Carbon Balance and Management

RESEARCH

Open Access

Nitrogen addition enhances soil carbon and nutrient dynamics in Chinese croplands: a machine learning and nationwide synthesis

Yu Li[1] and Yuan Li[2*]

Abstract

Nitrogen (N) addition is a critical driver of soil organic carbon (SOC) sequestration and nutrient cycling in croplands. However, its spatial variability and long-term effects under diverse environmental conditions remain poorly understood. We synthesised data from 479 cropland sites across China and apply machine learning models to evaluate the impacts of N addition on SOC and key soil nutrient indicators, including total nitrogen (TN), nitrate (NO_3^--N), ammonium (NH_4^+-N), the carbon-to-nitrogen ratio (C/N), and available phosphorus (AP). We further evaluated the moderating roles of climate zones, fertiliser types, and fertilisation duration. Our findings demonstrate that N addition significantly increased SOC, TN, NO_3^--N, NH_4^+-N, and AP contents, whereas the C/N ratio remains unaffected. SOC sequestration was greater in arid regions, whereas nutrient accumulation was more pronounced in humid zones. Organic and integrated (organic-inorganic) fertilisers outperformed chemical ones in enhancing SOC and nutrient cycling. Long-term N input (> 10 years) markedly intensified SOC storage and nutrient accumulation. We further developed the high-resolution (5 km) national-scale dataset that predicts the spatial responses of SOC and nutrient dynamics to nitrogen addition across China. This AI-derived dataset enables automated mapping of soil carbon and nutrient functions, capturing substantial spatial heterogeneity under varying environmental conditions. These results provide critical insights for optimising nitrogen management strategies, enhancing soil carbon sink functions, and informing precision agriculture policies in China.

Keywords Nitrogen addition, Soil organic carbon, Nutrient cycling, Machine learning, Carbon sequestration, China

Introduction

Nitrogen (N) is a critical macronutrient leading plant growth and driving soil biogeochemical cycles, thus central to terrestrial ecosystem productivity and sustainability. In managed agricultural systems, especially under intensified global food production, synthetic N fertilisers have become prevalent, particularly in countries like China where food security remains strategically important [1, 2]. Although synthetic N fertilisers have substantially improved crop yields, they simultaneously induce significant alterations in soil organic carbon (SOC) dynamics and nutrient cycling, with cascading consequences for soil health, greenhouse gas emissions, and long-term ecosystem stability [3]. Decades of research consistently show that increased N inputs from synthetic fertilisers and atmospheric deposition profoundly impact ecosystem functioning across biomes [4, 5]. Consequently, while N fertilisation underpins modern agricultural productivity, it concurrently reshapes SOC and

*Correspondence:
Yuan Li
yuan1028li@163.com
[1]Quality Foundation Development Institute of Xinjiang Uygur Autonomous Region, Urumqi 831399, China
[2]Institute of Soil, Fertilizer and Water Saving Agriculture, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China

soil nutrient dynamics, emphasising the urgent need to balance food production goals with environmental sustainability [6].

Although numerous studies have examined the effect of N inputs on SOC sequestration and nutrient cycling, their findings exhibit considerable variability due to differences in environmental and management conditions. On the one hand, N addition can stimulate plant growth and increase litter return, thereby increasing SOC stocks and modifying microbial activity in ways that enhance carbon stabilisation [7]. On the other hand, excessive or prolonged N inputs may inhibit microbial decomposition, reduce microbial diversity, acidify soils, or destabilise soil aggregates—processes that can suppress SOC accumulation or even promote carbon losses [8]. Similarly, the impact of N fertilisation on nutrient availability—particularly nitrate (NO_3^--N), ammonium (NH_4^+-N), total nitrogen (TN), and available phosphorus (AP)—is influenced by soil texture, pH, climatic factors, and biological interactions [9, 10].

Despite this growing body of knowledge, most empirical studies have been conducted at localised scales or under short-term experimental settings, making it difficult to generalise findings across regions with contrasting agroecological conditions [11, 12]. Global syntheses, such as the meta-analysis by Lu et al. [6], confirm that N effects on SOC vary across biomes, but nationwide, spatially explicit assessments remain scarce, particularly in intensively managed cropland systems like those in China. This constrains our ability to understand the spatial heterogeneity and cumulative responses of soil C and nutrients to N inputs under diverse environmental conditions.

Another key limitation in existing research is the reliance on conventional statistical methods, which may oversimplify the complex and nonlinear interactions among soil properties, climate variables, and management practices [13]. For example, linear regression models may fail to capture threshold responses or synergistic effects that emerge only under specific environmental contexts. The rise of machine learning (ML) techniques in soil science offers powerful tools to address this challenge, allowing for integrating multi-source datasets and modelling nonlinear relationships across large spatial scales [14, 15]. ML models have shown great promise in predicting SOC distributions, estimating nutrient fluxes, and supporting precision agriculture under climate change scenarios [16].

Our study uses the comprehensive global dataset presented by Elrys et al. [17], narrowing the focus to China due to its critical role in global N fertiliser consumption and strategic importance for food security. By specifically analysing China's croplands, we address significant knowledge gaps regarding how local environmental conditions affect soil N dynamics. Moreover, our approach integrates environmental covariates and advanced machine learning modelling, thus extending the original analysis by revealing spatially explicit patterns and driving mechanisms of nitrogen retention and loss unique to China's intensive agricultural systems. This targeted analysis provides a necessary complement to global-scale findings, enabling more precise N management recommendations tailored to China's environmental context.

In this study, we compiled a comprehensive national dataset from 479 cropland sites across China to systematically assess the effects of N addition on SOC and major soil nutrient indicators (TN, NO_3^--N, NH_4^+-N, C/N, and AP). Specifically, we aimed to (1) quantify the overall effects of N addition on SOC and soil nutrient concentrations across Chinese croplands; (2) evaluate how these effects are moderated by climatic zones, fertiliser types, and fertilisation durations; and (3) develop a machine learning model to predict the SOC sequestration potential of N addition and explore its spatial heterogeneity. By filling these knowledge gaps, our study provides critical insights into the optimisation of N management and contributes to carbon neutrality and sustainable agriculture goals in China and beyond.

Materials and methods

Literature search and data compilation

Our dataset was derived from the comprehensive global meta-analysis conducted by Elrys et al. [17] which systematically assessed the impacts of knowledge-based nitrogen (N) management practices on ecosystem nitrogen retention worldwide. Elrys et al. [17] performed their literature search using the Google Scholar database (http://scholar.google.com/) and included publications from previously published meta-analyses. The initial search conducted in January 2022 employed combinations of keywords such as "gross nitrogen transformations," "net nitrogen mineralisation," "nitrification rates," "nitrogen retention," "nitrogen fertilisation," "organic fertilisation," "straw," "nitrogen and phosphorus inputs," "nitrogen cycling," "knowledge-based N management practices," and "nitrification inhibitors." A subsequent search in February 2023 used refined keywords including "gross nitrogen transformations," "net nitrogen mineralisation," "nitrification rates," "nitrogen retention," "nitrogen loss," "nitrogen leaching," "gaseous emissions," "ammonia volatilisation," "nitrogen fertilisation," "NPK," "organic fertilisation," "straw," and "nitrification inhibitors." Studies were included if they met the criteria: (1) treatment and control plots were under identical biotic and abiotic conditions; (2) clear documentation of fertiliser type, application rate, experimental duration, and terrestrial ecosystem type; (3) direct application of organic or synthetic fertilisers in terrestrial ecosystems with measured soil N transformation rates or N loss pathways; and (4) measurements of soil gross N fluxes and pools taken from the topsoil layer (0–20 cm). For our analysis, we specifically extracted observations from cropland studies in China, resulting in 479 paired observations following rigorous screening, which involved field-based N addition experiments with clearly reported N rates and measurement of target variables during the crop growing season. Data not directly tabulated were extracted using Web Plot Digitizer (https://automeris.io/WebPlotDigitizer/) from published graphs and figures.

Response ratio calculation

To quantify the effects of N addition on SOC and key nutrient indicators—including TN, NO_3^--N, NH_4^+-N, the C/N, and AP—we used the natural log response ratio (lnRR), a standardized effect size commonly applied in ecological meta-analyses [18]. The lnRR is calculated as:

$$RR = ln\left(\frac{x_t}{x_c}\right) = ln\left(x_t\right) - ln\left(x_c\right)$$

where:

- x_t is the mean value of the treatment group (with N addition);
- x_c is the mean value of the control group (without N addition);
- $ln\left(\frac{x_t}{x_c}\right)$ reflects the relative change caused by N addition in natural log units.

Positive values of lnRR indicate an increase in the variable due to N addition, while negative values suggest a decrease. This metric facilitates comparisons across studies with different units or scales. The sampling variance of lnRR is estimated using the following equation:

$$v = \frac{s_t^2}{N_t x_t^2} + \frac{s_c^2}{N_c x_c^2}$$

where:

- s_t^2 and s_c^2 are the variances of the treatment and control groups, respectively;
- N_t and N_c are the sample sizes of the treatment and control groups;
- x_t^2 and x_c^2 are the squared means of the respective groups.

This variance estimate allows for inverse-variance weighting in further statistical analyses, such as mixed-effects models or meta-regression, ensuring that studies with higher precision contribute more to the overall effect estimates.

Environmental covariates

To assess the influence of environmental conditions on treatment effects, we compiled auxiliary environmental variables based on the geographic coordinates of each experimental site: Mean annual temperature (MAT) and mean annual precipitation (MAP) were extracted from the WorldClim database (version 2.1; https://www.worldclim.org/). Elevation was retrieved from the ASTER Global Digital Elevation Model (GDEM) (https://asterweb.jpl.nasa.gov/GDEM.asp). Soil properties, including fine texture (% clay + silt), bulk density (BD), soil pH, and soil water content, were obtained from the SoilGrids database (https://data.isric.org/). All site-specific environmen tal data were extracted using latitude and longitude coordinates. These variables were later used as predictors in machine learning modelling.

Machine learning modelling

To systematically evaluate the spatial response of farmland soil carbon storage and key nutrient indicators to N addition, we developed a machine learning model based on the Random Forest (RF) algorithm to predict six soil property indicators: SOC, TN, C/N, NO_3^--N, NH_4^+-N, and AP [19]. The model incorporated over twenty biotic and abiotic predictors, including climatic variables (e.g., mean annual temperature and precipitation), soil physicochemical properties (e.g., texture, pH, BD, and soil water content) and geographic factors. This approach allowed us to reconstruct and map the national-scale SOC sequestration potential with high accuracy and resolution, accounting for complex nonlinear interactions. The coefficient of determination (R^2) for each model was as follows: SOC (0.71), TN (0.68), C/N (0.63), NO_3^--N (0.60), NH_4^+-N (0.65), and AP (0.74), indicating high predictive accuracy across indicators.

Results

Spatial distribution of sampling sites and overall effects of nitrogen addition

The 479 cropland sampling sites analysed in this study span a wide range of climatic conditions across China, covering an extensive aridity gradient. Site locations ranged from arid and semi-arid regions in northern and northwestern China to humid and semi-humid areas in the southern and southeastern regions (Fig. 1(a)). Weighted response ratio (LnRR) analysis revealed that nitrogen (N) addition significantly increased soil organic carbon (SOC), total nitrogen (TN), nitrate nitrogen (NO_3^--N), ammonium nitrogen (NH_4^+-N), and available phosphorus (AP) contents ($P < 0.001$). In contrast, the soil carbon-to-nitrogen ratio (C/N) showed no significant

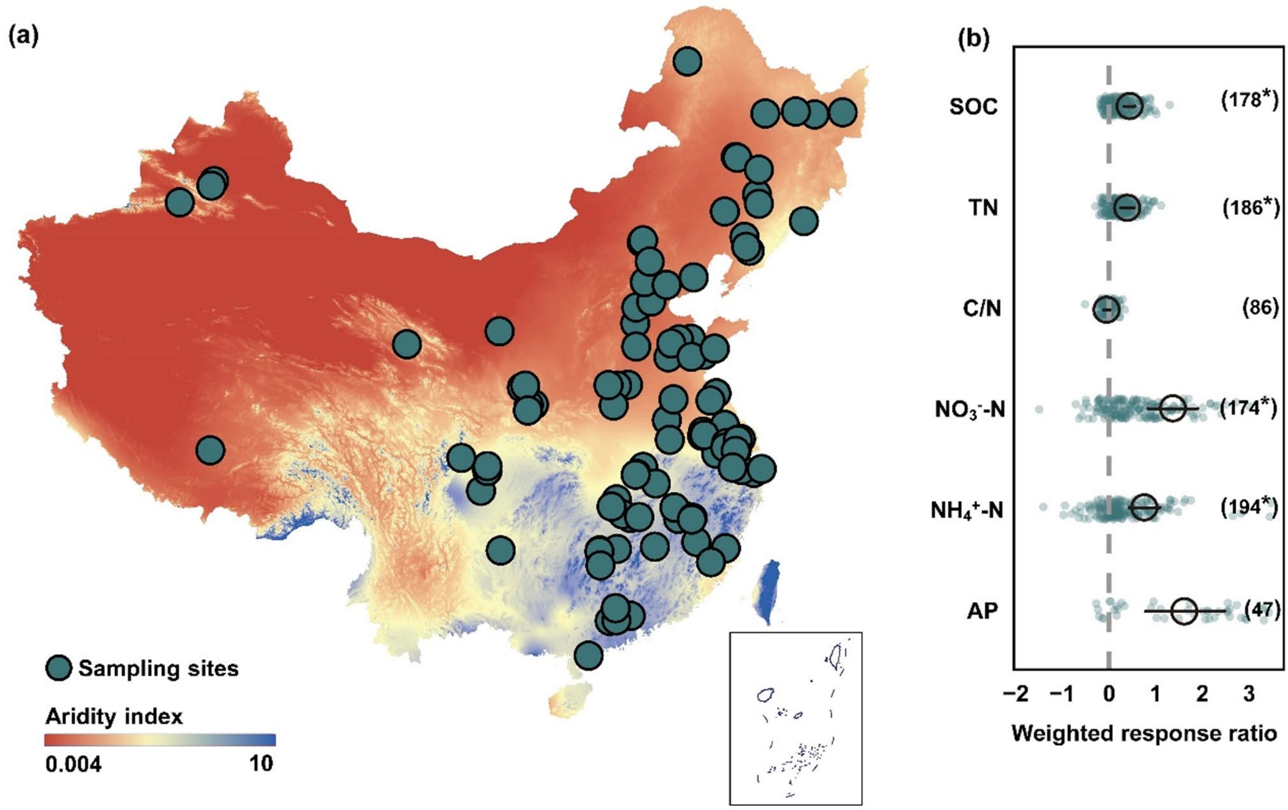

Fig. 1 Effects of nitrogen addition on soil carbon sequestration and nutrient cycling. **a**, Global distribution of study sites about nitrogen addition on soil carbon sequestration and nutrient cycling. **b**, Effects of nitrogen addition on soil carbon sequestration and nutrient cycling. Values are effect size ± 95% CI. The sample size in each category is given at the *Right*, and the symbol * indicates statistical significance. SOC, soil organic carbon; TN, total nitrogen; C/N, carbon to nitrogen ratio; NO_3^--N, nitrate nitrogen; NH_4^+-N, ammonium nitrogen; AP, available phosphorus

response to N addition ($P = 0.51$). Overall, nitrogen addition consistently enhanced soil carbon storage and nutrient accumulation across diverse environmental conditions (Fig. 1 (b)).

Climatic zone, fertiliser type, and duration influence nitrogen addition effects

The effects of N addition varied considerably across climatic zones, fertiliser types, and application durations (Fig. 2). Arid regions exhibited stronger SOC sequestration responses to N addition than humid regions, likely due to lower initial SOC levels. In contrast, increases in TN and NH_4^+-N were more pronounced in humid regions, while NO_3^--N and AP showed greater enhancement in arid zones. Fertiliser type also influenced nutrient dynamics. Organic fertilisers led to the greatest increases in SOC and TN, whereas organic-inorganic compound fertilisers induced the highest increases in NO_3^--N and AP. Inorganic fertilisers generally produced the lowest responses across most indicators (except for NH_4^+-N), suggesting that organic or mixed fertilisers are more effective in enhancing soil carbon and nutrient cycling. Additionally, longer fertilisation durations (> 10 years) were associated with progressively greater increases in SOC and nutrient indices, highlighting the cumulative benefits of sustained nitrogen input over time.

Positive dose–response relationships with nitrogen input rates

Across all evaluated indicators, the magnitude of soil response increased with the rate of nitrogen addition (Fig. 3). SOC, TN, NO_3^--N, NH_4^+-N, and AP all exhibited significantly positive correlations with N input rates, as reflected by increasing weighted response ratios (LnRRs) ($P < 0.001$). The C/N ratio, however, remained unresponsive to nitrogen input levels ($P = 0.51$). These results indicate that while nitrogen addition broadly enhances soil carbon and nutrient status, the magnitude of response varies among different indicators, with some nutrients (e.g., NH_4^+-N and AP) responding more strongly than others.

Random forest identification of key environmental predictors

Figure 4 presents the relative importance of environmental variables in predicting changes in six key soil indicators—SOC, TN, C/N ratio, NO_3^--N, NH_4^+-N,

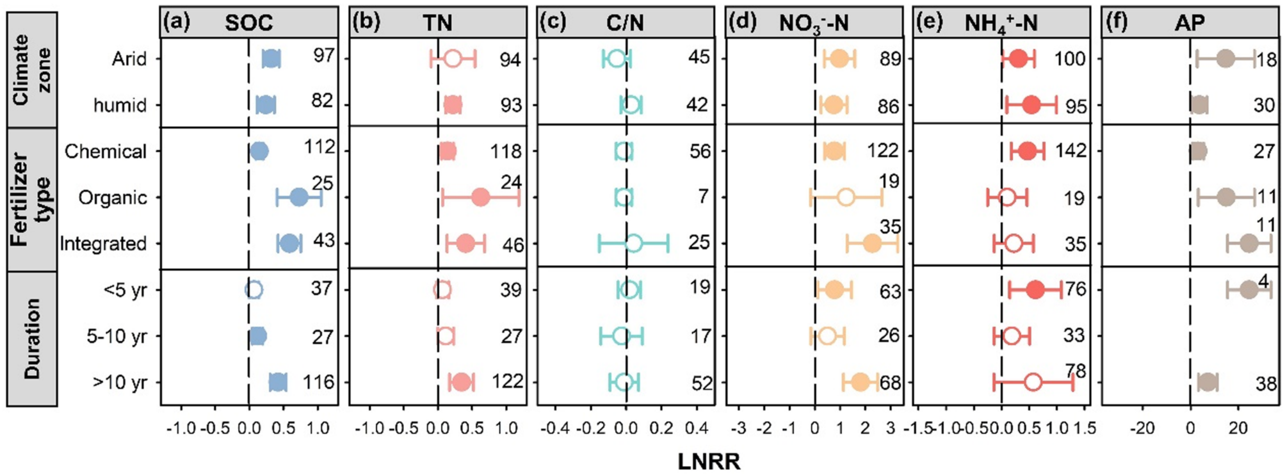

Fig. 2 Responses of soil carbon sequestration and nutrient cycling to nitrogen addition across different climatic zones, nitrogen fertiliser types, and durations of application. Values are effect size ± 95% CI. The sample size in each category is given at the Right. The closed symbols indicate significant effects, and the open symbols indicate nonsignificant effects. SOC, soil organic carbon; TN, total nitrogen; C/N, carbon to nitrogen ratio; NO_3^--N, nitrate nitrogen; NH_4^+-N, ammonium nitrogen; AP, available phosphorus

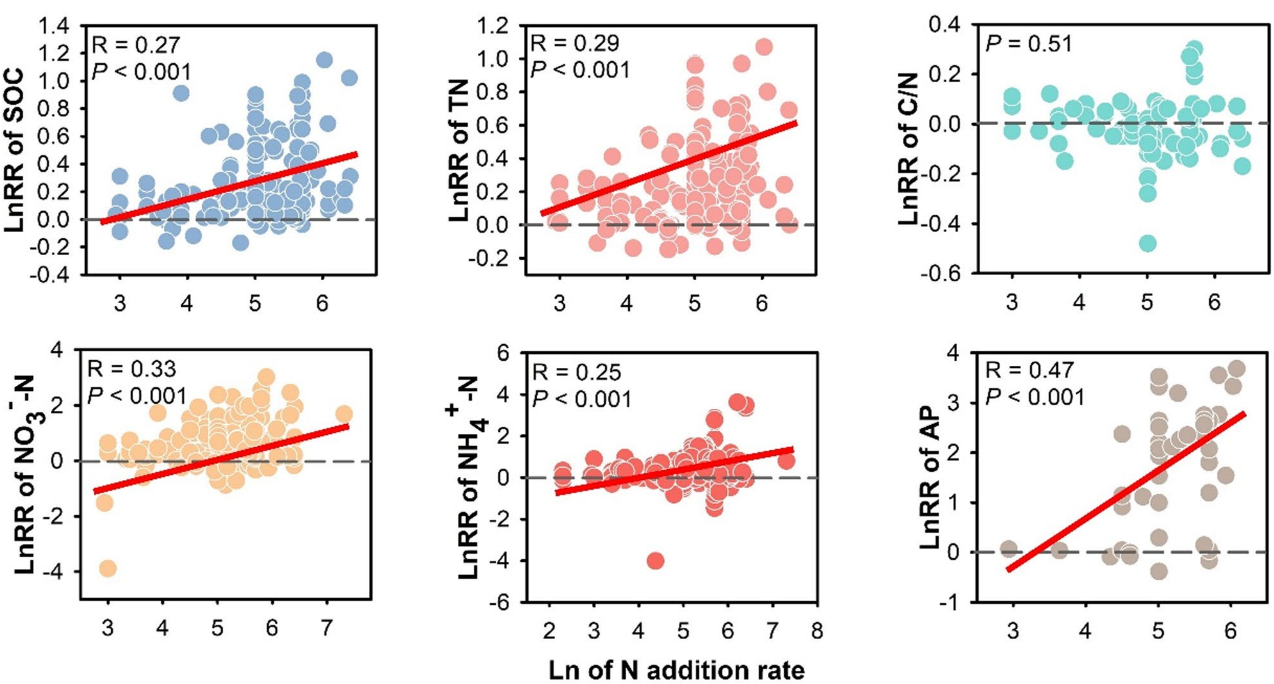

Fig. 3 Relationships between nitrogen addition rates and changes in indicators of soil carbon sequestration and nutrient cycling. SOC, soil organic carbon; TN, total nitrogen; C/N, carbon to nitrogen ratio; NO_3^--N, nitrate nitrogen; NH_4^+-N, ammonium nitrogen; AP, available phosphorus

and AP—under N addition across 479 cropland sites in China. Variable importance was assessed using the increase in mean squared error (MSE) in random forest models. Each soil indicators were influenced by a distinct set of predictors, reflecting divergence in environmental responses. Notably, the soil pH response ratio (pH_RR) was among the most important predictors for SOC (Fig. 4a), TN (Fig. 4b), NH_4^+-N (Fig. 4e), and AP (Fig. 4f), highlighting its overarching regulatory role. In contrast, the C/N ratio (Fig. 4c) was mainly influenced by bulk density and aridity index, whereas nitrate levels (NO_3^--N; Fig. 4d) were more sensitive to nitrogen addition rate and climatic factors. NH_4^+-N (Fig. 4e) showed the highest model performance (R^2 =0.77), predominantly shaped by elevation, pH, and texture. These findings underscore the context-dependent nature of soil biogeochemical responses and identify pH as a cross-cutting control across multiple soil functions.

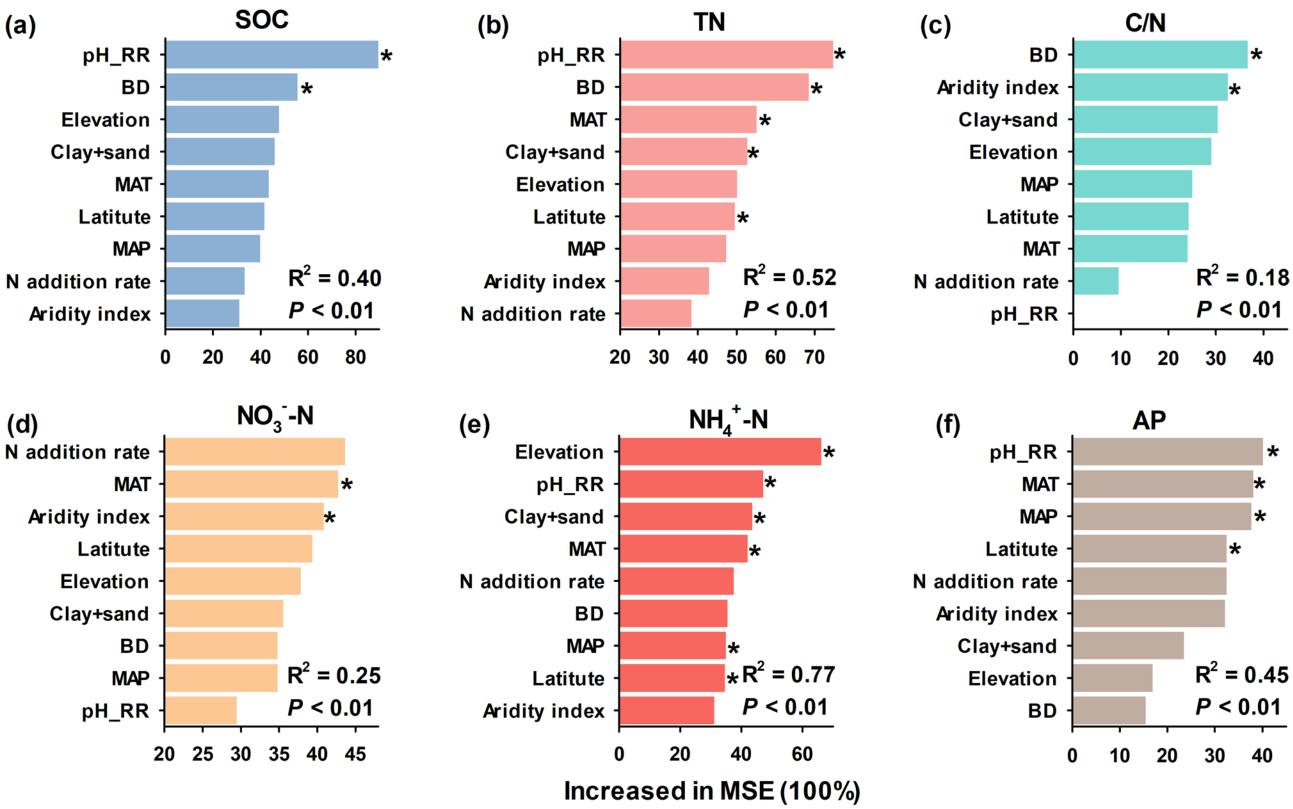

Fig. 4 Relative importance of environmental predictors for soil carbon and nutrient responses to nitrogen addition across Chinese croplands. Asterisks indicate significant predictors ($P < 0.05$). SOC, soil organic carbon; TN, total nitrogen; C/N, carbon to nitrogen ratio; NO_3^--N, nitrate nitrogen; NH_4^+-N, ammonium nitrogen; AP, available phosphorus; BD, bulk density; MAT, mean annual temperature; MAP, mean annual precipitation

Machine learning-based predictions of SOC and nutrient enhancements

Machine learning modelling provided high-resolution estimates of nitrogen addition effects at the national scale (Fig. 5). The model predicted an average increase in SOC of 30.25% (Fig. 5 (a)), and a 22.62% (Fig. 5 (b)) increase in TN. Notably, the C/N ratio was projected to decrease slightly by 2.28% (Fig. 5 (c)), reflecting disproportionate nitrogen enrichment relative to carbon gains. Among nutrient indicators, NH_4^+-N increased by 87.91% (Fig. 5 (e)), and NO_3^--N by 41.61% (Fig. 5 (f)). The most substantial change was observed for available phosphorus (AP), which increased by 651.81% (Fig. 5 (f)), suggesting strong N–P interactions in response to fertiliser inputs. These findings collectively underscore the substantial role of nitrogen addition in promoting soil carbon sequestration and nutrient cycling across China's croplands, while highlighting regional variability and indicator-specific response patterns.

Discussion

Climatic zones, fertiliser type, and application duration modulate nitrogen effects

Our results demonstrate that the effects of nitrogen addition on SOC and nutrient dynamics are context-dependent, varying significantly across climatic zones, fertiliser types, and application durations. In arid and semi-arid regions, N addition typically results in more pronounced increases in SOC compared to humid areas. This disparity arises primarily because plant productivity in drier regions tends to be more limited by nitrogen availability, making additional nitrogen inputs particularly effective in enhancing plant growth and, consequently, carbon sequestration in soils [20–22]. In contrast, Humid regions typically show smaller relative SOC gains after nitrogen additions due to higher baseline fertility, faster decomposition, and greater biological activity. Experimental drought in tropical forests increased soil CO_2 emissions, demonstrating strong moisture controls on decomposition (Cleveland et al., 2010). Similarly, temperature and moisture along tropical gradients significantly influence SOC turnover rates [23]. Nutrient dynamics also varied by climate zone. Several studies support the idea that total TN and NH_4^+-N increases are typically more pronounced in humid regions due to higher moisture availability and microbial activity. Moisture-rich conditions in humid climates enhance microbial decomposition and N mineralisation rates, leading to greater ammonium production and stabilisation [24, 25]. Conversely, NO_3^--N and AP showed larger increases

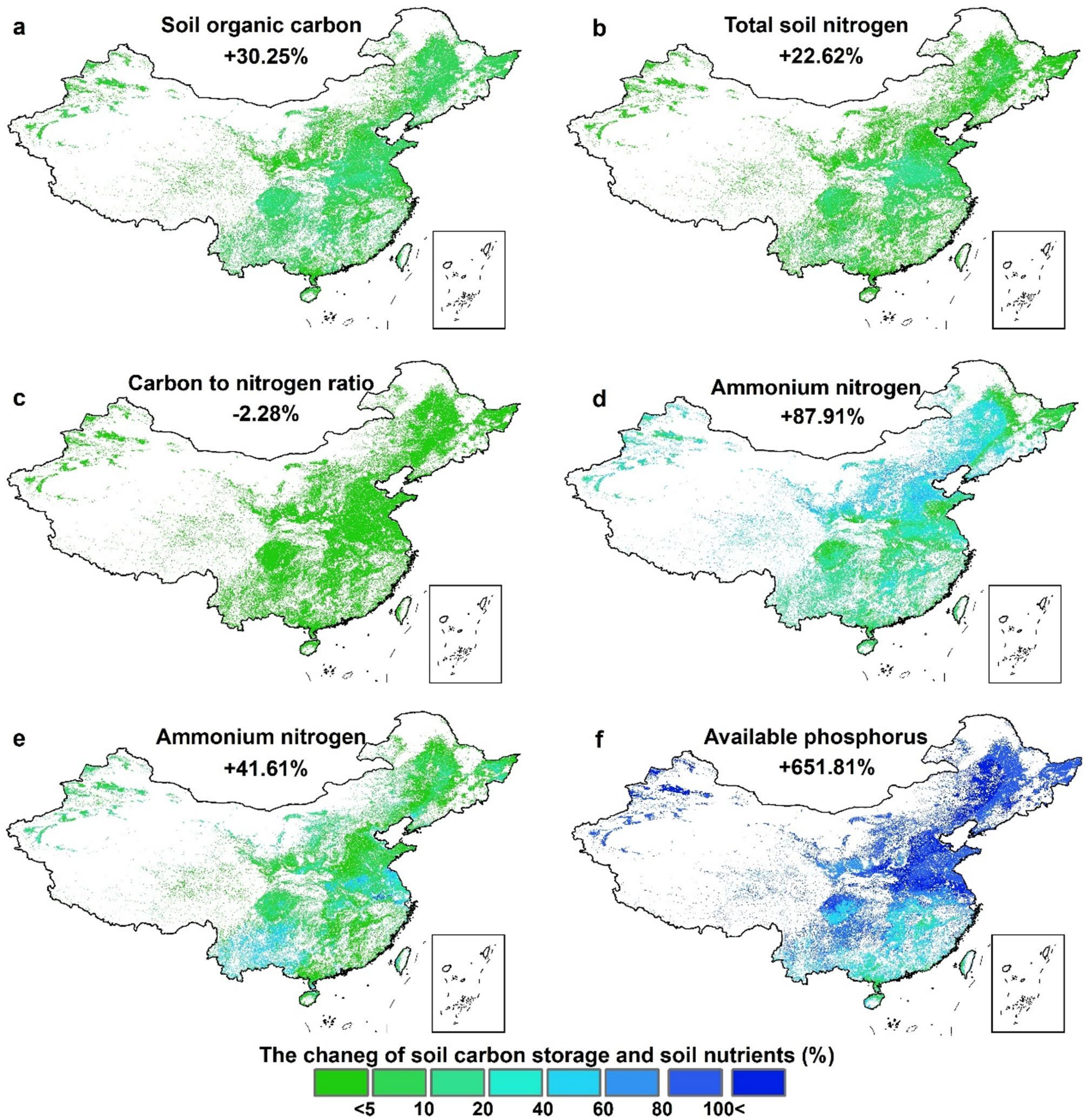

Fig. 5 Predicted spatial distribution of soil carbon sequestration and nutrient cycling potential under nitrogen addition across China

in arid zones, likely due to reduced plant uptake, slower leaching, and greater accumulation in low-rainfall soils.

Our findings highlight fertiliser type as a critical determinant influencing soil organic carbon (SOC), total nitrogen (TN), microbial diversity, and nutrient availability. Studies consistently demonstrates that organic fertilisers substantially enhance SOC and TN by directly adding organic matter, stimulating microbial biomass, and improving soil aggregation [26, 27]. The primary reason behind this positive response is that organic amendments provide abundant carbon sources and nutrients, which fuel microbial metabolism and activity, thereby enhancing microbial community growth and the stabilisation of organic carbon. Conversely, sole reliance on inorganic nitrogen fertilisers generally exhibits weaker effects on SOC accumulation and can negatively impact microbial diversity due to soil acidification and altered microbial habitats [28]. Integrated fertilisation strategies combining organic and inorganic fertilisers provide intermediate yet favourable outcomes, optimising both soil fertility and crop productivity, particularly enhancing nitrate NO_3^--N and available AP availability [29]. Therefore, implementing balanced fertilisation practices emerges as a promising approach for sustainable soil management and agricultural productivity enhancement.

Fertilisation duration had a clear cumulative effect. Long-term application (>10 years) was associated with significantly greater SOC and nutrient improvements than short-term treatments, corroborating results from long-term experiments [30]. These effects are likely due to enhanced microbial turnover, greater formation of stable organo-mineral complexes, and improved nutrient retention over time. Several studies also observed delayed SOC responses in the early years of fertilisation, underscoring the importance of sustained management.

Together, these findings suggest that tailoring fertiliser regimes to local climate, soil, and management histories is essential for maximising benefits and minimising unintended consequences. The strong performance of organic and integrated inputs across systems also supports growing calls for diversified fertilisation strategies in climate-smart agriculture.

Nitrogen addition enhances soil carbon and nutrient accumulation in a dose-dependent manner

Our findings from the meta-analysis align with numerous previous studies [7, 31], which reports that N fertilisation stimulates plant productivity and biomass return to the soil, enhancing belowground carbon inputs through litter, roots, and rhizodeposition [6, 32]. The elevated carbon inputs, in turn, fuel microbial activity and accelerate the stabilisation of organic matter via microbial transformation and aggregate formation [33]. Li et al. [34] similarly reported SOC increases of 10–25%, particularly under moderate to high N inputs, and in systems with organic matter incorporation. The absence of significant C/N change, despite parallel increases in SOC and TN, is consistent with long-term trials (e.g., Sun et al. [35]) and suggests synchronised carbon and nitrogen accumulation. This stoichiometric stability likely reflects co-regulation by plant inputs and microbial nitrogen immobilisation [36].

The observed increases in NO_3^--N and NH_4^+-N indicate that N addition stimulated internal nitrogen cycling, consistent with reports by Gao and Liu [37], which showed that N fertilisation enhanced mineralisation, ammonification, and nitrification rates. NO_3^--N accumulation may reflect microbial oxidation of NH_4^+ under favourable redox conditions, while elevated NH_4^+-N could result from increased mineralization and decreased plant uptake under N surplus [38]. Increased AP availability, despite the absence of phosphorus fertilisation in many studies, suggests a tight N–P coupling. This phenomenon has been attributed to nitrogen-induced stimulation of microbial phosphatase activity, increased root growth, or enhanced solubilization of soil P pools [39, 40]. Similar patterns were reported in several paddy and upland systems [41, 42], particularly under combined N and organic residue inputs.

Importantly, we detected a strong dose–response relationship between N input rates and the magnitude of soil responses across indicators. The weighted response ratios (LnRR) for SOC, TN, NO_3^--N, NH_4^+-N, and AP all increased significantly with increasing N addition. This cumulative effect is consistent with trends observed in multiple long-term studies [43, 44], which showed progressive SOC and nutrient accumulation with prolonged fertilisation. However, some studies reported that responses eventually plateau or reverse at very high N inputs due to microbial suppression, acidification, or N loss pathways [45, 46].

Random forest machine learning enables spatially explicit prediction of nitrogen-induced soil improvements

In addition to our meta-analysis, we employed ML models to predict the spatial distribution of nitrogen-induced changes in SOC and soil nutrients across Chinese croplands. We employed an RF machine learning model, which has been widely recognised for its ability to handle nonlinear relationships in complex environmental datasets [19]. Compared to conventional statistical techniques, ML offers advantages in capturing nonlinear interactions, integrating multiple data layers (e.g., climate, soil, and management), and generating high-resolution spatial predictions [47]. The RF model integrated both biotic and abiotic predictors, including climate, soil properties, N fertiliser type, fertilisation duration, and geographic features. This data-driven approach allowed us to reconstruct and map SOC sequestration potential at a national scale with high accuracy and resolution, capturing complex nonlinear interactions that conventional statistical techniques often fail to detect.

Our results from random forest analysis reveal that soil biogeochemical responses to nitrogen addition are governed by distinct environmental drivers, with pH emerging as a consistent and dominant predictor across multiple soil indicators (Fig. 4). This underscores the critical role of pH regulation in mediating nutrient transformations and organic carbon stabilization, aligning with prior findings that nitrogen-induced acidification significantly alters microbial processes and nutrient dynamics [48, 49]. The divergence in predictor importance across functions—such as the dominance of texture and elevation for NH_4^+-N, or aridity and N rate for NO_3^--N—highlights strong spatial heterogeneity and the need for site-specific nutrient management strategies [50, 51]. These insights emphasize that optimizing nitrogen use efficiency and sustaining soil multifunctionality require integrated approaches that account for underlying soil and climatic contexts.

The RF model predicted significant increases in SOC and soil nutrient concentrations in response to nitrogen fertilisation, with national-scale estimates indicating an average SOC increase of 30.25%. This result supports the widely observed trend that nitrogen addition enhances carbon sequestration by stimulating plant productivity, increasing organic matter inputs, and promoting microbial-driven stabilisation of SOC [34]. Alongside SOC changes, total TN increased by 22.62%, while NH_4^+-N and NO_3^-- exhibited contrasting patterns, with ammonium concentrations rising by 87.91% and nitrate levels increasing by a more moderate 41.61%. The disproportionate increase in ammonium relative to nitrate likely reflects regional differences in soil microbial activity, where limited nitrification under acidic or oxygen-limited conditions suppresses the conversion of NH_4^+-N to NO_3^--N. The relatively lower nitrate accumulation suggests potential losses through leaching and denitrification, particularly in areas with high precipitation and permeable soils [52].

Among the most striking findings, AP exhibited a dramatic increase of 651.81%, reinforcing concerns regarding phosphorus accumulation in intensively managed agricultural systems. This substantial rise in AP is likely driven by long-term phosphorus fertilisation, soil adsorption saturation, and organic manure applications, which contribute to the persistent buildup of bioavailable phosphorus in surface soils [53]. The low mobility of phosphorus further exacerbates its accumulation, increasing the risk of environmental contamination through surface runoff and eutrophication [54]. However, in the southern part of China, where the soil Ph is relatively lower, prolonged phosphorus fertilisation tends to a buildup of soil phosphorus fractions, increasing the risk of phosphorus loss and environmental contamination [29].

The RF model effectively captured regional heterogeneity in SOC and nutrient responses to nitrogen addition, providing valuable insights into the spatial variability of soil fertility dynamics. In arid and semi-arid regions, the model predicted the most pronounced SOC increases, likely due to initially low baseline carbon stocks, which make these areas particularly responsive to fertilisation inputs. In contrast, phosphorus accumulation was particularly pronounced in historically phosphorus-deficient soils, where long-term fertilisation and manure applications have led to soil phosphorus saturation, further reinforcing the need for site-specific phosphorus management strategies. Patterns of nitrate accumulation varied across agroecological zones, with regions characterised by high precipitation and sandy soils exhibiting lower NO_3^--N retention. This suggests that these areas may be more susceptible to nitrate leaching losses, highlighting the importance of integrating hydrological considerations into precision nitrogen management. By integrating climate, soil, and management data, the RF model offers a robust framework for optimising fertiliser application rates and spatially tailoring nutrient management practices. These predictive capabilities are particularly relevant for mitigating environmental risks associated with nitrogen losses, including groundwater contamination and nitrous oxide emissions.

Net GHG balance and policy implication

While our findings confirm that nitrogen addition can enhance SOC and nutrient availability, it is critical to evaluate these changes in the context of the full greenhouse gas (GHG) balance. SOC sequestration, when considered in isolation, may offer a partial or even misleading view of the climate impact of fertilisation. Given that N_2O has a global warming potential 298 times that of CO_2 over a 100-year horizon [55], even modest emission increases can negate the benefits of SOC storage. Shcherbak et al. [56] demonstrated a nonlinear relationship between N input and N_2O emissions, with sharp increases beyond agronomic optimums. Moreover, in flooded rice systems, nitrogen inputs—especially when paired with organic amendments like manure or straw—can stimulate methane (CH_4) emissions due to enhanced substrate availability for methanogenic microbes. Therefore, despite the observed benefits of organic and integrated fertilisation on enhancing soil carbon and nutrient dynamics, their application in anaerobic paddy soils may simultaneously increase CH_4 emissions [57, 58]. This trade-off was emphasised in Zou et al. [59] and van Groenigen et al. [60], where CH_4 fluxes offset or exceed gains from SOC accumulation. These findings reinforce the necessity of a whole-system GHG accounting framework, including CO_2, N_2O, and CH_4, to guide sustainable nitrogen use. Ignoring these interactions may lead to overly optimistic assessments of climate mitigation potential and encourage fertiliser practices that ultimately increase net GHG emissions. Integrating SOC gains with gas flux monitoring, yield impacts, and environmental co-benefits will be essential for climate-smart nitrogen strategies.

Despite the effectiveness of machine learning in predicting SOC and nutrient changes, the RF model has certain limitations that warrant further refinement. The current approach does not explicitly account for greenhouse gas emissions, such as N_2O, which are critical for evaluating the full environmental trade-offs of nitrogen fertilisation. Additionally, the model does not incorporate soil acidification risks, a key factor influencing long-term soil fertility and productivity under intensive nitrogen application. Microbial-mediated nutrient transformations, which play a central role in nitrogen cycling, are also not directly modelled, limiting our ability to fully capture the biogeochemical processes driving nutrient dynamics. Future research should integrate real-time greenhouse gas flux measurements, remote sensing data, and microbial community analyses to improve the predictive capacity of machine learning models in soil fertility assessments. Incorporating time-series modelling approaches could also enhance our understanding of the long-term sustainability of nitrogen-induced SOC and nutrient changes, allowing for more accurate predictions of future soil health trajectories. By advancing these methodologies, machine learning-driven soil modelling can further contribute to the development of precision agriculture strategies that balance productivity, soil conservation, and environmental sustainability.

From a carbon balance perspective, practical farmland management strategies should aim to reduce carbon emissions while enhancing soil carbon sequestration. Minimizing tillage intensity reduces CO_2 release from soil organic matter decomposition and protects soil structure, thereby lowering microbial respiration losses [61]. The retention of crop residues on-site serves as a continuous carbon input and fosters soil aggregation, which enhances the physical protection of organic carbon [62]. The use of cover crops during fallow periods increases belowground biomass input and improves soil carbon storage, while also reducing erosion-related carbon losses [63]. Application of organic amendments with a high C: N ratio—such as compost or biochar—not only contributes to long-term carbon stabilization in soil but also promotes microbial assimilation of nitrogen, thereby lowering N_2O emissions [64]. In parallel, optimizing fertilization timing and dose through precision agriculture technologies reduces energy-intensive inputs and associated upstream emissions. Together, these practices form an integrated pathway for climate-smart agriculture, contributing to both carbon neutrality goals and the long-term sustainability of agroecosystems.

Conclusion

This study provides comprehensive, nationwide evidence that nitrogen addition significantly enhances soil organic carbon (SOC) sequestration and nutrient accumulation in Chinese croplands. By synthesising data from 479 field sites and employing machine learning modelling, we demonstrate that nitrogen-induced increases in SOC, total nitrogen, nitrate, ammonium, and available phosphorus vary according to climate zone, fertiliser type, and application duration. The responses are dose-dependent and regionally heterogeneous, with particularly strong carbon sequestration potential observed in arid regions. Organic and integrated fertilisers confer greater benefits than chemical fertilisers, while long-term fertilisation further amplifies the positive effects on soil health. The application of machine learning enables precise predictions of nitrogen addition outcomes across environmental gradients, offering a powerful tool for guiding sustainable nitrogen management and optimising fertiliser strategies. While this study offers valuable insights, several limitations warrant consideration. The study does not explicitly account for greenhouse gas (GHG) emissions, such as nitrous oxide (N_2O), which play a critical role in assessing the full environmental trade-offs of nitrogen fertilisation. Given the potential for increased N_2O emissions to offset the carbon sequestration gains observed in SOC, future research should integrate GHG flux measurements to provide a more holistic evaluation of nitrogen management strategies. Addressing these emissions is essential to ensuring that fertilisation practices contribute to both enhanced soil health and broader climate mitigation goals. Overall, these findings offer a robust scientific foundation for enhancing the carbon sink capacity of croplands and developing precision agricultural strategies that align with carbon neutrality and food security objectives.

Acknowledgements
This work was supported by the Xinjiang Forestry and Grassland Administration (Fund No.: XJLCKJ-2025-08).

Author contributions
Conceptualization, Yu L.; methodology, Yu L.; software, Yu L.; formal analysis, Yu L.; investigation, Yu L.; resources, Yuan L.; data curation, Yuan L.; writing—original draft preparation, Yu L.; writing—review and editing, Yuan L.; visualization, Yu L.; project administration, Yuan L.; funding acquisition, Yuan L. All authors have read and agreed to the published version of the manuscript.

Funding
This work support by the Xinjiang Forestry and Grassland Administration (Fund No.: XJLCKJ-2025-08).

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics and consent to participate declarations
not applicable.

Consent to publish
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 April 2025 / Accepted: 13 June 2025
Published online: 19 June 2025

References

1. Vitousek PM, Naylor R, Crews T, David MB, Drinkwater L, Holland E, Johnes P, Katzenberger J, Martinelli LA, Matson P. Nutrient imbalances in agricultural development. Science. 2009;324:1519–20.
2. Zhang F, Chen X, Vitousek P, Shen J. Fertilizer nitrogen recovery efficiencies in crop production systems of China with and without consideration of gaseous losses. Glob Biogeochem Cycles 35 (2021).

3. Chen D, Li J, Lan Z, Hu S, Bai Y. Soil microbial community and functional responses to long-term nitrogen addition across a grassland biome. Glob Ecol Biogeogr. 2017;26:1450–61.
4. Aber JD, McDowell W, Nadelhoffer KJ, Magill A, Berntson G, Kamakea M, Melillo JM. Nitrogen saturation in temperate forest ecosystems. Bioscience. 1998;48:921–34.
5. LeBauer DS, Treseder KK. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecology. 2008;89:371–9.
6. Lu M, Yang Y, Luo Y. Responses of ecosystem carbon cycle to nitrogen addition: A meta-analysis. New Phytol. 2011;189:1040–50.
7. Zhang Y, Zhang J, Zhang W. Long-term nitrogen addition affects soil carbon and microbial community in a wheat–maize rotation system. Agric Ecosyst Environ. 2020;300:107019.
8. Zhou J, Xue K, Xie J, Deng Y, Wu L, Cheng X, Luo Y. Microbial mediation of soil carbon storage under nitrogen addition: evidence from long-term field experiments. ISME J. 2018;12:1289–99.
9. Li H, Zhang X, Wang Q, Luo Y. Interactions between nitrogen and phosphorus cycles in agricultural soils under long-term fertilization. Nutr Cycl Agrosyst. 2021;121:321–35.
10. Chen W, Zhang Q, Liu X, Wang C. Coupled responses of microbial community and enzyme activities to nitrogen addition and phosphorus limitation in agricultural soils. Soil Biol Biochem 176 (2023).
11. Xu X, Liu W, Zhang C, Chen J. Limited extrapolation from plot-level nitrogen addition to regional carbon models. Glob Change Biol. 2018;24:5014–26.
12. Zhao W, Li B, Chen X, Liu L. Temporal and spatial limitations in nitrogen addition studies: A critical review. Soil Tillage Res 199 (2020).
13. Sihi D, Dari B, Kuruvila AP, Jha G, Basu K. Explainable machine learning approach quantified the long-term (1981–2015) impact of climate and soil properties on yields of major agricultural crops across conus. Front Sustainable Food Syst. 2022;6:847892.
14. Shen Y, Wang Y, Chen Z, Liu L. Mapping soil organic carbon using machine learning and big data: A review. CATENA 208 (2022).
15. Zhang Y, Yu G. Advances in artificial intelligence for soil biogeochemistry modeling. Environ Model Softw 162 (2023).
16. Zayani H, Fouad Y, Michot D, Kassouk Z, Baghdadi N, Vaudour E, Lili-Chabaane Z, Walter C. Using machine-learning algorithms to predict soil organic carbon content from combined remote sensing imagery and laboratory vis-NIR spectral datasets. Remote Sens. 2023;15:4264.
17. Elrys AS, Wang J, Meng L, Zhu Q, El-Sawy MM, Chen Z, Tu X, El-Saadony MT, Zhang Y, Zhang J, Cai Z, Müller C, Cheng Y. Integrative knowledge-based nitrogen management practices can provide positive effects on ecosystem nitrogen retention. Nat Food. 2023;4:1075–89.
18. Hedges LV, Gurevitch J, Curtis PS. The meta-analysis of response ratios in experimental ecology. Ecology. 1999;80:1150–6.
19. Breiman L. Random forests. Mach Learn. 2001;45:5–32.
20. Huang S, Peng X, Huang Q, Zhang W. Soil aggregation and organic carbon fractions affected by long-term fertilization in a semi-arid cropland of Northwest China. Geoderma. 2018;319:113–20.
21. Wang J, Fu B, Qiu Y, Chen L. Soil carbon accumulation following nitrogen addition depends on aridity: a meta-analysis. Land Degrad Dev. 2020;31:1679–90.
22. Yue K, Peng Y, Peng C, Yang W, Wu F. Stimulation of terrestrial ecosystem carbon storage by nitrogen addition: a meta-analysis. Sci Rep. 2016;6:19895.
23. Zimmermann M, Leifeld J, Schmidt M, Smith P, Fuhrer J. Measured soil organic matter fractions can be related to pools in the RothC model. Eur J Soil Sci. 2007;58:658–67.
24. Booth MS, Stark JM, Rastetter E. Controls on nitrogen cycling in terrestrial ecosystems: a synthetic analysis of literature data. Ecol Monogr. 2005;75:139–57.
25. Knoepp JD, Swank WT. Using soil temperature and moisture to predict forest soil nitrogen mineralization. Biol Fertil Soils. 2002;36:177–82.
26. Wang J, Wang X, Xu M, Feng G, Zhang W, Lu Ca. Crop yield and soil organic matter after long-term straw return to soil in China. Nutr Cycl Agrosyst. 2015;102:371–81.
27. Zhang Y, Hao X, Alexander TW, Thomas BW, Shi X, Lupwayi NZ. Long-term and legacy effects of manure application on soil microbial community composition. Biol Fertil Soils. 2018;54:269–83.
28. Sun R, Guo X, Wang D, Chu H. Effects of long-term application of chemical and organic fertilizers on the abundance of microbial communities involved in the nitrogen cycle. Appl Soil Ecol. 2015;95:171–8.
29. Chen X, Yan X, Wang M, Cai Y, Weng X, Su D, Guo J, Wang W, Hou Y, Ye D, Zhang S, Liu D, Tong L, Xu X, Zhou S, Wu L, Zhang F. Long-term excessive phosphorus fertilization alters soil phosphorus fractions in the acidic soil of pomelo orchards. Soil Tillage Res. 2022;215:105214.
30. Ghosh A, Bhattacharyya R, Meena MC, Dwivedi BS, Singh G, Agnihotri R, Sharma C. Long-term fertilization effects on soil organic carbon sequestration in an inceptisol. Soil Tillage Res. 2018;177:134–44.
31. Xu X, Ouyang H, Kuzyakov Y. Response of soil microbial community to nitrogen fertilizer application in croplands: A meta-analysis. Glob Ecol Biogeogr. 2019;28:1336–48.
32. Treseder KK. Nitrogen additions and microbial biomass: A meta-analysis of ecosystem studies. Ecol Lett. 2008;11:1111–20.
33. Liang C, Schimel JP, Jastrow JD. The importance of anabolism in microbial control over soil carbon storage. Nat Microbiol. 2017;2:17105.
34. Li J, Wen Y, Li X, Li Y, Yang X, Lin Z, Song Z, Cooper JM, Zhao B. Soil labile organic carbon fractions and soil organic carbon stocks as affected by long-term organic and mineral fertilization regimes in the North China plain. Soil Tillage Res. 2018;175:281–90.
35. Sun Y, He XZ, Hou F, Wang Z, Chang S. Grazing increases litter decomposition rate but decreases nitrogen release rate in an alpine meadow. Biogeosciences. 2018;15:4233–43.
36. Bai E, Li S, Xu W. A meta-analysis of experimental warming effects on terrestrial nitrogen pools and dynamics. New Phytol. 2013;199:441–51.
37. Gao Y, Liu H. Responses of soil nitrogen fractions and microbial community to long-term fertilization in an arid region. Soil Tillage Res. 2017;165:68–74.
38. Templer PH, Mack MC, Hollinger DY. Influence of nitrogen additions on N cycling and retention in forest soils. Biogeochemistry. 2012;107:359–76.
39. Vitousek PM, Menge DNL, Reed SC, Cleveland CC. Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. Philosophical Trans Royal Soc B. 2010;365:1215–25.
40. Chen J, He J. Long-term nitrogen fertilization increases soil available phosphorus and decreases microbial diversity in paddy soils. Soil Biol Biochem. 2018;122:91–9.
41. Zhang J, Zhou S, Sun H, Lü F, He P. Three-year rice grain yield responses to coastal mudflat soil properties amended with straw Biochar. J Environ Manage. 2019;239:23–9.
42. van Paassen JG, Britton AJ, Mitchell RJ, Street LE, Johnson D, Coupar A, Woodin SJ. Legacy effects of nitrogen and phosphorus additions on vegetation and carbon stocks of upland heaths. New Phytol. 2020;228:226–37.
43. Wang Q-q, Liu L-l, Li Y, Qin S, Wang C-j, Cai A-d, Wu L, Xu M, Zhang W-j. Long-term fertilization leads to specific PLFA finger-prints in Chinese hapludults soil. J Integr Agric. 2020;19:1354–62.
44. Xu X, Wang J, Niu Y, Jiang W, Wang Y, Liu S, Wei W. 44-years of fertilization altered soil microbial community structure by changing soil physical, chemical properties and enzyme activity. J Soil Sci Plant Nutr. 2024;24:3150–61.
45. Lu M, Zhou X, Luo Y, Yang Y, Fang C, Chen J, Li B. Minor stimulation of soil carbon storage by nitrogen addition: a meta-analysis. Agric Ecosyst Environ. 2011;140:234–44.
46. Tian D, Niu S, Pan Q, Ren T, Chen S, Bai Y, Han X. Nonlinear responses of ecosystem carbon fluxes and stoichiometry to nitrogen addition: A meta-analysis. Glob Change Biol. 2016;22:934–43.
47. Jeong J, Resop JP. Random forest for Spatial prediction of soil properties using environmental covariates. Geoderma. 2019;338:49–60.
48. Guo JH, Liu XJ, Zhang Y, Shen JL, Han WX, Zhang WF, Christie P, Goulding KWT, Vitousek PM, Zhang FS. Significant acidification in major Chinese croplands. Science. 2010;327:1008–10.
49. Tian D, Niu S. A global analysis of soil acidification caused by nitrogen addition. Environ Res Lett. 2015;10:024019.
50. Zhang X, Davidson EA, Mauzerall DL, Searchinger TD, Dumas P, Shen Y. Managing nitrogen for sustainable development. Nature. 2019;568:516–20.
51. Zhou J, Delgado-Baquerizo M, Maestre FT, Gao Y. Climate and soil pH predict the Spatial variation in ecosystem multifunctionality in drylands across the Globe. Ecol Lett. 2017;20:540–50.
52. Quemada M, Baranski M, Nobel-de Lange M, Vallejo A, Cooper J. Meta-analysis of strategies to control nitrate leaching in irrigated agricultural systems and their effects on crop yield. Agric Ecosyst Environ. 2013;174:1–10.
53. Aulakh M, Garg A, Kabba B. Phosphorus accumulation, leaching and residual effects on crop yields from long-term applications in the subtropics. Soil Use Manag. 2007;23:417–27.
54. Lu X, Mahdi A-K, Han X-z, Chen X, Yan J, Biswas A. -x. Zou, Long-term application of fertilizer and manures affect P fractions in Mollisol. Sci Rep. 2020;10:14793.
55. IPCC CC. 2013: The physical science basis, Cambridge Univ Press, CAMBS, England, 2014.

56. Shcherbak I, Millar N, Robertson GP. Global meta-analysis of the nonlinear response of soil nitrous oxide (N_2O) emissions to fertilizer nitrogen. Proc Natl Acad Sci. 2014;111:9199–204.
57. You X, Wang S, Du L, Wu H, Wei Y. Effects of organic fertilization on functional microbial communities associated with greenhouse gas emissions in paddy soils. Environ Res. 2022;213:113706.
58. Ye R, Jin Q, Bohannan BJM, Keller JK, McAllister SA, Bridgham SD. Methane dynamics in wetland soils: the role of microbial communities and their interactions with environmental factors. Glob Change Biol. 2021;27:313–31.
59. Zou J, Huang Y, Jiang J. A 3-year field measurement of methane and nitrous oxide emissions from rice paddies in China. Glob Biogeochem Cycles 19 (2005).
60. van Groenigen JW, Velthof GL, Oenema O. Towards an agronomic assessment of N_2O emissions. Eur J Soil Sci. 2011;62:486–98.
61. Lal R. Soil carbon sequestration impacts on global climate change and food security. Science. 2004;304:1623–7.
62. Smith P, Martino D, Cai Z, Gwary D, Janzen H, Kumar P, McCarl B, Ogle S, O'Mara F, Rice C. Greenhouse gas mitigation in agriculture. Philosophical Trans Royal Soc B: Biol Sci. 2008;363:789–813.
63. Poeplau C, Don A. Carbon sequestration in agricultural soils via cultivation of cover crops– A meta-analysis. Agric Ecosyst Environ. 2015;200:33–41.
64. Lehmann J, Joseph S. Biochar for environmental management: science, technology and implementation. Taylor & Francis; 2024.

Publisher's note