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 Neutron-Gamma spectral analysis (NGSA) offers an alternative for rapid and non-destructive measurement of soil carbon. This paper explores the use of NGSA techniques to measure carbon levels in various soil types. We simulate in MCNP6.2 common soil types and apply different NGSA methods, including peak fitting, component fitting, singular value decomposition, and deep learning, to evaluate their effectiveness in measuring of 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 NGSA methods can provide accurate soil carbon measurements, with convolution techniques improving overall accuracy across all methods.

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

Email addresses: `jose.cortes@uta.edu` (Jose A. Cortes), `korzeniowski@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 [?]. In situ spectral analysis offers an alternative for rapid and non-destructive measurement of soil carbon [2]. This paper explores the use of NGSA techniques to measure carbon levels in various soil types. We simulate in MCNP6.2 common soil types and apply different NGSA 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 NGSA 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 NGSA techniques.

Table 1: Material compositions for soil simulation

Material	C %	H %	O %	Si %	Na %	Al %	K %
Carbon	100.0	0.0	0.0	0.0	0.0	0.0	0.0
Water	0.0	11.2	88.8	0.0	0.0	0.0	0.0
Quartz	0.0	0.0	53.3	46.7	0.0	0.0	0.0
Feldspar	0.0	0.0	48.8	32.1	8.8	10.3	0.0
Mica	0.0	0.5	48.2	21.2	0.0	20.3	9.8

Table 1 presents the elemental composition of common soil materials used in the simulations. Mechanical mixing is used to combine materials based on

their weight proportions, such that a 50% carbon and 50% water mix would have a composition of 50% carbon, 5.6% hydrogen, and 44.4% oxygen. To measure the effectiveness of NGSA methods for carbon measurement, we simulate combinations of soil materials with varying carbon (C) and moisture (Water) content.

2.2. Simulation in MCNP

MCNP6 [3] was used to simulate gamma-ray spectra resulting from neutron activation of soil samples. Each simulation modeled a soil sample undergoing neutron activation 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 [4].

In figure 1, the side view of the MCNP simulation setup is shown, highlighting the arrangement of the neutron source, soil slab, and detector. The geometry is designed to accurately represent the in situ conditions under which the soil samples are analyzed.
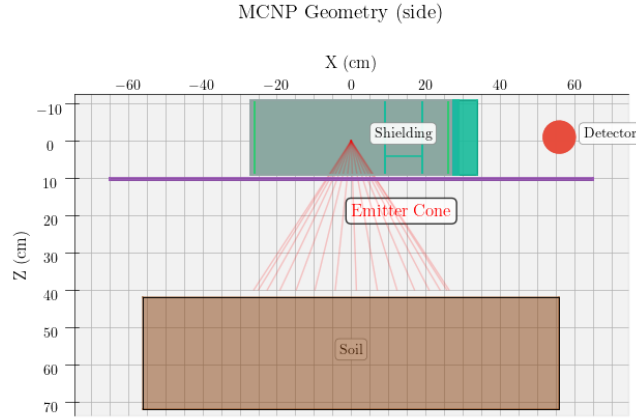

Figure 1: Geometry of MCNP simulation setup showing neutron source, soil slab, and detector configuration (Side View)

In figure 2, the top view of the MCNP simulation setup is presented, pro-

3

viding a clear view of the spatial arrangement of the components, including the shielding around the detector to minimize background radiation.
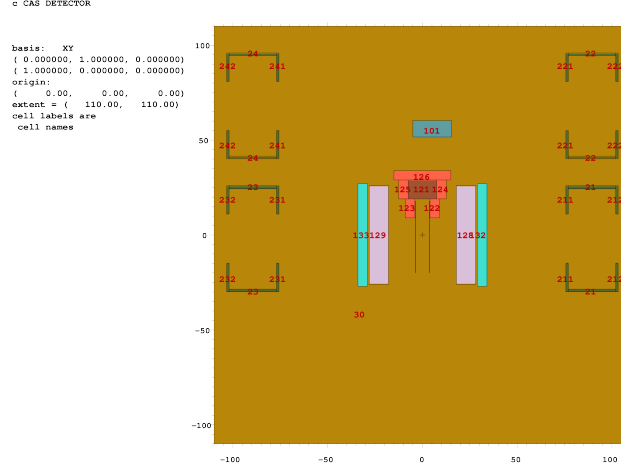

Figure 2: Cross section of MCNP simulation setup showing wheels and shielding (Top View)

Key simulation parameters included:

- **Neutron source energy:** API120 portable neutron (D-T generator) generator [5]

- **Soil slab dimensions:** 112 cm × 90 cm × 30 cm

- **Detector type:** NaI detector [6]

- **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 NGSA techniques. Table 2 lists the MCNP chemical identifiers for the elements considered in the simulations.

4

	Table 2: MCNP Chemical Identifiers						
Element	C	H	O	Si	Na	Al	K
MCNP Identifier	6000	1001	8016	14000	11023	13027	19000

2.3. Spectral Readings

The spectral readings obtained from the MCNP simulations provide a detailed distribution of the gamma-ray emissions from the soil samples per neutron, similar to what would be measured in a real-world scenario.
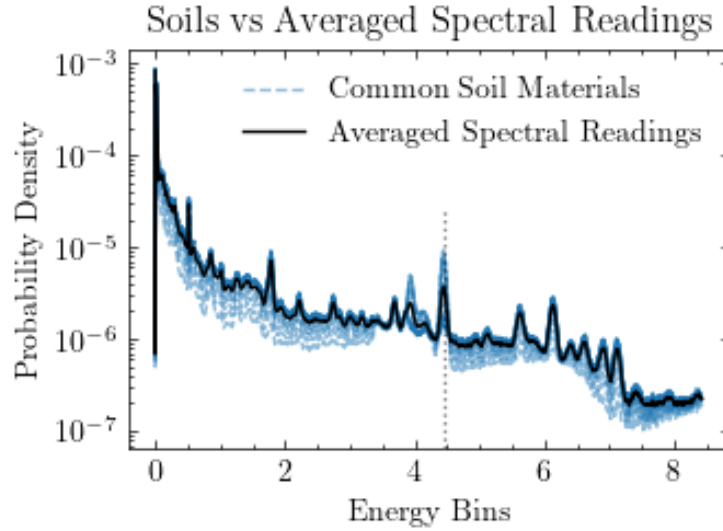

Figure 3: Example MCNP spectral reading showing gamma-ray energy distribution

2.4. Training and Testing Data

The training data for the NGSA 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. The data also ranges from 0% to 80% moisture content.

Table 3: Carbon level classifications

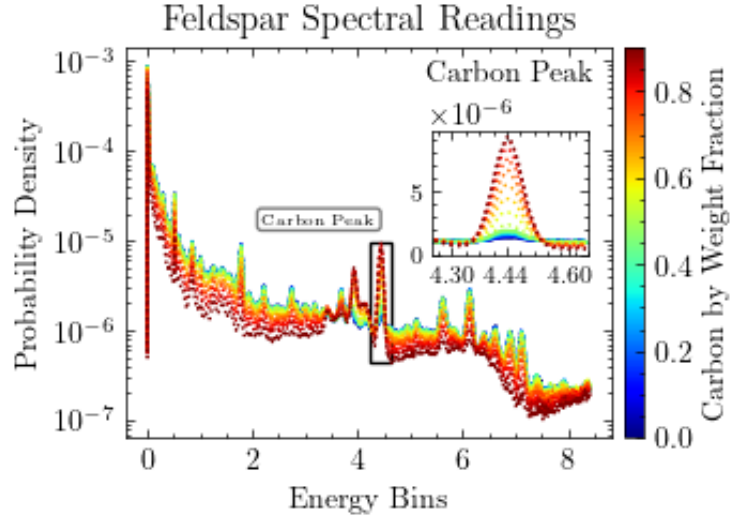| Carbon Level	Associated Amount
Natural	0%-15% Carbon
High	15%-80% Carbon

Figure 4: Feldspar spectral reading by carbon level showing the variation in spectral signatures with different carbon concentrations

2.5. Data Convolution

In the context of NGSA, 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.
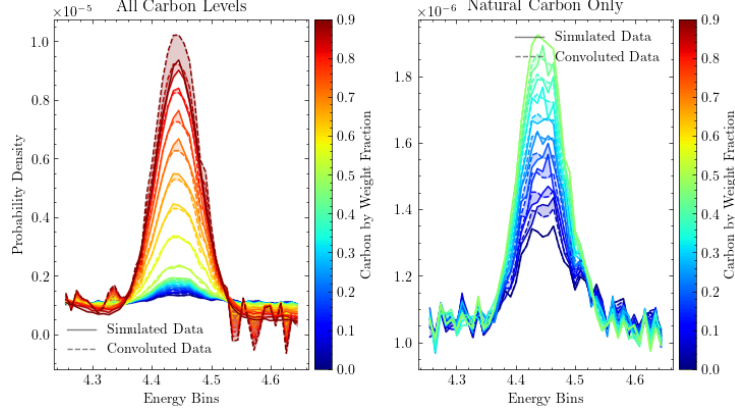
Figure 5: Comparison of simulated vs convoluted data for Feldspar spectral readings

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. Calibration Layer

All models undergo a calibration process to align their predictions with the carbon measurements. This involves a regression model between a key characteristic and the predicted values. A linear regression model is used for this purpose.

The Scipy python package is used for the fitting process[7], leveraging its curve fitting capabilities to refine the initial parameter estimates. All fitting problems are taken as fitting a curve f(x, p0) where p0 are the initial parameters. The fitting process iteratively adjusts these parameters to minimize the difference between the predicted and actual values, using a least-squares approach. The Levenberg-Marquardt algorithm is employed to optimize the fitting

7

process[8]. When the fitting is bounded, Trust Region Reflective optimization [9] is used.

3.2. Peak Baseline 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 [10]. 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 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.
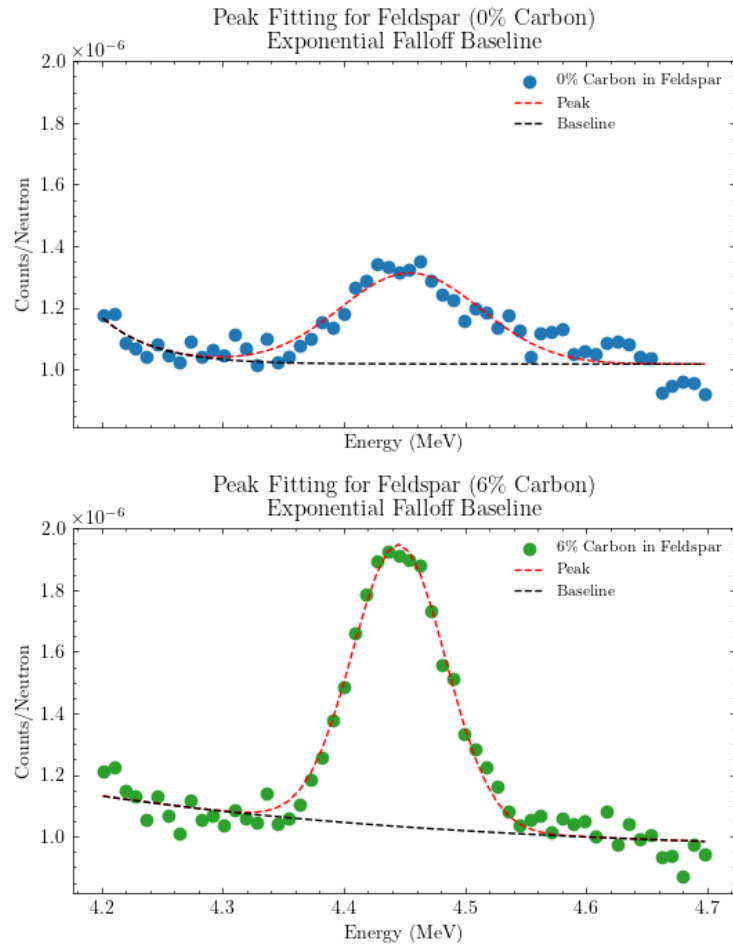
Figure 6: Peak fitting example showing fitted peak and baseline for feldspar spectrum

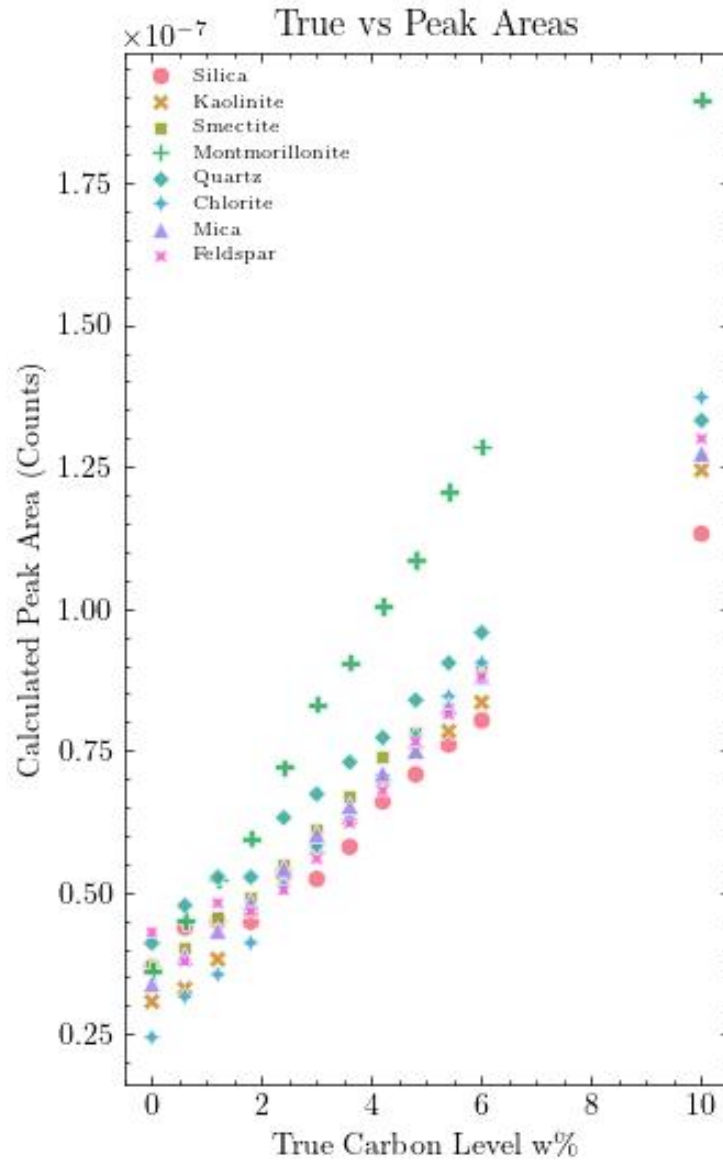The final prediction is calibrated using the peak areas.

Figure 7: Peak fitting prediction results showing carbon level vs predicted values

3.3. 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

10

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.
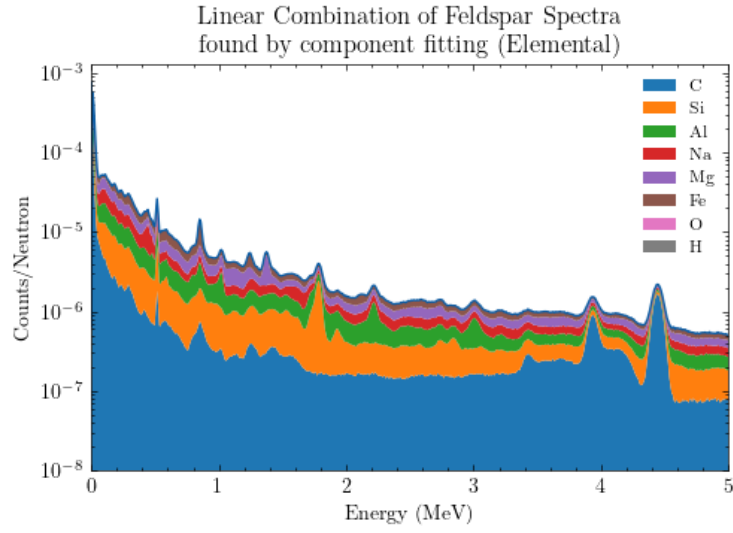

Figure 8: Component fitting process showing linear combination of elemental spectral components for feldspar analysis
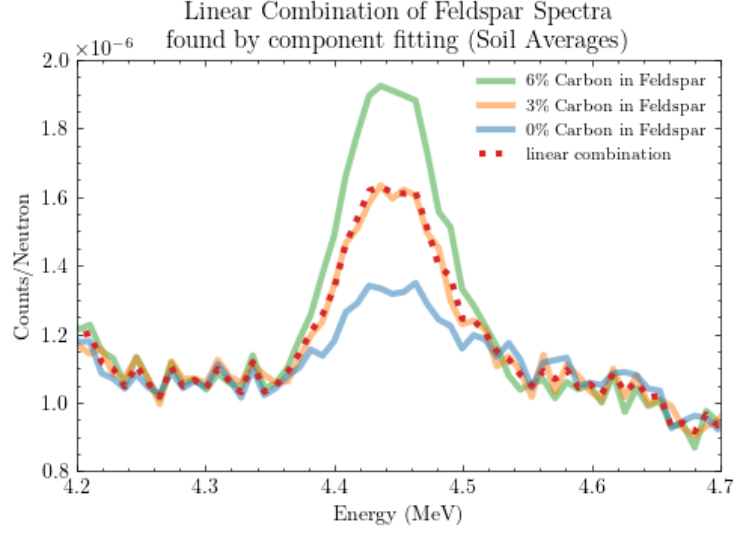
Figure 9: Component fitting process showing linear combination of spectral components for feldspar analysis

Components can be any known spectral signature, either from pure elemental samples [11] as shown in figure 8 or derived from soil samples similar to the target soil 9. 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.

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.
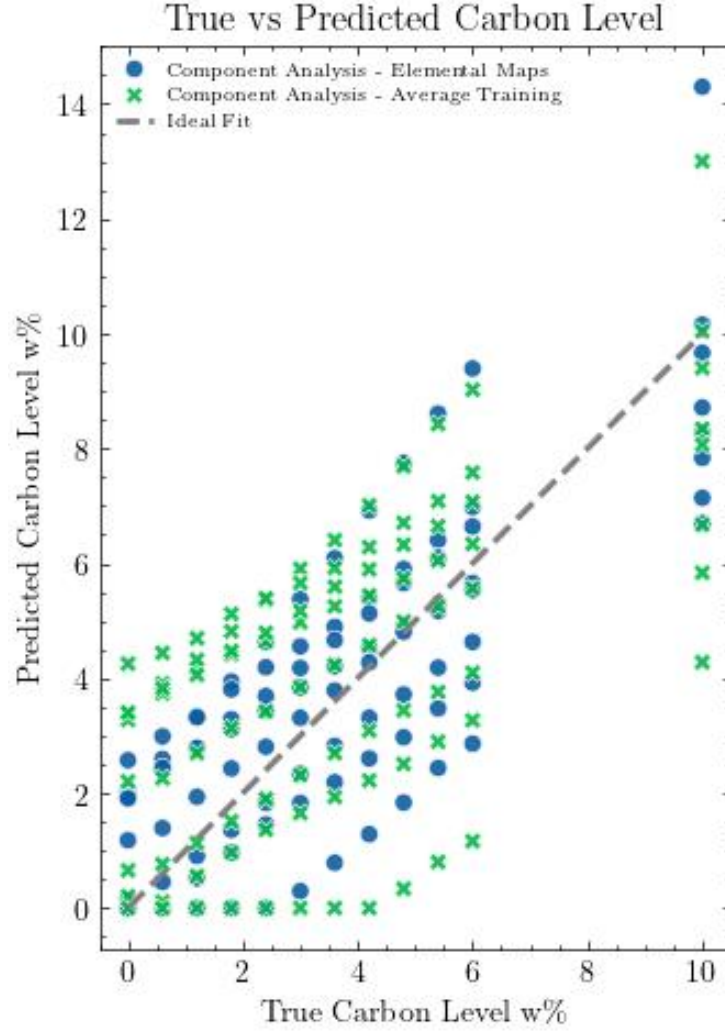
Figure 10: Component fitting prediction results showing carbon level vs predicted values

3.4. Convex Optimization

Convex optimization techniques can be applied to the spectral data to iden-
tify and quantify the contributions of different soil components. By formulating
the problem as a convex optimization task, it is possible to find the optimal
coefficients for the spectral components that minimize the difference between
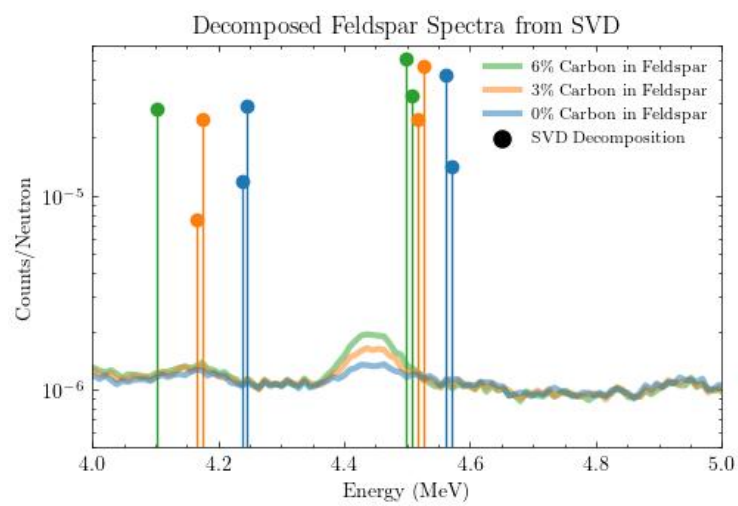the observed and modeled spectra.

13

Figure 11: Convex optimization process showing SVD decomposition of feldspar spectral data
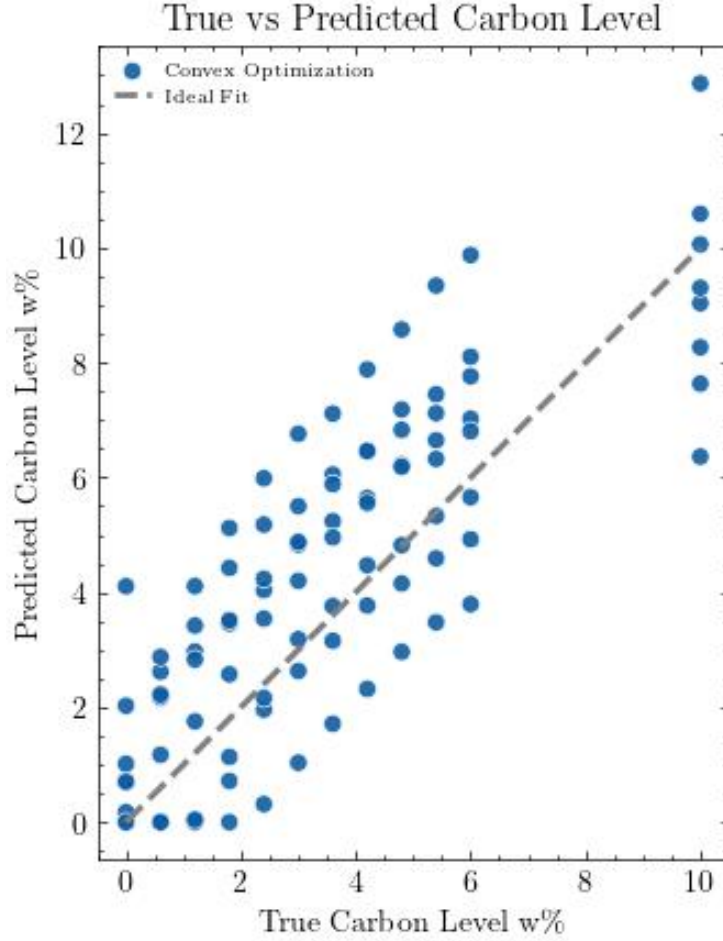
Figure 12: Convex optimization predictions showing carbon level vs predicted values using SVD

3.5. 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

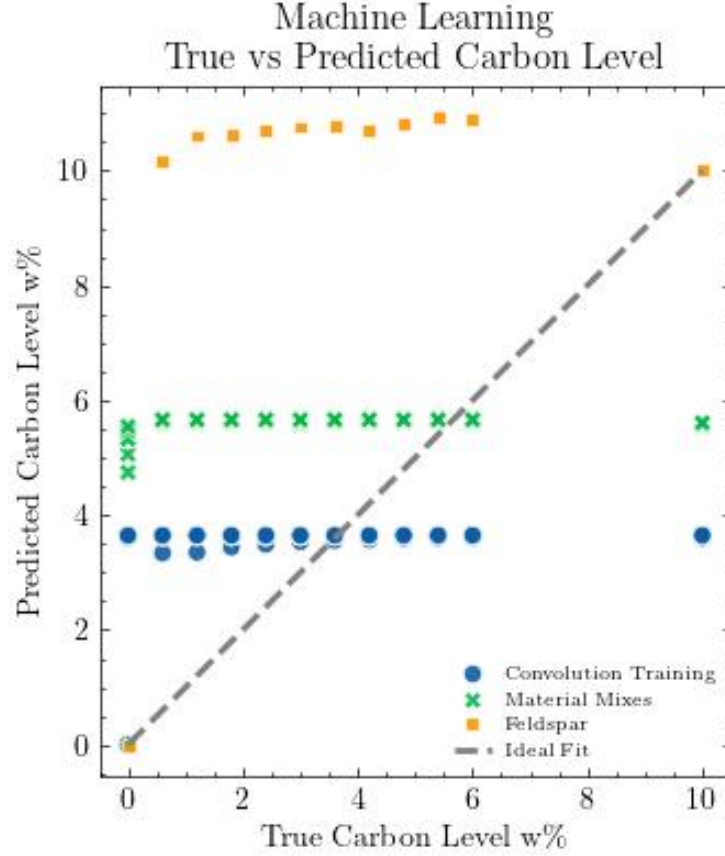deep learning model to predict existence, concentration and carbon peak areas.

Figure 13: 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.

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

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] 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`

[3] C. J. Werner, J. C. Armstrong, F. B. Brown, J. S. Bull, L. Casswell, L. J. Cox, D. A. Dixon, R. A. Forster, III, J. T. Goorley, H. G. Hughes, III, J. A. Favorite, R. L. Martz, S. G. Mashnik, M. E. Rising, C. J. Solomon, Jr., A. Sood, J. E. Sweezy, A. J. Zukaitis, C. A. Anderson, J. S. Elson, J. W. Durkee, Jr., R. C. Johns, G. W. McKinney, G. E. McMath, J. S. Hendricks, D. B. Pelowitz, R. E. Prael, T. E. Booth, M. R. James, M. L. Fensin, T. A. Wilcox, B. C. Kiedrowski, MCNP User's Manual Code Version 6.2, Tech. Rep. LA-UR-17-29981, Los Alamos National Laboratory, Los Alamos, NM, USA (Oct. 2017).
URL `http://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-17-29981`

[4] A. Kavetskiy, G. Yakubova, S. A. Prior, H. A. Torbert, 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`.
URL `https://www.intechopen.com/chapters/54779`

[5] 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`.
URL `https://www.sciencedirect.com/science/article/pii/S0168583X18304476`

[6] 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`

[7] 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 (3) (2020) 261–272, publisher: Nature Publishing Group. `doi:10.1038/s41592-019-0686-2`.
URL `https://www.nature.com/articles/s41592-019-0686-2`

[8] J. J. Moré, The Levenberg-Marquardt algorithm: Implementation and theory, in: Numerical Analysis, Springer, Berlin, Heidelberg, 1978, pp. 105–

116, iSSN: 1617-9692. `doi:10.1007/BFb0067700`.
URL `https://link.springer.com/chapter/10.1007/BFb0067700`

[9] M. A. Branch, T. F. Coleman, Y. Li, A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems, SIAM Journal on Scientific Computing 21 (1) (1999) 1–23, publisher: Society for Industrial and Applied Mathematics. `doi:10.1137/S1064827595289108`.
URL `https://epubs.siam.org/doi/10.1137/S1064827595289108`

[10] 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 (1) (2011) 544–549. `doi:10.1016/j.nima.2010.12.224`.
URL `https://www.sciencedirect.com/science/article/pii/S0168900210030500`

[11] 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`.
URL `https://www.sciencedirect.com/science/article/pii/S2352186423002158`

[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`.
URL `https://www.sciencedirect.com/science/article/pii/S0969806X24007539`