Spatial Prediction of Soil Organic Carbon and Controlling Factors in the Central Tianshan Mountains: Postprint

Authors: He Jianlan, Yan Qingwu, Chen Yiyun, Li Keqi, Bai Junping, Wu Zihao

Date: 2025-10-23T12:01:28+00:00

Abstract

Accurately characterizing the spatial pattern of soil organic carbon (SOC) in the middle Tianshan Mountains and revealing its main influencing factors are of great significance for assessing soil quality, achieving carbon sequestration and sink enhancement, and ensuring ecological security. However, due to significant differences in topography, rainfall, evaporation, vegetation cover, and soil pH within the Tianshan region, especially between areas near and far from the Tianshan Mountains, as well as the overall complex ecological conditions, regional SOC exhibits high spatial heterogeneity, which poses a great challenge for achieving accurate SOC mapping using digital soil mapping techniques. This study, based on 463 soil sampling data points and combining multiple machine learning models such as eXtreme Gradient Boosting (XGB), obtained spatial distribution maps of SOC, and employed the SHAP (Shapley Additive exPlanations) algorithm to reveal the main influencing factors of SOC spatial differentiation. The results demonstrate that the XGB model achieved the optimal fitting performance, with an R^2 of 0.716, LCCC of 0.824, and RMSE of only $1.554\ \mathrm{g \cdot kg^{-1}}$. The spatial distribution of SOC in this region exhibits a pattern of higher values in the north and lower values in the south. Further analysis using the SHAP algorithm revealed that SOC content in this region is primarily influenced by pH, elevation, and mean annual precipitation. Moreover, partial dependence plot results indicated the existence of threshold effects between environmental variables and SOC. When pH exceeds 7.8, mean annual evaporation exceeds 620 mm, or mean annual rainfall falls below 300 mm and vegetation net primary productivity falls below $130\ \mathrm{g\ C \cdot m^{-2} \cdot a^{-1}}$, SOC content drops sharply, indicating that controlling soil pH, evapotranspiration, and vegetation growth conditions is beneficial for regional soil carbon sequestration and sink enhancement. Measures such as using acidic amendments, applying organic fertilizers, and ensuring agricultural irrigation should be implemented to regulate the values of key variables affecting SOC in this region, thereby achieving the goal of increasing SOC content. This study confirms the importance of considering variables such as soil pH, elevation, and mean annual precipitation, and the research results provide data support for soil carbon sink management in the middle Tianshan Mountains and offer a reference for SOC digital mapping in other arid and semi-arid regions.

Full Text

Abstract

Accurately characterizing the spatial patterns of soil organic carbon (SOC) in the middle section of the Tianshan Mountains and identifying its primary influencing factors are crucial for assessing soil quality, enhancing carbon sequestration, and ensuring ecological security. However, significant variations in terrain, rainfall, evaporation, vegetation cover, and soil pH—particularly between proximal and distal mountain areas—create complex ecological conditions that result in high spatial heterogeneity of regional SOC. This poses substantial challenges for precise SOC mapping using digital soil mapping techniques. Based on 463 soil sampling points and combining multiple machine learning models including eXtreme Gradient Boosting (XGBoost, XGB), this study generated spatial distribution maps of SOC and employed the Shapley Additive exPlanations (SHAP) algorithm to reveal the dominant factors controlling SOC spatial variation. Results demonstrate that the XGB model achieved optimal performance, with a fitting R^2 of 0.716, Lin' s concordance correlation coefficient (LCCC) of 0.824, and root mean square error (RMSE) of 1.554 $g \cdot kg^{-1}$. The spatial distribution of SOC exhibited a distinct north-high, south-low pattern. SHAP analysis further revealed that SOC content is primarily influenced by soil pH, elevation, and mean annual precipitation. Partial dependence plots identified significant threshold effects in the relationships between environmental variables and SOC. SOC content declined sharply when pH exceeded 7.8, mean annual evaporation surpassed 620 mm, mean annual precipitation fell below 300 mm, or net primary productivity dropped below 130 $g\ C \cdot m^{-2} \cdot yr^{-1}$. These findings indicate that controlling soil pH, evapotranspiration, and vegetation growth conditions can enhance regional soil carbon sequestration. Management strategies including application of acidic amendments, organic fertilizers, and guaranteed agricultural irrigation can regulate key variables affecting SOC and thereby increase SOC content. This study confirms the importance of considering soil pH, elevation, and mean annual precipitation, providing data support for soil carbon sink management in the middle Tianshan Mountains and offering a valuable reference for digital SOC mapping in other arid and semi-arid regions.

Keywords: soil organic carbon; arid and semi-arid areas; spatial mapping; dominant factors; middle section of Tianshan Mountains

1. Materials and Methods

1.1 Study Area Overview

The middle section of the Tianshan Mountains is located in central Xinjiang (39°34 -46°12 N, 75°31 -91°54 E), covering a total area of 56.86$×10^{4}$ km^{2}. The overall terrain slopes higher in the south and lower in the north. Situated in the interior of the Eurasian continent and far from oceanic moisture sources, the region receives limited precipitation despite some water vapor from the Atlantic Ocean reaching the area via westerly airflow. This moisture is blocked by the Tianshan range, resulting in relatively abundant precipitation on the northern slopes but arid conditions on the southern slopes. Overall, the region experiences limited rainfall with strong evaporation, leading to low vegetation coverage and severe soil salinization. Soils in the middle Tianshan Mountains are coarse-textured with high gravel content, loose structure, and weak water and nutrient retention capacity [Figure 1: see original paper].

1.2 Soil Sample Collection and Chemical Analysis

Using a random sampling method, 463 soil samples (0–30 cm depth) were collected from the study area in 2021. Soil samples were air-dried, ground, and sieved in the laboratory before SOC content was determined using the Walkley-Black method.

1.3 Environmental Variables and Preprocessing

Based on the SCORPAN model framework and considering the study area characteristics, a total of 17 environmental variables potentially influencing SOC content were selected, including soil pH, elevation, slope, climate factors, land use type, vegetation net primary productivity (NPP), and soil physicochemical properties. Detailed information on variable categories, spatial resolution, and data sources is provided in Table 1, with spatial distributions illustrated in Figure 2 [Figure 2: see original paper].

Elevation data were derived from ASTER GDEM and used to calculate slope in ArcGIS. All environmental variables were resampled to a 0.05° spatial resolution. Categorical variables were converted to dummy variables using one-hot encoding prior to model fitting.

1.4 Model Implementation

1.4.1 Random Forest (RF) Random Forest is an ensemble learning algorithm that uses decision trees as base learners, combining predictions through bagging to improve accuracy. It builds decision trees independently on multiple bootstrap samples, selects features randomly at each node for splitting, and averages the predictions across all trees. RF excels at handling high-dimensional and large-scale data, captures nonlinear relationships, and exhibits strong generalization capability with reduced overfitting risk.

In this study, the RandomForestRegressor from the scikit-learn library in Python was implemented. The model was configured with 200 decision trees, a maximum depth of 10, and max_{features} set to the square root of the total number of input features.

1.4.2 Gradient Boosting Decision Tree (GBDT) GBDT is a boosting-based ensemble algorithm that also employs decision trees as base learners. It constructs a loss function and uses its gradient as the boosting direction. During each iteration, new decision trees are added to the existing model to minimize the loss function. GBDT can flexibly handle different data types, demonstrates robustness to outliers and missing values, and achieves relatively high prediction accuracy.

The GradientBoostingRegressor from scikit-learn was used, with 200 decision trees, maximum depth of 5, max_{features} set to the square root of input features, and a learning rate of 0.1.

1.4.3 eXtreme Gradient Boosting (XGB) XGBoost is a tree-based ensemble algorithm that introduces regularization terms to prevent overfitting and supports parallel computation, enhancing algorithm efficiency. The XGBoost library in Python was implemented with key parameters optimized through careful tuning. The model used 200 decision trees with maximum depth of 5, and categorical data were specified as “category” type to avoid additional one-hot encoding.

1.4.4 Deep Neural Network (DNN) DNN is a deep learning model with multiple hidden layers where adjacent layers are fully connected. This architecture enables automatic feature extraction and learning, reducing dependence on manual feature engineering. A DNN with three fully connected layers was designed, each followed by batch normalization and Dropout layers (rate = 0.3) to enhance generalization. The ReLU activation function introduced nonlinearity, and training was conducted over 200 epochs to ensure adequate learning.

1.4.5 Convolutional Neural Network (CNN) CNN is a deep learning architecture suitable for grid-structured data such as images, but also effective for non-image data. Through convolutional layers that automatically extract local features and pooling layers that reduce spatial dimensions, CNN effectively captures hierarchical structures and spatial correlations, improving prediction accuracy. The implemented model included two convolutional layers (3\$$\times 3 kernels, 32 and 64 filters respectively), each followed by batch normalization and max pooling (2\times$\$2 window, stride 2). The output was flattened, passed through fully connected layers with Dropout, and mapped to a one-dimensional output for regression.

1.4.6 Model Interpretation with SHAP SHAP (Shapley Additive exPlanations) is a game theory-based algorithm for interpreting machine learning predictions. Shapley values quantify each feature' s marginal contribution to predictions, representing feature importance. In regression, a SHAP value of zero indicates no influence on the prediction, positive values increase predictions above the average, and negative values decrease them. The SHAP library was used to calculate SHAP values and generate feature importance plots and partial dependence plots.

1.5 Model Evaluation and Uncertainty Analysis

Bootstrap sampling was used to partition the 463 samples into calibration (70%) and validation (30%) sets. Model performance was assessed using three metrics: R^2, RMSE, and Lin' s concordance correlation coefficient (LCCC). LCCC measures agreement between two datasets, with values closer to 1 indicating stronger consistency:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(A_i - B_i)^2}{\sum_{i=1}^{n}(A_i - \bar{A})^2}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A_i - B_i)^2}$$

$$\text{LCCC} = \frac{2r\theta_A\theta_B}{\theta_A^2 + \theta_B^2 + (\bar{A} - \bar{B})^2}$$

where n is sample size, A and B are measured and predicted SOC values, $\bar{A}$ and B are means, r is Pearson's correlation coefficient, and and are coefficients of variation. Higher R^2 and LCCC with lower RMSE indicate better performance.

Uncertainty maps were generated by training the best-performing model 50 times and calculating the mean prediction and standard deviation for each pixel. Uncertainty was quantified as the standard deviation across predictions.

2. Results and Analysis

2.1 Descriptive Statistics

SOC content across sampling points ranged from 0.05 to 24.33 $\text{g} \cdot \text{kg}^{-1}$, with a mean of 2.86 $\text{g} \cdot \text{kg}^{-1}$ and standard deviation of 2.11 $\text{g} \cdot \text{kg}^{-1}$. The distribution exhibited positive skewness (2.11) and leptokurtic characteristics (8.76), indicating a non-normal distribution with heavy tails [Figure 3: see original paper].

2.2 Model Performance Comparison

Among all models, XGB achieved the highest predictive performance with $R^2 = 0.624$ and the lowest RMSE = 1.554 $\text{g} \cdot \text{kg}^{-1}$ in the validation set. Machine learning models (RF, GBDT, XGB) consistently outperformed deep learning models (DNN, CNN) and spatial regression models (GWR, MGWR). While GWR showed reasonable calibration performance ($R^2 = 0.583$), its predictive performance was substantially lower ($R^2 = 0.412$). MGWR performed poorly with negative predictive R^2, indicating unreliable predictions. The superior performance of XGB demonstrates its effectiveness for this dataset .

Variable importance analysis revealed that the most critical factors for XGB were land use type, vegetation NPP, mean annual evaporation, mean annual precipitation, mean annual temperature, population density, total potassium, total nitrogen, silt content, sand content, and elevation.

2.3 Global and Local Controlling Factors of SOC

Based on mean absolute SHAP values, soil pH emerged as the dominant factor controlling SOC spatial variation, with an average absolute SHAP value of 0.5 $g \cdot kg^{-1}$, significantly higher than other variables. Elevation ranked second with an average absolute SHAP value of approximately 0.35 $g \cdot kg^{-1}$. Mean annual precipitation and vegetation NPP showed similar importance (0.23–0.25 $g \cdot kg^{-1}$), while population density and mean annual evaporation contributed moderately (around 0.1 $g \cdot kg^{-1}$). Remaining variables including total potassium, total nitrogen, land use type, silt content, mean annual temperature, and sand content had relatively weak contributions [Figure 4: see original paper].

Local analysis identified pH as the primary controlling factor at 72.79% of sampling points, followed by vegetation NPP (13.17%). Secondary controlling factors included elevation, mean annual precipitation, and SOC content itself, which appeared as secondary factors at 89.85% of sampling points [Figure 5: see original paper].

2.4 Nonlinear Relationships and Threshold Effects

Partial dependence plots for six numerical variables (pH, elevation, mean annual precipitation, mean annual evaporation, total potassium, and total nitrogen) revealed complex nonlinear relationships with SOC. All variables exhibited significant threshold effects:

- **Soil pH**: When $pH < 6.5$, SHAP values were positive, promoting SOC accumulation. Between pH 6.5–7.8, SHAP values approached zero, indicating minimal influence. When $pH > 7.8$, SHAP values turned negative, significantly reducing SOC content.
- **Elevation**: Below 1400 m, SHAP values remained stable around -0.1 $g \cdot kg^{-1}$, showing negative or negligible effects. Above 1400 m, SHAP values increased positively with elevation, likely due to increased snowmelt water and vertical climate zonation effects.
- **Mean annual precipitation**: SHAP values increased positively with precipitation up to 300 mm, beyond which the positive effect stabilized.

- **Mean annual evaporation**: SHAP values decreased with increasing evaporation up to 500 mm, then stabilized around -0.1 $\mathrm{g \cdot kg^{-1}}$, indicating consistent negative impacts at high evaporation rates.
- **Total potassium and nitrogen**: Both showed negative relationships with SOC, with SHAP values decreasing as nutrient content increased [Figure 6: see original paper].

Optimal SOC accumulation occurred when: pH = 6.5-7.8, elevation > 1400 m, mean annual precipitation = 300-460 mm, mean annual evaporation < 620 mm, vegetation NPP > 130 $\mathrm{g\ C \cdot m^{-2} \cdot yr^{-1}}$, total potassium < 12.3 $\mathrm{g \cdot kg^{-1}}$, and total nitrogen > 0.5 $\mathrm{g \cdot kg^{-1}}$.

2.5 Spatial Distribution and Uncertainty of SOC

The predicted SOC distribution showed a clear north-high, south-low pattern, with higher values concentrated in the northwestern and central regions. High SOC areas aligned spatially with water bodies, forming a " "shaped pattern. This distribution corresponded closely with low evaporation zones and high-elevation, high-precipitation, high-NPP corridors [Figure 7: see original paper].

Prediction uncertainty ranged from 0.088 to 2.948 $\mathrm{g \cdot kg^{-1}}$, with higher uncertainty in the southern region, likely due to data gaps. The superior performance of XGB over deep learning models may be attributed to the relatively small sample size and tabular nature of the data, where tree-based models often exhibit more stable and robust performance.

3. Discussion

3.1 Superiority of Interpretable Machine Learning Models

This study compared local spatial regression models with machine learning and deep learning models, demonstrating that the latter achieve significantly higher prediction accuracy, particularly in capturing complex nonlinear relationships that traditional spatial regression models cannot effectively model. While deep learning models showed strong generalization ability (with minimal difference between calibration and validation R^2), their performance was limited by the small dataset size and tabular data structure. In contrast, tree-based models like XGB demonstrated superior efficiency and stability for this data type, consistent with findings that deep learning models may underperform classical methods on tabular data without extensive parameter tuning.

3.2 Influence of Environmental Variables on SOC Spatial Patterns

Results confirm that soil pH, elevation, and climate factors are the primary controls on SOC distribution in the study area. Soil pH affects microbial activity and enzyme function; moderately acidic conditions (pH 6.5-7.8) inhibit microbial decomposition and promote SOC accumulation, while alkaline conditions (pH $>$ 7.8) accelerate SOC loss. Although previous studies have reported negative correlations between pH and SOC in alkaline soils, our study area' s pH range (6.34–9.24) suggests that targeted acid amendments could be effective, though careful dosage control is needed to avoid over-acidification.

Elevation influences SOC through vertical climate zonation. The positive relationship above 1400 m reflects temperature-driven reductions in microbial decomposition and enhanced vegetation growth from snowmelt water. The mid-slope position likely receives optimal precipitation for plant growth, further promoting SOC accumulation.

Climate factors, particularly precipitation and evaporation, significantly affect SOC. Precipitation $>$300 mm promotes vegetation growth and carbon fixation, while high evaporation reduces soil moisture, enhances soil respiration, and depletes SOC. Given the difficulty of directly altering precipitation and evapotranspiration, protecting and restoring vegetation to improve soil water retention represents a practical management strategy.

3.3 Recommendations for Carbon Sequestration

Based on variable importance and threshold analysis, soil pH management offers the greatest potential for SOC enhancement. Application of acidic amendments (e.g., gypsum), organic fertilizers (e.g., manure, crop residues), and ensuring adequate irrigation can optimize pH and moisture conditions. Since grasslands showed the highest SOC content among land use types, establishing artificial grasslands with drought- and salt-tolerant species could further improve soil carbon stocks, particularly given that evaporation far exceeds precipitation in this region.

3.4 Limitations and Future Directions

A key limitation is the lack of high-elevation samples; while the Tianshan peaks exceed 7000 m, our highest sampling point was only 3084 m, potentially introducing uncertainty in high-altitude predictions. Future work should prioritize sampling across broader elevation gradients. Additionally, using 2021 data may not reflect current conditions; incorporating more recent data would improve the timeliness and applicability of SOC maps. Despite these limitations, this study provides a robust framework for understanding SOC controls and guiding management in arid and semi-arid regions.

4. Conclusion

This study identified XGB as the optimal model for predicting SOC in the middle Tianshan Mountains, revealing a distinct north-high, south-low spatial pattern. SHAP analysis demonstrated that soil pH, elevation, and mean annual precipitation are the dominant controlling factors, with critical thresholds identified at pH = 7.8, evaporation = 620 mm, precipitation = 300 mm, and NPP

$= 130$ g C $\cdot$ m^{-2} $\cdot$ yr^{-1}. Management strategies including acidic amendments, organic fertilizer application, and irrigation can regulate these key variables to enhance SOC accumulation. These findings provide scientific evidence and actionable pathways for precision soil carbon management in the middle Tianshan Mountains and offer an important reference for SOC research and practice in other arid and semi-arid regions.

References

[1] Chappell E, Easher T H, Saurette D, et al. Soil Organic Carbon: Past, Present, and Future Research[M]. Germany: Springer, 2021: 1-10.

[2] Kuht J, Eremeev V, Talgre L, et al. Changes in the soil microbial hydrolytic activity and the content of organic carbon and total nitrogen by growing spring barley undersown with red clover in different farming systems[J]. Agriculture, 2019, 9(7): 146.

[3] Zhang Jie, Li Min, Ao Ziqiang, et al. Estimation of soil organic carbon storage of terrestrial ecosystem in arid western China[J]. Journal of Arid Land Resources and Environment, 2018, 32(9): 132-137.

[4] Zhuo Z, Chen Q, Zhang X, et al. Soil organic carbon storage, distribution, and influencing factors at different depths in the dryland farming regions of Northeast and North China[J]. Catena, 2022, 210: 105934.

[5] Liao Qi, Shi Xuezheng, Yu Dongsheng, et al. Scale effect of climatic factors on forest soil organic carbon[J]. Acta Pedologica Sinica, 2010, 47(2): 270-278.

[6] Yang L, Jia W, Shi Y, et al. Spatiotemporal differentiation of soil organic carbon of grassland and its relationship with soil physicochemical properties on the northern slope of Qilian Mountains, China[J]. Sustainability, 2020, 12(22): 9396.

[7] Chang Shuai, Yu Hongbo, Cao Congming, et al. Distribution characteristics of soil organic carbon in Xilin Gol steppe and its influencing factors[J]. Arid Zone Research, 2021, 38(5): 1355-1366.

[8] Han Ya, Zhu Wenbo, Li Shuangcheng. Modelling relationship between NDVI and climatic factors in China: Using geographically weighted regression[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2016, 52(6): 1125-1133.

[9] Song Ying, Gao Mingxiu, Wang Jiafan, et al. Spatial prediction and influencing factors analysis of soil salinization in coastal area based on MGWR[J]. Environmental Science, 2024, 45(7): 4293-4301.

[10] Hagenauer J, Helbich M. A geographically weighted artificial neural network[J]. International Journal of Geographical Information Science, 2022, 36(2): 215-235.

[11] Wu Z, Chen Y, Yang Z, et al. Mapping soil organic carbon in low relief farmlands based on stratified heterogeneous relationship[J]. Remote Sensing,

2022, 14(15): 3575.

[12] Zhang C, Chen Y Y, Wei Y J, et al. Unraveling the threshold and interaction effects of environmental variables on soil organic carbon mapping in plateau watershed[J]. Geoderma, 2024, 450: 117032.

[13] Zheng Y T, Zhao X N, Li X Y, et al. Mapping soil organic carbon density via geographically weighted regression with smooth terms: A case study in Shanxi Province[J]. Ecological Indicators, 2024, 166: 112588.

[14] Chakraborty S K, Chandel N S, Jat D, et al. Deep learning approaches and interventions for futuristic engineering in agriculture[J]. Neural Computing & Applications, 2022, 34(23): 20539-20573.

[15] Padarian J, Minasny B, Mcbratney A B. Machine learning and soil sciences: A review aided by machine learning tools[J]. Soil, 2020, 6(1): 35-52.

[16] Biau G, Scornet E. A random forest guided tour[J]. Test, 2016, 25(2): 197-227.

[17] Breiman L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[18] Friedman J H. Greedy function approximation: A gradient boosting machine[J]. Annals of Statistics, 2001, 29(5): 1189-1232.

[19] Liang W, Luo S, Zhao G, et al. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms[J]. Mathematics, 2020, 8(5): 765.

[20] Chen T Q, Guestrin C. XGBoost: A scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). San Francisco, CA: Assoc Computing Machinery, 2016: 785-794.

[21] Pathy A, Meher S, Balasubramanian P. Predicting algal biochar yield using eXtreme Gradient Boosting (XGB) algorithm of machine learning methods[J]. Algal Research-Biomass Biofuels and Bioproducts, 2020, 50: 102006.

[22] Passafaro T L, Lopes F B, Dórea J R R, et al. Would large dataset sample size unveil the potential of deep neural networks for improved genome enabled prediction of complex traits? The case for body weight in broilers[J]. Bmc Genomics, 2020, 21(1): 771.

[23] Shruti P, Rekha R. A review of convolutional neural networks, its variants and applications[C]//2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), 2023: 31-36.

[24] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS). Long Beach, CA: Neural Information Processing Systems (NIPS), 2017: 4768-4777.

[25] Islam M T, Xing L. Deciphering the feature representation of deep neural networks for high performance AI[J]. Ieee Transactions on Pattern Analysis and

Machine Intelligence, 2024, 46(8): 5273-5285.

[26] Liu Q, Wang X D. Designing an efficient framework for large scale data processing and analysis based on deep learning technology[C]//Assoc Comp M. Proceedings of the 3rd International Conference on Cyber Security, Artificial Intelligence and Digital Economy (CSAIDE). Nanjing: Assoc Computing Machinery, 2024: 269-276.

[27] Wang S, Zhang Z, Geng S, et al. Research on optimization of random forest algorithm based on Spark[J]. Cmc Computers Materials & Continua, 2022, 71(2): 3721-3731.

[28] Shao Q, Xu Y M, Wu H Y. Spatial prediction of COVID-19 in China based on machine learning algorithms and geographically weighted regression[J]. Computational and Mathematical Methods in Medicine, 2021, 2021: 7196492.

[29] Anding K, Haar L, Polte G, et al. Comparison of the performance of innovative deep learning and classical methods of machine learning to solve industrial recognition tasks[C]//Proceedings of the Joint TC1-TC2 International Symposium on Photonics and Education in Measurement Science. Germany: Spie Int SOC Optical Engineering, 2019: UNSP 111440R.

[30] Shwartz Ziv R, Armon A. Tabular data: Deep learning is not all you need[J]. Information Fusion, 2021, 81: 84-90.

[31] Hong Yan, Yuan Lin, Li Mengxue, et al. Progress in the application of artificial intelligence technology in the garment industry from the perspective of product life cycle[J]. Journal of Silk, 2024, 61(12): 1-22.

[32] Li Shengbo, Guan Yang, Hou Lian, et al. Key technique of deep neural network and its applications in autonomous driving[J]. Journal of Automotive Safety and Energy, 2019, 10(2): 119-145.

[33] Tian Fang, Zhou Xiaoxin, Shi Dongyu, et al. A preventive control method of power system transient stability based on a convolutional neural network[J]. Power System Protection and Control, 2020, 48(18): 1-8.

[34] Nie Xiangqin, Chen Hanyue, Niu Zheng, et al. Digital SOC mapping in croplands using agricultural activity factors derived from time series data in western Fujian[J]. Journal of Geo-information Science, 2022, 24(9): 1835-1852.

[35] Li Jing, Wu Huijun, Wu Xueping, et al. Long term conservation tillage enhanced organic carbon and nitrogen contents of particulate organic matter in soil aggregates[J]. Scientia Agricultura Sinica, 2021, 54(2): 334-344.

[36] Pan Yiyi, He Yunfeng, Shan Linan, et al. A review on soil organic carbon during desertification progresses[J]. Chinese Journal of Soil Science, 2011, 42(6): 1515-1521.

[37] Iversen C, Norby R J S N. Global Environmental Change[M]. Germany: Springer Reference, 2014: 297-316.

[38] Li Yulu, Wu Xueqing, Yuan Ye, et al. Spatial temporal variation and driving mechanism of vegetation carbon and water use efficiency in Heilongjiang Province, China[J]. Chinese Journal of Applied Ecology, 2024, 35(12): 3349-3358.

[39] Feng Jin, Cui Dong, Sun Guojun, et al. Soil organic carbon in relation to soil physicochemical properties in Xinjiang[J]. Pratacultural Science, 2017, 34(4): 692-697.

[40] Ma Huiying, Li Xinzhu, Ma Xinyu, et al. Characteristics and driving factors of soil organic carbon fractions under different vegetation types of the mid northern piedmont of the Tianshan Mountains, Xinjiang[J]. Ecology and Environmental Sciences, 2022, 31(6): 1124-1131.

[41] Fu Hua, Chen Yaming, Wang Yanrong, et al. Organic carbon content in major grassland types in Alxa, Inner Mongolia[J]. Acta Ecologica Sinica, 2004, 24(3): 469-476.

[42] He Zhibin, Zhao Wenzhi, Liu Hu, et al. Characteristic of soil organic carbon and relationship with environment factors under Picea crassifolia forest in the Qilian Mountain[J]. Acta Ecologica Sinica, 2006, 26(8): 2572-2577.

[43] Liu Wei, Cheng Jimin, Gao Yang, et al. Distribution of soil organic carbon in grassland on the Loess Plateau and its influencing factors[J]. Acta Pedologica Sinica, 2012, 49(1): 68-76.

[44] Tavakkoli E, Uddin S, Rengasamy P, et al. Field applications of gypsum reduce pH and improve soil C in highly alkaline soils in southern Australia’ s dryland cropping region[J]. Soil Use and Management, 2022, 38(1): 466-477.

[45] Zuo W, Xu L, Qiu M, et al. Effects of different exogenous organic materials on improving soil fertility in coastal saline-alkali soil[J]. Agronomy-Basel, 2023, 13(1): 61.

[46] Zhang Y, Li F, Lu Z, et al. Organic amendments effects on soil aggregation and carbon sequestration in saline-alkaline croplands in China[J]. Agronomy Journal, 2023, 115(4): 2083-2095.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.