In Situ Spectral Analysis for Soil Carbon Measurement

Jose A. Cortes[a], Andrzej Korzeniowski[a], Galina Yakubova[b], Aleksandr Kavetskiy[b], H. Allen Torbert[b]

[a] *University of Texas at Arlington, Arlington, 76019, TX, USA*
[b] *USDA ARS Auburn Lab, Auburn, AL, USA*

Abstract

Soil carbon is a key component of soil health and plays a crucial role in the global carbon cycle. Accurate measurement of carbon is essential for measuring soil quality and its impact on the environment. Traditional methods for measuring carbon are often time-consuming, expensive, and require laboratory analysis. In situ spectral analysis offers an alternative for rapid and non-destructive measurement of soil carbon. This paper explores the use of spectral analysis techniques to measure carbon levels in various soil types. We simulate common soil types and apply different spectral analysis methods, including peak fitting, component fitting, singular value decomposition, and deep learning, to evaluate their effectiveness in measuring soil carbon. The results demonstrate that peak fitting with exponential falloff baseline achieves the lowest mean squared error (7.66×10^{-5}), followed by component analysis methods. The study shows that spectral analysis methods can provide accurate soil carbon measurements, with convolution techniques improving overall accuracy across all methods.

Keywords: soil carbon, spectral analysis, neutron activation, gamma-ray spectroscopy, MCNP simulation, peak fitting, component analysis, deep learning

Email addresses: jose.cortes@mavs.uta.edu (Jose A. Cortes), akorzeni@uta.edu (Andrzej Korzeniowski), galina.yakubova@usda.gov (Galina Yakubova), aleksandr.kavetskiy@usda.gov (Aleksandr Kavetskiy), allen.torbert@usda.gov (H. Allen Torbert)

1. Introduction

Soil carbon is a key component of soil health and plays a crucial role in the global carbon cycle. Accurate measurement of carbon is essential for measuring soil quality, and its impact on the environment [1]. Traditional methods for measuring carbon are often time-consuming, expensive, and require laboratory analysis [2]. In situ spectral analysis offers an alternative for rapid and non-destructive measurement of soil carbon [3].

This paper explores the use of spectral analysis techniques to measure carbon levels in various soil types. We simulate common soil types and apply different spectral analysis methods, including peak fitting, component fitting, singular value decomposition, and deep learning, to evaluate their effectiveness in measuring soil carbon.

2. Data Generation

2.1. Common Soil Types

To investigate the effectiveness of spectral analysis methods for soil carbon measurement, we simulate a range of common soil types. The simulated data includes spectral readings across different wavelengths, capturing the unique spectral signatures of each soil type. This data serves as a foundation for applying various spectral analysis techniques.

Table 1: Elemental compositions used in MCNP simulations

Element	MCNP Identifier	Density (g/cm^3)
Si	14028	2.33
Al	13027	2.7
H	1001	0.001
Na	11023	0.97
O	8016	0.00143
Fe	26000	7.87
Mg	12024	1.74
C	6000	2.33

Table 2: Material compositions for soil simulation

Material	Compound Makeup (by weight)	Density (g/cm^3)
Silica	SiO_2 (76.4%), Al_2O_3 (23.6%)	2.32
Kaolinite	SiO_2 (46.5%), Al_2O_3 (39.5%), H_2O (14.0%)	3.95
Smectite	SiO_2 (66.7%), Al_2O_3 (28.3%), H_2O (5.0%)	2.785
Montmorillonite	SiO_2 (73.7%), Al_2O_3 (24.6%), H_2O (1.7%)	2.7
Quartz	SiO_2 (100.0%)	2.62
Chlorite	SiO_2 (30.0%), Al_2O_3 (24.0%), Fe_2O_3 (23.3%), H_2O (22.7%)	2.6
Mica	SiO_2 (48.9%), Al_2O_3 (40.3%), H_2O (10.8%)	2.7
Feldspar	SiO_2 (68.0%), Al_2O_3 (32.0%)	2.55
Coconut	C (100.0%)	0.53

2.2. Simulation in MCNP

MCNP6 [4] was used to simulate gamma-ray spectra resulting from neutron activation of soil samples. Each simulation modeled a soil matrix with varying concentrations of carbon and other common soil constituents. The geometry was set up to mimic in situ measurement conditions, with a neutron source placed above a soil slab and a detector positioned to capture emitted gamma

rays [5].

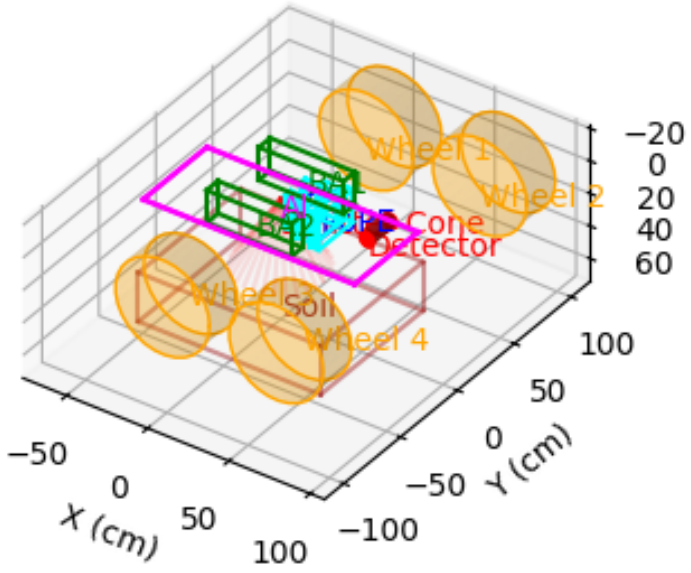

Figure 1: Geometry of MCNP simulation setup showing neutron source, soil slab, and detector configuration

Key simulation parameters included:

- **Neutron source energy:** API120 portable neutron (D-T generator) generator [6]
- **Soil slab dimensions:** 112 cm × 90 cm × 30 cm
- **Detector type:** Geiger-Mueller (G-M) detector [7]
- **Tally:** F8 (pulse height tally) for gamma spectra

This approach enables the generation of realistic spectral data for a variety of soil compositions, forming the basis for evaluating different spectral analysis techniques.

2.3. Spectral Readings

The spectral readings obtained from the MCNP simulations provide a detailed representation of the gamma-ray emissions from the soil samples. Mathematically it is a probability density function (PDF) of the energy distribution of the emitted gamma rays.

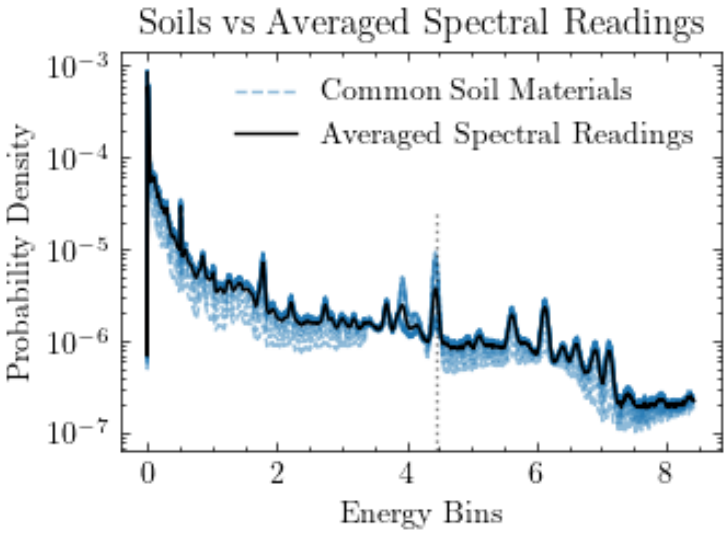

Figure 2: Example MCNP spectral reading showing gamma-ray energy distribution

2.4. Training and Testing Data

The training data for the spectral analysis methods is picked from the edge cases of the simulated data. This includes the highest and lowest carbon levels both as would be found in simulation as well as natural soils. The testing data is all cases of the simulated data, excluding the training data.

Table 3: Carbon level classifications

Carbon Level	Associated Amount
Natural	0%-6% Carbon
High	6%-100% Carbon

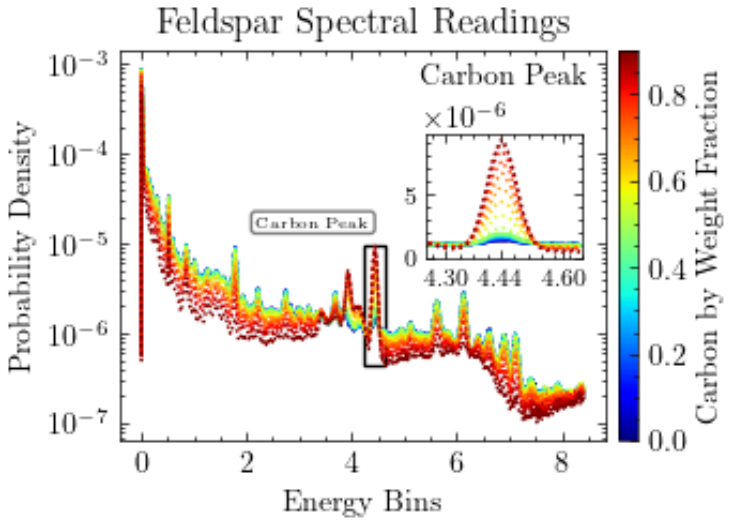

Figure 3: Feldspar spectral reading by carbon level showing the variation in spectral signatures with different carbon concentrations

2.5. Data Convolution

In the context of spectral analysis, MCNP can be used to simulate the interaction of radiation with soil materials, providing spectrums to analyze. Linear Convolution is used to quickly predict spectral readings for material mixtures by combining the spectral signatures of individual components. This does not account for the complex interactions between materials, but it provides a simplified approach to generate spectral data for analysis. The error metric for this convolution method is based on the difference between the simulated spectral readings and the readings obtained from MCNP simulations. The effects of convolution as training data on the analysis results will be investigated in the results section.

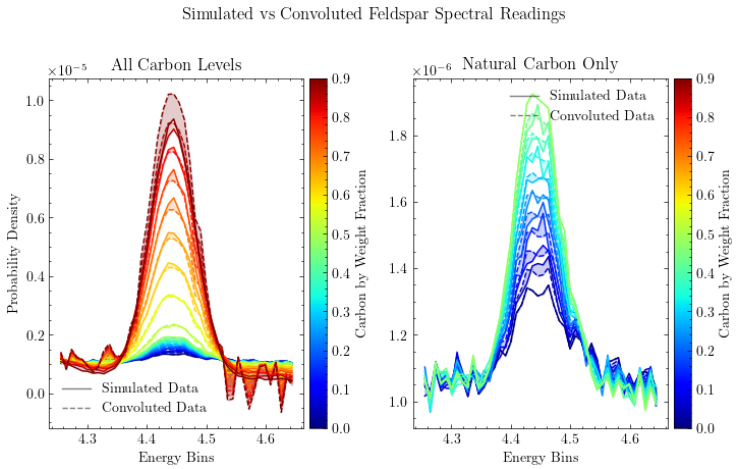

Figure 4: Comparison of simulated vs convoluted data for Feldspar spectral readings

3. Analysis Methods of Spectral Readings

This section explores various spectral analysis methods applied to the simulated spectral readings. Each method is evaluated for its effectiveness in measuring Carbon levels. Error is calculated using mean squared error (MSE) between the predicted and actual carbon levels in the test data.

3.1. Peak Fitting

Peak fitting involves using the least-squares method in identifying and quantifying the baseline and peaks in the spectral data that correspond to specific soil components [8]. This method is useful for extracting information about the concentration of individual elements or compounds in the soil. For effective peak fitting, the data is filtered to focus on the peak area.

The fitting function is defined as:

$$F_f = F_p + F_b \tag{1}$$

where F_p is the peak function (e.g., Gaussian) and F_b is the baseline function (e.g., linear or exponential falloff).

Table 4: Function parameterizations for peak fitting

Function Type	Example Expression
Linear	$ax + b$
Exp Falloff	$a \cdot \exp(-b \cdot x) + c$
Gaussian	$a \cdot \exp(-((x-b)^2)/c^2) + d$

The SciPy Python package is used for the fitting process [9], leveraging its curve fitting capabilities to refine the initial parameter estimates. The baseline function is subtracted from the fitted function to isolate the peak, and the area under the peak is calculated to quantify the concentration of the corresponding element or compound in the soil.

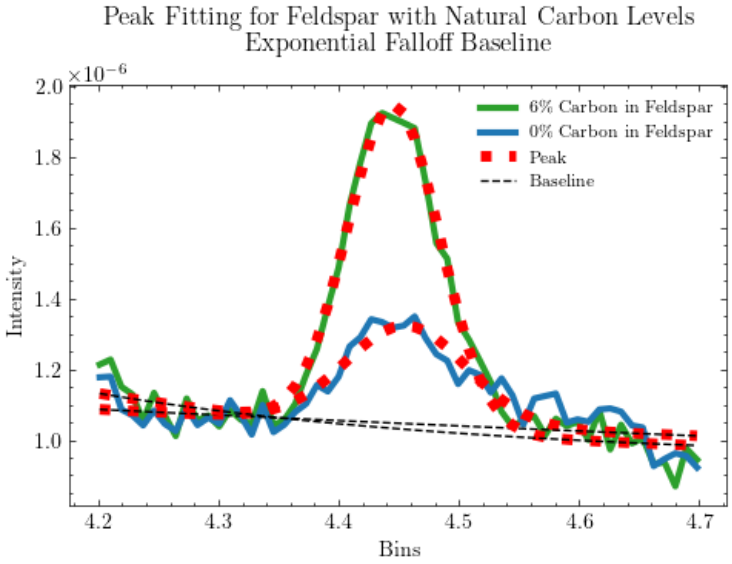

Figure 5: Peak fitting example showing fitted peak and baseline for feldspar spectrum

An activation layer of linear regression is used to compare the peak areas to known soil carbon concentrations, allowing for the calibration of the model's predictions.

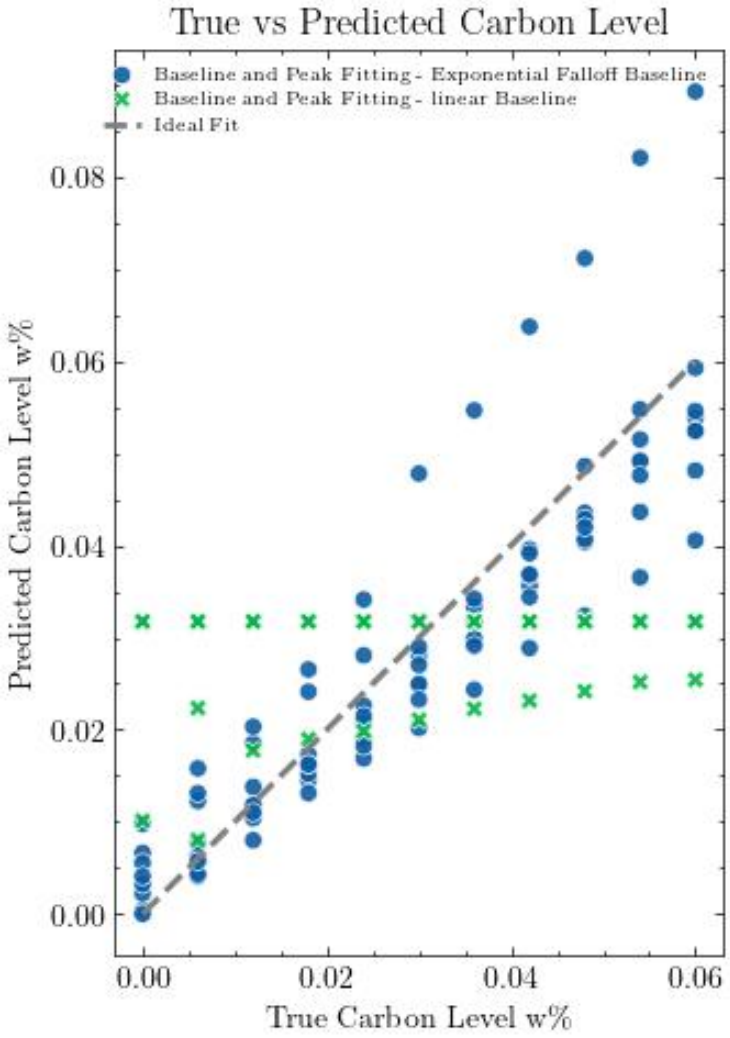

Figure 6: Peak fitting prediction results showing carbon level vs predicted values

3.2. Component Fitting

Component fitting involves modeling the spectral data as a combination of known spectral signatures of soil components. This method allows for the estimation of the concentration of multiple components in the soil based on their spectral contributions.

The combined spectral function is defined as:

$$F_c = \sum_i A_i \cdot F_i \tag{2}$$

where F_c is the combined spectral function, A_i are the coefficients representing the concentration of each component, and F_i are the spectral functions of individual components.

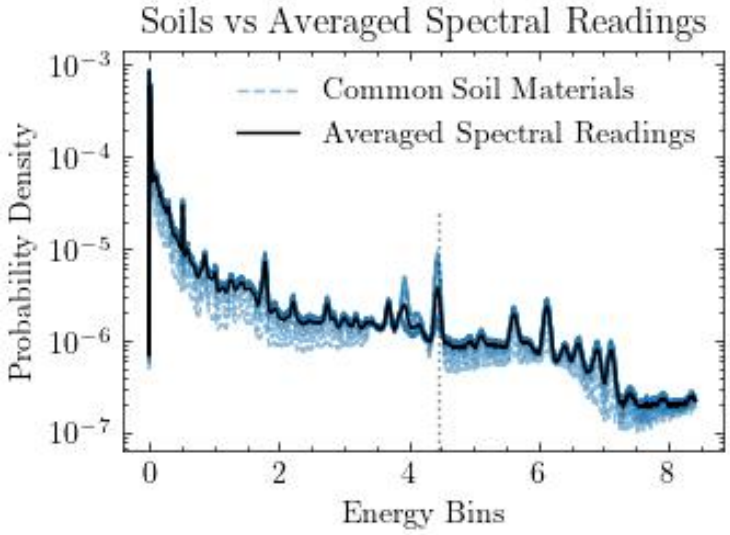

Figure 7: Common soil spectra vs average soil spectrum showing the spectral signatures of different soil components

Components can be any known spectral signature, either from pure elemental samples [10] or from the average of a set of soil samples. The fitting process involves adjusting the coefficients A_i to minimize the difference between the combined spectral function F_c and the observed spectral data. This method also benefits from filtering of low energy signals which are generally more likely to be caused by noise.

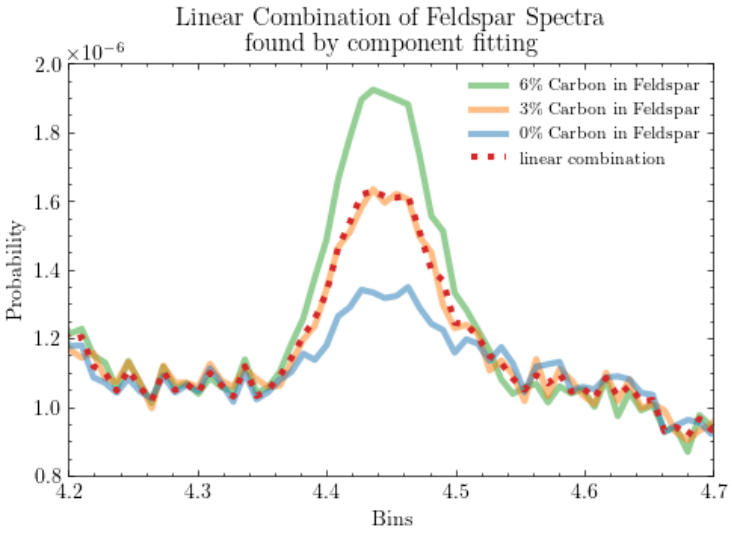

Figure 8: Component fitting process showing linear combination of spectral components for feldspar analysis

The carbon coefficient A_C is then used to estimate the Carbon level in the soil. This method is particularly useful for analyzing complex soil mixtures where multiple known components contribute to the spectral signature. This method is also generalizable to study other elements or compounds.

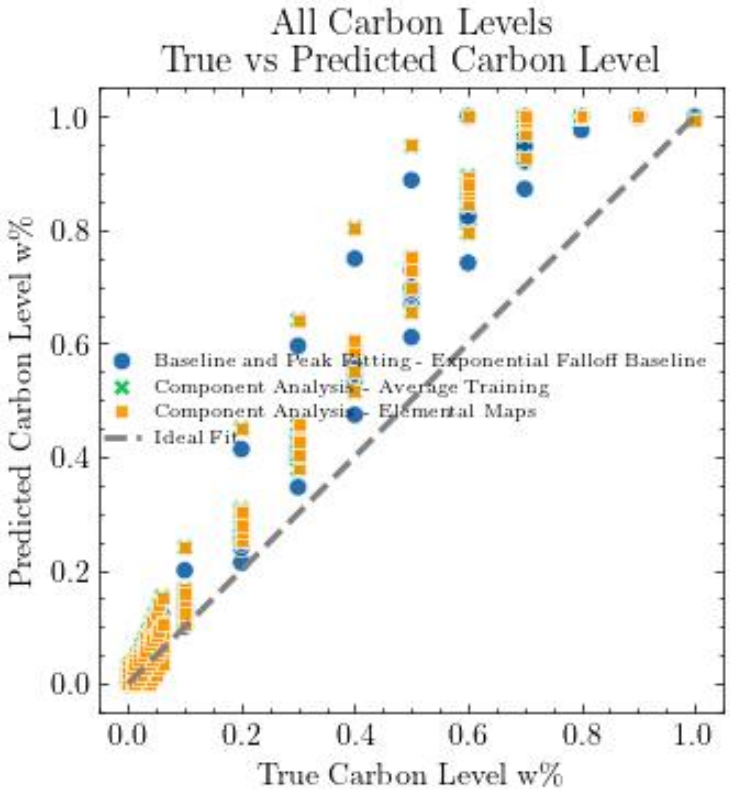

Figure 9: Component fitting prediction results showing carbon level vs predicted values

3.3. Singular Value Decomposition (SVD)

Singular Value Decomposition is a mathematical technique used to decompose the spectral data into convex components [11]. The resulting singular values inside the strong window can be summed to provide a measure of the concentration of carbon in the soil.

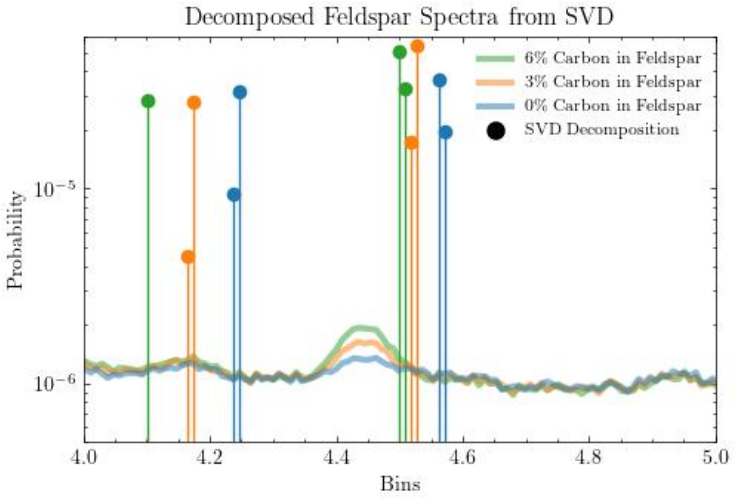

Figure 10: Convex optimization process showing SVD decomposition of feldspar spectral data

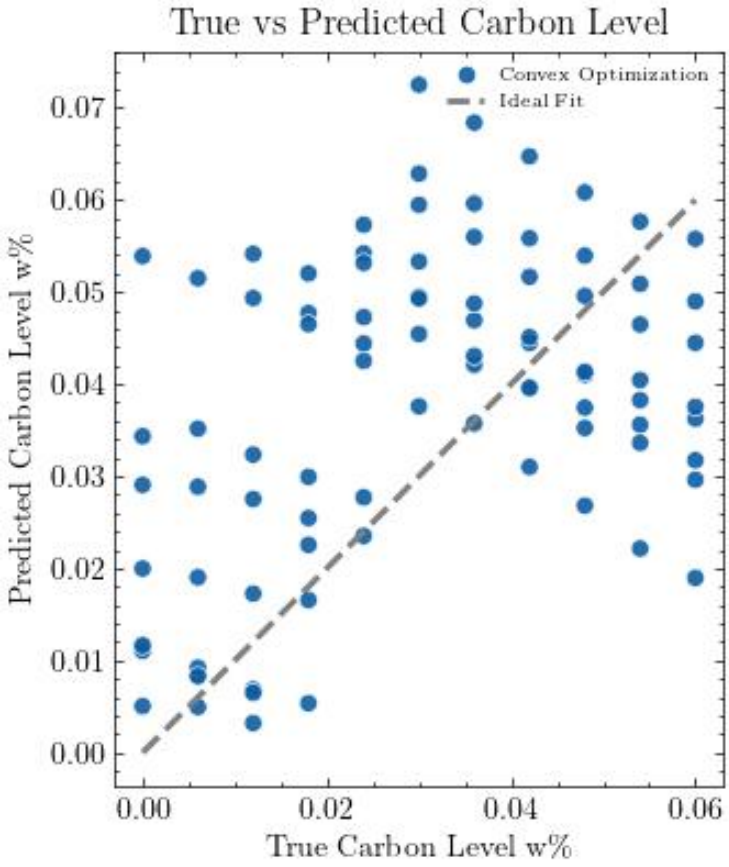

Figure 11: Convex optimization predictions showing carbon level vs predicted values using SVD

3.4. Deep Learning

Deep learning techniques, such as convolutional neural networks (CNNs), can be applied to spectral data for feature extraction and classification. These methods can learn complex relationships in the data and provide robust predictions of carbon levels based on spectral readings. The most important difference between deep learning and the previous methods is that it requires a large amount of training data to be effective. One method by Kim et al. [12] uses a deep learning model to predict existence, concentration and carbon peak areas.

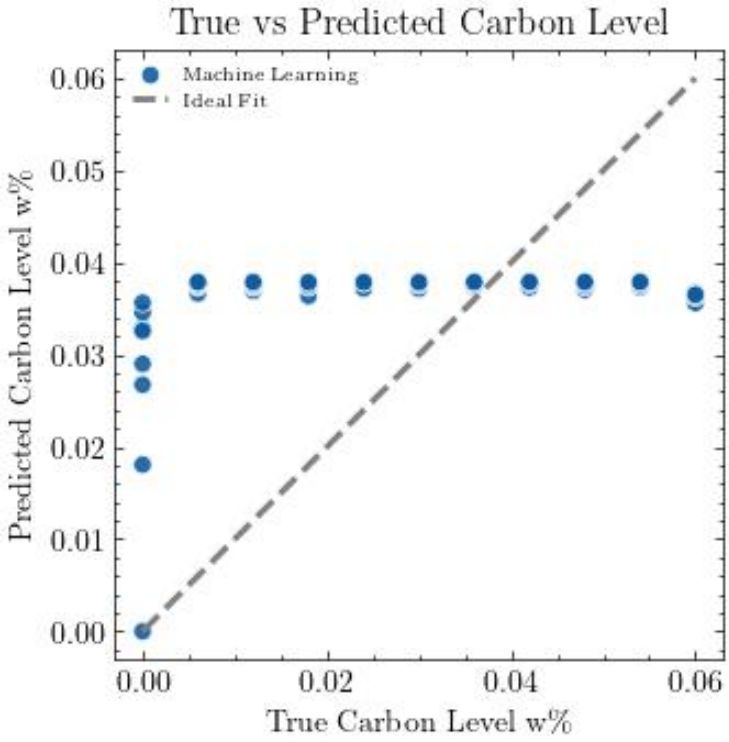

Figure 12: Deep learning prediction results showing carbon level vs predicted values using machine learning optimization

4. Results

The effectiveness of each method in measuring carbon levels is evaluated based on accuracy using mean squared error (MSE) as the metric. The results are summarized in Table 5.

Table 5: Performance comparison of spectral analysis methods

Method Group	Method	MSE
Peak Fitting	Exponential Falloff Baseline	7.66×10^{-5}
Component Analysis	Elemental Maps	2.10×10^{-4}
Component Analysis	Average Training	3.43×10^{-4}
Peak Fitting	Linear Baseline	3.52×10^{-4}
Machine Learning	Deep Learning	3.67×10^{-4}
Convex Optimization	SVD	4.27×10^{-4}

4.1. Comparing Analysis Methods

Peak fitting with exponential falloff baseline is the most effective method for measuring carbon levels in soil, achieving the lowest MSE (7.66×10^{-5}). Component analysis methods also perform well, with elemental maps achieving the second-best performance. The deep learning method shows promise but requires further optimization to improve its performance.

4.2. Effects of Carbon Levels on Results

Table 6: Method performance by carbon level

Carbon Level	MSE Values						
	Peak Fitting (Exp Falloff)	Peak Fitting (Linear)	Component (Average)	Component (Elemental)	Convex Optimization	Filtered ML	
Agricultural	7.66×10^{-5}	3.52×10^{-4}	3.43×10^{-4}	2.10×10^{-4}	4.27×10^{-4}	3.70×10^{-3}	3.
All	1.42×10^{-2}	3.48×10^{-2}	1.92×10^{-2}	1.92×10^{-2}	2.64×10^{-2}	1.33×10^{-1}	1.

Lower carbon levels tend to result in higher MSE values across all methods, indicating that the spectral signatures of low-carbon soils are less distinct and more challenging to analyze accurately. The methods generally perform better with higher carbon concentrations, where the spectral features are more pronounced.

4.3. Effects of Convolution on Results

Table 7: Method performance by training dataset type

Dataset	MSE Values					
	Peak Fitting (Exp Falloff)	Peak Fitting (Linear)	Component (Average)	Component (Elemental)	Convex Optimization	Filtered
Convolution Training	7.26×10^{-5}	3.50×10^{-4}	2.95×10^{-4}	1.91×10^{-4}	2.93×10^{-4}	3.65×1
Feldspar	3.43×10^{-5}	3.60×10^{-4}	7.48×10^{-7}	7.92×10^{-7}	2.14×10^{-3}	1.10×1
Material Mixes	7.66×10^{-5}	3.52×10^{-4}	3.43×10^{-4}	2.10×10^{-4}	4.27×10^{-4}	3.70×1

Convolution generally improves the accuracy of spectral analysis methods by smoothing out noise and enhancing the signal-to-noise ratio. The results show that convolution leads to lower MSE values across all methods, indicating that it is beneficial for spectral analysis in soil carbon measurement.

5. Discussion

5.1. Conclusions

The study demonstrates the potential of spectral analysis methods for measuring soil carbon levels. Peak fitting and component fitting methods show the best performance, while deep learning techniques require further refinement. Convolution is beneficial for improving the accuracy of spectral analysis.

The peak fitting method with exponential falloff baseline achieved the highest accuracy, suggesting that proper baseline correction is crucial for effective spectral analysis. Component analysis methods provide a good balance between accuracy and interpretability, making them suitable for practical applications.

5.2. Future Work

Future work will focus on:

- Optimization of deep learning architectures for spectral analysis

- Investigation of hybrid methods combining multiple analysis techniques
- Field validation of the proposed methods
- Extension to other soil properties beyond carbon content

6. Acknowledgments

We acknowledge the contributions of the USDA scientists for their guidance and support in this research. The spectral data generated in this study is available for further research and validation.

References

[1] R. Lal, J. M. Kimble, R. F. Follett, B. A. Stewart (Eds.), Soil Processes and the Carbon Cycle, CRC Press, Boca Raton, 2018. doi:10.1201/9780203739273.

[2] A. De Ros, I. Piccoli, L. Sartori, B. Portelli, G. Serra, N. Dal Ferro, F. Morari, Advancing soil texture and organic carbon spatial variability assessment: Integrating proximal gamma-ray spectroscopy and electromagnetic induction via data fusion for site-independent analysis, CATENA 254 (2025) 108980. doi:10.1016/j.catena.2025.108980.
URL https://www.sciencedirect.com/science/article/pii/S0341816225002826

[3] G. Yakubova, A. Kavetskiy, S. A. Prior, H. A. Torbert, Tagged neutron method for carbon analysis of large soil samples, Applied Radiation and Isotopes 150 (2019) 127–134. doi:10.1016/j.apradiso.2019.05.028.
URL https://www.sciencedirect.com/science/article/pii/S0969804319301381

[4] C. J. Werner, J. C. Armstrong, F. B. Brown, J. S. Bull, L. Casswell, L. J. Cox, D. A. Dixon, R. A. Forster, J. T. III, Goorley, H. G. Hughes, J. A. III, Favorite, R. L. Martz, S. G. Mashnik, M. E. Rising, C. J. Solomon, A. Jr.,

Sood, J. E. Sweezy, A. J. Zukaitis, MCNP User's Manual Code Version 6.2, Tech. Rep. LA-UR-17-29981, Los Alamos National Laboratory, Los Alamos, NM, USA (Dec. 2017).

[5] A. Kavetskiy, G. Yakubova, S. A. Prior, H. A. Torbert, Neutron-Stimulated Gamma Ray Analysis of Soil, in: New Insights on Gamma Rays, IntechOpen, 2017. doi:10.5772/68014.

[6] A. Kavetskiy, G. Yakubova, S. A. Prior, H. A. Torbert, Energy correlated timing spectra in target neutron techniques, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 433 (2018) 80–86. doi:10.1016/j.nimb.2018.07.028.

[7] G. Yakubova, A. Kavetskiy, S. A. Prior, H. A. Torbert, Measuring and mapping moisture content in agricultural fields by neutron-gamma analysis, Soil and Tillage Research 248 (2025) 106444. doi:10.1016/j.still.2024.106444.
URL https://www.sciencedirect.com/science/article/pii/S0167198724004458

[8] R. P. Gardner, X. Ai, C. R. Peeples, J. Wang, K. Lee, J. L. Peeples, A. Calderon, Use of an iterative convolution approach for qualitative and quantitative peak analysis in low resolution gamma-ray spectra, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 652 (2011) 544–549. doi:10.1016/j.nima.2010.12.224.

[9] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0: fundamental algorithms

for scientific computing in Python, Nature Methods 17 (2020) 261–272. doi:10.1038/s41592-019-0686-2.

[10] A. Kavetskiy, G. Yakubova, S. A. Prior, H. A. Torbert, Neutron gamma analysis of soil carbon: Post-irradiation physicochemical effects, Environmental Technology & Innovation 31 (2023) 103219. doi:10.1016/j.eti.2023.103219.

[11] B. Liu, H. Yang, H. Lv, F. Jing, X. Gao, M. Yan, A deconvolution method for scintillator gamma-ray spectrum analysis based on convex optimization, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 957 (2020) 163399. doi:10.1016/j.nima.2020.163399.

[12] W. Kim, K. Ko, J. Park, S. Lee, H. Yun, G. Cho, Deep learning-based gamma spectroscopic analysis considering multiple variables for in situ applications, Radiation Physics and Chemistry 226 (2025) 112261. doi:10.1016/j.radphyschem.2024.112261.