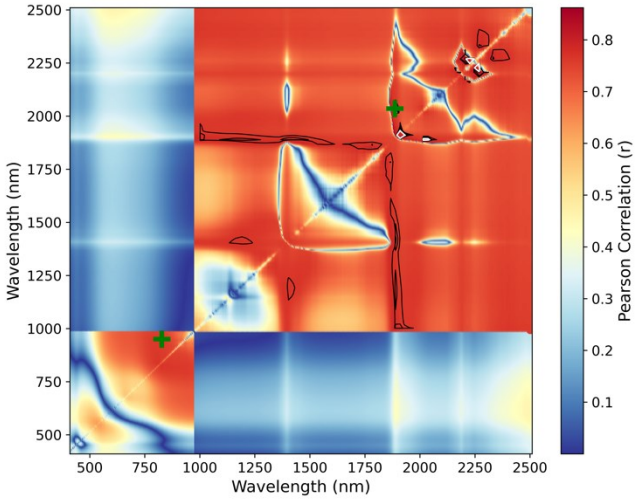

Fig. 7: Heat map of absolute value of Pearson's correlation coefficient (PCC) between NDSI and SOC values. The green '+' symbols indicate the NDSI with highest PCC in each region. The black contour lines indicate PCC from 0.75 to 0.8 while the white contour lines indicate PCC above 0.8.

performed to identify the most relevant spectral regions related to SOC, which provided a basis for understanding the spectral regions affecting SOC rather than refining the models' inputs. Hyperparameter tuning for the models, including RFR, SVR, and PLSR was conducted using Scikit-learn [55] and grid search functionality, while DNN tuning was facilitated by the Hyperband package [56]. A 10-fold cross-validation process was integrated within the hyperparameter tuning phase for all models.

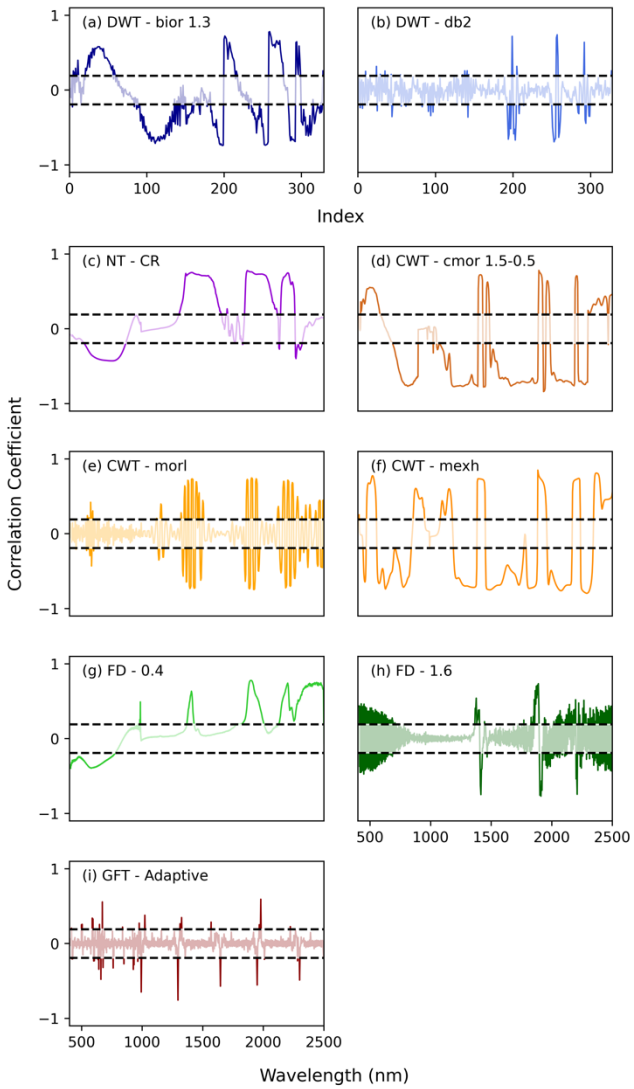

Fig. 8: Pearson's Correlation Coefficients between Transformed Spectra and Soil Organic Carbon (SOC). Correlations outside dashed lines are significant ($p < 0.01$).

IV. Results

A. Spectral Sensitivity and Feature Importance

Fractional derivative analysis significantly transformed soil reflectance spectra, revealing subtle variations indicative of soil carbon content (Fig. 6). As derivative orders increased, more pronounced spectral peaks and troughs emerged, with orders 0.4 and 1.6 (Fig. 6c, 6i) showing significant differences in maximum and minimum derivative values, suggesting increased sensitivity to spectral features linked to soil carbon. Normalized Difference Spectral Indices (NDSI) analysis highlighted strong correlations between soil organic carbon (SOC) and specific spectral regions within the SWIR and VNIR bands (Fig. 7). The SWIR region showed the highest correlations, particularly around 1900-2100 nm and 2200-2400 nm, while the VNIR region exhibited a significant correlation band around 800-900 nm.

Feature importance analysis (Fig. 8 and Fig. 9) further underscored the significance of these spectral ranges. PCC Scores (Fig. 8) indicated that certain transformations, especially CWT-mexh, emphasized critical regions in both

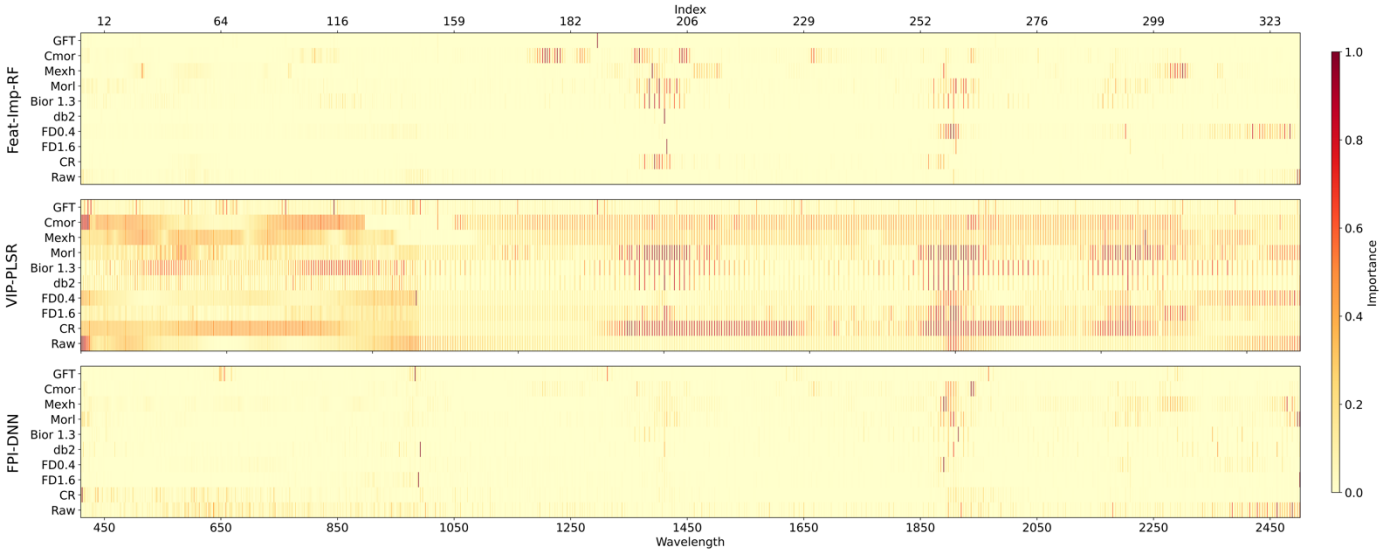

Fig. 9: Spectral Feature Importance for SOC Estimation.

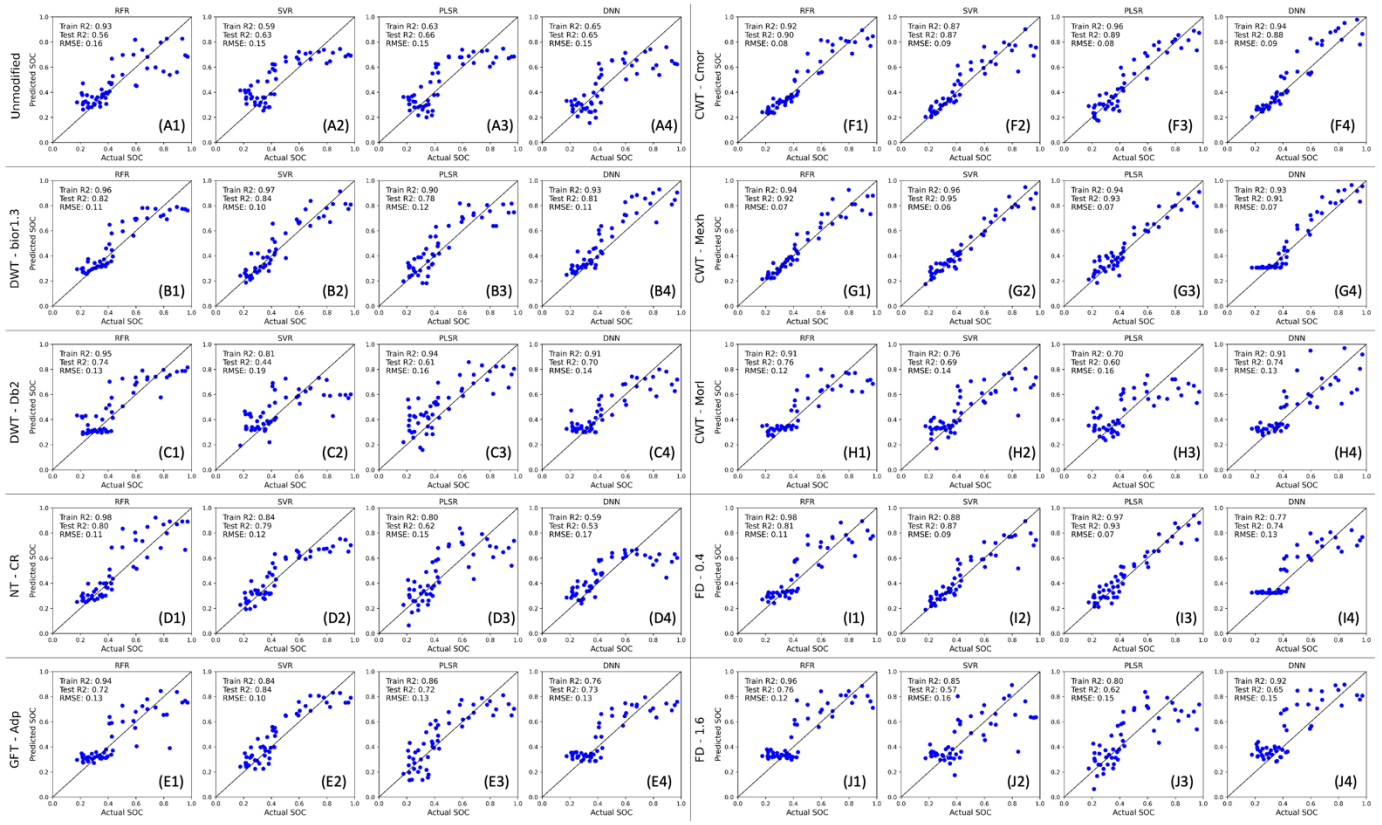

Fig. 10: Scatter plot showing the actual vs predicted SOC for each of the pipelines. The first character (A-J) indicates the spectral transformation used and the second character (1-4) indicates the ML model for the estimation.

VNIR and SWIR bands. In Fig. 9, the Feature Importance subplot revealed significant concentrations of important features around 1300-1500 nm, 1830-1950 nm, and 2150-2300 nm, primarily within the SWIR region. The Variable Importance in Projection (VIP) subplot highlighted a broader range of important features, spanning both VNIR (400-1000 nm) and SWIR regions. The Feature Permutation Importance (FPI) subplot displayed fewer bands of high importance but consistently identified the 1850-2000 nm region as significant across all spectra. Overall, the 1850-2000 nm region within the SWIR demonstrated notable importance across all feature importance methods and spectral transformations.

B. Model Results for SOC Estimation

Model performance varied significantly depending on the modeling algorithm and spectral transformation (Fig. 10). RFR models consistently outperformed others, particularly with CWT-mexh, CWT-cmor, FD-0.4, and DWT-bior1.3, achieving test R^2 values of 0.8-0.92 and RMSE between 0.07-

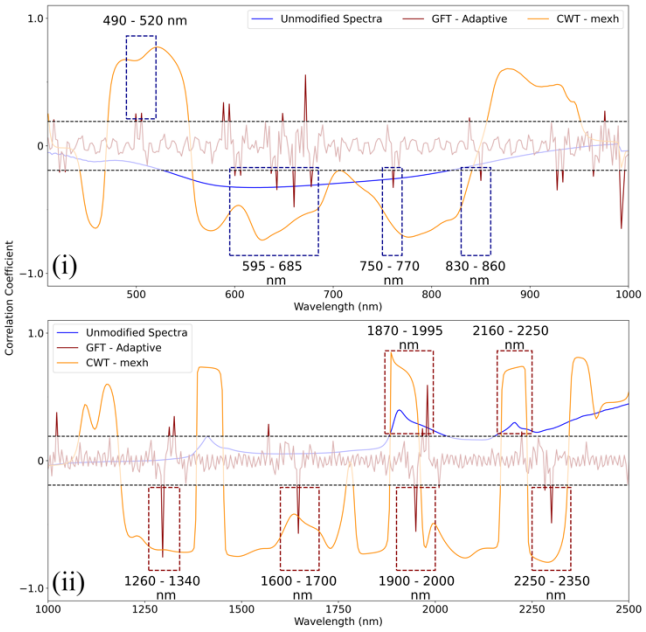

Fig. 11: Key spectral regions based on the correlation between soil organic carbon and the spectra. Dashed lines indicate significance at $p < 0.01$.

0.11. These transformations demonstrated strong potential for SOC estimation, though some models, especially RFR and PLSR, showed overfitting. PLSR models performed well, with test R^2 up to 0.93, while SVR results were more variable. DNN models, although better than SVR with unmodified spectra, generally performed similarly or slightly worse with transformed data. Transformations like CWT-mexh, with a high density of significant features, were linked to better model robustness, whereas GFT-Adaptive, with fewer important features, led to more frequent overfitting.

V. Discussion

A. Key Spectral Regions for SOC Estimation

Pearson's Correlation Coefficient analysis identified several key spectral regions with significant correlations to SOC content across the VNIR and SWIR spectra (Fig. 11). In the VNIR, positive correlations in the 490-520 nm range, influenced by humic substances, and negative correlations in the 595-685 nm, 750-770 nm, and 830-860 nm bands, related to iron oxides and moisture interactions, were observed. The SWIR revealed robust SOC-sensitive regions, notably at 1870-1995 nm and 2160-2250 nm, linked to O-H and C-H bonds in organic matter [57]. Negative correlations in bands like 1260-1340 nm and 1600-1700 nm suggest indirect SOC reflections through water and clay minerals, with the 2200-2400 nm range also showing sensitivity to organic molecules [58]. The consistent identification of these regions across both GFT-Adaptive and CWT-mexh transformations underscores their robustness as SOC indicators.

Further analysis emphasized the significance of specific spectral regions for SOC estimation across various transformation methods (Fig. 12). DWT methods highlighted the NIR region, particularly around 910-912 nm, while NT-CR and CWT techniques underscored the importance of the SWIR, especially between 1865-2382 nm. Fractional

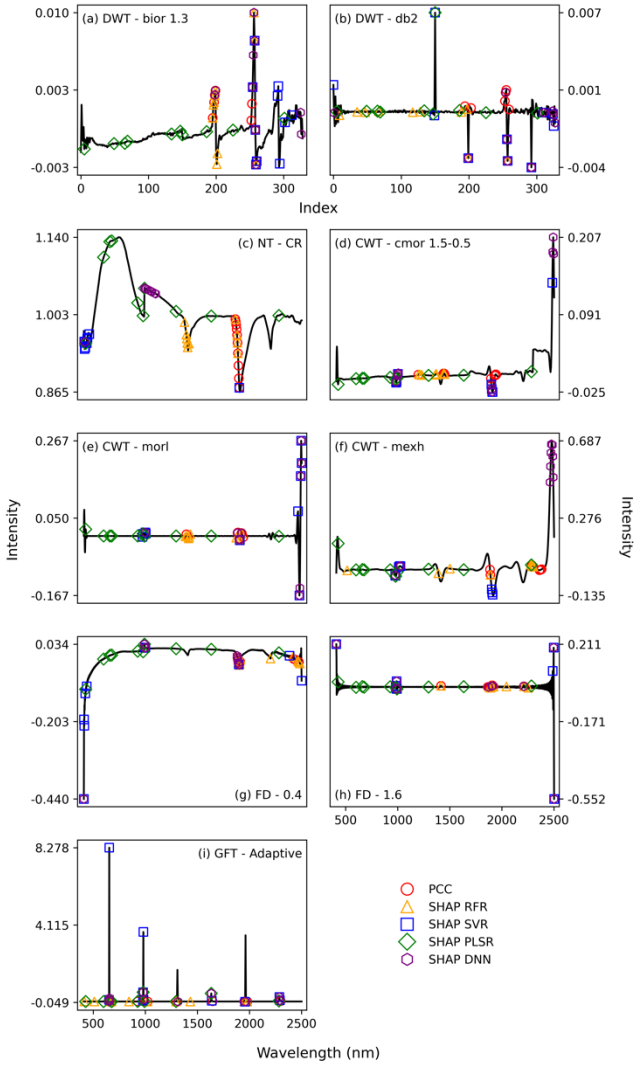

Fig. 12: Feature Importances across all models and transformations techniques. Markers indicate significant features based on Pearson Correlation Coefficient (PCC) and SHAP values.

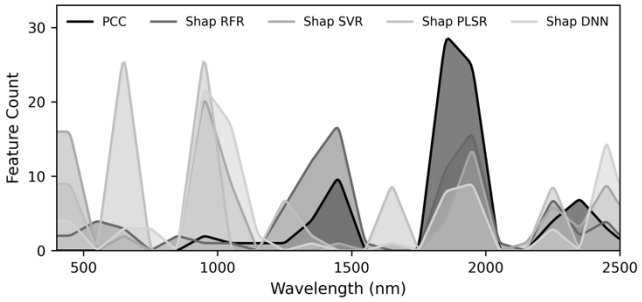

Fig. 13: Distribution of important features for estimating SOC in the VNIR-SWIR spectra.

derivatives revealed sensitivity across both VNIR and SWIR, with notable correlations from 400-800 nm and 2200-2400 nm. GFT-Adaptive pinpointed unique wavelengths, such as 1296 nm and 1980 nm, suggesting the potential of subtle spectral features for SOC analysis, with further investigation warranted to understand these findings.

Overall, feature importance analysis (Fig. 13) demonstrated that significant SOC-related features were predominantly located in the SWIR region (67%), particularly in the 1800-

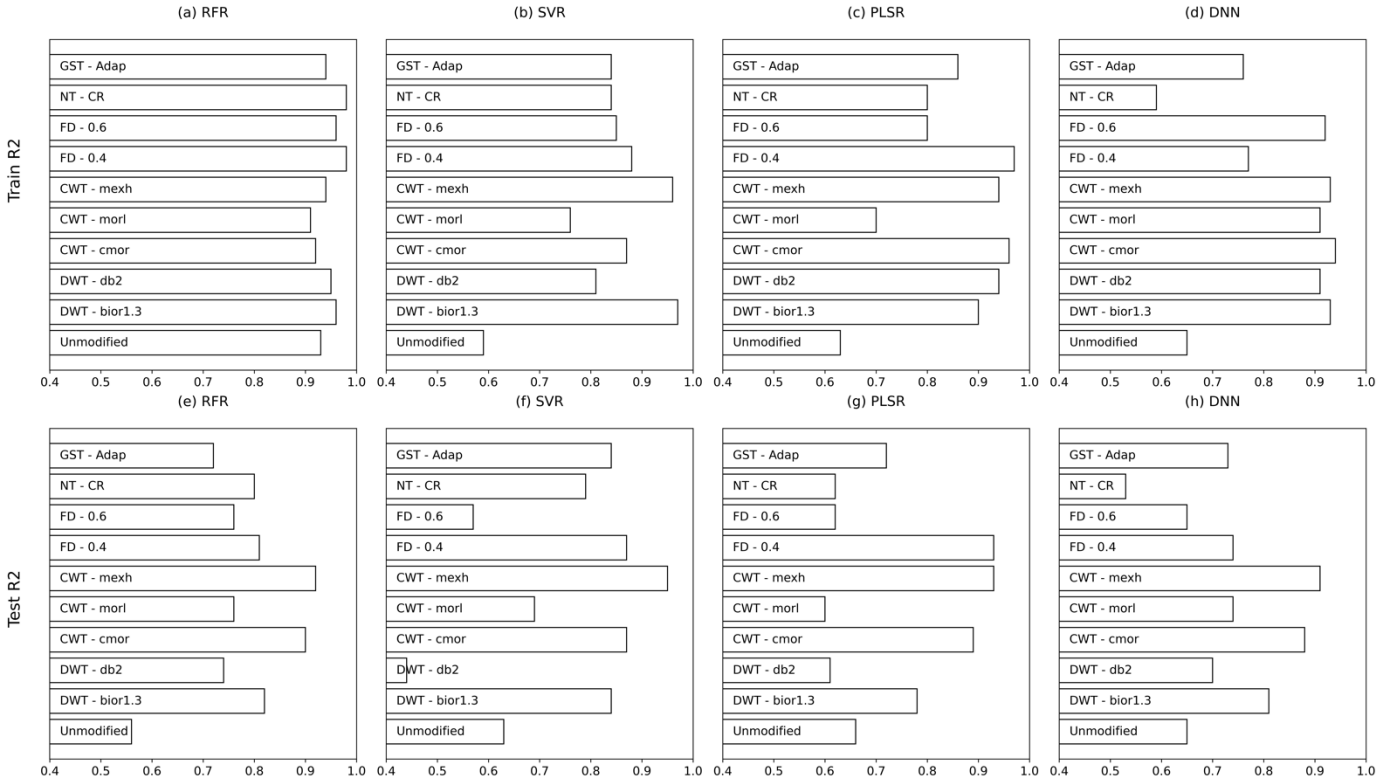

Fig. 14 Evaluation of models and spectral transformations for SOC prediction. (a-d) Training R^2 scores and (e-h) test R^2 scores for RFR, SVR, PLSR, and DNN models, respectively.

2000 nm range, aligning with known absorption features linked to organic matter, moisture, and clay minerals [59]. The VNIR also contributed essential features, particularly in the 900-1000 nm and 600-700 nm ranges. This comprehensive analysis reaffirms the critical role of both VNIR and SWIR regions in SOC estimation and highlights underexplored spectral variations that could advance SOC characterization.

C. Model Performance Analysis

Model performance across RFR, SVR, PLSR, and DNN algorithms (Fig. 14) revealed that RFR consistently outperformed others, achieving high R^2 scores in both training and testing, with minimal overfitting. SVR and DNN models often showed larger gaps between training and testing R^2 scores, indicating higher overfitting, particularly with DWT-db2 and FD-1.6 transformations. PLSR exhibited moderate performance, with R^2 scores ranging from 0.6 to 0.78 and less overfitting. CWT-mexh, FD, and GFT-Adaptive transformations significantly improved model performance, with CWT-mexh delivering the highest test R^2 scores across all models (up to 0.95 with SVR). GFT-Adaptive, despite identifying fewer significant features, consistently achieved test R^2 scores above 0.7 across all algorithms, demonstrating its efficiency and potential in scenarios with limited spectral resolution, making it valuable for SOC prediction with fewer spectral bands.

VI. Conclusion

This study advances methodologies for SOC estimation by integrating hyperspectral imaging with signal processing techniques. Notably, CWT-mexh models achieved R^2 scores up to 0.95, with the RFR outperforming SVR, PLSR, and DNN models in both accuracy and reduced overfitting. Key findings include:

1) The SWIR region, especially 1800-2000 nm, shows the strongest correlation with SOC content (above 0.8).
2) CWT (Mexican Hat and Complex Morlet), Fractional Derivatives (Order 0.4 and 1.6), and Frame Transformation methods significantly improve SOC prediction accuracy.
3) Random Forest Regressor, paired with selected spectral transformations, offers the most accurate and robust SOC estimation framework.
4) The Frame Transformation method's success with fewer features suggests its potential where spectral resolution is limited, while CWT (Mexh) excels with high spectral resolution.
5) Consistent spectral regions (1800-2000 nm, 900-1000 nm, and weaker correlations in 400-700 nm) warrant further investigation regarding specific SOC constituents.

This study primarily focused on lab-based spectroscopy and did not address depth-sensitive analysis. Future research should explore spectral responses across soil depths,

scalability for airborne and spaceborne platforms, and assessments across diverse soil types to validate the robustness and generalizability of the findings. Additionally, integrating GFT-Adaptive with other spectral methods may further enhance SOC prediction accuracy.

ACKNOWLEDGMENT

The authors thank the members of the Remote Sensing Lab at Saint Louis University and the Shakoor Lab at Donald Danforth Plant Science Center for providing data collection, curation, and analytics support.

REFERENCES

[1] G. Gebeyehu, T. Soromessa, and D. Teketay, "Organic Carbon Stocks, Dynamics and Restoration in Relation to Soils of Agroecosystems in Ethiopia: A Review," *International Journal of Environment,* vol. 6, no. 1, pp. 1-22, 2017/02/28 2017, doi: 10.3126/ije.v6i1.16864.

[2] D. Angers *et al.*, "A well-established fact: Rapid mineralization of organic inputs is an important factor for soil carbon sequestration," *European Journal of Soil Science,* vol. 73, no. 3, p. e13242, 2022/05/01 2022, doi: https://doi.org/10.1111/ejss.13242.

[3] H. Karabcova, L. Pospisilova, K. Fiala, P. Skarpa, and M. Bjelkova, "Effect of organic fertilizers on soil organic carbon and risk trace elements content in soil under permanent grassland," *Soil and Water Research,* vol. 10, no. 4, pp. 228-235, / 2015. [Online]. Available: http://dx.doi.org/10.17221/5/2015-SWR.

[4] C. Jansson *et al.*, "Crops for Carbon Farming," *Frontiers in Plant Science,* Perspective vol. 12, 2021. [Online]. Available: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.636709.

[5] E. Goidts, B. Van Wesemael, and M. Crucifix, "Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales," *European Journal of Soil Science,* vol. 60, no. 5, pp. 723-739, 2009/10/01 2009, doi: https://doi.org/10.1111/j.1365-2389.2009.01157.x.

[6] L. Henneron *et al.*, "Bioenergetic control of soil carbon dynamics across depth," *Nature Communications,* vol. 13, no. 1, p. 7676, 2022/12/12 2022, doi: 10.1038/s41467-022-34951-w.

[7] C. S. Zhang, Y. Tang, X. L. Xu, and G. Kiely, "Towards spatial geochemical modelling: Use of geographically weighted regression for mapping soil organic carbon contents in Ireland," (in English), *Appl Geochem,* vol. 26, no. 7, pp. 1239-1248, Jul 2011, doi: 10.1016/j.apgeochem.2011.04.014.

[8] E. Ben-Dor, S. Chabrillat, J. A. M. Demattê, G. R. Taylor, J. Hill, M. L. Whiting, and S. Sommer, "Using Imaging Spectroscopy to study soil properties," *Remote Sensing of Environment,* vol. 113, pp. S38-S55, 2009/09/01/ 2009, doi: https://doi.org/10.1016/j.rse.2008.09.019.

[9] L. Guo *et al.*, "Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas," *Geoderma,* vol. 398, p. 115118, 2021/09/15/ 2021, doi: https://doi.org/10.1016/j.geoderma.2021.115118.

[10] D. Hong *et al.*, "SpectralGPT: Spectral Remote Sensing Foundation Model," *IEEE Transactions on Pattern Analysis & Machine Intelligence,* vol. 46, no. 08, pp. 5227-5244, 2024, doi: 10.1109/TPAMI.2024.3362475.

[11] M. Zhao, S. Shi, J. Chen, and N. Dobigeon, "A 3-D-CNN Framework for Hyperspectral Unmixing With Spectral Variability," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 60, pp. 1-14, 2022, doi: 10.1109/TGRS.2022.3141387.

[12] Y. Ding *et al.*, "Multi-feature fusion: Graph neural network and CNN combining for hyperspectral image classification," *Neurocomputing,* vol. 501, pp. 246-257, 2022/08/28/ 2022, doi: https://doi.org/10.1016/j.neucom.2022.06.031.

[13] L. Xu *et al.*, "Estimation of Organic Carbon in Anthropogenic Soil by VIS-NIR Spectroscopy: Effect of Variable Selection," *Remote Sensing*, vol. 12, no. 20, doi: 10.3390/rs12203394.

[14] Y. Hong *et al.*, "Transferability of Vis-NIR models for Soil Organic Carbon Estimation between Two Study Areas by using Spiking," *Soil Science Society of America Journal,* vol. 82, no. 5, pp. 1231-1242, 2018/09/01 2018, doi: https://doi.org/10.2136/sssaj2018.03.0099.

[15] Y. Liu *et al.*, "Transferability of a Visible and Near-Infrared Model for Soil Organic Matter Estimation in Riparian Landscapes," *Remote Sensing*, vol. 6, no. 5, pp. 4305-4322doi: 10.3390/rs6054305.

[16] K. Gröchenig, "Localization of Frames,Banach Frames, and theInvertibility of the FrameOperator," *Journal of Fourier Analysis and Applications,* vol. 10, no. 2, pp. 105-132, 2004/03/01 2004, doi: 10.1007/s00041-004-8007-1.

[17] P. Balazs, N. Holighaus, T. Necciari, and D. Stoeva, "Frame Theory for Signal Processing in Psychoacoustics," in *Excursions in Harmonic Analysis, Volume 5: The February Fourier Talks at the Norbert Wiener Center*, R. Balan, J. J. Benedetto, W. Czaja, M. Dellatorre, and K. A. Okoudjou Eds. Cham: Springer International Publishing, 2017, pp. 225-268.

[18] J. Kovačević and A. Chebira, "An Introduction to Frames," *Foundations and Trends® in Signal Processing,* vol. 2, no. 1, pp. 1-94, 2007, doi: 10.1561/2000000006.

[19] A. Aldroubi, A. Baskakov, and I. Krishtal, "Slanted matrices, Banach frames, and sampling," *Journal of Functional Analysis,* vol. 255, no. 7, pp. 1667-1691, 2008/10/01/ 2008, doi: https://doi.org/10.1016/j.jfa.2008.06.024.

[20] A. J. Sykes *et al.*, "Characterising the biophysical, economic and social impacts of soil carbon sequestration as a greenhouse gas removal technology," *Global Change Biology,* vol. 26, no. 3, pp. 1085-1108, 2020/03/01 2020, doi: https://doi.org/10.1111/gcb.14844.

[21] P. Smith *et al.*, "How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal," *Global Change Biology,* vol. 26, no. 1, pp. 219-241, 2020/01/01 2020, doi: https://doi.org/10.1111/gcb.14815.

[22] R. P. Sishodia, R. L. Ray, and S. K. Singh, "Applications of Remote Sensing in Precision Agriculture: A Review," *Remote Sensing*, vol. 12, no. 19, doi: 10.3390/rs12193136.

[23] K. Lenhard, A. Baumgartner, P. Gege, S. Nevas, S. Nowy, and A. Sperling, "Impact of Improved Calibration of a NEO HySpex VNIR-1600 Sensor on Remote Sensing of Water Depth," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, no. 11, pp. 6085-6098, 2015, doi: 10.1109/TGRS.2015.2431743.

[24] S. Saini and L. Dewan, "Application of discrete wavelet transform for analysis of genomic sequences of Mycobacterium tuberculosis," (in eng), *Springerplus,* vol. 5, p. 64, 2016, doi: 10.1186/s40064-016-1668-9.

[25] B. K. Pandey, H. Tiwari, and D. Khare, "Trend analysis using discrete wavelet transform (DWT) for long-term precipitation (1851–2006) over India," *Hydrological Sciences Journal,* vol. 62, no. 13, pp. 2187-2208, 2017/10/03 2017, doi: 10.1080/02626667.2017.1371849.

[26] J. Yang, X. Li, and X. Ma, "Improving the Accuracy of Soil Organic Carbon Estimation: CWT-Random Frog-XGBoost as a Prerequisite Technique for In Situ Hyperspectral Analysis," *Remote Sensing*, vol. 15, no. 22, doi: 10.3390/rs15225294.

[27] G. Lee, R. Gommers, F. Waselewski, K. Wohlfahrt, and A. O'Leary, "PyWavelets: A Python package for wavelet analysis," *Journal of Open Source Software,* vol. 4, no. 36, p. 1237, 2019/04/12 2019, doi: 10.21105/joss.01237.

[28] X. Xu, S. Chen, Z. Xu, Y. Yu, S. Zhang, and R. Dai, "Exploring Appropriate Preprocessing Techniques for Hyperspectral Soil Organic Matter Content Estimation in Black Soil Area," *Remote Sensing*, vol. 12, no. 22, doi: 10.3390/rs12223765.

[29] T. Guo, T. Zhang, E. Lim, M. Lopez-Benitez, F. Ma, and L. Yu, "A Review of Wavelet Analysis and Its Applications: Challenges and Opportunities," *IEEE Access,* vol. 10, pp. 58869-58903, 2022, doi: 10.1109/access.2022.3179517.

[30] L. Shen, M. Gao, J. Yan, Z.-L. Li, P. Leng, Q. Yang, and S.-B. Duan, "Hyperspectral Estimation of Soil Organic Matter Content using Different Spectral Preprocessing Techniques and PLSR Method," *Remote Sensing*, vol. 12, no. 7, doi: 10.3390/rs12071206.

[31] L. P. A. Arts and E. L. van den Broek, "The fast continuous wavelet transformation (fCWT) for real-time, high-quality, noise-

resistant time–frequency analysis," *Nature Computational Science,* vol. 2, no. 1, pp. 47-58, 2022/01/01 2022, doi: 10.1038/s43588-021-00183-z.

[32] S. Kaewpijit, J. L. Moigne, and T. El-Ghazawi, "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 41, no. 4, pp. 863-871, 2003, doi: 10.1109/TGRS.2003.810712.

[33] W. Shao, J. Barras, and P. Kosmas, "A novel wavelets method for cancelling time-varying interference in NQR signal detection," *Signal Processing,* vol. 154, pp. 238-249, 2019/01/01/ 2019, doi: https://doi.org/10.1016/j.sigpro.2018.08.016.

[34] S. J. Ward, R. Layouni, S. Arshavsky-Graham, E. Segal, and S. M. Weiss, "Morlet Wavelet Filtering and Phase Analysis to Reduce the Limit of Detection for Thin Film Optical Biosensors," *ACS Sensors,* vol. 6, no. 8, pp. 2967-2978, 2021/08/27 2021, doi: 10.1021/acssensors.1c00787.

[35] A. Biswas, H. P. Cresswell, R. A. Viscarra Rossel, and B. C. Si, "Characterizing scale- and location-specific variation in non-linear soil systems using the wavelet transform," *European Journal of Soil Science,* vol. 64, no. 5, pp. 706-715, 2013/10/01 2013, doi: https://doi.org/10.1111/ejss.12063.

[36] H. M. Srivastava, A. Singh, A. Rawat, and S. Singh, "A family of Mexican hat wavelet transforms associated with an isometry in the heat equation," *Mathematical Methods in the Applied Sciences,* vol. 44, no. 14, pp. 11340-11349, 2021/09/30 2021, doi: https://doi.org/10.1002/mma.7492.

[37] S. Lovejoy and D. Schertzer, "Haar wavelets, fluctuations and structure functions: convenient choices for geophysics," *Nonlin. Processes Geophys.,* vol. 19, no. 5, pp. 513-527, 2012, doi: 10.5194/npg-19-513-2012.

[38] P. Vielva, "A Comprehensive Overview of the Cold Spot," *Advances in Astronomy,* vol. 2010, p. 592094, 2010/11/04 2010, doi: 10.1155/2010/592094.

[39] Y. F. Pu, "Fractional-Order Euler-Lagrange Equation for Fractional-Order Variational Method: A Necessary Condition for Fractional-Order Fixed Boundary Optimization Problems in Signal Processing and Image Processing," *IEEE Access,* vol. 4, pp. 10110-10135, 2016, doi: 10.1109/ACCESS.2016.2636159.

[40] J. Wang, J. Ding, A. Abulimiti, and L. Cai, "Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS–NIR) spectroscopy, Ebinur Lake Wetland, Northwest China," *PeerJ,* vol. 6, p. e4703, 2018/05/03 2018, doi: 10.7717/peerj.4703.

[41] J. Guan, J. Ou, Z. Lai, and Y. Lai, "Medical Image Enhancement Method Based on the Fractional Order Derivative and the Directional Derivative," *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 32, no. 03, p. 1857001, 2018/03/01 2017, doi: 10.1142/S021800141857001X.

[42] M. F. Noomen, A. K. Skidmore, F. D. van der Meer, and H. H. T. Prins, "Continuum removed band depth analysis for detecting the effects of natural gas, methane and ethane on maize reflectance," *Remote Sensing of Environment,* vol. 105, no. 3, pp. 262-270, 2006/12/15/ 2006, doi: https://doi.org/10.1016/j.rse.2006.07.009.

[43] M. Loum, M. Diack, N. Y. B. Ndour, and D. Masse, "Effect of the Continuum Removal in Predicting Soil Organic Carbon with Near Infrared Spectroscopy (NIRS) in the Senegal Sahelian Soils," *Open Journal of Soil Science,* vol. 06, no. 09, pp. 135-148, 2016, doi: 10.4236/ojss.2016.69014.

[44] W. L. Hwang, P. T. Huang, B. C. Kung, J. Ho, and T. L. Jong, "Frame-Based Sparse Analysis and Synthesis Signal Representations and Parseval K-SVD," *IEEE Transactions on Signal Processing,* vol. 67, no. 12, pp. 3330-3343, 2019, doi: 10.1109/TSP.2019.2916105.

[45] P. G. Casazza and R. G. Lynch, "A brief introduction to Hilbert space frame theory and its applications," 2015 2015, doi: 10.48550/ARXIV.1509.07347.

[46] H. G. Feichtinger and T. Strohmer, *Gabor Analysis and Algorithms: Theory and Applications*. Birkhäuser Boston, 2012.

[47] F. Bergeaud and S. Mallat, "Matching Pursuit of Images," in *Wavelet Analysis and Its Applications*, vol. 7, Y. Zeevi and R. Coifman Eds.: Academic Press, 1998, pp. 285-300.

[48] Y. Belov, A. Kulikov, and Y. Lyubarskii, "Gabor frames for rational functions," *Inventiones mathematicae,* vol. 231, no. 2, pp. 431-466, 2023/02/01 2023, doi: 10.1007/s00222-022-01151-8.

[49] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. Velasco, "Theory, implementation and applications of nonstationary Gabor frames," *Journal of Computational and Applied Mathematics,* vol. 236, no. 6, pp. 1481-1496, 2011/10/15/ 2011, doi: https://doi.org/10.1016/j.cam.2011.09.011.

[50] C. Rasmussen *et al.*, "Beyond clay: towards an improved set of variables for predicting soil organic matter content," *Biogeochemistry,* vol. 137, no. 3, pp. 297-306, 2018/02/01 2018, doi: 10.1007/s10533-018-0424-3.

[51] V. V. Moca, H. Bârzan, A. Nagy-Dăbâcan, and R. C. Mureşan, "Time-frequency super-resolution with superlets," *Nature Communications,* vol. 12, no. 1, p. 337, 2021/01/12 2021, doi: 10.1038/s41467-020-20539-9.

[52] A. J. E. M. Janssen, "The duality condition for Weyl-Heisenberg frames," in *Gabor Analysis and Algorithms: Theory and Applications*, H. G. Feichtinger and T. Strohmer Eds. Boston, MA: Birkhäuser Boston, 1998, pp. 33-84.

[53] M. Dörfler, "Quilted Gabor frames – A new concept for adaptive time-frequency representation," *Advances in Applied Mathematics,* vol. 47, no. 4, pp. 668-687, 2011/10/01/ 2011, doi: https://doi.org/10.1016/j.aam.2011.02.007.

[54] C. R. Harris *et al.*, "Array programming with NumPy," *Nature,* vol. 585, no. 7825, pp. 357-362, 2020/09/01 2020, doi: 10.1038/s41586-020-2649-2.

[55] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research,* vol. 12, no. 85, pp. 2825-2830, 2011 2011. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html.

[56] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, "Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization," *Journal of Machine Learning Research,* vol. 18, no. 185, pp. 1-52, 2018 2018. [Online]. Available: http://jmlr.org/papers/v18/16-558.html.

[57] B. Koirala, Z. Zahiri, and P. Scheunders, "A Robust Supervised Method for Estimating Soil Moisture Content From Spectral Reflectance," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 60, pp. 1-13, 2022, doi: 10.1109/TGRS.2022.3212600.

[58] T. Shi, Y. Chen, H. Liu, J. Wang, and G. Wu, "Soil Organic Carbon Content Estimation with Laboratory-Based Visible–Near-Infrared Reflectance Spectroscopy: Feature Selection," *Applied Spectroscopy,* vol. 68, no. 8, pp. 831-837, 2014/08/01 2014, doi: 10.1366/13-07294.

[59] A. Bayer, M. Bachmann, A. Müller, and H. Kaufmann, "A Comparison of Feature-Based MLR and PLS Regression Techniques for the Prediction of Three Soil Constituents in a Degraded South African Ecosystem," *Applied and Environmental Soil Science,* vol. 2012, p. 971252, 2012/09/17 2012, doi: 10.1155/2012/971252.