Guo *et al. Biochar* (2026) 8:134
https://doi.org/10.1007/s42773-026-00652-6

Biochar

REVIEW

Open Access

Application and prospect of artificial intelligence in biochar for acidic soil amelioration

Linyu Guo[1,2], Kewei Li[1] and Ren-kou Xu[1,2*]

Abstract

Effective amelioration of acidic soils with biochar is critical for mitigating soil degradation and promoting sustainable agriculture, yet this requires robust mechanistic understanding and predictive modeling. This review synthesizes the key biogeochemical pathways by which biochar neutralizes soil acidity and enhances soil pH buffering capacity. We identify critical data gaps—notably the overlooked relative contributions of organic and inorganic alkalis—that currently limit the predictive performance of artificial intelligence (AI) models. A statistical literature analysis reveals that ensemble learning algorithms, particularly Random Forest, are predominant across various application scenarios, such as crop yield prediction, due to their robustness against overfitting and their capacity for feature importance ranking. However, no single algorithm is universally optimal. The choice of algorithm depends on several factors, including the amount of available data, the complexity of the problem, and the computational resources. Both multi-source data fusion and multi-model ensembling can enhance the predictive performance of AI models when appropriately configured. Future data fusion should prioritize integrating remote sensing with proximal sensor, multi-modal microbial, and imaging-chemical data to elucidate interactions across the biochar–soil–microbe–plant continuum. In conclusion, the predictive accuracy for biochar's performance in acidic soil amelioration can be enhanced by: (1) integrating sensing technologies to develop new methods with high resolution and low detection limits; (2) applying AI while ensuring model interpretability and avoiding overclaims of causality; (3) constructing benchmark datasets to expand the usability of biochar data; (4) improving the representation of microbial responses to biochar in AI models; and (5) strengthening interdisciplinary collaboration among soil scientists, model developers, and sensing technology experts.

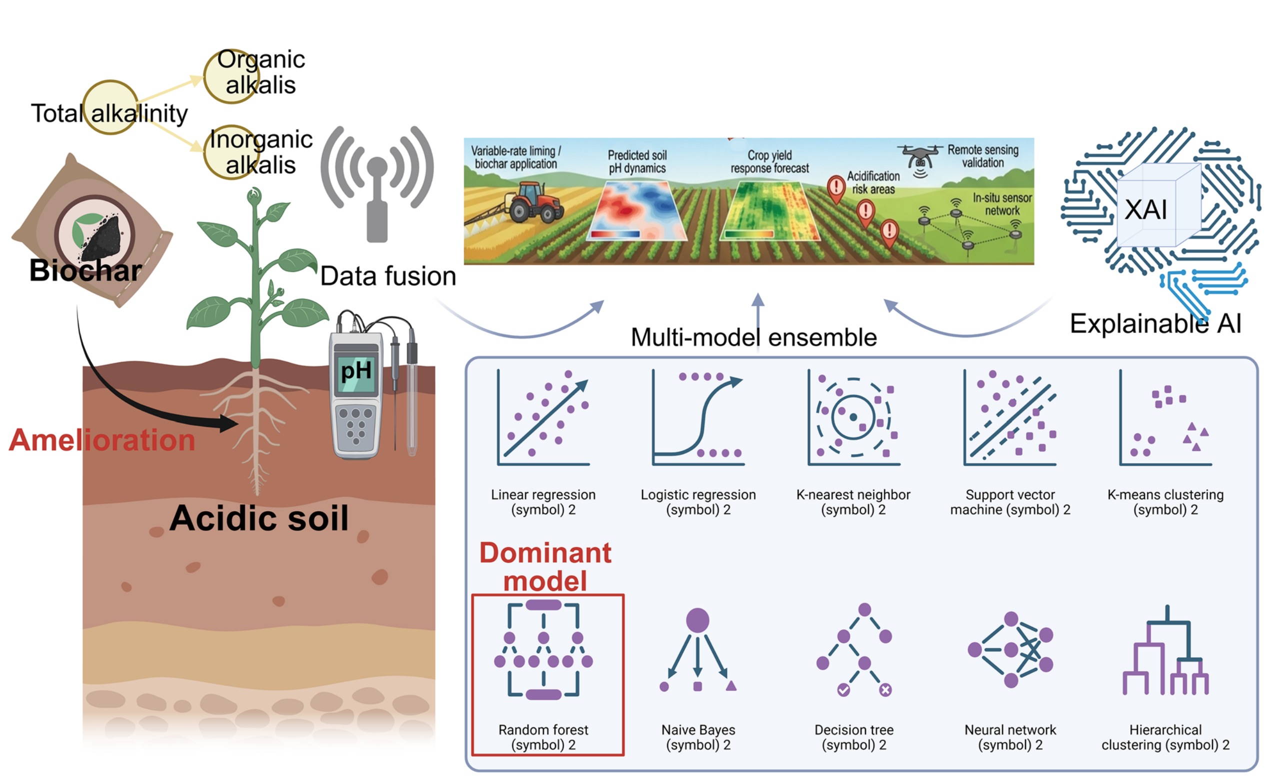
Highlights

- Ensemble learning, especially Random Forest, dominates AI modeling for predicting biochar efficacy in acidic soils.
- Decoupling organic and inorganic alkalis is critical for advancing mechanism-based AI predictions.
- Multi-source data fusion integrates sensing technologies to overcome data limitations and enhance model performance.

Keywords Biochar, Soil acidification, Artificial intelligence, Random Forest, Multi-source data fusion

*Correspondence:
Ren-kou Xu
rkxu@issas.ac.cn
Full list of author information is available at the end of the article

Graphical Abstract



1 Introduction

Soil acidification poses a major threat to the sustainability of global agriculture. As early as 1995, Von Uexküll and Mutert (1995) reported that approximately 3.95 billion hectares of global land have a pH of less than 5.5, accounting for about 30% of the world's arable land. These acidic soils are predominantly distributed in the cold, humid temperate zones and the high-rainfall tropical and subtropical regions. Soil acidification has been further exacerbated by excessive application of chemical nitrogen fertilizers, intensive land use, and industrial nitrogen oxide emissions. The essence of soil acidification is the accumulation of acidic substances at a rate exceeding the soil's acid buffering capacity—a phenomenon particularly pronounced in China (Zhang et al. 2025a; Guo et al. 2010; Zhao et al. 2023). Over the past few decades, acidification has even emerged in the deeper soil layers of some regions in China. Soil acidification not only leads to the leaching loss of base cations, including calcium (Ca^{2+}), magnesium (Mg^{2+}), potassium (K^{+}), and sodium (Na^{+}), and phosphorus (P) immobilization (Alekseeva et al. 2011; Jiang et al. 2015). It also induces aluminum (Al) and manganese toxicity to plant roots and soil microorganisms, while increasing the bioavailability and toxicity of heavy metals such as lead and cadmium. These factors severely inhibit crop growth and lead to yield reductions (Reddy et al. 1995; Rout et al. 2001; Li et al. 2024b; Guo et al. 2025). At a more fundamental level, acidification profoundly alters the biogeochemical cycles of carbon (C), nitrogen (N), P, and sulfur in soils by suppressing enzyme activity and disturbing the structure and function of microbial communities (Ali et al. 2025; Hu et al. 2024; Lehmann et al. 2011; Lee et al. 2022; Li et al. 2024e; Waheed et al. 2025). Developing efficient and sustainable technologies for ameliorating acidic soils is therefore an urgent necessity for ensuring food security and advancing sustainable agricultural development. The application of alkaline amendments—such as lime, biochar, crop straws, and manure—is a primary strategy for mitigating soil acidification and alleviating aluminum toxicity (Dai et al. 2017). However, traditional lime amendment has several disadvantages, including dust pollution, limited effectiveness in deep soil layers, a tendency for soil compaction, a short duration of action, and a propensity for re-acidification (Tang et al. 2013). Industrial by-products often contain heavy metals, increasing the risk of secondary pollution (Li et al. 2010). Organic wastes frequently contain excessive nitrogen and phosphorus, which may lead to water eutrophication (Case et al. 2017). Furthermore, the effectiveness of straw return in ameliorating acid soils remains debated (Butterly et al. 2013). In contrast, biochar can reduce soil active and exchangeable acidity while also significantly enhancing soil pH buffering capacity (pHBC), thereby strengthening the soil's resistance to re-acidification (Shi et al. 2017, 2019c; Xu et al. 2012). Its high C content, porous structure, negatively charged surface, and highly alkaline nature endow biochar with the ability to modify soil physicochemical properties. This includes reducing bulk density, raising soil pH, and improving nutrient availability, water-holding capacity, and cation exchange capacity (CEC) (Liu et al. 2015; Hossain et al. 2020). Moreover, biochar promotes soil–microbe interactions, enhancing soil functions and stimulating soil enzyme activities, thereby providing nutrients and energy for plant growth (Ali et al. 2025; Lehman et al. 2011; Liao et al. 2019; Pokharel et al. 2020). These properties distinguish biochar from traditional amendments, positioning it as an acid ameliorant capable of providing both short-term remediation and long-term soil health maintenance.

The mechanisms by which biochar neutralizes soil acidity have been extensively studied. Its abundant base cations and oxygen-containing functional groups can lower soil active acidity (He et al. 2022a, 2022b), while the sustained dissolution of carbonates and silicates enhances the soil's pHBC (Yuan et al. 2011b; Xu et al. 2012; Shi et al. 2017). However, biochar-mediated amelioration of acidic soils involves a complex system characterized by multi-interface, multi-scale, and multi-process interactions. The nonlinear interactions among biochar properties and production conditions (feedstock, pyrolysis temperature, functional group abundance, mineral composition), inherent soil conditions (initial pH, CEC, clay content, organic matter), and environmental factors (precipitation, temperature, management practices) render traditional statistical models inadequate for accurately predicting amelioration outcomes (Zhang et al. 2025b). Dai et al. (2021) systematically summarized how four key properties of biochar—porous structure, volatile organic carbon, pH, and electrochemical properties—directly and indirectly influence soil bacterial, fungal, and protist communities. Furthermore, Lei et al. (2023) confirmed through machine learning (ML) that microorganisms serve as a critical mediator linking biochar amendment to soil ecological functions, emphasizing that their quantitative characterization is paramount for developing interpretable and generalizable predictive models. Consequently, artificial intelligence (AI) offers a novel methodological toolkit for unraveling the complexities inherent in biochar-mediated amelioration of acidic soils.

ML models, which are based on statistical learning, can capture the nonlinear relationships between biochar properties, soil conditions, and the improvement effect. This enables the prediction of changes in key soil properties, such as pH, soil organic carbon (SOC) content, and the concentration of available heavy metals, following biochar application. Deep learning (DL), a subset of ML, can automatically extract abstract features from high-dimensional data, including biochar characterization spectra and microscopic images. For instance, Random Forest (RF) has been shown to provide highly accurate predictions of SOC content (Ding et al. 2025). The gradient boosting family (e.g., Gradient Boosting Decision Tree (GBDT) and eXtreme Gradient Boosting (XGBoost)) exhibits outstanding performance in predicting soil properties due to enhanced predictive accuracy and capacity to handle heterogeneous features. Through a pot experiment combined with an RF model, Ke et al. (2022) identified the most critical biochar properties influencing soil acidity amelioration as alkaline functional groups, Ca, Mg, P, and K contents, and specific surface area—with contribution rates of 13%, 13%, 12%, 10%, 9%, and 10%, respectively. However, current research remains predominantly "data-correlation-driven"—while models can establish statistical correlations between biochar properties and ameliorative effects, correlation does not equate to causation. Deeper questions regarding the mechanisms and efficacy of biochar in acidic soil amelioration remain to be elucidated. Although interpretability tools such as SHapley Additive exPlanations (SHAP) and Partial Dependence Plots (PDP) can quantify feature contributions, they cannot distinguish correlation from causality (Xiong et al. 2014; Wadoux 2025). Furthermore, most models are trained on short-term incubation data, rendering them incapable of capturing long-term dynamics such as biochar aging and soil re-acidification (Cheng et al. 2008; Zimmerman 2010).

The advancement of proximal soil sensing (PSS) technologies—including visible and near-infrared spectroscopy (Vis–NIR), portable X-ray fluorescence spectrometry (PXRF), fiber-optic pH sensors, and planar optode imaging—has enabled a technological leap from laboratory determinations to in situ field measurements (Kahlert et al. 2004; Rudolph-Mohr et al. 2017; Greenberg et al. 2022a, b). Characterization techniques such as Attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR), X-ray photoelectron spectroscopy (XPS), X-ray diffraction (XRD), and scanning/transmission electron microscopy coupled with energy dispersive X-ray spectroscopy (SEM/TEM-EDS) reveal and quantify the physicochemical mechanisms underlying biochar–soil interactions across atomic, molecular, and microscopic scales (Cheng et al. 2008; Keiluweit et al. 2010; Mukherjee et al. 2011). These advances facilitate the transformation of such mechanisms into structured, quantifiable, high-quality datasets (Wilkinson et al. 2016; IBI 2015; EBC 2022). These developments provide a rich data source for AI-based prediction of biochar amelioration efficacy, provided that appropriate calibration and workflows are implemented (Wilkinson et al. 2016). Such frameworks can integrate data on climate, soil properties, biochar characteristics, and management practices to predict ameliorative effects under diverse scenarios (Jeffery et al. 2011; Li et al. 2022; Jiao et al. 2025). Therefore, this review aims to address the following questions: (1) What are the biogeochemical mechanisms by which biochar neutralizes soil acidity and enhances pHBC and what are the mechanistic gaps that AI modeling needs to resolve? (2) How can AI assist in designing, predicting, and elucidating these processes? (3) How can multi-sensor technologies provide the necessary data support? (4) How can we transcend the current data-correlation paradigm towards mechanism-guided intelligent prediction? Our overarching objectives are twofold: (1) to identify synergies and complementarities among different technological approaches, and (2) to explore viable pathways for transitioning from a "data-correlation-driven" approach to "mechanism-guided intelligent prediction"—thereby providing a scientific foundation and technical support for the global management of acidified soils and the advancement of sustainable agricultural development.

2 Key processes in biochar-mediated acidic soils and implications for artificial intelligence modeling

2.1 Biochar decreased soil acidity

2.1.1 Direct chemical neutralization mechanisms

The alkaline nature of biochar is attributed to its constituent carbonates (predominantly amorphous, with some crystalline phases like calcite and dolomite), surface oxygen-containing functional groups (OFGs, e.g., carboxyl –COOH and phenolic hydroxyl –OH), and soluble organic compounds (e.g., conjugate bases of organic acids) (Noble et al. 1996; Yan And Schubert 2000; Xu And Coventry 2003). During pyrolysis, particularly from feedstocks like crop straws, these components, such as inherent organic anions and inorganic carbonates, are retained or even concentrated (Yuan et al. 2011a, 2011b; Yuan and Xu 2012). The primary mechanisms by which biochar ameliorates soil acidity include:

(1) Base cation exchange and direct neutralization by carbonates

Abundant base cations (e.g., Ca^{2+}, Mg^{2+}, K^{+}, Na^{+}) in biochar exchange with exchangeable Al^{3+} in acidic soils. The released Al^{3+} hydrolyzes to generate H^{+}, which are then neutralized by the biochar's alkaline components, primarily carbonates. This process, known as the "liming effect" of biochar, raises soil pH (Yuan et al. 2011b; Shi et al. 2019a). Carbonates are the principal source of inorganic alkalinity, contributing 20–73% of the biochar's total alkalinity (Yuan et al. 2011b).

(2) Deprotonation of surface OFGs

Biochar itself is alkaline with a high pH. The weakly acidic OFGs abundantly present on its surface, such as carboxyl (–COOH) and phenolic hydroxyl (–OH) groups, partially exist as organic anions (e.g., $–COO^{-}$, $–O^{-}$) when the soil pH is high. In acidic soils, these anions can bond directly to H^{+}, thereby consuming H^{+} and directly increasing the soil pH (Berek And Hue 2016; Shi et al. 2017, 2018a, 2018b). Feedstock type and pyrolysis temperature determine the abundance of these OFGs, with higher temperatures generally reducing their concentration. Nevertheless, biochar produced at moderate temperatures (e.g., 500 ℃) retains sufficient OFGs to provide a balance between alkalinity and structural stability (Yuan et al. 2011b).

(3) Complexation and buffering by organic anions

Soluble organic anions, originating from the biochar surface or released into the soil solution, can complex with H^{+} to form neutral molecules, thereby reducing soil acidity (He et al. 2022a, 2022b). Furthermore, these organic components mitigate Al^{3+} toxicity by forming stable organo-aluminum complexes, effectively immobilizing Al^{3+} into more inert, less bioavailable forms (Xu et al. 2012).

(4) Enhancement of soil surface charge and pHBC

Biochar application significantly increases the soil's CEC. This enhanced negative surface charge improves the retention of essential base cations and simultaneously elevates the soil's pHBC, thereby strengthening its resilience against future acidification (Xu et al. 2012).

2.1.2 Microbially-mediated geochemical mechanisms

Biochar enhances microbial functions by acting as both a physical scaffold and a functional amplifier. Its high specific surface area and porosity create protective niches—or microbial "hotspots"—that shield functional microorganisms (e.g., urease-producing bacteria, carbon-fixing autotrophs) from environmental stress (Yu

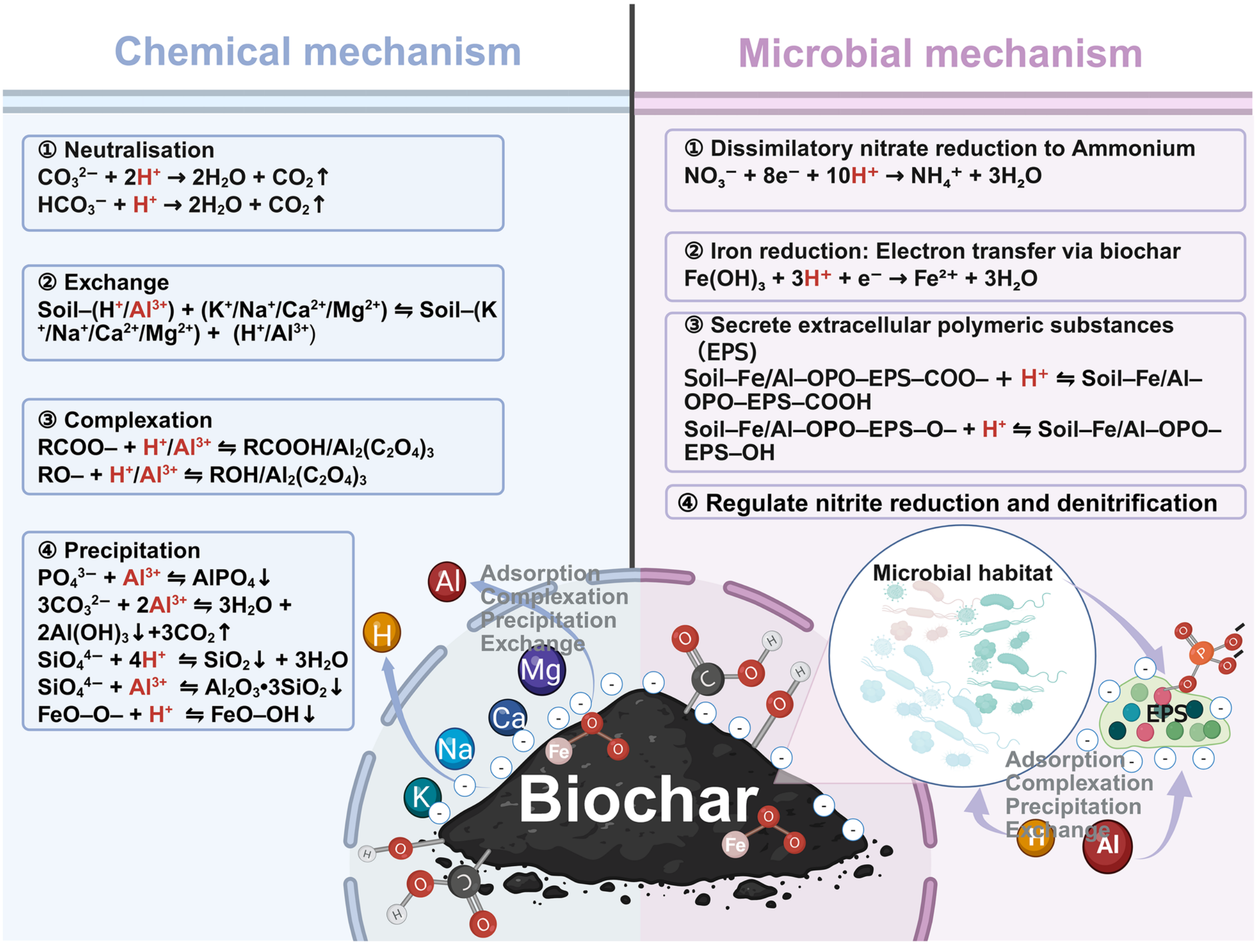


Fig. 1 The mechanisms for biochar remediation of acidic soil: Reduce soil acidity and increase soil acid buffering capacity

et al. 2020; Viana et al. 2024), while also providing carbon and nutrient sources. This "carrier effect" promotes the colonization and metabolic activity of key functional groups, thereby stimulating enzymatic processes (e.g., urease activity) and accelerating the biogeochemical cycling of carbon, nitrogen, and phosphorus. Beyond general increases in microbial diversity (Deng et al. 2026; Liu et al. 2023; Li et al. 2026b), long-term biochar amendment selectively enriches key phyla such as Proteobacteria and Actinobacteria (Ameloot et al. 2014; Palansooriya et al. 2019). This community shift establishes a functional foundation for subsequent microbially-driven alkali-generating processes. Furthermore, biochar acts as a "micro-reactor" and nucleation template, facilitating microbially- and enzyme-induced carbonate precipitation. This process synthesizes new carbonate minerals, providing a mechanism for the long-term neutralization of soil acidity (Martinez et al. 2013; Almajed et al. 2019; Li et al. 2024a, b, c, d, e). Concurrently, the secretion of extracellular polymeric substances (EPS) by microorganisms represents another crucial pathway for soil acidity amelioration. The mechanism is driven by abundant functional groups within the EPS matrix—notably phosphoryl, carboxyl, and phenolic hydroxyl groups—which engage in complexation and buffering reactions at soil mineral interfaces, mitigating proton (H^+) activity and Al^{3+} toxicity (Kayoumu et al. 2025). The significance of this pathway is underscored by evidence showing that EPS addition can raise the pH of an acidic red soil by 1.23 units, a process dominated by the activity of phosphoryl groups (Nkoh et al. 2022).

2.2 Biochar amendment enhances soil pHBC

pHBC is a critical indicator for evaluating soil resistance to acidification. Figure 1 illustrates the multi-process synergistic mechanisms by which biochar reduces soil acidity and enhances soil pHBC.

2.2.1 Key chemical processes

(1) Protonation of surface organic anions: A rapid and direct "first buffering line of defense"

The organic anions generated from the dissociation of weakly acidic oxygen-containing functional groups on the biochar surface act as conjugate bases, capable of directly undergoing protonation reactions with H^+ in the soil solution (Shi et al. 2019a, b, c):

$$\text{Biochar-COO}^- / \text{Biochar-O}^- + H^+ \\ \rightarrow \text{Biochar} - \text{COOH}/\text{Biochar} - \text{OH}$$

This process is characterized by its rapid response and direct action, representing the primary pathway for enhancing pHBC in the initial stages following biochar application (Shi et al. 2017). Its pHBC is strongly dependent on biochar production conditions: low-temperature pyrolysis (~350–500 ℃) favors the retention of more oxygen-containing functional groups, thereby endowing such biochars with a more pronounced protonation pHBC (Yuan et al. 2011b). Spectroscopic evidence confirmed that biochars produced at low temperatures (300–500 ℃) exhibit significant $-COO^-$ characteristic peaks at ~1590 cm^{-1} and ~1400 cm^{-1} (Yuan et al. 2011b). ATR-FTIR further revealed that as the system pH decreases, the absorption peak of $-COO^-$ weakens while the –COOH peak intensifies, providing direct evidence for the occurrence of protonation (Shi et al. 2017). From a modeling perspective, this mechanism provides well-defined, quantifiable reaction sites (e.g., via Boehm titration or FTIR peak intensity) that can serve as key input features for predictive AI models.

(2) Enhancing soil CEC: Constructing a robust "ion exchange buffering reservoir"

Soil CEC is a fundamental chemical property governing soil fertility and ecological function. It determines the soil's ability to retain and supply cationic nutrients (Ca^{2+}, Mg^{2+}, K^+, NH_4^+) and, critically, influences the tolerance of crops to aluminum toxicity in acidic soils (Zhao et al. 2022; Xu et al. 2012; Shi et al. 2019c). For a given Al^{3+} activity in soil solution, soils with higher CEC and greater exchangeable base cations exhibit higher Ca^{2+}/Al^{3+}, Mg^{2+}/Al^{3+}, and K^+/Al^{3+} ratios, which mitigate Al toxicity and result in higher critical Al^{3+} activity thresholds for crop growth (Zhao et al. 2022). Biochar possesses a substantial CEC, typically 10 times greater than that of acidic soils (Munera-Echeverri et al. 2018), and its application increases the net negative charge in soil (Alling et al. 2014; Yuan et al. 2011a). Zeta potential measurements indicate that biochar particles carry negative charge across pH 3–7, with the electronegativity increasing as pH rises (He et al. 2020).

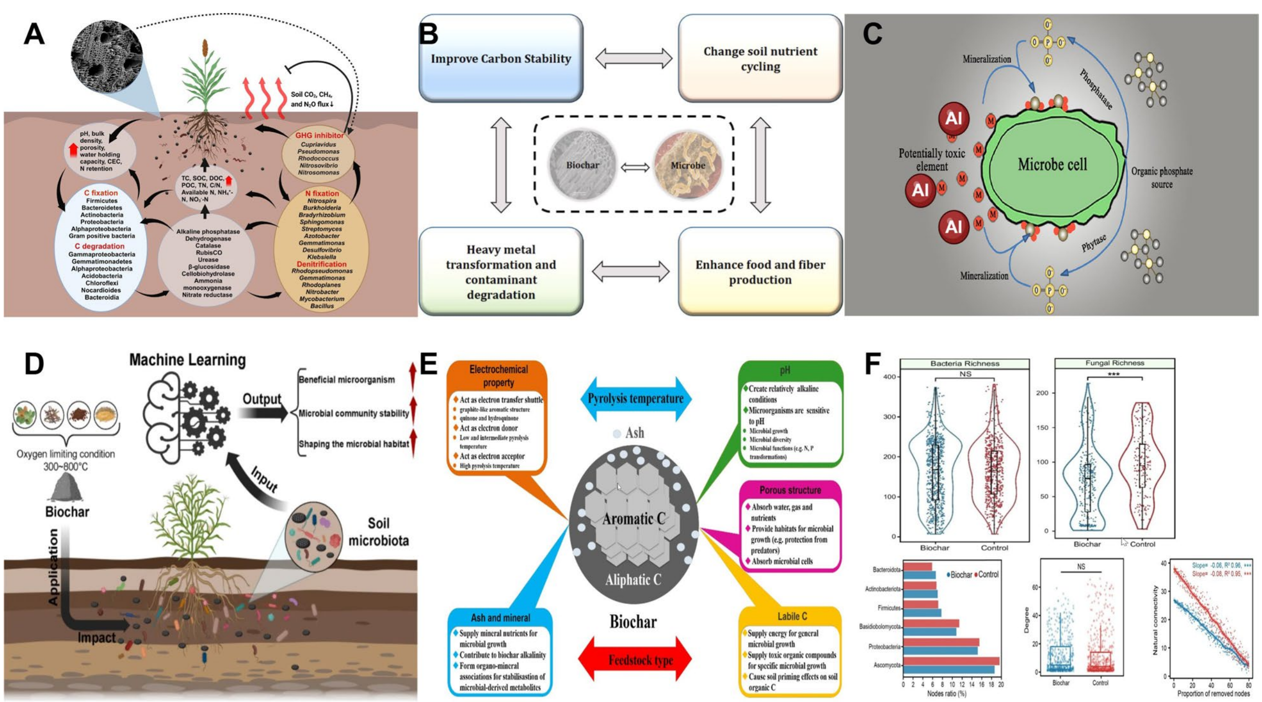


Fig. 2 Mechanisms of biochar-mediated microbial regulation in the amelioration of acidic soils. **A–F** are adapted from Parasar et al. (2025), He et al. (2021), Jiang et al. (2020), Lei et al. (2023), Dai et al. (2021), Lei et al. (2023)

The high CEC of biochar is primarily attributable to the abundant oxygen-containing functional groups on its surface, which dissociate to generate negative charge sites (Shi et al. 2017; Yuan et al. 2011b). The increase in CEC enhances the exchange reactions between base cations and H^+ by providing more exchange sites, thereby reducing the concentration of H^+ in the soil solution (Xu et al. 2012):

$$[\text{Soil colloid}] - M^{n+} + nH^+ \rightarrow [\text{Soil colloid}] - H^+ + M^{n+} \left(M^{n+} = Ca^{2+}, Mg^{2+}, K^+\right)$$

However, the effect of biochar on soil CEC depends not only on biochar properties but also on inherent soil properties. Amendment with 1% biochar increased the CEC of soils with relatively low initial CEC but did not significantly affect soils with high initial CEC (Yuan and Xu 2012). The impact of this mechanism is most pronounced in soils with low initial CEC and weak pHBC (Huff and Lee 2016). Conversely, the magnitude of improvement is less significant in soils with high CEC (Xu et al. 2012). Biochar-induced CEC increase is positively correlated with pHBC enhancement, as CEC provides exchange sites for buffering proton inputs (Shi et al. 2017, 2018a; Xu et al. 2012). This dependency highlights the necessity of including initial soil CEC as a critical interactive variable in predictive AI models.

(3) Dissolution and precipitation of silicates and carbonates: Providing sustained and long-term "alkaline substance replenishment"

Inorganic components, primarily carbonates and silicates, provide biochar with its long-term acid-neutralizing capacity. In high-temperature (> 500°C) biochars, carbonates such as calcite and dolomite are the dominant inorganic alkalis, with their abundance increasing directly with pyrolysis temperature, as confirmed by XRD and acid–base titration (Yuan et al. 2011b). The gradual dissolution of these carbonates creates a distinct buffering plateau between pH 5.5 and 8.0 in titration curves, signifying their sustained acid-neutralizing activity. Biochars derived from specific feedstocks, such as rice husk and corn straw, are also rich in amorphous silicates. These amorphous silicates slowly hydrolyze in acidic soils, consuming H^+ ($H_3SiO_4^- + H^+ \rightarrow H_4SiO_4$) and providing an additional, slower-release neutralizing mechanism (Shi et al. 2017). This process can contribute up to 20% of the increase in soil pHBC and indirectly mitigates aluminum toxicity by promoting the formation of stable aluminosilicate minerals (Qian et al. 2016).

2.2.2 Microbially-mediated biogeochemical cycles: Constructing a dynamic "biological buffering engine"

The restructuring of soil microbial community structure and function by biochar has been demonstrated to play a pivotal role in enhancing the long-term and dynamic pHBC of soils (Shi et al. 2017). Figure 2 illustrates the key microbial mechanisms through which biochar application influences soil acidity and nutrient cycling. By modulating microbial communities and their functions, particularly within the carbon, nitrogen, and iron cycles, biochar stimulates critical biochemical reactions that either produce alkali or consume acid. These processes can enhance the soil's long-term resilience against soil re-acidification (Quilliam et al. 2013; Lei et al. 2023; Wu et al. 2025).

(1) Direct biogeochemical pathway: Driving alkali-producing metabolic processes

Biochar stimulates microbial alkali-generating processes by altering the soil microenvironment and promoting the metabolic activity of specific functional microorganisms. These processes include three primary mechanisms: First, biochar promotes ammonification while concurrently inhibiting nitrification. By elevating soil pH and organic nitrogen availability, biochar stimulates microbial ammonification (the conversion of organic N to NH_4^+), a process that consumes protons (H^+) and generates net alkalinity (Shi et al. 2019b). Simultaneously, biochar inhibits autotrophic nitrification ($NH_4^+ \rightarrow NO_3^- + 2H^+$), a strongly acidifying process. This inhibition is attributed to mechanisms including competitive NH_4^+ adsorption onto biochar surfaces and shifts in microbial competition, which often result in the suppression of ammonia-oxidizing bacteria (AOB). This shift can favor alternative, less acidifying pathways like heterotrophic nitrification, ultimately reducing the net rate of soil acidification (Lin et al. 2017; Xie et al. 2023). Second, biochar promotes dissimilatory nitrate reduction to ammonium (DNRA), a process which can be directly facilitated by acting as an electron shuttle. Within anaerobic microsites created by biochar (e.g., internal pores), microorganisms like *Clostridium* and *Geobacter* reduce NO_3^- to NH_4^+. This pathway ($NO_3^- + 8e^- + 10H^+ \rightarrow NH_4^+ + 3H_2O$) provides a dual benefit: it consumes substantial protons to mitigate acidity and re-immobilizes easily leached nitrate as more stable ammonium, thus conserving soil nitrogen (Pandey et al. 2020). Third, biochar regulates denitrification pathways, promoting complete denitrification to dinitrogen (N_2). By altering the denitrifying community—for example, by increasing the abundance of microbes carrying the *nosZ* gene (which encodes the enzyme for nitrous oxide (N_2O) reduction)—biochar drives the process towards its final product (N_2).

This not only ensures proton consumption throughout the pathway but also mitigates the emission of the potent greenhouse gas N_2O, thereby altering nitrogen fluxes in a manner that contributes favorably to soil acid–base balance (Wang et al. 2023).

(2) Indirect mineral transformation pathway: Consuming acidity and immobilizing active aluminum

Biochar mediates the biogeochemical transformation of iron (Fe) and aluminum (Al) minerals, leading to net proton consumption by altering microbial electron transfer pathways. On one hand, biochar acts as an electron shuttle, facilitating the microbial reduction of insoluble Fe(III) oxides to more soluble Fe(II) by microorganisms such as *Geobacter* and *Shewanella* (Nealson And Saffarini 1994; Weber et al. 2006). This reductive dissolution process—simplified as $Fe(OH)_3 + 3H^+ + e^- \rightarrow Fe^{2+} + 3H_2O$ (with electrons derived from concurrent oxidation of organic electron donors)—directly consumes protons, thereby lowering soil solution acidity (Long et al. 2024). On the other hand, while this iron reduction releases previously occluded or co-precipitated Al^{3+}, biochar concurrently provides a mechanism for its immobilization. Silicates (H_4SiO_4) and phosphates (PO_4^{3-}) released from the biochar, along with microbially-produced organic ligands, complex with this free Al^{3+} (Shi et al. 2017). This interaction promotes the formation of stable secondary minerals, such as hydroxy-aluminosilicates and

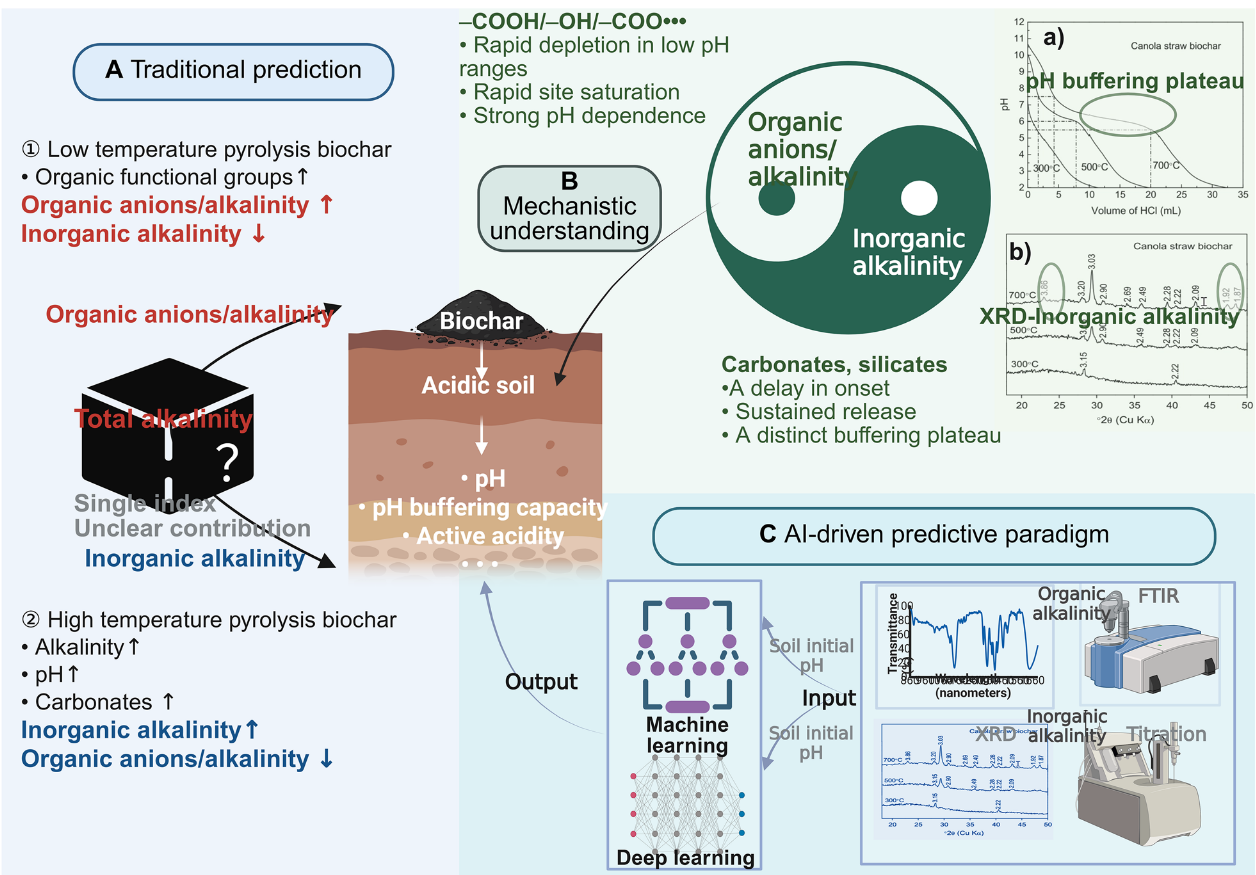


Fig. 3 Conceptual comparison of traditional modeling approaches versus AI-based frameworks for biochar-mediated acidic soil amelioration, highlighting the importance of decoupling organic and inorganic alkalinity. **A** Traditional model predictions that overlook the decoupling of organic anions/alkalinity from inorganic alkalinity, which amplifies the black-box effect and obscures the underlying mechanisms. **B** Differences in the mechanisms of acidic soil remediation between organic alkalinity (e.g., carboxylate, hydroxyl, and carbonyl groups) and inorganic alkalinity (e.g., carbonates, silicates): **a** pH buffering plateau (observed mainly for biochars produced at 500 and 700 °C, except for corn straw biochar) illustrating the dominant role of inorganic carbonates in long-term acid neutralization, while organic anions contribute primarily to rapid initial pH increase; **b** X-ray diffraction (XRD) patterns identifying inorganic alkaline mineral phases (e.g., calcite, dolomite, and sylvite) in biochar, adapted from Yuan and Xu (2011b). **C** Prospects and a proposed workflow for incorporating the organic-inorganic alkalinity decoupling into AI-based predictive modeling, aiming to enhance model interpretability and mechanistic transparency. Panels (**a**) and (**b**) are adapted from Yuan and Xu (2011b) with permission

amorphous aluminum phosphates. Consequently, potentially acidic Al^{3+}—which would otherwise hydrolyze to produce H^+ ($Al^{3+} + H_2O \rightleftharpoons Al(OH)^{2+} + H^+$)—is effectively removed from the soil solution. This immobilization serves as a critical, indirect pathway for mitigating soil acidification (Qian and Chen 2013).

(3) Secretion of EPS rich in functional groups

EPS provide a direct buffering effect via the protonation of their organic anions dissociated from weak acidic functional groups. The abundant organic anions (e.g., carboxylates, $-COO^-$) within the EPS matrix, particularly when adsorbed onto soil mineral surfaces, act as proton sinks. In response to an acid influx, these groups readily protonate (e.g., $-COO^- + H^+ \rightarrow -COOH$), consuming protons and buffering the soil against a pH decline (Flemming And Wingender 2010). This buffering action is significant; studies have demonstrated that EPS addition increased soil pHBC by 17–22%, effectively delaying re-acidification (Nkoh et al. 2022). Furthermore, EPS regulate the mobility and toxicity of aluminum through complexation. Functional groups within the EPS matrix also complex with highly toxic, exchangeable Al^{3+}, transforming it into less bioavailable forms, such as adsorbed hydroxy-aluminum species and solid-phase organo-aluminum complexes (Nkoh et al. 2020b, 2020c). Finally, EPS help maintain soil chemical stability by retaining essential base cations. Negatively charged sites within the EPS matrix electrostatically adsorb base cations (e.g., Ca^{2+}, Mg^{2+}). This interaction mitigates their leaching loss during acidification events, thereby preserving soil nutrient status and buffering reserves (Nkoh et al. 2020a, 2020c).

2.3 Insights and considerations for AI modeling: The need to decouple organic and inorganic alkaline components

Predictive models of biochar-mediated acidic soil amelioration often oversimplify alkalinity by treating total alkalinity as a single empirical feature for predicting changes in soil pH and pHBC (Brewer et al. 2011; He et al. 2022a, 2022b). While convenient, this approach conflates chemically distinct sources of alkalinity that differ profoundly in origin, reaction pathways, and kinetics, fundamentally limiting the model's ability to distinguish material-specific behavior or predict temporal dynamics. Direct experimental evidence substantiates this limitation. Using selective acid washing and titration, He et al. (2022b) successfully partitioned total alkalinity into its organic (organic anions) and inorganic (carbonates and other mineral alkalis) components. Despite comparable total alkalinities, two biochars exhibited fundamentally different alkaline profiles: corn-straw biochar (~ 93 $cmol_c\ kg^{-1}$) was dominated by inorganic non-carbonate alkalis (≈ 61%), whereas rapeseed-straw biochar (~ 284 $cmol_c\ kg^{-1}$) contained nearly balanced organic (46%) and inorganic (54%) fractions. Functionally, inorganic alkalis provided sustained resistance to acidity, while organic alkalis induced transient pH increases that depleted rapidly. These findings reveal that using total alkalinity as a single feature assumes false functional equivalence and masks mechanistic divergence. As illustrated in Fig. 3, decoupling these components is essential for building mechanistically robust AI models.

From a ML perspective, collapsing two chemically distinct alkalinity classes—each with unique origins, kinetics, and soil interactions—into a single scalar value constitutes and a semantic collapse of the feature space. This compression prevents models from identifying distinct causal pathways for organic versus inorganic

Table 1 Effects of pyrolysis temperature on biochar properties and their implications for acidic soil amendment

Pyrolysis temperature	pH	Major alkaline component	Functional group abundance	Carbonate content	Buffering characteristics	Recommended application
Low (~ 300 ℃)	Moderate (6–9)	Organic functional groups dominant	High ($-COO^-$, –OH)	Low	Rapid but limited buffering capacity	Slightly acidified soils requiring immediate pH correction
Medium (~ 500 ℃)	High (9–11)	Organic–inorganic synergy	Moderate	Significantly increased	Balanced buffering (rapid + sustained)	Optimal for most acidic soil amendment scenarios
High (~ 700 ℃)	Very high (> 11)	Carbonates dominant	Low	Very high	Sustained buffering with pronounced plateau region; slow response	Long-term remediation and carbonate-rich soil applications

The "Recommended application" column reflects expert synthesis based on the empirically generalized from Yuan et al. (2011b), supplemented by the soil incubation and field trial evidence from Deng et al. (2025), Sulaiman et al. (2026), Yuan et al. (2011a), Yuan and Xu (2012), and Xu et al. (2012). For feedstocks or soil conditions outside those tested, these recommendations require further validation components and precludes capturing their contrasting temporal profiles (instantaneous vs. sustained responses) or interactions with soil properties such as CEC, initial pH, and clay content. This conceptual flaw underlies the systematic biases often observed when predictive models for biochar-amended acidic soils are extrapolated across feedstocks and soil types (Li et al. 2024a; Aurangzeib et al. 2025). Mechanistically, organic and inorganic alkalis display distinctive chemical behaviors. Organic alkalinity arises from oxygen-containing functional groups—carboxyl, phenolic hydroxyl, and lactone—on biochar surfaces. Their protonation follows Langmuir-type surface adsorption equilibria, exhibiting rapid onset, site saturation, and short-term buffering in low-pH soils. These functional groups also act as cation-complexation sites (Qian et al. 2015; Wang et al. 2024b), yet are inherently pH-sensitive. In contrast, inorganic alkalinity is governed by mineral components such as carbonates, silicates, and phosphates, whose buffering capacity depends on dissolution kinetics (Rajapaksha et al. 2015), producing delayed but sustained neutralization manifested as a plateau at pH 5.5–8.0. Pyrolysis conditions exert a decisive regulatory control over these two alkaline classes. Characterization of biochars from 17 feedstocks by Brewer et al. (2011) demonstrated that pyrolysis temperature and pressure drive aromaticity and functional-group depletion. With increasing temperature, biochar pH and ash content rise sharply while oxygen-containing functional groups decline—a widely observed pattern (Yuan et al. 2011b; Wang et al. 2014; Bai et al. 2017; Sarfraz et al. 2019; An et al. 2025). Accordingly, design recommendations for biochar selection based on the relative abundance and buffering kinetics of organic and inorganic alkalinity are summarized in Table 1. Notably, these recommendations are independently supported by a recent global meta-analysis of 286 field studies (Zhang et al. 2025c), which reported that biochar produced at <500 °C from wood feedstock achieved the greatest yield gains in strongly acidic soils—consistent with the medium-temperature recommendation in Table 1. It should be noted that the recommendations in Table 1 are not derived from formal meta-analysis or statistically optimized thresholds. These recommendations—including the assignment of pyrolysis temperature categories to specific amendment scenarios—are empirically generalized from the systematic characterization of four crop straw biochars (canola, corn, soybean, and peanut) at 300, 500, and 700 °C by Yuan et al. (2011b), supplemented by the soil incubation and field trial evidence from Deng et al. (2025), Sulaiman et al. (2026), Yuan et al. (2011a), Yuan and Xu (2012), and Xu et al. (2012). The pH ranges, carbonate contents, and buffering characteristics represent observed trends under these specific experimental conditions, rather than statistically derived universal thresholds. The recommended applications reflect expert synthesis based on available evidence and should be interpreted as a general reference framework, with the understanding that optimal selection remains context-dependent on feedstock type, soil properties, and specific amendment objectives.

Mechanistic decoupling of these alkaline components provides a far more informative and physically meaningful feature set for data-driven modeling. In current AI applications—such as RF models predicting soil properties—parameters like biochar C/N ratio or pyrolysis temperature frequently emerge as dominant predictors (Li et al. 2024a, 2025). However, these are merely latent proxies for the underlying ratio of organic to inorganic alkalinity. Recasting these proxies into directly measurable, component-specific descriptors is essential for transitioning models from superficial correlation identification to true causal representation of soil buffering mechanisms. The barrier to this transformation is the scarcity of such component-resolved data. Although numerous publications report total alkalinity, few provide detailed measures such as carboxyl-group density or carbonate content—indicators indispensable for mechanistic AI modeling. This deficiency, repeatedly noted in recent meta-analyses (Liu et al. 2019; He et al. 2021; Li et al. 2024a, 2025; Aurangzeib et al. 2025), remains the field's chief bottleneck. Progress therefore depends on collective action toward standardized data-reporting practices that explicitly include component-specific indicators rather than aggregated totals, providing the mechanistic foundation required for next-generation AI predictive modeling in biochar-amended acidic soils.

(See figure on next page.)

Fig. 4 Global analysis of studies (total N = 119) employing AI for acidic soil remediation. **A** Geographic distribution of the reviewed literature. **B** Word cloud of keywords extracted from the literature, where larger words indicate higher frequency. **C** Word cloud of models and analytical methods, where larger words indicate higher frequency. **D** Annual publication volume over time, showing a distinct upward trend in biochar simulation researches. Data sourced from the Web of Science core collection, accessed on January 12, 2026. *RF* Random Forest, *ANN* Artificial Neural Network, *XGBoost* eXtreme Gradient Boosting, *MLR* Multiple Linear Regression, *GBM* Gradient Boosting Machine. Details of literature retrieval and data generation are provided in Text S1 of the Supporting Information

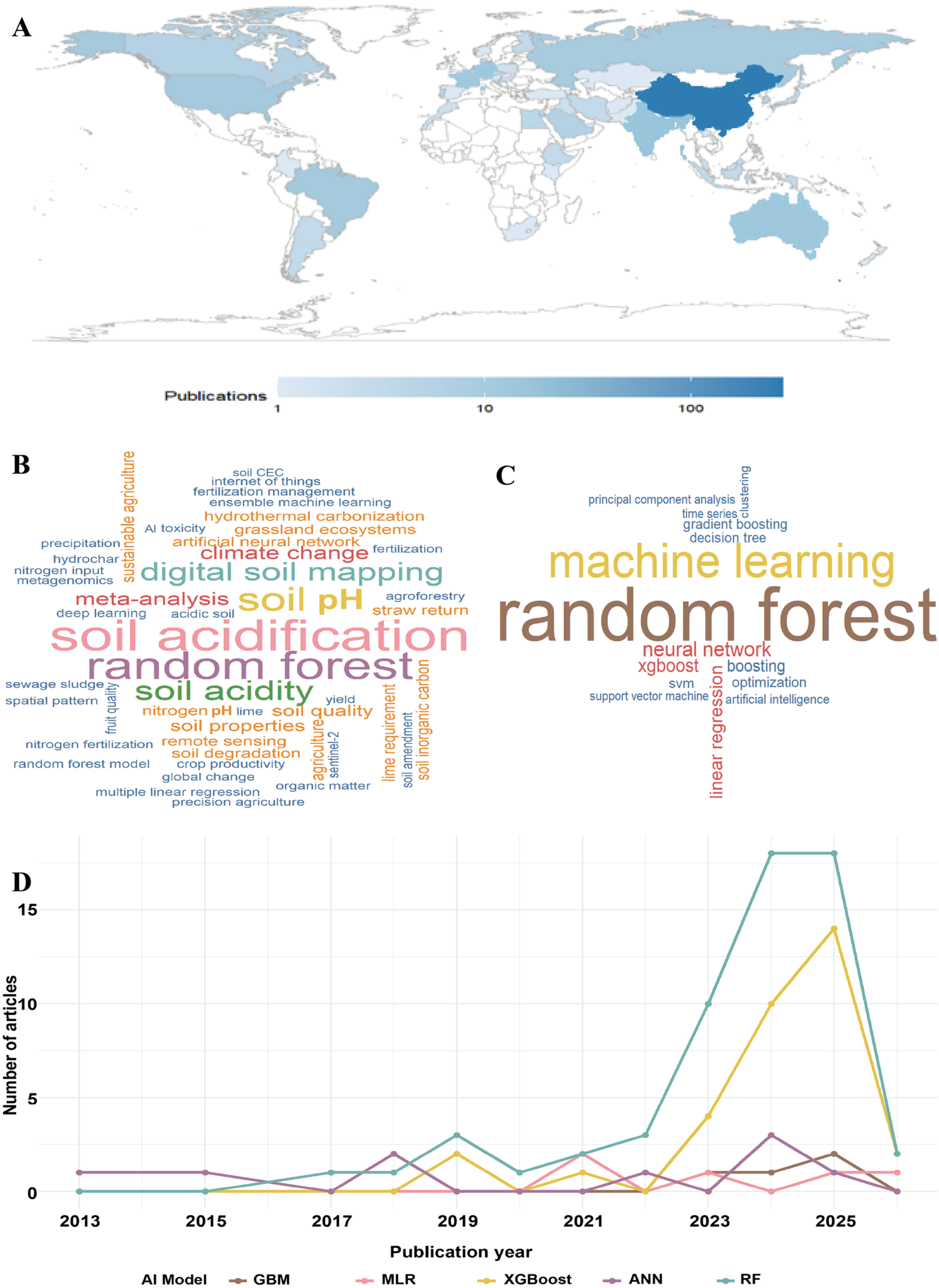


Fig. 4 (See legend on previous page.)

3 Advancing biochar-mediated amelioration of acidic soils using AI

3.1 Research trends and current status of AI applications in evaluating biochar efficacy for acidic soil amelioration

The three-factor framework of soil acidification—active acidity (pH), exchangeable acidity (exchangeable H^+/Al^{3+}), and pHBC—provides well-defined targets for evaluating amelioration efficacy (Liu et al. 2025b). Biochar reduces soil acidity through liming effects, cation exchange, complexation, and precipitation reactions, while enhancing pHBC via protonation of OFGs and release of base cations (Dai et al. 2017). However, the complex nonlinear interactions among biochar properties, inherent soil characteristics, and environmental conditions render traditional statistical models inadequate for accurately predicting amelioration outcomes. Data-driven models, including ML and deep learning approaches, are capable of capturing the intricate nonlinear relationships between biochar properties, soil conditions, and ameliorative effects. Figure 4A illustrates the global distribution of researches in this field, with the majority originating from the Northern Hemisphere, particularly China (78), India (17), Germany (14), France (13), and the United States (11) (for literature search methods, see Text S1, Supporting Information). The predominance of studies from China is consistent with the country's extensive acidic soil area (approximately 32.4% of its total land, Zhao et al. 2023) and its high research output in biochar-related fields over the past decade (Zhang et al. 2025d), but we acknowledge that this distribution does not imply that acidic soil amelioration is a uniquely Chinese concern. Rather, soil acidification is a global agricultural challenge affecting diverse pedoclimatic regions, and the methodological frameworks discussed herein are intended to be transferable across regions, provided that site-specific calibration and validation are performed. To understand the models employed in predicting biochar efficacy for acidic soil amelioration, we conducted a literature review. As shown in Fig. 4C, our analysis reveals a distinct upward trend in biochar simulation researches over the past decade. This trend can be attributed to the continuous development and optimization of models, the progressive improvement in simulation accuracy, and significant advances in data acquisition and sharing. Among these approaches, RF—a prominent ML method—has emerged as the dominant technique in this field (Fig. 4B–D; discussed in detail below). Furthermore, increased attention driven by government policy support and growing recognition of soil acidification and carbon-related issues across various sectors has provided strong impetus for biochar simulation researches.

3.1.1 Designing biochar efficiently using AI

AI models can assist in identifying optimal designs while avoiding over-optimization. First, AI can identify key feature factors. Palansooriya et al. (2022) employed ML models, which suggested that biochar nitrogen content and application rate were the two most significant features influencing heavy metal immobilization. Regarding inverse design, Lyu et al. (2024) proposed using ML to design biochars with high phosphorus adsorption capacity for water pollution remediation. Xiang et al. (2024) further elucidated the threshold effect of metal loading, demonstrating that Fe/Ca modification significantly enhanced Cd adsorption efficiency; however, the effect tended to plateau at excessively high loading levels, suggesting that over-modification should be avoided. AI has also provided new tools for investigating biochar aging effects and structural stability. Liao et al. (2025) constructed multi-factor aging prediction models (RF, GBDT and XGBoost), predicting that optimal biochar structural stability occurred at carbon contents of 60–80%, oxygen contents of 10–30%, and hydrogen contents of 4–6%. Adsorption performance peaked after 35–40 freeze–thaw cycles coupled with 25–35 wet-dry cycles, simulating an "optimal performance window" equivalent to 2–3 years of natural aging. These findings provide quantitative targets for the directed preparation of aging-resistant biochars. In the context of multi-objective optimization design, Wang et al. (2024a) employed interpretable ML (RF, XGBoost) to simultaneously predict heavy metal adsorption capacity and biochar specific surface area (SSA). Their results indicated SSA and adsorption capacity were not linearly correlated; excessively high pyrolysis temperatures, while increasing SSA, compromised surface functional groups. This finding challenges the conventional empirical assumption that "high SSA necessarily ensures high efficacy", suggesting the necessity of multi-objective trade-offs in biochar design.

3.1.2 Predicting biochar efficacy for acidic soil amelioration using AI

Elevating soil pH is the primary indicator of successful acidic soil amelioration, and ML models have proven highly effective in predicting this outcome. For instance, Zhao et al. (2026a) utilized Light Gradient Boosting Machine (LightGBM) and Deep Neural Network (DNN) models to accurately predict soil pH enhancement, identifying biochar pH, CEC, electrical conductivity, and initial soil pH as the critical determinants. Furthermore, because acidic soils often suffer from elevated heavy metal bioavailability, predicting metal immobilization is equally crucial. Using an optimized RF model, Du et al. (2024) demonstrated that biochar properties dictate Cd immobilization efficiency (60.96% contribution), substantially outweighing the influence of experimental conditions (19.6%) or inherent soil properties (19.44%). Ultimately, the goal of soil amelioration is to improve crop yield. Using interpretable ML on 447 observational datasets, Zhai et al. (2026) found that RF models provided the most robust predictions of biochar-induced plant growth. Moreover, SHAP analysis suggested that biochar-related variables dominated the prediction (72.3% contribution—primarily biochar type and application rate) compared to soil-related variables (27.7%, with initial pH contributing only 7.79%). Addressing crop yield via a binary classification approach, Jiao et al. (2025) corroborated these findings. By applying random under sampling to mitigate data imbalance, their RF model (recall = 0.817) confirmed that biochar pH, application rate, initial soil pH, and total nitrogen are the foremost predictors of positive yield responses. Despite these benefits, AI models also highlight the "dual identity" of biochar as both an ameliorant and a potential stressor. Dong et al. (2026) applied an RF classification model to a large dataset, achieving 79% accuracy in predicting biochar's net impact on soil biota. Their model identified high biochar pH (> 6), excessive application rates (> 50 g kg^{-1}), and high pyrolysis temperatures (> 550 °C) as primary risk factors for adverse ecological effects. This presents a critical trade-off for acidic soil management: while highly alkaline biochars offer superior pHBC, their unoptimized application risks detrimental impacts on soil micro- and macro-fauna. Synthesizing these AI-driven insights enable the establishment of optimized, safe application frameworks. Based on predictive modeling, Zhai et al. (2026) recommend high-temperature (℃) straw- or husk-derived biochars applied at approximately 25% (w/w) for acidic soils. To prevent secondary environmental stresses—such as osmotic shock, ionic imbalance, and micronutrient immobilization—AI models explicitly caution against exceeding critical safety thresholds. Specifically, application rates should not surpass 60% (w/w), the electrical conductivity of biochar must remain below 6250 μS cm^{-1}, and biochar nitrogen and phosphorus contents should not exceed 4% and 6%, respectively.

3.1.3 Elucidating microbial responses in biochar-amended acidic soils using AI

Traditional analytical approaches struggle to extract generalizable patterns from high-dimensional, sparse amplicon data, particularly in complex acidic soils driven by overlapping stresses (e.g., low pH, aluminum toxicity). Consequently, ML provides robust frameworks to decipher biochar–microbe interactions by efficiently identifying key responsive taxa, structural network shifts, and functional trajectories (Dai et al. 2021). A primary application of ML lies in the robust identification of microbial biomarkers that indicate successful soil amelioration. Algorithms, particularly RF, excel at mining expansive microbiome datasets to pinpoint core taxa that exhibit consistent responses to biochar's dual role as a carbon source and a liming agent. By constructing classification models, these approaches elucidate biochar-driven community restructuring (e.g., the transition from acid-tolerant oligotrophs to copiotrophs). For instance, Li et al. (2026a) integrated 843 16S rRNA (16S ribosomal RNA (rRNA)) samples across 24 studies using an RF model coupled with recursive feature elimination. This approach successfully screened 138 biomarker genera exhibiting consistent responses to biochar across taxonomic levels, achieving a robust classification accuracy of 85.9% (AUC = 0.859; AUC, area under the curve). Expanding this scope, Lei et al. (2023) integrated 1288 16S rRNA and 525 ITS (the internal transcribed spacer) samples from 53 studies, revealing that fungal communities exhibited a more pronounced response to biochar than bacteria. This finding is highly pertinent to acidic soils, where fungi often dominate prior to remediation. Their RF model achieved an impressive 99.54% classification accuracy in identifying characteristic taxa, including 14 specific bacterial orders (e.g., Gemmatimonadetes) and 7 fungal families (e.g., Erythrobasidiaceae), effectively capturing the structural succession driven by biochar. Furthermore, ML extends beyond taxonomic identification to quantify biomarker contributions to ecosystem functions, facilitating predictions of microbial assembly and the restoration of nutrient cycling previously inhibited by extreme acidity (Lei et al. 2023; Liu et al. 2023; Li et al. 2026a). Rehman et al. (2024), for example, integrated soil enzyme activities and microbial biomass into a binary ML model, accurately delineating "stressed" from "remediated" soils (sensitivity: 80.1%, specificity: 80.5%). Nevertheless, integrating ML into biochar-microbiome research presents ongoing challenges. First, reliance on relative abundance from amplicon sequencing obfuscates absolute microbial population changes (Dai et al. 2021). Relative abundance—expressed as the proportion of each taxon within a community—does not reflect absolute cell numbers; an observed increase in the relative abundance of a given taxon could result either from its absolute proliferation or from the decline of other community members (Vandeputte et al. 2017; Beule et al. 2021). These scenarios carry fundamentally different implications for interpreting microbial responses to biochar amendment. To address this limitation, future models must incorporate absolute quantification methods. Quantitative PCR (qPCR) targeting functional genes (e.g., ammonia monooxygenase subunit A gene (*amoA*) for ammonia

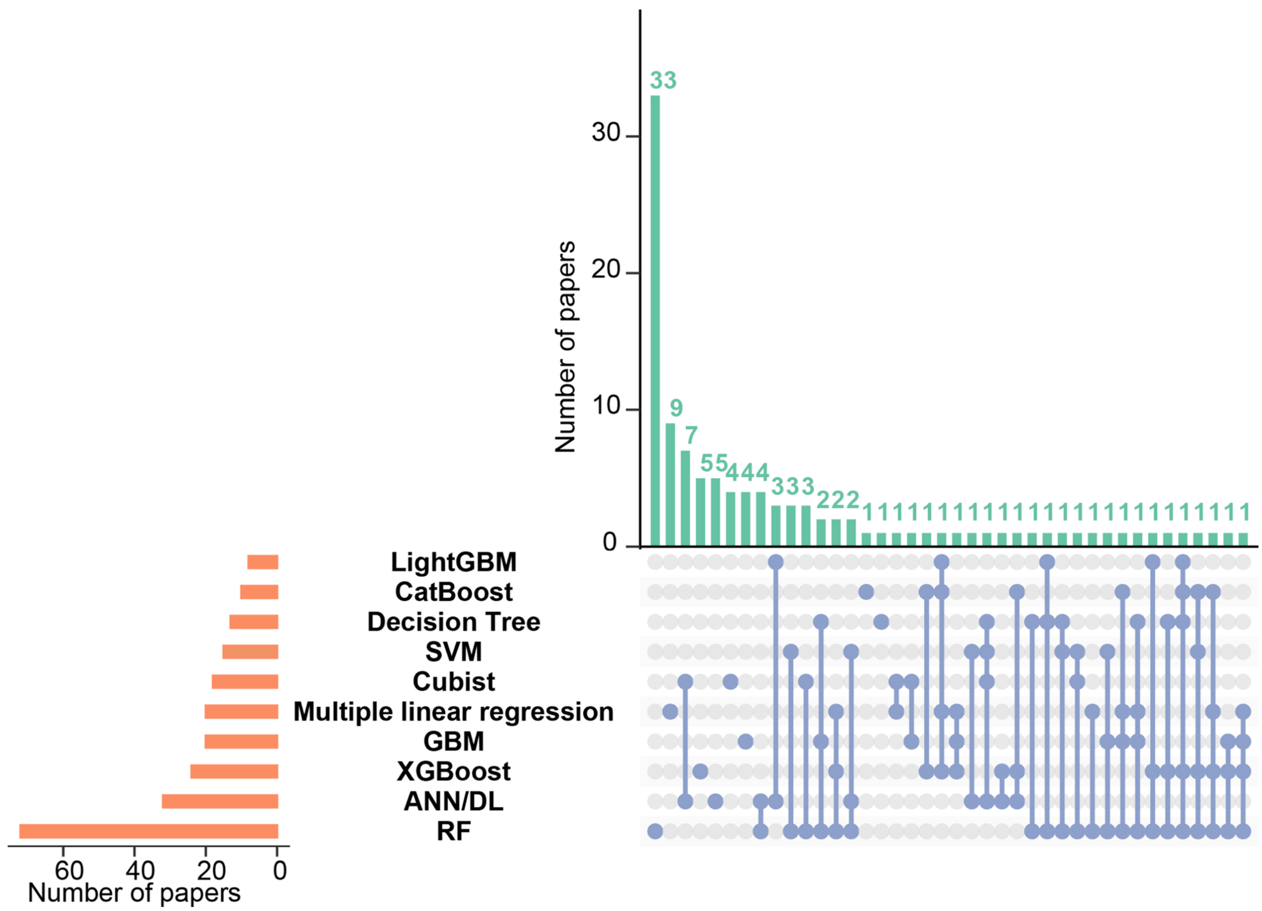


Fig. 5 Publications using AI for predicting acidic soil remediation effects. Bar chart on the left shows total publications for each algorithm. The dot chart below illustrates intersections between algorithms, with connecting lines indicating instances where two or more algorithms are applied together in a single study. Top bar chart displays number of publications associated with these algorithmic intersections. *RF* Random Forest, *SVM* Support Vector Machine, *GBM* Gradient Boosting Machine, *XGBoost* eXtreme Gradient Boosting, *LightGBM* Light Gradient Boosting Machine, *CatBoost* Categorical Boosting, *ANN* Artificial Neural Network, *DL* Deep Learning. Data sourced from Web of Science 12 January 2026. Details of literature retrieval and data generation are provided in Text S1 of the Supporting Information

oxidizers, nitrite reductase gene (*nirS*) and nitrous oxide reductase gene (*nosZ*) for denitrifiers) provides a widely adopted approach for absolute quantification of key functional groups (Jian et al. 2020). For community-level absolute quantification, methods such as synthetic chimeric DNA spike-ins added directly to samples prior to DNA extraction, allowing calculation of absolute abundance based on the known spike-in copy number (Tkacz et al. 2018), the integrated high-throughput absolute abundance quantification (iHAAQ) approach combining high-throughput sequencing with qPCR (Wang et al. 2022), or droplet digital PCR (ddPCR) for absolute quantification without standard curves (Quan et al. 2018) can be adopted. Recent advances have expanded this toolkit: the Accu16S/AccuITS method employs synthetic internal spike-in DNA added prior to PCR, enabling standard-curve-based absolute quantification of bacterial and fungal amplicons in a single sequencing run (Bai et al. 2026a; Wen et al. 2023). Complementarily, Tourlousse and Sekiguchi (2025) developed 12 synthetic rRNA operons (rDNA-mimics) as spike-in standards for cross-domain absolute quantification of fungal and bacterial microbiomes. Integrating absolute abundance data into ML frameworks will substantially enhance the ecological interpretability and practical relevance of predictive models for biochar-amended soils. Second, marker gene-based functional predictions (e.g., FAPROTAX, Functional Annotation of Prokaryotic Taxa) risk false positives and require validation through multi-omics approaches including metagenomics and metatranscriptomics. Third, current models largely omit viruses and protists. Integrating trans-kingdom data into ML frameworks is essential for untangling complex food web responses to biochar application.

Table 2 Applications of different AI models in biochar-mediated acidic soil amelioration

Prediction target	Machine learning model	Key features identified	Performance metrics	Reference
Soil pH improvement	LightGBM, DNN	Biochar pH, biochar CEC, biochar EC, initial soil pH	$R^2 = 0.92$, MAE = 0.291, RMSE = 0.539	Zhao et al. (2026a)
Crop yield response (Classification)	RF	Biochar pH, biochar addition rate, initial soil pH, initial soil total nitrogen	Recall = 0.817, Accuracy = 0.687	Jiao et al. (2025)
Crop yield (regression)	RF, MLP-NN	Biochar application rate, initial SOC, biochar pH, biochar total phosphorus	Test $R^2 = 0.83$ (RF), 0.84 (MLP-NN)	Xu et al. (2025)
Heavy metal immobilization (Cd, Pb, Zn, etc.)	RF	Biochar properties (especially N content), biochar application rate, soil properties, heavy metal type	Training $R^2 = 0.90$, Testing $R^2 = 0.85$	Guo et al. (2023)
	RF, XGBoost	Biochar properties (pH, CEC, ash), soil pH, soil CEC, initial metal concentration	$R^2 > 0.8$ for Cr, Cu, Ni, Pb, Zn	Chen et al. (2025)
	RF	Biochar properties (60.96%), experimental conditions (19.6%), soil properties (19.44%)	–	Du et al. (2024)
Heavy metal adsorption capacity (Cd, Pb, etc.)	XGBoost, LightGBM, DNN	Biochar pH, CEC, EC, H/C, O/C; adsorption conditions (pH, time, concentration)	$R^2 = 0.89–0.92$	Zhao et al. (2026b); Jiang et al. (2025)
	RF, XGBoost, CatBoost	Metal type, initial concentration, biochar feedstock, pyrolysis temperature, solution pH	$R^2 = 0.92–0.96$	Yaseen & Alhalimi (2025); Duan et al. (2024)
Biochar yield prediction	XGBoost, LightGBM, DNN	Pyrolysis temperature, biomass three-component parameters (cellulose, hemicellulose, lignin), ash content	$R^2 = 0.89–0.92$	Zhao et al. (2024, 2025)
	RF, XGBoost	Pyrolysis temperature, residence time, fixed carbon, nitrogen content, ash content	$R^2 = 0.87–0.98$	Leng et al. (2022); Zhu et al. (2023)
Biochar carbon content	XGBoost, RF	Pyrolysis temperature, biomass composition (C, H, O, N), ash content	$R^2 = 0.90–0.96$	Leng et al. (2022); Zhu et al. (2019)
Biochar specific surface area	XGBoost, RF, DNN	Pyrolysis temperature, ash content, plant organ type	$R^2 = 0.84–0.92$	Leng et al. (2022); Jiang et al. (2025)
Greenhouse gas emissions (N_2O, CH_4, CO_2)	XGBoost, RF, LSTM	Mineral N fertilizer rate, organic fertilizer application, irrigation mode, soil pH, soil organic carbon	$R^2 = 0.75–0.87$ (for CH_4/N_2O)	Wu et al. (2023); Hamrani et al. (2020)
Soil organic carbon (SOC) sequestration	RF	Biochar application rate, initial SOC, latitude, biochar pH, biochar total phosphorus	Test $R^2 = 0.87$	Xu et al. (2025)
	RF, XGBoost	Soil properties (60.27%), biochar properties, climate factors, Cd stress level	–	Sun et al. (2026)
Biochar cation exchange capacity (CEC)	CatBoost, LightGBM, XGBoost	Feedstock composition (cellulose, lignin), pyrolysis temperature	$R^2 = 0.76–0.95$	Xiang et al. (2025)
Biochar pH	CatBoost, LightGBM, XGBoost	Pyrolysis temperature, feedstock composition (cellulose, lignin)	$R^2 = 0.76–0.95$	Xiang et al. (2025)

Details of literature retrieval are provided in Text S2 of the Supporting Information. *LightGBM* Light Gradient Boosting Machine, *DNN* Deep Neural Network, *RF* Random Forest, *MLP-NN* Multilayer Perceptron Neural Network, *XGBoost* eXtreme Gradient Boosting, *CatBoost* Categorical Boosting, *LSTM* Long Short-Term Memory

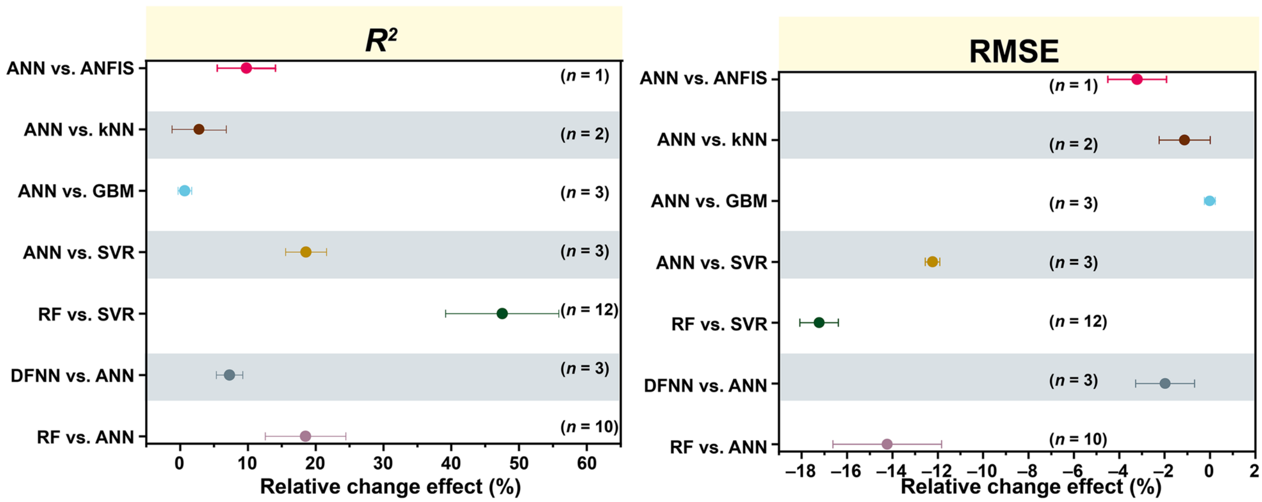


Fig. 6 Comparative performance of machine learning algorithms and other modeling approaches for predicting the efficacy of biochar in acidic soil remediation. Relative changes are expressed as percentages for R^2 (left panel) and RMSE (right panel). Each comparison is made against a reference model listed on the right side of the pairing (e.g., "RF vs. ANN" indicates that Random Forest (RF) is compared against Artificial Neural Network (ANN)). Other comparator approaches include statistical methods, such as Response Surface Methodology (RSM), and computational intelligence techniques, such as Adaptive Neuro-Fuzzy Inference System (ANFIS). Positive values indicate improvements in R^2 (increase) or reductions in RMSE (decrease) for the left-side model relative to the right-side reference. Error bars denote 95% confidence intervals estimated via bootstrapping; the number in parentheses for each comparison indicates the number of studies included (*n*). *RF* Random Forest, *ANN* Artificial Neural Network, *DFNN* Deep Feedforward Neural Network, *SVR* Support Vector Regression, *kNN* k-Nearest Neighbor, *RSM* Response Surface Methodology, *ANFIS* Adaptive Neuro-Fuzzy Inference System. Detailed literature retrieval and data generation procedures are provided in Text S3 of the Supporting Information

3.2 Dominant models and their applications

AI and ML techniques have emerged as indispensable tools for predicting the efficacy of biochar-mediated soil remediation through correlating biochar properties, initial soil conditions, environmental variables, and field observations. Based on a systematic analysis of 119 eligible studies, our review has identified the top 10 AI algorithms employed in biochar for acidic soil amelioration as shown in Fig. 5 (for literature search methods, see Text S1, Supporting information). RF was the most frequently used algorithm, involved in 72 studies, including those combining RF with other algorithms. Among the other ensemble techniques, XGBoost (23 studies), Gradient Boosting Machine (GBM, 18), Categorical Boosting (CatBoost, 8), and LightGBM (8) were also frequently employed. Among DL techniques, artificial neural network (ANN) was the most commonly employed. These frequency counts are descriptive tallies of our literature sample (n = 119) and are not accompanied by uncertainty intervals, as they represent complete enumeration of the reviewed studies rather than statistical estimates from a sampled population. RF has emerged as the preferred choice for feature screening and attribution analysis, owing to its robust resistance to overfitting and its built-in utility for feature importance ranking, a finding that corroborates the conclusions of Ding et al. (2025) derived from large-scale meta-analyses of SOC prediction. Biochar amelioration of acidic soil can increase crop yields (Zhang et al. 2024a; Xu et al. 2025), immobilize labile heavy metals such as copper, raise SOC content (Shi et al. 2019a; Sun et al. 2026), and mitigate greenhouse gas emissions (Bai et al. 2026b; He et al. 2024; Khan et al. 2026; Wu et al. 2023). Yang et al. (2025) reported that long-term (≥4 years) continuous biochar application increased crop yields by an average of 10.8%, significantly reduced methane (CH_4) emissions in paddy fields by 13.5% and N_2O emissions across all soils by 21.4%, and increased SOC content by an average of 52.5%. Furthermore, we performed a detailed frequency analysis across these application scenarios and found that RF was the most frequently used model in all cases (Fig. S1; Table 2; for literature search methods, see Text S1, Supporting information). To elucidate the drivers behind the preferential adoption of RF across these four scenarios, we conducted a detailed analysis of dataset characteristics and algorithmic adaptability for each sub-domain. The complete rationale and supporting evidences are provided in Text S4 (Supporting Information).

To systematically evaluate algorithm performance, we conducted a meta-analysis of the most commonly deployed models (Fig. 6). We compiled all eligible studies into a unified random-effects meta-analytic framework to assess overall algorithmic performance trends (Meta-analysis methods and data extraction protocols are described in Text S3, Supporting Information). We acknowledge that this aggregation introduces heterogeneity in terms of target variables, datasets, and experimental designs—a limitation that must be considered when interpreting the pooled results. Across the evaluated studies—which predominantly utilized cross-validation techniques such as k-fold or leave-one-out cross-validation—the coefficient of determination (R^2) and root mean square error (RMSE) were the primary metrics for goodness-of-fit and predictive accuracy, respectively. Meta-analytic results indicated that RF exhibited significantly higher predictive accuracy than Support Vector Regression (SVR, a variant of Support Vector Machine), as evidenced by non-overlapping 95% confidence intervals (CIs) and effect sizes positioned to the right of the zero line. In the RF vs. ANN comparison, the latter performed comparably to RF, although RF maintained a statistically significant advantage in the majority of included studies. Overall, both RF and ANN demonstrated higher predictive accuracy (higher R^2, lower RMSE) than SVR within the compiled acidic-soil biochar literature. These results represent a meta-analytic summary across multiple application scenarios (e.g., SOC, yield, heavy metal immobilization); the inherent heterogeneity in target variables and experimental designs across primary studies warrant cautious interpretation. Furthermore, the algorithm usage frequencies reported earlier in this review (Fig. 5) are descriptive literature counts and should not be conflated with the statistically tested performance rankings presented here. The variable importance identified by RF was further validated by Xia et al. (2023), whose path analysis confirmed that key predictors (e.g., soil pH, organic matter) correlated significantly with maize physiological indicators. Looking forward, deep learning architectures (e.g., Convolutional Neural Networks (CNNs), Transformers) hold promise for soil property inversion and microbial function prediction, particularly as spectral libraries and field monitoring datasets continue to expand (Ding et al. 2025).

3.3 Model interpretability and mechanistic linking

Despite the formidable predictive capabilities of AI models, their inherent "black-box" opacity has historically constrained their utility in mechanistic research. To improve model transparency, researchers are increasingly adopting Explainable AI (XAI) tools that help bridge computational outputs with underlying physicochemical and biological principles. We emphasize a critical distinction: XAI tools provide predictive association and facilitate hypothesis generation; they do not, by themselves, substitute for experimentally validated causal inference (Novielli et al. 2025).

In addition to quantifying individual feature contributions, SHAP—which originated from game theory—is applicable to both regression and classification tasks and provides variable importance ranking as well as prediction interpretability (Grunwald 2022). Lu et al. (2025) reported that copper adsorption by biochar was most strongly associated with operational experimental conditions, accounting for 53.5% of the total contribution—specifically, initial concentration (30.41%), ash content

Table 3 Comparison of the advantages and disadvantages of machine learning and deep learning in biochar efficacy prediction in acidic soil remediation, including performance features and preferable applications

Comparison dimension	Machine learning	Deep learning	Reference
Training data requirement	Performs well on small to medium-sized datasets; low data demand	Typically requires very large amounts of training data to achieve optimal performance	LeCun et al. 2015
Computational resource requirement	Low; can run on standard computers	High; requires GPU or high-performance computing clusters	Alom et al. 2019
Model structure	Shallow architecture (1–2 functional layers) or tree-based structure	Deep network architecture (5–100 + hidden layers)	Reichstein et al. 2019
Feature extraction	Requires manual feature engineering; relies on domain knowledge	Automatically learns hierarchical feature representations	Padarian et al. 2019; Li and Hong 2023
Model interpretability	High (feature importance, SHAP, PDP, etc.)	Low ("black box" models; difficult to interpret)	Shen et al. 2024; Xiang et al. 2025
Training time	Short to moderate	Long; requires extensive iterations	Alom et al. 2019
Overfitting risk	Moderate; controllable via regularization	High; requires large datasets and regularization strategies	LeCun et al. 2015
Preferred applications	Lab/field point-scale prediction; small-to-medium scale modeling; scenarios requiring mechanistic interpretation	Multi-source remote sensing data fusion; high-dimensional spectral processing; spatiotemporal sequence modeling	Li and Hong 2023

(16.86%), and biochar dosage (9.36%). In their model, the contributions of biochar's physical properties (9.5%) and pyrolysis conditions (4.4%) were substantially lower than anticipated. Shen et al. (2024) similarly identified initial concentration as a dominant feature driving heavy metal adsorption. Palansooriya et al. (2022) ranked biochar nitrogen content and application rate as the most influential predictors of heavy metal immobilization efficiency in soil systems, whereas intrinsic soil properties exhibited relatively less influence in their model.

While SHAP quantifies how much a feature contributes, PDP visualize how that contribution manifests by displaying nonlinear thresholds and feature interactions derived from the model. Lyu et al. (2024) employed PDP to illustrate a model-derived critical threshold in biochar design, indicating that higher carbon content does not infinitely equate to higher adsorption capacity within the model's predictive framework.

Feature importance derived from a single algorithm may reflect structural biases inherent to that model. To address this, multi-model consensus and data-preprocessing strategies are employed. Ke et al. (2021) utilized Fuzzy C-means clustering integrated with a Back Propagation Neural Network (FCM-BPNN) for heavy metal adsorption prediction ($R^2 = 0.987$). Dashti et al. (2023) applied multi-model consensus and found that the "heavy metal-to-biochar concentration ratio" and "total carbon content" were consistently ranked among the top predictors across different algorithmic architectures. In summary, SHAP provides global contribution rankings, PDP visualize model-derived nonlinear relationships, and multi-model consensus offers a means of comparing variable rankings across algorithms. We reiterate that these analyses are model-dependent and should be interpreted as associative patterns rather than causal truths (Ding et al. 2025). Their primary value lies in generating testable mechanistic hypotheses and prioritizing factors for subsequent experimental validation, thereby directing future researches toward targeted causal experiments rather than providing definitive mechanistic conclusions.

3.4 Methodological selection under different conditions: Machine learning versus deep learning

Selecting the optimal predictive architecture for biochar-mediated acidic soil amelioration requires a profound understanding of the distinctions between traditional ML and DL. This methodological crossroad should not be navigated blindly; rather, algorithm selection must be rigorously guided by data volume, specific prediction tasks, available computational resources, and the prerequisite level of model interpretability, as detailed in Table 3.

3.5 Challenges in AI applications

3.5.1 Data availability and heterogeneity: The dilemma of indicator comparability

The predictive robustness of any AI model is intrinsically chained to the representativeness and quality of its training data. However, severe data heterogeneity in current biochar research complicates the direct cross-study comparison of model performance. As Ding et al. (2025) explicitly noted in their investigation of SOC prediction, evaluation metrics such as R^2 and RMSE reported across different studies frequently lack baseline comparability. The root cause lies in disparate dataset architectures—variations in sampling density, spatial scales, and covariate quality substantially distort cross-model evaluations. This poses a critical challenge: when researchers compile meta-datasets from diverse literature, they inevitably inherit systematic biases introduced by heterogeneous experimental designs and preprocessing protocols. Consequently, establishing unified benchmarking platforms and standardized data-sharing protocols is urgently required to ensure fair model evaluations and robust meta-analyses.

3.5.2 Correlation does not imply causation (The persistent "black box" problem)

Fundamentally, current machine learning algorithms operate as sophisticated correlational engines rather than causal inference frameworks. While XAI tools like SHAP and PDP have immensely enhanced interpretability, the inherent "black box" problem remains fundamentally unresolved (Xiong et al. 2014; Wadoux 2025). Bridging this gap requires transitioning from statistical correlation to mechanistic causality. On one front, this necessitates integrating structural equation modeling or developing causal forest frameworks. On another front, it demands advancements in novel characterization techniques, particularly in digitalizing microbial systems. For instance, relying solely on relative microbial abundance can obscure true ecological responses—an apparent percentage increase in a specific taxon may result merely from the decline of competitors, rather than its own proliferation (Zhang et al. 2025e). To extract true causal mechanistic insights, ML models must be fed absolute quantification data, which can be achieved by integrating internal standard strains or qPCR.

3.5.3 Limited extrapolation capacity: Biochar aging and temporal dynamics

Most AI models in this domain are trained on static, "snapshot" historical data, severely limiting their capacity to extrapolate temporal changes under future land-use or climate scenarios. In contrast, traditional process-based models possess an intrinsic mechanistic understanding of critical phenomena like biochar aging and continuous soil re-acidification, making them more robust for long-term forecasting in dynamic environments (Cheng et al. 2008; Zimmerman 2010). In acidic soils, biochar continuously "ages"—its localized pHBC gradually diminishes, surface oxygen-containing functional groups proliferate, and cation exchange sites evolve. Wang et al. (2020) highlighted this paradigm, emphasizing the urgent need to transition from qualitative descriptions to quantitative simulations of biochar aging. Because soils frequently face re-acidification risks as biochar ages, future research must establish long-term monitoring systems to capture these temporal successions. Consequently, AI frameworks should increasingly incorporate time-series architectures—such as Long Short-Term Memory (LSTM) networks (Wu 2021)—to simulate the coupling between biochar aging, dynamic nitrification processes, and the prediction of critical time windows for re-acidification, ultimately providing dynamic early warnings for field management.

3.5.4 Deficits in data paradigms and feature standardization

While relentless algorithmic iteration drives AI forward, the absence of a standardized, reusable "data ecosystem" has emerged as the primary bottleneck in biochar informatics. A unified metadata dictionary is imperative for cross-study compilation; yet, standardized conventions for critical descriptive dimensions are severely lacking. For instance, Shen et al. (2024) highlighted that disparate studies report biochar elemental composition based on entirely inconsistent baselines (e.g., dry-basis vs. ash-free basis), necessitating arduous harmonization prior to modeling. Furthermore, critical gaps frequently plague the input feature space. While most publications report basic C, H, oxygen (O), and N contents, very few explicitly detail the macromolecular structure of the feedstocks—such as cellulose, hemicellulose, and lignin content—despite their irreplaceable explanatory power in predicting biochar CEC and Fixed Carbon (Xiang et al. 2025). The scarcity of complete feature reporting fundamentally cripples model accuracy. The challenges Lyu et al. (2024) encountered while predicting phosphorus adsorption perfectly illustrate this issue: numerous studies detailed their exact methods for Fe, Ca, and Mg modification, but completely failed to report the final, actual elemental loadings on the engineered biochar, resulting in the total absence of this crucial mechanistic feature from the model's input layer. Overcoming these "usable but suboptimal" data silos requires a paradigm shift toward open-source raw data reporting and comprehensive feature documentation.

3.6 Enhancing AI predictive performance through data fusion

In the researches on biochar-mediated acidic soil amelioration, single data sources are often insufficient to fully characterize the complex continuum spanning from macroscopic ecosystem responses to microscopic interfacial processes. Although remote sensing (RS) technologies provide large-scale vegetation indices and surface attributes, they are constrained by surface signal interference and the inability to penetrate the soil profile. Conversely, PSS enables rapid in situ measurements at point scales but suffers from limited spatial coverage. Laboratory spectroscopy and microscopic characterizations elucidate intricate mechanisms yet cannot support regional extrapolation. Similarly, microbial molecular data elucidate community succession but are frequently decoupled from dynamic physicochemical processes. Consequently, multi-source data fusion—integrating data across different platforms, scales, and modalities into a unified feature space—has emerged as a core strategy to overcome the limitations of isolated datasets and unlock the full predictive potential of artificial intelligence (Fig. 7).

3.6.1 Remote sensing-PSS fusion: Bridging macroscopic proxies and ground truth

Integrating RS information with PSS data through advanced fusion techniques represents a robust approach for enhancing predictive accuracy (Ding et al. 2025). Within this framework, RS imagery delivers large-scale, spatiotemporally continuous vegetation indices and surface reflectance data (Fig. 8), while PSS techniques (e.g., Vis–NIR spectroscopy, PXRF) provide precise, point-scale "ground truth" soil attributes (Azizi et al. 2026). These multi-source observations can be leveraged through feature-level or decision-level fusion strategies to calibrate model inputs, optimize feature spaces, and enhance model generalizability, thereby significantly improving the prediction of key soil properties. Studies integrating RS and PSS data typically employ two main paradigms: direct substitution and advanced model fusion/assimilation (Bao et al. 2025). The direct substitution approach involves using RS-derived spectral indices (e.g., Normalized Difference Vegetation Index and Enhanced Vegetation Index) or environmental covariates directly as input features for machine learning models, thereby supplementing traditional sparse soil sampling. In contrast, advanced fusion techniques dynamically bridge these two scales. A prime example of feature-level spectral fusion is demonstrated by Bao et al. (2023), who fused Sentinel-2A satellite imagery with laboratory Vis–NIR spectral data. They successfully generated hyperspectral-resolution images through spectral reconstruction methods, substantially improving

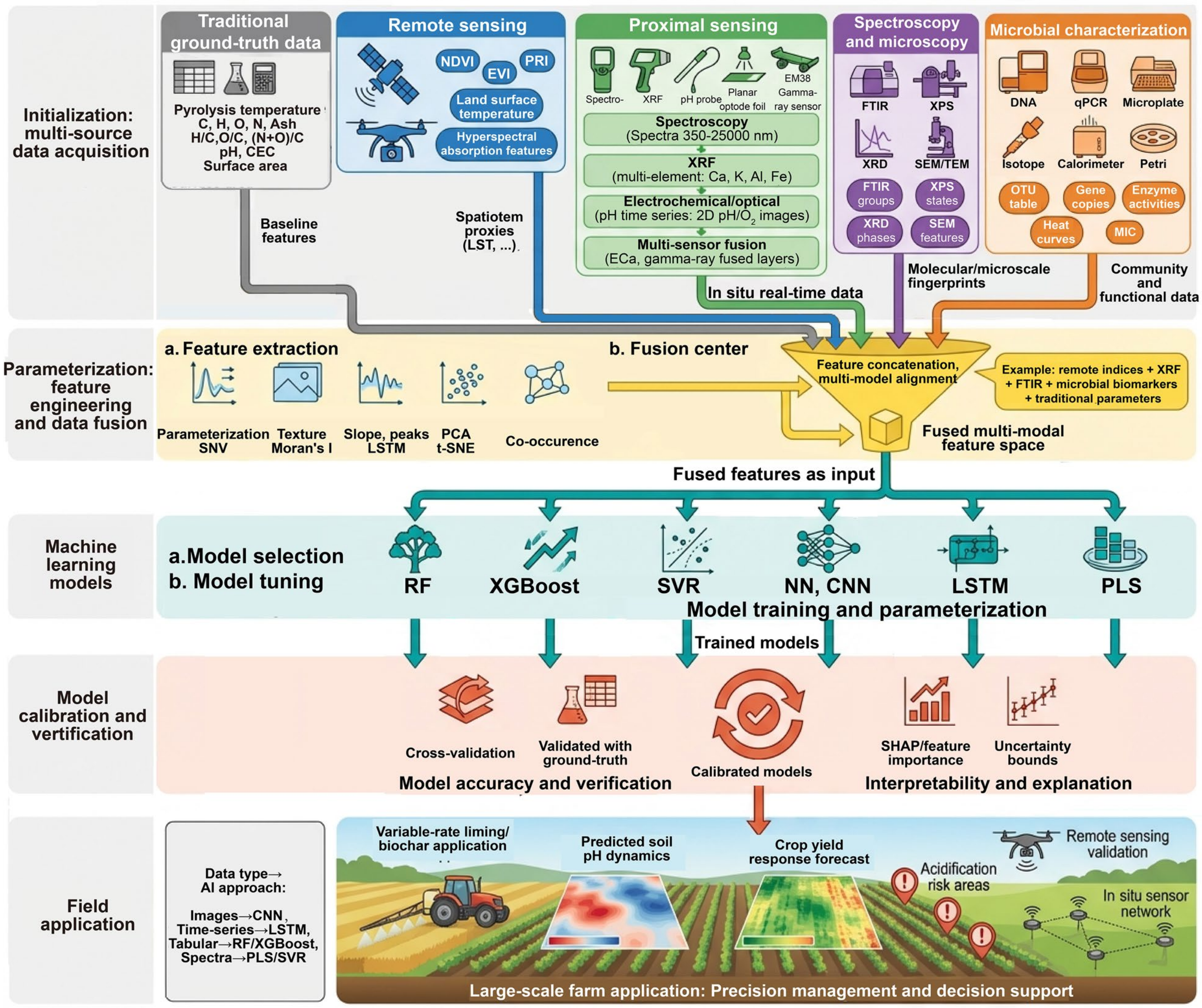


Fig. 7 Schematic overview of multi-source data fusion for AI modeling of biochar efficacy in acidic soil remediation. The modeling framework integrates multi-source input data, including soil properties, environmental conditions, and management practices, alongside data acquired from satellites, aerial platforms, unmanned aerial vehicles, and proximal soil sensing. The data processing pipeline involves feature extraction and a fusion center, followed by model selection and training. Model calibration involves sensitivity analysis and parameter tuning, leading to model validation. The biochar efficacy simulation process integrates multi-source data and generates prediction outputs, which are refined through AI modeling and data fusion. Model performance is validated through comparisons with laboratory analysis data, remote sensing observations, and proximal sensing measurements. This framework is expected to enable large-scale farmland applications. *NDVI* Normalized Difference Vegetation Index, *EVI* Enhanced Vegetation Index, *PRI* Photochemical Reflectance Index *XRF* X-ray Fluorescence Spectroscopy, *FTIR* Fourier Transform Infrared Spectroscopy, *XRD* X-ray Diffraction, *XPS* X-ray Photoelectron Spectroscopy, *SEM/TEM* Scanning Electron Microscopy/ Transmission Electron Microscopy, *RF* Random Forest, *XGBoost* eXtreme Gradient Boosting, *SVR* Support Vector Regression, *NN* Neural Network, *CNN* Convolutional Neural Network, *LSTM* Long Short-Term Memory, *PLS* Partial Least Squares, *MIC* Maximum Information Coefficient

SOC prediction accuracy. Specifically, they first identified soil texture classes in remote sensing imagery using soil type maps, established relationships between satellite reflectance spectra and corresponding PSS data at sampling locations, and ultimately reconstructed hyperspectral images via spectral response functions. The fused model achieved an R^2 value of 0.78, significantly outperforming independent models fed exclusively with RS data ($R^2=0.57$) or PSS data ($R^2=0.69$). Extending this approach into precision agriculture, Dehkordi et al. (2020) pioneered the use of unmanned aerial vehicles (UAVs) equipped with RGB and multispectral sensors to continuously assess biochar effects on crop performance throughout the growing season. By fusing aerial data with ground-based measurements, their findings quantified that century-old historical biochar had a limited

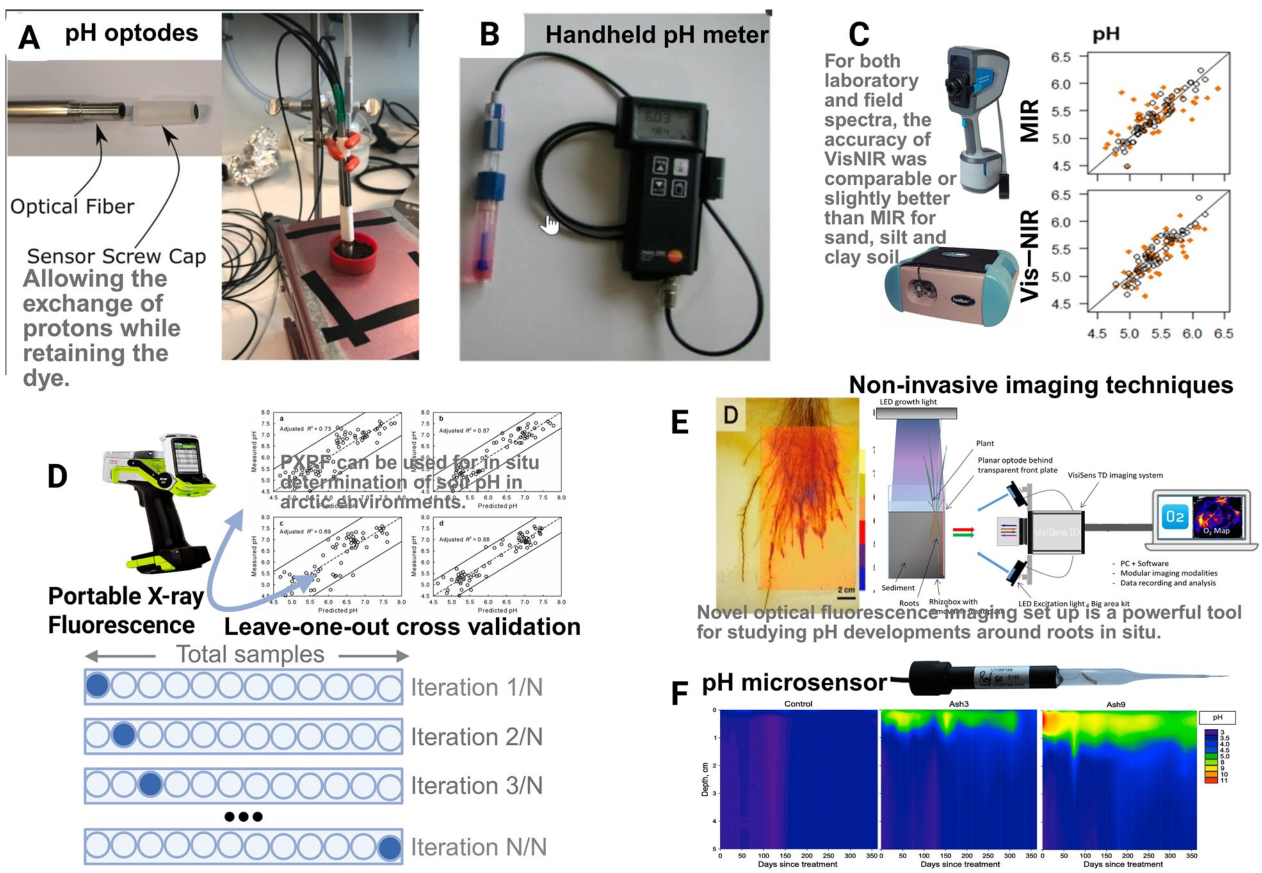


Fig. 8 Recent advances in proximal soil sensing in situ. **A** pH-optodes (Merl et al. 2022); **B** pH meters (von Cossel et al. 2019); **C** Visible/near- and mid-infrared spectroscopy in situ (Greenberg et al. 2022a, b); **D** Portable X-ray Fluorescence (PXRF) Spectrometry (Chakraborty et al. 2016); **E** Planar optode sensor for rhizosphere, adapted from Maisch et al. (2019) and Koop-Jakobsen et al. (2021); **F** pH microsensor (Hansen et al. 2017). *PXRF* Portable X-ray Fluorescence, *MIR* Mid-Infrared, *Vis–NIR* Visible and Near-Infrared (spectroscopy)

impact on immediate crop yields compared to inherent soil pedological properties. This research highlights that RS-PSS data fusion not only evaluates the long-term field efficacy of biochar but also provides critical spatial decision support for variable-rate fertilization.

The rapid advancement of PSS technologies has mitigated numerous bottlenecks in field monitoring and data-driven modeling. Benefiting from continuous sensor miniaturization, techniques such as Vis–NIR spectroscopy, PXRF spectrometry, and novel electrochemical/optical sensors are transitioning from laboratory-based ex-situ analyses toward real-time, in situ field deployments with vastly improved limits of detection. For instance, monitoring pH—the most critical indicator of successful acidic soil amelioration—has witnessed the development of various innovative in situ techniques (Fig. 8). Novel fiber-optic pH sensors (pH optodes), utilizing fluorescent materials whose reversible intensity changes correlate dynamically with pH, can be directly inserted into the soil profile, offering exceptional durability and portability (Kahlert et al. 2004; Merl et al. 2022). Furthermore, biosensors employing betalain as a pH-responsive indicator, when coupled with smartphone imaging and ML algorithms (Ichsan et al. 2026), have enabled the two-dimensional visualization of pH and O_2 micro-distributions at the root–soil interface via planar optode foils (Blossfeld et al. 2013; Hu et al. 2025). Complementing these chemical sensors, PXRF spectrometers now enable the in situ scanning of challenging matrices like frozen soils (Chakraborty et al. 2016). Simultaneously, portable Vis–NIR and mid-infrared spectrometers effectively capture the spectral fingerprints of soil organic matter, minerals, and moisture content in the field (Greenberg et al. 2022a, b). Together, these tools provide the high-density technical support necessary for rapid, large-scale mapping. However, when feeding these multimodal datasets into AI models, caution must be exercised. Establishing independent, high-quality in situ observational datasets remains crucial for rigorous model validation. Notably, simply accumulating multi-source data does not inherently guarantee improved model performance; without proper feature selection and dimensionality reduction, indiscriminate fusion can introduce severe multicollinearity, algorithmic redundancy, and amplified noise.

3.6.2 Multi-modal data fusion: Linking community structure with functional processes

Multi-modal microbial data fusion comprehensively integrates taxonomic composition from high-throughput sequencing targeting the 16S rRNA gene and ITS region, absolute abundances of functional genes via qPCR, metabolic potential characterized by enzyme activity profiling, and real-time metabolic vigor captured by isothermal microcalorimetry. This integration constructs a holistic data continuum spanning from mere microbial composition to functional cooperation, niche partitioning, and life-strategy trade-offs. Consequently, it supplies AI models with feature dimensions that far exceed the predictive capacity of isolated datasets, particularly for assessing biochar efficacy in acidic soil systems. Recent studies underscore the power of this approach. For instance, Qi et al. (2025) integrated 16S rRNA sequencing with soil physicochemical properties, revealing that biochar-driven shifts in rare microbial taxa intricately correlated with wheat growth. This multi-modal approach successfully identified critical biological indicators that remain elusive to conventional, standalone sequencing datasets. Similarly, Chen et al. (2024) systematically assessed the ecological threats of nanomaterials to soil microbiomes by coupling meta-analysis with machine learning. Advancing this spatial resolution, Xu et al. (2023) integrated 16S sequencing, enzyme activity imaging, and physicochemical data to precisely pinpoint microbial colonization patterns within enzyme hotspots in the maize rhizosphere during organic waste biotransformation. This generated spatially explicit training data, elevating predictions from "point-scale means" to "microdomain distributions." Transitioning to functional validation, Shahare et al. (2023) systematically proved that ML algorithms could achieve high-accuracy predictions of soil health based on multi-modal data, validating enzyme activities as robust functional indicators of biochar amelioration. At the ecosystem level, Shu et al. (2023) deployed RF models to integrate physicochemical properties, bacterial structures, and organic carbon mineralization rates. Their model successfully isolated the synergistic environmental and microbial drivers of carbon mineralization during the restoration of degraded alpine grasslands. Expanding on this, Deng et al. (2024) integrated multi-decadal physicochemical, microbial, and crop yield data from long-term fertilization trials. Using RF models, they demonstrated that soil ecosystem multifunctionality strongly correlated with crop yield, pinpointing bacterial community diversity as the paramount predictor. In summary, multi-modal microbial data fusion transcends superficial correlations, enabling a profound mechanistic elucidation of soil-microbe interactive responses to biochar application.

3.6.3 Imaging-chemical data fusion: Deep mining of spatially explicit information

Advanced imaging technologies—including hyperspectral imaging (HSI), planar optodes, and electron microscopy—yield rich, spatially explicit datasets vital for biochar-soil researches. However, the high dimensionality and complexity of these images far exceed the processing limits of traditional point-measurement analytics. AI—particularly DL algorithms equipped with powerful autonomous feature extraction capabilities—has therefore emerged as a pivotal tool. By fusing image-derived spatial features with quantitative chemical data, these algorithms enable the deep mining and accurate prediction of complex biochar-soil interactions.

HSI fuses spectral signatures with spatial distribution, wherein each pixel contains a continuous spectral curve, creating a highly complex "3D data cube". While traditional analyses rely on manually engineered spectral indices, DL architectures autonomously learn the most discriminative spectral-spatial variables. As comprehensively reviewed by Wang et al. (2021), CNNs and their variants can simultaneously extract spectral features and spatial textures, drastically improving the predictive accuracy for SOC and heavy metal contents. Providing demonstrative applications, Weng et al. (2025) and Yao et al. (2025) coupled HSI with DNN to achieve the nondestructive detection of heavy metal contaminants, demonstrating the feasibility of direct end-to-end modeling from raw images to toxicity levels. Weng et al. (2025) and Yao et al. (2025) coupled hyperspectral imaging (HSI) with deep neural networks (DNN) for the non-destructive detection of heavy metal contaminants, demonstrating the feasibility of direct end-to-end modeling from spectral images to toxicity levels. In contrast to this macroscopic and rapid-screening approach, SEM/TEM-EDS provides high-resolution imagery with corresponding elemental distribution maps, offering pixel-scale characterization of biochar–soil–microbe–plant interfaces. Although AI-driven analysis of such microscopic imaging data in acidic soil-biochar systems remains limited (Yang et al. 2023; Xiang et al. 2024; Zhang et al. 2025f), pioneering studies such as Liu et al. (2025a) have begun to apply convolutional neural networks (CNNs) to high-resolution soil images for assessing hydrological properties. This emerging line of research suggests that, given the availability of sufficiently annotated image datasets, deep learning architectures could be adapted to extract features from SEM/TEM-EDS micrographs, enabling quantitative linkage between interfacial elemental distributions and biochar performance—thereby supporting a more data-driven understanding of ameliorative mechanisms at the micro-scale.

Crucially for acidic soils, planar optode technology visualizes two-dimensional pH and O_2 dynamics at the root–soil interface at submillimeter resolutions. Features extracted from these images—such as localized pH gradients, heterogeneity patch areas, hotspot distributions, and spatial statistics (e.g., Moran's I, semivariance)—serve as novel inputs for ML models predicting root health, microbial vigor, or the mitigation of aluminum toxicity. Highlighting the power of feature fusion, Li et al. (2024d) substantially boosted soil carbon prediction accuracy via the weighted fusion of multi-scale spatial and spectral features. Similarly, Zhao et al. (2026b) demonstrated that fusing spectral and textural metrics vastly outperformed isolated spectral features in complex material analyses. Ultimately, integrating spectral fingerprints (chemical composition) with morphological textures (e.g., biochar pore structure) facilitates seamless cross-scale predictions from microscopic images to macroscopic soil functions.

Executing this imaging-chemical fusion requires rigorous modeling strategies: feature-level fusion (concatenating image-extracted features with chemical vectors), decision-level fusion (weighted ensembling of separately trained models), and end-to-end fusion (deep feature integration within multi-input neural networks). Validating this, Zou et al. (2025) synthesized multi-source data and HSI to construct comprehensive evaluation indices, underscoring the necessity of multi-modal fusion for complex biological systems. As imaging resolution improves and DL architectures evolve, fusing multi-modal imaging data promises the creation of "digital twins" spanning from the nanoscale to the field level, enabling the comprehensive monitoring and precise forecasting of biochar-mediated acidic soil amelioration.

3.6.4 Multi-model ensemble: mitigating algorithmic preferences and structural biases

Individual ML models inherently harbor algorithmic biases, meaning that predictions generated by different algorithms on the identical dataset can diverge substantially. Multi-model ensemble strategies mathematically integrate the outputs of numerous independent models, effectively neutralizing model-specific errors and enhancing predictive robustness (Jose et al. 2022). While such ensemble strategies have shown preliminary promise in several effect endpoints relevant to acidic soil amelioration—such as general SOC prediction and labile heavy

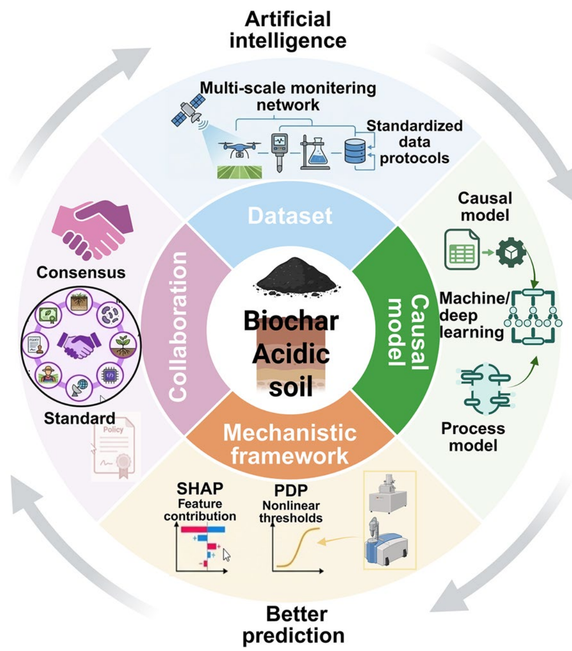


Fig. 9 Future perspectives for enhancing AI-based predictions of biochar efficacy in acidic soil amendment

metal immobilization (Shi et al. 2018a, 2018b; Xu et al. 2025; Shi et al. 2019a; Sun et al. 2026; Bai et al. 2026b; Khan et al. 2026; Wu et al. 2023; Zhang et al. 2024b)—these applications serve primarily as methodological templates for ensemble strategies, rather than as validated implementations targeting the core mechanisms of biochar-amended acidic soils, such as increasing soil pHBC and reducing active Al^{3+} concentration. Nevertheless, drawing on insights from broader environmental AI research, a rationally designed ensemble framework could, in principle, comprise base learners with complementary strengths: RF for capturing complex nonlinearities, XGBoost for handling sparse environmental covariates, and CatBoost for automatically processing categorical variables such as biochar feedstock types (Zhu et al. 2019; Xiang et al. 2025). The outputs of these learners could then be synthesized via weighted averaging or Stacking to potentially yield predictions more stable than those from any single algorithm. Importantly, we acknowledge that none of the studies reviewed have experimentally validated these ensemble architectures specifically in biochar-amended acidic soil systems. We therefore regard the application of ensemble strategies to this particular pedological context as a methodological hypothesis and an emergent research frontier, rather than an established finding. Currently, the primary challenge impeding robust implementation is the absence of standardized criteria for base-learner selection and meta-learner combination, making it difficult to guarantee genuine complementarity among sub-models. To bridge this gap, we suggest that future work prioritize the establishment of open-source benchmarking platforms and shared experimental datasets tailored to acidic soil conditions, which would enable systematic stress-testing of various ensemble configurations and guide rational algorithm selection for this specific agricultural challenge.

4 Future perspectives and recommendations

4.1 Enhancing high-quality training datasets

The predictive capacity of AI models in biochar-mediated acidic soil amelioration is governed primarily by the representativeness, richness, and overall quality of available training datasets. Pioneering studies, exemplified by Bao et al. (2023) and Dehkordi et al. (2020), have established promising paradigms for multi-source data fusion and long-term monitoring. Yet, despite these advances, significant heterogeneity persists across studies in experimental design, data resolution, and metadata reporting. Compounding this challenge, many researchers assume that analytical workflows—covering data preprocessing, feature engineering, model calibration, and validation—are inherently consistent and, as such, often omit explicit documentation of these procedures, regarding them as self-evident. In reality, inconsistencies in workflow implementation or inadequate documentation can substantially undermine the scientific rigor and defensibility of the resulting model predictions. Addressing these intertwined issues, future research must prioritize the establishment of standardized, multi-scale monitoring networks that seamlessly integrate remote sensing, proximal sensing, and laboratory analyses to capture dynamic soil system responses at relevant spatial and temporal scales. Equally important is the adoption of harmonized data-sharing protocols that conform to the FAIR principles—Findable, Accessible, Interoperable, and Reusable—to ensure comparability across studies and facilitate the creation of unified benchmarking datasets. Such open and standardized data ecosystems will not only enhance reproducibility and cross-study synthesis but also provide a robust foundation for training next-generation AI models capable of accurately predicting biochar performance and long-term soil recovery trajectories (Fig. 9).

4.2 Advancing the deep integration of AI with causal inference models

While machine learning has proven highly effective at identifying correlations, it remains limited in revealing the causal mechanisms underpinning biochar-mediated amelioration of acidic soils. To progress beyond correlation-based inference, future studies should focus on integrating causal reasoning and mechanistic understanding into data-driven frameworks. Zhang et al. (2024a) and Chang et al. (2024) demonstrated that embedding process-based model variables as dynamic covariates within RF architectures can substantially elevate predictive accuracy—improving R^2 by 59–80% compared with standalone process-based models—by capturing the temporal evolution of SOC. Building upon such advances, next-generation frameworks should emphasize knowledge-guided machine learning (KGML) and physics-informed AI, wherein chemical and biological laws explicitly constrain model structure and loss functions. Incorporating structural causal models, causal forests, or Bayesian networks can further distinguish genuine causal relationships from spurious correlations. Concurrently, coupling process-based submodels (e.g., pH buffering kinetics and aluminum speciation thermodynamics) with machine learning frameworks will strengthen mechanistic consistency. In addition, integrating active learning with Bayesian optimization can enable the design of computational "counterfactual" experiments to strategically explore parameter spaces, supporting the targeted design of biochar variants and their precision application across heterogeneous acidic landscapes.

4.3 Strengthening model interpretability and coupling with mechanistic frameworks

To pierce the "black-box" opacity that currently inhibits trust in AI-derived biochar strategies, multi-level XAI pipelines must be rigorously adopted. The true value of interpretability tools extends beyond calculating variable weights; it lies in anchoring statistical outputs to established pedological and chemical frameworks. To operationalize this, future studies should focus on: (1) Constructing "feature–mechanism" mapping dictionaries: For instance, explicitly comparing the nonlinear inflection points captured by PDP with the critical thermodynamic thresholds for mechanism switching (e.g., the transition from purely electrostatic adsorption to surface precipitation); (2) Dimensionality expansion via spatial feature engineering: Fusing high-resolution microspectroscopic outputs (e.g., extended X-ray absorption fine structure (EXAFS), SEM–EDS elemental mapping) with macroscopic data to visually and computationally validate mechanism dominance at the pore scale; (3) Establishing multi-tiered interpretability: Developing holistic XAI toolchains that span global feature importance (identifying universal drivers), local dependencies (evaluating interactions), and individual sample explanations (diagnosing specific field anomalies).

4.4 Fostering interdisciplinary collaborative research paradigms

Advancing the understanding and application of biochar-mediated acidic soil amelioration demands close collaboration among experts across multiple disciplines, including soil chemistry, microbial ecology, agronomy, hydrology, climatology, and computer science. Such cross-domain integration is essential to address core AI challenges—overfitting, limited generalizability, and a persistent lack of interpretability—while ensuring that model outputs remain mechanistically meaningful. Soil chemists are needed to elucidate processes such as pH buffering and aluminum speciation dynamics; microbial ecologists must interpret functional community data; agronomists should validate crop yield responses; and computer scientists can optimize algorithm architectures and uncertainty estimates. Equally important is engagement with practitioners and stakeholders—including farmers, policymakers, and carbon market regulators—to align acceptable uncertainty thresholds, interpretability expectations, and model performance criteria with practical decision-making needs. Through coordinated construction of standardized benchmark datasets, unified evaluation metrics, and physically constrained predictive models, such interdisciplinary cooperation will build transparency and trust in AI-assisted environmental management. This collaborative paradigm will accelerate the deployment of precision biochar applications, strengthen integration into carbon-crediting frameworks, and support global climate-smart agriculture initiatives, ensuring that AI advances contribute directly to sustainable soil management and carbon sequestration goals (Fig. 9).

4.5 Operational guidance for mechanism-guided intelligent prediction

To translate this framework into practice, we propose the following operational priorities: (i) Core input features for model development. Essential features should include biochar pyrolysis temperature, feedstock type, pH, total alkalinity, carbonate content (via acid digestion or XRD), carboxyl group density (via Boehm titration), and soil initial pH, CEC, clay content, and organic matter—each directly corresponding to the buffering mechanisms elaborated in Sect. 2.2.1. KGML has been shown to significantly outperform pure data-driven approaches when training data are limited (Liu et al. 2024). (ii) Standardized reporting variables. Beyond the above, reporting should cover application rate, pyrolysis duration, experimental duration, climate zone, and crop type to enable cross-study synthesis and meta-analysis (Wang et al. 2025; Storey et al. 2025). (iii) Experimental metadata for model transferability. Full documentation of pyrolysis conditions, soil sampling depth, and analytical methods is required to support model extrapolation—a critical need given that purely data-driven models often fail when applied outside their training domain (Jin et al. 2026). (iv) Validation designs. Model validation should incorporate independent datasets from geographically distinct regions to assess spatial generalizability (Oneto And Chicco 2025), complemented by field-scale trials with extended monitoring durations to capture temporal dynamics (Storey et al. 2025). (v) Quantitative performance benchmarks. Rather than prescribing fixed thresholds—which remain context-dependent and lack consensus in the literature—we recommend that future studies establish domain-specific benchmarks through systematic intercomparison across multiple datasets and modeling frameworks. Provisional targets may be defined relative to baseline models (e.g., total-alkalinity-only benchmarks) rather than as absolute values. (vi) Protocols for addressing the alkalinity gap. Routine separation of total alkalinity into organic and inorganic fractions should be adopted using established methods (He et al. 2022b), enabling component-specific feature engineering as discussed in Sect. 2.3. (vii) Integration of mechanism knowledge with ML architectures. This can be achieved through multiple KGML strategies: knowledge-guided architectures that embed mechanistic constraints into network structures, pre-training with process-based model simulations, knowledge-guided loss functions that penalize outputs violating biogeochemical principles, and knowledge-guided inference that incorporates mechanistic priors into parameter estimation (Jin et al. 2026).

5 Conclusions

The efficacy of biochar in ameliorating acidic soils fundamentally depends on complex biogeochemical mechanisms. However, current research struggles to decouple the distinct contributions of organic and inorganic alkalinity—a critical gap that constrains mechanism-based forecasting. Concurrently, existing AI applications face critical bottlenecks, including heavy reliance on data quality, insufficient interpretability, and limited dynamic predictive capacity. To overcome these barriers, multi-source data fusion—integrating remote and proximal sensing, multi-modal microbial data, imaging-chemical features, and multi-model ensembling—emerges as a pivotal pathway to unlock AI's full predictive potential. Nevertheless, this review acknowledges several limitations. Although we endeavored to cover key methodological advances, this review was not conducted under a formal systematic framework such as PRISMA, and literature retrieval was primarily based on targeted keyword searches in Web of Science Core Collection; relevant studies using alternative terminology or published in other databases may have been omitted. The comparison of model performance was based on published case studies that vary considerably in geographical scale, data quality, and evaluation metrics, limiting direct comparability. Furthermore, the fusion strategies and operational protocols discussed here remain largely conceptual; empirical testing and benchmarking are required to validate their practical effectiveness across diverse soil systems. Moving forward, future research must prioritize establishing unified, high-quality training datasets. By advancing the deep integration of AI with causal inference models, enhancing mechanistic interpretability, and fostering interdisciplinary collaboration, AI can evolve from a correlational tool into a transformative engine for the precise, mechanism-driven management of global acidified ecosystems.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s42773-026-00652-6.

Supplementary Material 1.

Acknowledgements

This study was financially supported by Key Science and Technology Projects of Jiangxi Province (Grant No. JXNK202607-02).

Author contributions

All authors contributed to the study conception and design. Conceptualization, literature search and data collection, and critical revision were performed by Linyu Guo, Kewei Li, and Ren-kou Xu. The first draft of the manuscript was written by Linyu Guo, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This study was funded by Key Science and Technology Projects of Jiangxi Province (Grant No. JXNK202607-02).

Data availability

The datasets used or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Competing interests

Renkou Xu is an EBM of the journal Biochar, and he was not involved in the peer-review or handling of the manuscript. The authors have no relevant financial or non-financial interests to disclose.

Author details

[1]State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 211135, China. [2]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing 100049, China.

Received: 22 April 2026 Revised: 6 July 2026 Accepted: 20 July 2026
Published online: 28 August 2026

References

Alekseeva T, Alekseev A, Xu RK, Zhao AZ, Kalinin P (2011) Effect of soil acidification induced by a tea plantation on chemical and mineralogical properties of Alfisols in eastern China. Environ Geochem Health 33(2):137–148

Ali A, Jabeen N, Chachar Z, Chachar S, Ahmed S, Ahmed N, Laghari AA, Sahito ZA, Rasulov R, Yang ZP (2025) The role of biochar in enhancing soil health & interactions with rhizosphere properties and enzyme activities in organic fertilizer substitution. Front Plant Sci 16:1595208

Alling V, Hale SE, Martinsen V, Mulder J, Smebye A, Breedveld GD, Cornelissen G (2014) The role of biochar in retaining nutrients in amended tropical soils. J Plant Nutr Soil Sci 177:671–680

Almajed A, Tirkolaei HK, Kavazanjian E Jr, Hamdan N (2019) Enzyme induced biocementated sand with high strength at low carbonate content. Sci Rep 9:1135

Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, Van Essen BC, Awwal AAS, Asari VK (2019) A state-of-the-art survey on deep learning theory and architectures. Electronics 8(3):292

Ameloot N, Sleutel S, Case SDC, Alberti G, McNamara NP, Zavalloni C, Vervisch B, delle Vedove G, De Neve S (2014) C mineralization and microbial activity in four biochar field experiments several years after incorporation. Soil Biol Biochem 78:195–203

An X, Zhu Z, Luo X, Chen C, Liu T, Zou L, Li S, Liu Y (2025) Effects of raw materials and pyrolysis temperatures on physicochemical properties of biochars derived from hemp stalks. Plants (Basel) 14(16):2564

Aurangzeib M, Zhang S, Yan S, Wei Z, Zhang H, Niu X, Yan P (2025) Does biochar improve acidic soil physical properties and crop yield under varying climatic and soil conditions? A global comprehensive meta-analysis. Soil Use Manage 41(2):e70081

Azizi K, Zhao S, Ayoubi S, Kamangarpour A, Nabiollahi K, Ge Y, Zhao K, Rodrigo-Comino J, Dematte JAM (2026) Integrating proximal sensing data (soil spectra and magnetic susceptibility) with the common covariates to predict soil carbon pools in semiarid regions. Soil Tillage Res 256:106840

Bai DF, Fang O, Li CH, Cai B, Tan XY, Jiang MM, Gan B, Fu JX, Gao YY, Wang Y, Liu YX (2026a) Accu16S/AccuITS: accurate and broadly applicable amplicon sequencing for absolute microbiome quantification. IMETA 5:e70116

Bai JZ, Zhang FQ, Haider S, Song JJ, Feng YZ, De Almeida Moreira BR (2026b) Biochar's impact on carbon sequestration and yield-scaled global warming potential in China: a machine learning-enhanced meta-analysis. J Clean Prod 540:147459

Bai XF, Zhou XQ, Li ZF, Ni JW, Bai X (2017) Properties and applications of biochars derived from different biomass feedstock sources. Int J Agric Biol Eng 10(2):242–250

Bao Y, Meng X, Liu H, Xu M, Wang M (2025) A novel method for soil organic carbon prediction using integrated `ground-air-space' multimodal remote sensing data. Geoderma 460:117453

Bao Y, Yao F, Meng X, Zhang J, Liu H, Mouazen AM (2023) Predicting soil organic carbon in cultivated land across geographical and spatial scales: integrating Sentinel-2A and laboratory Vis-NIR spectra. ISPRS J Photogramm Remote Sens 203:1–18

Berek AK, Hue NV (2016) Characterization of biochars and their use as an amendment to acid soils. Soil Sci 181(9–10):412–426

Beule L, Lehtsaar E, Rathgeb A, Karlovsky P (2021) Relative abundances of species or sequence variants can be misleading: soil fungal communities as an example. Microorganisms 9(3):589

Blossfeld S, Schreiber CM, Liebsch G, Kuhn AJ, Hinsinger P (2013) Quantitative imaging of rhizosphere pH and CO_2 dynamics with planar optodes. Ann Bot 112(2):267–276

Brewer CE, Unger R, Schmidt-Rohr K, Brown RC (2011) Criteria to select biochars for field studies based on biochar chemical properties. Bioenergy Res 4(4):312–323

Butterly CR, Baldock JA, Tang C (2013) The contribution of crop residues to changes in soil pH under field conditions. Plant Soil 366(1–2):185–198

Case SDC, Oelofse M, Hou Y, Oenema O, Jensen LS (2017) Farmer perceptions and use of organic waste products as fertilisers – a survey study of potential benefits and barriers. Agric Syst 151:84–95

Chakraborty S, Weindorf DC, Michaelson GJ, Ping CL, Choudhury A, Kandakji T, Acree A, Sharma A, Dandan W (2016) In-situ differentiation of acidic and non-acidic tundra via portable X-ray fluorescence (PXRF) spectrometry. Pedosphere 26(4):549–560

Chang NJ, Chen D, Cai YR, Li JZ, Zhang MX, Li H, Wang LG (2024) Enhancing crop yield and carbon sequestration and greenhouse gas emission mitigation through different organic matter input in the Bohai Rim region: an estimation based on the DNDC-RF framework. Field Crop Res 319:109624
Chen JF, Zhang W, Peng Y, Wang JL, Wu M (2025) Machine learning prediction of the heavy metal concentration in soil. J Soils Sediments 25(11):3369–3380
Chen SS, Teng Y, Luo YM, Kuramae E, Ren WJ (2024) Threats to the soil microbiome from nanomaterials: a global meta and machine-learning analysis. Soil Biol Biochem 188:109248
Cheng CH, Lehmann J, Engelhard MH (2008) Natural oxidation of black carbon in soils: changes in molecular form and surface charge along a climosequence. Geochim Cosmochim Acta 72(6):1598–1610
Dai ZM, Xiong XQ, Zhu H, Xu HJ, Leng P, Li JH, Tang C, Xu JM (2021) Association of biochar properties with changes in soil bacterial, fungal and fauna communities and nutrient cycling processes. Biochar 3(3):239–254
Dai ZM, Zhang XJ, Tang C, Muhammad N, Wu JJ, Brookes PC, Xu JM (2017) Potential role of biochars in decreasing soil acidification – a critical review. Sci Total Environ 581:601–611
Dashti A, Raji M, Harami HR, Zhou JL, Asghari M (2023) Biochar performance evaluation for heavy metals removal from industrial wastewater based on machine learning: application for environmental protection. Sep Purif Technol 312:123399
Dehkordi RH, Burgeon V, Fouche J, Gomez EP, Cornelis JT, Nguyen F, Denis A, Meersmans J (2020) Using UAV collected RGB and multispectral images to evaluate winter wheat performance across a site characterized by century-old biochar patches in Belgium. Remote Sens 12(15):2504
Deng HY, Ma XZ, Liu ZK, Hu HW, Di HJ, Liu YJ, Shi SJ, Hao XY, Zhao Y, He JZ, Shen JP (2024) Soil ecosystem multifunctionality is strongly linked with crop yield after four decades chemical fertilization in black soil. Agr Ecosyst Environ 368:109007
Deng R, Lan ZY, Xiao KY, Shang XL, Fang SZ (2025) Biochar additions improve soil nutrient availability and biomass production of *Cyclocarya paliurus* via regulating soil bacterial community and pH value on the acidic soil. J Soil Sci Plant Nutr 25(3):5751–5765
Deng YY, Yang Y, Tu L, Liao YF, Li AY, Zhai HQ, Deng H, Li Y, Hu LN (2026) Coupled regulation mechanisms of carbon sequestration, microbial communities, and manganese stabilization in manganese-contaminated soil by phosphorus-modified biochar. Appl Soil Ecol 217:106616
Ding ZJ, Liu K, Grunwald S, Smith P, Ciais P, Wang B, Wadoux AMJC, Ferreira C, Karunaratne S, Shurpali N, Yin XG, Roberts D, Madgett O, Duncan S, Zhou MX, Liu ZY, Harrison MT (2025) Advancing soil organic carbon prediction: a comprehensive review of technologies, AI, process-based and hybrid modelling approaches. Adv Sci 12(31):e04152
Dong YC, Tunali M, Nowack B (2026) Fertilizer or pollutant: analyzing the effects of biochar on soil organisms using machine learning. Biochar 8(1):28
Du ZL, Sun X, Zheng SN, Wang SY, Wu LN, An Y, Luo YM (2024) Optimal biochar selection for cadmium pollution remediation in Chinese agricultural soils via optimized machine learning. J Hazard Mater 476:135065
Duan QN, Yan PW, Feng YC, Wan QR, Zhu XL (2024) Machine learning assisted adsorption performance evaluation of biochar on heavy metal. Front Environ Sci Eng 18(5):55
EBC (2022) European Biochar Certificate - Guidelines for a Sustainable Production of Biochar. Version 10.1. European Biochar Foundation, Arbaz, Switzerland. Available at: http://european-biochar.org (Accessed: date).
Flemming HC, Wingender J (2010) The biofilm matrix. Nat Rev Microbiol 8(9):623–633
Greenberg I, Seidel M, Vohland M, Koch HJ, Ludwig B (2022a) Performance of in situ vs laboratory mid-infrared soil spectroscopy using local and regional calibration strategies. Geoderma 409:115614
Greenberg I, Seidel M, Vohland M, Ludwig B (2022b) Performance of field-scale lab vs in situ visible/near- and mid-infrared spectroscopy for estimation of soil properties. Eur J Soil Sci 73(1):e13180
Grunwald S (2022) Artificial intelligence and soil carbon modeling demystified: power, potentials, and perils. Carbon Footprints 1:5
Guo GM, Lin LY, Jin FM, Masek O, Huang Q (2023) Application of heavy metal immobilization in soil by biochar using machine learning. Environ Res 231(2):116098
Guo JH, Liu XJ, Zhang Y, Shen JL, Han WX, Zhang WF, Christie P, Goulding KWT, Vitousek PM, Zhang FS (2010) Significant acidification in major Chinese croplands. Science 327(5968):1008–1010
Guo LY, Yan J, Shi YXX, Li KW, Guan P, Xu RK (2025) Active aluminum promoted copper uptake by Chinese cabbage grown in an acidic Cu-contaminated soil: a new insight with the diffusive gradients in thin-films technique (DGT). Plant Soil 513(1–2):473–487
Hansen M, Bang-Andreasen T, Sorensen H, Ingerslev M (2017) Micro vertical changes in soil pH and base cations over time after application of wood ash on forest soil. For Ecol Manage 406:274–280
Hamrani A, Akbarzadeh A, Madramootoo CA (2020) Machine learning for predicting greenhouse gas emissions from agricultural soils. Sci Total Environ 741:140338
He DB, Dong ZX, Zhu B (2024) An effective global biochar application strategy for reducing global cropland nitrogen emissions without compromising crop yield: findings from a global meta-analysis and density functional theory calculation. Environ Technol Innov 33:103546
He MJ, Xiong XN, Wang L, Hou DY, Bolan NS, Ok YS, Rinklebe J, Tsang DCW (2021) A critical review on performance indicators for evaluating soil biota and soil health of biochar-amended soils. J Hazard Mater 414:125378
He X, Hong ZN, Shi RY, Cui JQ, Lai HW, Lu HL, Xu RK (2022a) The effects of H_2O_2- and HNO_3/H_2SO_4-modified biochars on the resistance of acid paddy soil to acidification. Environ Pollut 293:118588
He X, Jiang J, Hong ZN, Pan XY, Dong Y, Xu RK (2020) Effect of aluminum modification of rice straw-based biochar on arsenate adsorption. J Soils Sediments 20(8):3073–3082
He X, Lu HL, Wu CM, Xu RK (2022b) Effects of inorganic alkalis and organic anions in biochars on acidic paddy soil resistance to acidification. J Soils Sediments 22(4):1201–1213
Hossain MZ, Bahar MM, Sarkar B, Donne SW, Ok YS, Palansooriya KN, Kirkham MB, Chowdhury S, Bolan N (2020) Biochar and its importance on nutrient dynamics in soil and plant. Biochar 2(4):379–420
Hu X, Liang CG, Sun HT, Liu X, Chen YC, Han C, Fang Z, Liu ZD, Zhang H, Jones KC, Liu C, Garcia A, Luo J (2025) A ratiometric fluorescent sensing foil for high resolution 2D pH measurement based on a novel hydroxy-pyrene green fluorophore. Talanta 284:127230
Hu ZK, Delgado-Baquerizo M, Fanin N, Chen XY, Zhou Y, Du GZ, Hu F, Jiang L, Hu SJ, Liu MQ (2024) Nutrient-induced acidification modulates soil biodiversity-function relationships. Nat Commun 15(1):2858
Huff MD, Lee JW (2016) Biochar-surface oxygenation with hydrogen peroxide. J Environ Manage 165:17–21
Ichsan C, Nisa K, Ramadhan N (2026) Comparing ACES AP0 and AP1 color primaries for smartphone colorimetry in detecting soil CO_2 leakage with a beetroot (*Beta vulgaris*) biosensor. Anal Bioanal Chem 418(5):1417–1429
International Biochar Initiative (IBI). (2015). Standardized Product Definition and Product Testing Guidelines for Biochar That Is Used in Soil (v2.1). IBI.
Jeffery S, Verheijen FGA, van der Velde M, Bastos AC (2011) A quantitative review of the effects of biochar application to soils on crop productivity using meta-analysis. Agr Ecosyst Environ 144(1):175–187
Jian C, Luukkonen P, Yki-Järvinen H, Salonen A, Korpela K (2020) Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. PLoS ONE 15(1):e0227285
Jiang J, Yuan M, Xu RK, Bish DL (2015) Mobilization of phosphate in variable-charge soils amended with biochars derived from crop straws. Soil till Res 146:139–147
Jiang LH, Liu XD, Yin HQ, Liang YL, Liu HW, Miao B, Peng QQ, Meng DL, Wang SQ, Yang JJ, Guo ZW (2020) The utilization of biomineralization technique based on microbial induced phosphate precipitation in remediation of potentially toxic ions contaminated soil: a mini review. Ecotox Environ Safe 191:110009
Jiang ZH, Xia Q, Lu XY, Yue WJ, Chen AH, Liu XG, Chen JH, Zhao CX (2025) Machine learning prediction of biochar-specific surface area based on plant characterization information. Renew Energy 243:122633
Jiao WQ, Li KC, Zhou M, Zhou NA, Chen QS, Hu T, Qi CC (2025) Determining whether biochar can effectively increase crop yields: a machine learning model development with imbalanced data. Environ Technol Inno 38:104154
Jin ZN, Liu LC, Yang Q, Jia XW, Tao SL, Guo YK, Ghosh R, Wang S, Zhu Q, Jung M, Guan KY, Kumar V, Reichstein M, Fang JY, Luo YQ (2026)

Knowledge-guided machine learning for global change ecology research. Global Change Biol 32(2):1–12
Jose DM, Vincent AM, Dwarakish GS (2022) Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. Sci Rep 12:4678
Kahlert H, Steinhardt T, Behnert J, Scholz F (2004) A new calibration free pH-probe for in situ measurements of soil pH. Electroanalysis 16(24):2058–2064
Kayoumu M, Wang HL, Duan GL (2025) Interactions between microbial extracellular polymeric substances and biochar, and their potential applications: a review. Biochar 7(1):62
Ke B, Nguyen H, Bui XN, Bui HB, Nguyen-Thoi T (2021) Prediction of the sorption efficiency of heavy metal onto biochar using a robust combination of fuzzy C-means clustering and back-propagation neural network. J Environ Manage 293:112808
Ke XL, Wang Y, Liu ML, Yun ZZ, Bian RJ, Cheng K, Liu XY, Zheng JF, Zhang XH, Li LQ, Pan GX (2022) Screening major properties of biochar affecting acid soil amelioration based on pot experiments and random forest model. J Soil Sci Plant Nut 22(4):4103–4115
Keiluweit M, Nico PS, Johnson MG, Kleber M (2010) Dynamic molecular structure of plant biomass-derived black carbon (biochar). Environ Sci Technol 44(4):1247–1253
Khan MN, Han TF, Huang J, Shah A, Hao XY, Núñez-Delgado A, Sial TA, Liu WJ, Singh AK, Ma XZ, Zhang HM (2026) Relationships between soil phosphorus fractions and greenhouse gas emissions in degraded red and black soils amended with magnesium treated biochar. GCB Bioenergy 18(2):e70102
Koop-Jakobsen K, Meier RJ, Mueller P (2021) Plant-mediated rhizosphere oxygenation in the native invasive salt marsh grass Elymus athericus. Front Plant Sci 12:669751
LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
Lee S, Sorensen JW, Walker RL, Emerson JB, Nicol GW, Hazard C (2022) Soil pH influences the structure of virus communities at local and global scales. Soil Biol Biochem 166:108569
Lehmann J, Rillig MC, Thies J, Masiello CA, Hockaday WC, Crowley D (2011) Biochar effects on soil biota – a review. Soil Biol Biochem 43(9):1812–1836
Lei CT, Lu T, Qian HF, Liu YX (2023) Machine learning models reveal how biochar amendment affects soil microbial communities. Biochar 5(1):89
Leng LJ, Yang LH, Lei XN, Zhang WJ, Ai ZJ, Yang ZQ, Zhan H, Yang JP, Yuan XZ, Peng HY, Li HL (2022) Machine learning predicting and engineering the yield, N content, and specific surface area of biochar derived from pyrolysis of biomass. Biochar 4(1):63
Li YT, Hong HY (2023) Modelling flood susceptibility based on deep learning coupling with ensemble learning models. J Environ Manage 325:116450
Li BZ, Guo YL, Liang F, Liu W, Wang YJ, Cao WC, Song H, Chen JS, Guo JH (2024a) Global integrative meta-analysis of the responses in soil organic carbon stock to biochar amendment. J Environ Manage. https://doi.org/10.1016/j.jenvman.2023.119745
Li JW, Peduruhewa JH, Brown RW, Chadwick DR, Griffiths RI, Fu HR, Bian HF, Sheng LX, Ma QX, Jones DL (2026a) Biochar reshapes soil bacterial community composition and survival strategies: a meta-analysis revealing trade-offs between microbial stability and functional complexity. Biol Fertil Soils 62:251–266
Li JY, Wang N, Xu RK, Tiwari D (2010) Potential of industrial byproducts in ameliorating acidity and aluminum toxicity of soils under tea plantation. Pedosphere 20(5):645–654
Li KW, Xu RK, Cai ZJ (2024b) Aluminum mobilization characteristics in four typical soils from different climate zones during their acidification. Plant Soil 504(1–2):29–46
Li L, Yang JK, Yu YL, Shakoor A, Virk AL, Li FM, Yang HS, Kan ZR (2025) Crop straw converted to biochar increases soil organic carbon but reduces available carbon. Eur J Agron 164:127499
Li MD, Tao XQ, Lang CP, Zhang JW, Xu GZ, Zhu LP, Yin J (2024c) Enzyme-induced carbonate precipitation using yellow soybean for eco-friendly treatment of expansive soils. Int J Geomech 24(3):04024007
Li XY, Li ZM, Qiu HM, Chen GY, Fan PP, Liu Y (2024d) Multi-scale spatial and spectral feature fusion for soil carbon content prediction based on hyperspectral images. Ecol Indic 160:111843
Li X, Wu D, Liu X, Huang YP, Cai AD, Xu H, Ran JW, Xiao J, Zhang WJ (2024e) A global dataset of biochar application effects on crop yield, soil properties, and greenhouse gas emissions. Sci Data 11:57
Li YZ, Gupta R, You SM (2022) Machine learning assisted prediction of biochar yield and composition via pyrolysis of biomass. Bioresour Technol 359:127511
Li ZT, Pan Y, Meng J, Zhao MQ, Yao ZY, Hu MJ, Shan SD, Chen HH (2026b) Plant-derived biochars ameliorate acidity and reshape microbial communities to enhance tea productivity and quality in acidic soils. Appl Soil Ecol 219:106805
Liao HK, Li YY, Yao HY (2019) Biochar amendment stimulates utilization of plant-derived carbon by soil bacteria in an intercropping system. Front Microbiol 10:1361
Liao W, Zhang X, Yang RC, Yang HP, Wang J, Ding HG, Zhang SH, Chen HP, Jiang JC (2025) Machine learning prediction of biochar structure stability and adsorption efficiency based on biomass characteristics and aging factors. J Energy Inst 123:102245
Lin YX, Ding WX, Liu DY, He TH, Yoo G, Yuan JJ, Chen ZM, Fan JL (2017) Wheat straw-derived biochar amendment stimulated N_2O emissions from rice paddy soils by regulating the *amoA* genes of ammonia-oxidizing bacteria. Soil Biol Biochem 113:89–98
Liu GS, Tian SK, Wang Q, Wang HZ, Kong LW (2025a) High-resolution measurement of moisture filed at soil surface with interfered image processing method and machine learning techniques. J Hydrol 652:132623
Liu LC, Zhou W, Guan KY, Peng B, Xu SM, Tang JY, Zhu Q, Till J, Jia XW, Jiang CY, Wang S, Qin ZQ, Kong H, Grant R, Mezbahuddin S, Kumar V, Jin ZN (2024) Knowledge-guided machine learning can improve carbon cycle quantification in agroecosystems. Nat Commun 15:357
Liu Q, Liu BJ, Zhang YH, Hu TL, Lin ZB, Liu G, Wang XJ, Ma J, Wang H, Jin HY, Ambus P, Amonette JE, Xie ZB (2019) Biochar application as a tool to decrease soil nitrogen losses (NH_3 volatilization, N_2O emissions, and N leaching) from croplands: options and mitigation strength in a global perspective. Glob Change Biol 25(6):2077–2093
Liu SW, Cen BT, Yu ZN, Qiu RL, Gao T, Long XX (2025b) The key role of biochar in amending acidic soil: reducing soil acidity and improving soil acid buffering capacity. Biochar 7(1):52
Liu WJ, Jiang H, Yu HQ (2015) Development of biochar-based functional materials: toward a sustainable platform carbon material. Chem Rev 115(22):12251–12285
Liu YG, Shi MX, Chen ZH, Tang ZJ, Tan JY, Tian JH, Elsgaard L, Lu Y, Li B (2023) Nonlinear effects of biochar on greenhouse gas emissions and soil microbial community composition under intensive vegetable production in South China. Appl Soil Ecol 189:104975
Long JM, Liu FS, Ouyang LZ, Chen CL, Han N, Tan YT, Xiao Y, Zhou DS (2024) Biochar changes iron-reducing bacteria community in paddy soils and promotes the bacterial iron reduction. J Soils Sediments 25(1):3656–3667
Lu HN, Li P, Huang YJ, Liu F, Gao BS, Yang J, Liang YH, Feng ZJ (2025) Predicting the Cu adsorption performance of original biochar in aquatic environments via interpretable machine learning models. J Water Process Eng 75:108079
Lyu H, Xu ZM, Zhong J, Gao WH, Liu JX, Duan M (2024) Machine learning-driven prediction of phosphorus adsorption capacity of biochar: insights for adsorbent design and process optimization. J Environ Manage 369:122405
Maisch M, Lueder U, Kappler A, Schmidt C (2019) Iron lung: how rice roots induce iron redox changes in the rhizosphere and create niches for microaerophilic Fe(II)-oxidizing bacteria. Environ Sci Technol Lett 6(10):600–605
Martinez BC, DeJong JT, Ginn TR, Montoya BM, Barkouki TH, Hunt C, Tanyu B, Major D (2013) Experimental optimization of microbial-induced carbonate precipitation for soil improvement. J Geotech Geoenviron Eng 139(4):587–598
Merl T, Rasmussen MR, Koch LR, Sondergaard JV, Bust FF, Koren K (2022) Measuring soil pH at in situ like conditions using optical pH sensors (pH-optodes). Soil Biol Biochem 175:108862
Mukherjee A, Zimmerman AR, Harris W (2011) Surface chemistry variations among a series of laboratory-produced biochars. Geoderma 163(3–4):247–255

Munera-Echeverri JL, Martinsen V, Strand LT, Zivanovic V, Cornelissen G, Mulder J (2018) Cation exchange capacity of biochar: an urgent method modification. Sci Total Environ 642:190–197
Nealson KH, Saffarini D (1994) Iron and manganese in anaerobic respiration: environmental significance, physiology, and regulation. Annu Rev Microbiol 48:311–343
Nkoh JN, Hong ZN, Xu RK (2022) Laboratory studies on the effect of adsorbed microbial extracellular polymeric substances on the acidity of selected variable-charge soils. Soil Sci Soc Am J 86(2):162–180
Nkoh JN, Yan J, Hong ZN, Xu RK (2020a) Effects of citrate, oxalate, and phosphate on the sorption of Cr(VI) by extracellular polymeric substances. J Water Process Eng 37:101510
Nkoh JN, Yan J, Xu RK, Shi RY, Hong ZN (2020b) The mechanism for inhibiting acidification of variable charge soils by adhered *Pseudomonas fluorescens*. Environ Pollut 260:114049
Nkoh NJ, Liu ZD, Yan J, Cai SJ, Hong ZN, Xu RK (2020c) The role of extracellular polymeric substances in bacterial adhesion onto variable charge soils. Arch Agron Soil Sci 66(13):1780–1793
Noble AD, Zenneck I, Randall PJ (1996) Leaf litter ash alkalinity and neutralisation of soil acidity. Plant Soil 179(2):293–302
Novielli P, Magarelli M, Romano D, Di Bitonto P, Stellacci AM, Monaco A, Amoroso N, Bellotti R, Tangaro S (2025) Leveraging explainable AI to predict soil respiration sensitivity and its drivers for climate change mitigation. Sci Rep 15(1):12527
Oneto L, Chicco D (2025) Eight quick tips for biologically and medically informed machine learning. PLoS Comput Biol 21(1):e1012711
Padarian J, Minasny B, McBratney AB (2019) Using deep learning to predict soil properties from regional spectral data. Geoderma Reg 16:e00198
Palansooriya KN, Li J, Dissanayake PD, Suvarna M, Li LY, Yuan XZ, Sarkar B, Tsang DCW, Rinklebe J, Wang XN, Ok YS (2022) Prediction of soil heavy metal immobilization by biochar using machine learning. Environ Sci Technol 56(7):4187–4198
Palansooriya KN, Wong JTF, Hashimoto Y, Huang LB, Rinklebe J, Chang SX, Bolan N, Wang HL, Ok YS (2019) Response of microbial communities to biochar-amended soils: a critical review. Biochar 1(1):3–22
Pandey CB, Kumar U, Kaviraj M, Minick KJ, Mishra AK, Singh JS (2020) DNRA: a short-circuit in biological N-cycling to conserve nitrogen in terrestrial. Sci Total Environ 738:139710
Parasar BJ, Agarwala N (2025) Unravelling the role of biochar-microbe-soil tripartite interaction in regulating soil carbon and nitrogen budget: a panacea to soil sustainability. Biochar 7(1):37
Pokharel P, Ma ZL, Chang SX (2020) Biochar increases soil microbial biomass with changes in extra- and intracellular enzyme activities: a global meta-analysis. Biochar 2(1):65–79
Qi JQ, Yuan HY, Sun SC, Zama EF, Tao BX, Liu J, Zhang BH 2025 Biochar amendment alters rare microbial taxa and enhances wheat growth in alkaline farmland: insights into soil microbiome dynamics. Front Microbiol 16
Qian LB, Chen BL (2013) Dual role of biochars as adsorbents for aluminum: the effects of oxygen-containing organic components and the scattering of silicate particles. Environ Sci Technol 47(15):8759–8768
Qian LB, Chen BL, Chen MF (2016) Novel alleviation mechanisms of aluminum phytotoxicity via released biosilicon from rice straw-derived biochars. Sci Rep 6:29346
Qian LB, Chen MF, Chen BL (2015) Competitive adsorption of cadmium and aluminum onto fresh and oxidized biochars during aging processes. J Soils Sediments 15(5):1130–1138
Quan PL, Sauzade M, Brouzes E (2018) dPCR: a technology review. Sensors-Basel 18(4):1271
Quilliam RS, Glanville HC, Wade SC, Jones DL (2013) Life in the 'charosphere' – does biochar in agricultural soil provide a significant habitat for microorganisms? Soil Biol Biochem 65:287–293
Rajapaksha AU, Ahmad M, Vithanage M, Kim KR, Chang JY, Lee SS, Ok YS (2015) The role of biochar, natural iron oxides, and nanomaterials as soil amendments for immobilizing metals in shooting range soil. Environ Geochem Health 37(6):931–942
Reddy KJ, Wang L, Gloss SP (1995) Solubility and mobility of copper, zinc and lead in acidic environments. Plant Soil 171(1):53–58
Rehman H, Khan AHA, Butt TA, Toqeer M, Bilal M, Ahmad M, Al-Naghi AAA, Latifee ER, Algassem OAS, Iqbal M (2024) Synergistic biochar and *Serratia marcescens* tackle toxic metal contamination: a multifaceted machine learning approach. J Environ Manage 370:122575
Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat, (2019) Deep learning and process understanding for data-driven Earth system science. Nature 566(7743):195–204
Rout GR, Samantaray S, Das P (2001) Aluminium toxicity in plants: a review. Agronomie 21(1):3–21
Rudolph-Mohr N, Tötzke C, Kardjilov N, Oswald SE (2017) Mapping water, oxygen, and pH dynamics in the rhizosphere of young maize roots. J Plant Nutr Soil Sci 180(3):336–346
Sarfraz R, Li SW, Yang WH, Zhou BQ, Xing SH (2019) Assessment of physicochemical and nutritional characteristics of waste mushroom substrate biochar under various pyrolysis temperatures and times. Sustainability 11(1):277
Shahare Y, Singh MP, Singh P, Diwakar M, Singh V, Kadry S, Sevcik L (2023) A comprehensive analysis of machine learning-based assessment and prediction of soil enzyme activity. Agriculture 13(7):1323
Shen T, Peng HY, Yuan XZ, Liang YS, Liu SQ, Wu ZB, Leng LJ, Qin PF (2024) Feature engineering for improved machine-learning-aided studying heavy metal adsorption on biochar. J Hazard Mater 466:133442
Shi RY, Hong ZN, Li JY, Jiang J, Abdulaha-Al Baquy M, Xu RK, Qian W (2017) Mechanisms for increasing the pH buffering capacity of an acidic Ultisol by crop residue-derived biochars. J Agric Food Chem 65(37):8111–8119
Shi RY, Hong ZN, Li JY, Jiang J, Kamran MA, Xu RK, Qian W (2018a) Peanut straw biochar increases the resistance of two Ultisols derived from different parent materials to acidification: a mechanism study. J Environ Manage 210:171–179
Shi RY, Li JY, Jiang J, Kamran MA, Xu RK, Qian W (2018b) Incorporation of corn straw biochar inhibited the re-acidification of four acidic soils derived from different parent materials. Environ Sci Pollut Res 25(10):9662–9672
Shi RY, Li JY, Ni N, Xu RK (2019c) Understanding the biochar's role in ameliorating soil acidity. J Integr Agr 18(7):1508–1517
Shi RY, Liu ZD, Li Y, Jiang TM, Xu MG, Li JY, Xu RK (2019a) Mechanisms for increasing soil resistance to acidification by long-term manure application. Soil till Res 185:77–84
Shi RY, Ni N, Nkoh JN, Li JY, Xu RK, Qian W (2019b) Beneficial dual role of biochars in inhibiting soil acidification resulting from nitrification. Chemosphere 234:43–51
Shu XY, Hu YF, Liu WJ, Xia LL, Zhang YY, Zhou W, Liu WL, Zhang YL (2023) Linking between soil properties, bacterial communities, enzyme activities, and soil organic carbon mineralization under ecological restoration in an alpine degraded grassland. Front Microbiol 14:1131836
Storey VC, Parsons P, Bueso AC, Tremblay MC, Lukyanenko R, Castillo A, Maaß W (2025) Domain knowledge in artificial intelligence: using conceptual modeling to increase machine learning accuracy and explainability. Data Knowl Eng 160:102482
Sulaiman S, Hernandez-Ramirez G, Navaranjan N, Sulaiman Z (2026) How does wheat straw-derived biochar influence the nutrient pool of a site-specific Luvisols in a laboratory incubation? J Environ Qual 55(1):70121
Sun X, Du ZL, Ding J, Zheng SN, Yao YP, Wu LN, Chen HG, An Y, Luo YM (2026) Machine learning integrated with a causal pathway framework unravels differential mechanisms of biochar-driven soil organic carbon dynamics under cadmium stress. Environ Sci Technol 60(3):41556313
Tang C, Weligama C, Sale P (2013) Subsurface soil acidification in farming systems: its possible causes and management options. In: Xu J, Sparks DL (eds) Molecular environmental soil science. Springer, Dordrecht, pp 389–412
Tkacz A, Hortala M, Poole PS (2018) Absolute quantitation of microbiota abundance in environmental samples. Microbiome 6(1):110
Tourlousse DM, Sekiguchi Y (2025) Synthetic DNA spike-in standards for cross-domain absolute quantification of microbiomes by rRNA gene amplicon sequencing. ISME Commun 5 1 ycaf028.
Vandeputte D, Kathagen G, D'hoe K, Vieira-Silva S, Valles-Colomer M, Sabino J, Wang J, Tito RY, De Commer L, Darzi Y, Vermeire S, Falony G, Raes J, (2017) Quantitative microbiome profiling links gut community variation to microbial load. Nature 551:507–511
Viana RSR, de Figueiredo CC, Chagas JKM, Paz-Ferreiro J (2024) Combined use of biochar and phosphate rocks on phosphorus and heavy metal availability: a meta-analysis. J Environ Manage 353:120204
von Cossel M, Druecker H, Hartung E (2019) Low-input estimation of site-specific lime demand based on apparent soil electrical conductivity and in situ determined topsoil pH. Sensors 19(23):5280

Von Uexküll HR, Mutert E (1995) Global extent, development and economic impact of acid soils. Plant Soil 171(1):1–15
Wadoux AMJ-C (2025) Artificial intelligence in soil science. Eur J Soil Sci 76(2):e70080
Waheed A, Xu HL, Qiao X, Aili A, Yiremaikebayi Y, Haitao D, Muhammad M (2025) Biochar in sustainable agriculture and climate mitigation: mechanisms, challenges, and applications in the circular bioeconomy. Biomass Bioenerg 193:107531
Wang CY, Liu BH, Liu LP, Zhu YJ, Hou JL, Liu P, Li X (2021) A review of deep learning used in the hyperspectral image analysis for agriculture. Artif Intell Rev 54(7):5205–5253
Wang CY, Lou J, Yan K, Wang HZ, Xu JM (2022) Analysis of soil microbial community structure via integrated high-throughput absolute abundance quantification (iHAAQ) method. Acta Pedol Sin 59(5):1432–1446 (**(in Chinese)**)
Wang CJ, Zhao YX, Gao YR, Chen HB, Li XF, Zhou BL, Fan DP, Fang Z, Liu J (2024a) Interpretable machine learning for predicting heavy metal removal and optimizing biochar characteristics. J Water Process Eng 68:106484
Wang GJ, Hu JJ, Zhou J, Liu S, Li QJ, Sun ZM (2025) Knowledge-guided large language model for material science. Annu Rev Mater Res 1(2):100007
Wang LW, O'Connor D, Rinklebe J, Ok YS, Tsang DCW, Shen ZT, Hou DY (2020) Biochar aging: mechanisms, physicochemical changes, assessment, and implications for field applications. Environ Sci Technol 54(23):14797–14814
Wang XL, Ding SJ, Cheng T, Wang S, Qing ZT, Hu J, Wang XB (2024b) Mechanism of Cd adsorption by sulphur modified biochar and its application in Cd-contaminated soil. Pol J Environ Stud 33:5901–5912
Wang Y, Yin RZ, Liu RH (2014) Characterization of biochar from fast pyrolysis and its effect on chemical properties of the tea garden soil. J Anal Appl Pyrolysis 110:375–381
Wang Z, Tao TT, Wang H, Chen J, Small GE, Johnson D, Chen JH, Zhang YJ, Zhu QC, Zhang SM, Song YT, Kattge J, Guo P, Sun X (2023) Forms of nitrogen inputs regulate the intensity of soil acidification. Glob Change Biol 29(14):4044–4055
Weber KA, Achenbach LA, Coates JD (2006) Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction. Nat Rev Microbiol 4(10):752–764
Wen T, Ding Z, Thomashow LS, Hale L, Yang S, Xie P, Liu X, Wang H, Shen Q, Yuan J (2023) Deciphering the mechanism of fungal pathogen-induced disease-suppressive soil. New Phytol 238(6):2634–2650
Weng SZ, Zhang QQ, Han KX, Pan MJ, Tan YJ, Chen Q, Wu FH, Wang C, Zheng L, Lei Y, Sha W (2025) Reference-based image super-resolution of hyperspectral and red-green-blue image for determination of wheat kernel quality using deep learning networks. Eng Appl Artif Intell 139:109513
Wilkinson MD, Dumontier M, Aalbersberg IJsbrand Jan and Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, Santos LB da S, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, `t Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher Scott J. and Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B, (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3:160018
Wu HT (2021) Developing an intelligent agricultural system based on long short-term memory. Mobile Netw Appl 26(3):1397–1406
Wu Z, Dong YB, Zhang X, Xu XT, Xiong ZQ (2023) Biochar single application and reapplication decreased soil greenhouse gas and nitrogen oxide emissions from rice-wheat rotation: A three-year field observation. Geoderma 435:116498
Wu HW, Huo YX, Qi FY, Zhang YQ, Li R, Qiao M (2025) Biochar-supported microbial systems: a strategy for remediation of persistent organic pollutants. Biochar 7(1):113
Xia H, Riaz M, Liu B, Li YX, El-Desouki Z, Jiang CC (2023) Peanut shell biochar in acidic soil increases nitrogen absorption and photosynthesis characteristics of maize under different nitrogen levels. Environ Dev Sustain 25(9):8957–8974
Xiang K, Chen HL, Yao CQ, Wan YY, Ren H, Zhou JM (2025) Data-driven machine learning with SHAP for predicting multiple biochar properties from agricultural and forestry waste. Biomass Bioenerg 200:107958
Xiang X, Li XQ, Yang TT, Cao JG, Yang ZZ (2024) A novel Fe–Ca composite modified biochar prepared for enhanced Cd adsorption: experimental, DFT calculation and mechanism unveiling. J Water Process Eng 68:106565
Xie J, Wang ZF, Wang Y, Xiang SJ, Xiong ZY, Gao M (2023) Manure combined with biochar reduces rhizosphere nitrification potential and *amoA* gene abundance of ammonia-oxidizing microorganisms in acid purple soil. Appl Soil Ecol 181:104660
Xiong X, Grunwald S, Myers DB, Kim J, Harris WG, Comerford NB (2014) Holistic environmental soil-landscape modeling of soil organic carbon. Environ Modell Softw 57:202–215
Xu RK, Coventry DR (2003) Soil pH changes associated with lupin and wheat plant materials incorporated in a red-brown earth soil. Plant Soil 250(1):113–119
Xu RK, Zhao AZ, Yuan JH, Jiang J (2012) pH buffering capacity of acid soils from tropical and subtropical regions of China as influenced by incorporation of crop straw biochars. J Soils Sediments 12(4):494–502
Xu SW, Yuan M, Chapman SJ, Zheng NG, Yao HY, Kuzyakov Y (2023) Bioconverted organic wastes shape microbiota in maize rhizosphere: localization and identification in enzyme hotspots. Soil Biol Biochem 184:109105
Xu ZC, Li MW, Zhou R, Xu GR (2025) Machine learning predicted crop yield and soil organic carbon variation after biochar application. Soil Use Manage 41(2):e70092
Yan F, Schubert S (2000) Soil pH changes after application of plant shoot materials of faba bean and wheat. Plant Soil 220(1–2):279–287
Yang HW, Liu XR, Liu YL, Cui JH, Xiao Y (2023) Revolutionizing biochar synthesis for enhanced heavy metal adsorption: harnessing machine learning and Bayesian optimization. J Environ Chem Eng 11(5):110593
Yang JR, Xia LL, van Groenigen KJ, Zhao X, Ti CP, Wang WL, Zhangliu D, Fan M, Zhuang MH, Smith P, Lal R, Butterbach-Bahl K, Han XR, Meng J, Liu J, Cai HG, Cheng YH, Liu XR, Shu XY, Jiao XY, Pan ZD, Tang GM, Yan XY (2025) Sustained benefits of long-term biochar application for food security and climate change mitigation. P Natl Acad Sci USA 122(33):e2509237122
Yao KS, Zhang YB, Sun J, Xu YX, Zhou B, Wang KZ, Zhang B, Du XJ, Li Y (2025) Nondestructive detection of heavy metal lead in eggs using hyperspectral imaging combined with deep learning-based feature extraction method. J Food Compos Anal 146:107994
Yaseen ZM, Alhalimi FL (2025) Heavy metal adsorption efficiency prediction using biochar properties: a comparative analysis for ensemble machine learning models. Sci Rep 15(1):13434
Yu FM, Li CM, Dai CL, Liu KH, Li Y (2020) Phosphate: coupling the functions of fertilization and passivation in phytoremediation of manganese-contaminated soil by *Polygonum pubescens* blume. Chemosphere 260:127651
Yuan JH, Xu RK (2012) Effects of biochars generated from crop residues on chemical properties of acid soils from tropical and subtropical China. Soil Res 50(7):570–578
Yuan JH, Xu RK, Qian W, Wang RH (2011a) Comparison of the ameliorating effects on an acidic Ultisol between four crop straws and their biochars. J Soils Sediments 11(5):741–750
Yuan JH, Xu RK, Zhang H (2011b) The forms of alkalis in the biochar produced from crop residues at different temperatures. Bioresour Technol 102(3):3488–3497
Zhai YH, Zhao WY, Dong K, Wang YW, Lin H, Dong B, Wang DQ, Xu YF, Liu JB (2026) Explainable machine learning reveals key determinants of biochar soil-application effects on plant growth. Environ Technol Inno 41:104829
Zhang GX, Huang FY, Shi YS, Chen K, Wu XF, Zhou H, Wu WC, Wang ZY, Wang LJ, Tang SG, Yuan M (2025f) Integration of experimental and computational techniques to elucidate the Cu(II) adsorption mechanisms of nitrogen-doped chitosan microsphere biochars. Chem Eng J 521:166840
Zhang L, Heuvelink GBM, Mulder VL, Chen SC, Deng XF, Yang L (2024b) Using process-oriented model output to enhance machine learning-based soil organic carbon prediction in space and time. Sci Total Environ 922:170778

Zhang NH, Xing JM, Wei LJ, Liu C, Zhao WJ, Liu ZH, Wang YH, Liu EK, Ren XL, Jia ZK, Wei T, Siddique KHM, Zhang P (2025b) The potential of biochar to mitigate soil acidification: a global meta-analysis. Biochar 7(1):49
Zhang NH, Xing JM, Wei LJ, Liu C, Zhao WJ, Liu ZH, Wang YH, Liu EK, Ren XL, Jia ZK, Wei T, Siddique KHM, Zhang P (2025d) The potential of biochar to mitigate soil acidification: a global meta-analysis. Biochar 7:49
Zhang SJ, Xu GY, Quan XL, Tan XD, Zhang RX, Fu X, Peng H, Luo S (2025e) Biochar accelerates straw decomposition and reduces greenhouse gas emissions by driving microbial community dynamics. Chem Biol Technol Agric 12(1):150
Zhang WN, Pang JY, Qi JF, Lu Y, Liu JH, Yu MF, Li HG, Wang EL, Lambers H (2025c) Biochar's dual impact on soil acidity management and crop yield enhancement in acidic soils: a meta-analysis. Plant Soil 514(2):1849–1865
Zhang WB, Wei CL, Li JY, Zhou Q, Ma JQ, Li YL, He YX, Zou Y, Luo YM, Yao YJ (2025a) Stabilization of acidification in China's cropland soils. Nat Geosci 18(11):1125–1132
Zhang XT, Feng XY, Chai N, Kuzyakov Y, Zhang F, Li FM (2024a) Biochar effects on crop yield variability. Field Crop Res 316:109518
Zhao CX, Jiang ZH, Lu XY, Yue WJ, Chen JH, Liu XG (2024) Prediction of biochar yield based on machine learning model of "enhanced data" training. Biomass Bioenerg 182:107089
Zhao CX, Yang H, Ma H, Chen AH, Liu XG, Chen JH (2025) Machine learning prediction of biochar yield based on different classification methods. J Renew Sustain Energy 17(5):053103
Zhao CX, Yang H, Zhang YM, Xia Q, Yue WJ, Chen AH, Liu XG (2026a) Prediction of soil pH improvement through biochar: a machine learning based solution. Land Degrad Dev 37(4):1479–1490
Zhao HQ, Mao JH, Wang MM, Xie Y, Wang P, Zhao YS, Shi YN, Huangfu XD (2026b) Combining spectral analysis and machine learning methods for rapid determination of coal quality indices from hyperspectral images. Measurement 257:118539
Zhao WR, Shi RY, Hong ZN, Xu RK (2022) Critical values of soil solution Al3+ activity and pH for canola and maize cultivation in two acidic soils. J Sci Food Agr 102(15):6984–6991
Zhao XQ, Pan XZ, Ma HY, Dong XY, Che J, Wang C, Shi Y, Liu K, Shen RF (2023) Scientific issues and strategies of acid soil use in China. Acta Pedlogica Sinica 60(5):1248–1263 (**(in Chinese)**)
Zhu XZ, Li YN, Wang XN (2019) Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions. Bioresour Technol 288:121527
Zhu SD, Preuss N, You FQ (2023) Advancing sustainable development goals with machine learning and optimization for wet waste biomass to renewable energy conversion. J Clean Prod. 422:138606
Zimmerman AR (2010) Abiotic and microbial oxidation of laboratory-produced black carbon (biochar). Environ Sci Technol 44(4):1295–1301
Zou ZY, Yuan DY, Wu QS, Xiao YC, Li MH, Wang QL, Zhen JB, Li J, Cui Q, Zhou RH, Zhou M, Xu LJ (2025) The comprehensive index for assessing the freshness of salmon using hyperspectral imaging technology combined with multisource data fusion method. J Food Compos Anal 147:108025