Received: 8 July 2021 | Revised: 28 February 2022 | Accepted: 28 February 2022

DOI: 10.1111/gcb.16154

Global Change Biology WILEY

RESEARCH ARTICLE

Decipher soil organic carbon dynamics and driving forces across China using machine learning

Huiwen Li[1,2] | Yiping Wu[1,3] | Shuguang Liu[4] | Jingfeng Xiao[5] | Wenzhi Zhao[6] | Ji Chen[7] | Georgii Alexandrov[8] | Yue Cao[9]

[1]Department of Earth & Environmental Science, Xi'an Jiaotong University, Xi'an, China

[2]Key Laboratory of Degraded and Unused Land Consolidation Engineering, The Ministry of Natural Resources of China, Xi'an, China

[3]Technology Innovation Center for Land Engineering and Human Settlements, Shaanxi Land Engineering Construction Group Co. Ltd, Xi'an Jiaotong University, Xi'an, China

[4]National Engineering Laboratory for Applied Technology of Forestry and Ecology in South China, Central South University of Forestry and Technology, Changsha, China

[5]Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, New Hampshire, USA

[6]Key Laboratory of Ecohydrology and River Basin Science, Northwest Institute of Eco-environment and Resources, Chinese Academy of Sciences, Lanzhou, China

[7]Department of Civil Engineering, The University of Hong Kong, Hong Kong, China

[8]A. M. Obukhov Institute of Atmospheric Physics, Russian Academy of Sciences, Moscow, Russian Federation

[9]Xi'an Institute for Innovative Earth Environment Research, Xi'an, China

Correspondence
Yiping Wu, Technology Innovation Center for Land Engineering and Human Settlements, Shaanxi Land Engineering Construction Group Co. Ltd, Xi'an Jiaotong University, Xi'an, China.
Email: yipingwu@xjtu.edu.cn

Funding information
National Science Foundation of China, Grant/Award Number: U20A2089; the National Key Research and Development Program of China, Grant/Award Number: 2019YFC0507403; the National Science Foundation of China, Grant/Award Number: 31961143011; the Fundamental Research Funds for the Central Universities, Grant/Award Number: xzy022020008; the Strategic Priority Research Program of Chinese Academy of Sciences, Grant/Award Number: XDB40020205; Shaanxi Key Research and Development Program, Grant/Award Number: 2022ZDLSF06-04

Abstract

The dynamics of soil organic carbon (SOC) play a critical role in modulating global warming. However, the long-term spatiotemporal changes of SOC at large scale, and the impacts of driving forces remain unclear. In this study, we investigated the dynamics of SOC in different soil layers across China through the1980s to 2010s using a machine learning approach and quantified the impacts of the key factors based on factorial simulation experiments.Our results showed that the latest (2000–2014) SOC stock in the first meter soil (SOC_{100}) was 80.68 ± 3.49 Pg C, of which 42.6% was stored in the top 20 cm, sequestrating carbon with a rate of 30.80 ± 12.37 g C m^{-2} yr^{-1}since the 1980s. Our experiments focusing on the recent two periods (2000s and 2010s) revealed that climate change exerted the largest relative contributions to SOC dynamics in both layers and warming or drying can result in SOC loss. However, the influence of climate change weakened with soil depth, while the opposite for vegetation growth. Relationships between SOC and forest canopy height further confirmed this strengthened impact of vegetation with soil depth and highlighted the carbon sink function of deep soil in mature forest. Moreover, our estimates suggested that SOC dynamics in 71% of topsoil were controlled by climate change and its coupled influence with environmental variation (CE). Meanwhile, CE and the combined influence of climate change and vegetation growth dominated the SOC dynamics in 82.05% of the first meter soil. Additionally, the national cropland topsoil organic carbon increased with a rate of 23.6 ± 7.6 g C m^{-2} yr^{-1}since the 1980s, and the widely applied nitrogenous fertilizer was a key stimulus. Overall, our study extended the knowledge

about the dynamics of SOC and deepened our understanding about the impacts of the primary factors.

KEYWORDS
climate change, factorial simulation experiments, fertilization, random forest, SOC, vegetation growth

1 | INTRODUCTION

Soil organic carbon (SOC) is the largest component of the terrestrial organic carbon pool (Batjes, 1996), playing an essential role in the global carbon cycle and exerting profound feedback to climate change (Bradford et al., 2016; Davidson & Janssens, 2006; Prietzel et al., 2016). Furthermore, abundant organic carbon in soil is beneficial for soil fertility and quality, safeguarding food security and water retention (Lal, 2004; Zhao et al., 2018). Hence, increasing SOC stocks is widely considered as an effective approach to mitigate global warming and improve ecosystem services (Meyer et al., 2018; Prietzel et al., 2016). SOC dynamics are determined by the long-term balance between inputs (i.e., aboveground litter or root exudates) through the vegetation net primary production and removals via a series of geophysical and biochemical processes (Plante & Conant, 2014). These processes closely interact with climate change, vegetation growth, environmental variation, and even human activities (Jobbágy & Jackson, 2000), making SOC (especially that in topsoil) be highly sensitive to changes in abiotic and biotic factors (Gaitan et al., 2019). As the largest organic carbon pool, the variations in SOC, even with a slight magnitude, may lead to critical disturbance to the global carbon cycle and climate (Paustian et al., 2019; Ramesh et al., 2019). However, our understanding about the long-term spatiotemporal distributions of SOC and contributions of different drivers over large spatial domains are still limited.

Assessment of the SOC dynamics at large scales is pivotal to understand the global carbon cycle, which has attracted tremendous efforts in investigating SOC densities and stocks at global, national, and regional scales. Traditionally, two major approaches have been used to estimate SOC stocks over large spatial domains. The first way is to upscale the site-level soil profiles to regional stocks based on the area-weighted average method under the spatial constraints of soil taxonomic units (or other scope constraints). Substantial differences were found in different studies—SOC stock in the global first meter was estimated to vary from 504 to 3000 petagrams of carbon (Pg C; 1 Pg C = 10^{15} g C) (Batjes, 1996; Jobbágy & Jackson, 2000; Scharlemann et al., 2014). Large differences were also reported at regional scales. For example, SOC stock in the first meter soil in China was estimated to be in the range of 69.1 to 185.7 Pg C (Ni, 2001; Tang et al., 2018; Wang & Zhou, 1999; Wang et al., 2000; Xu et al., 2018; Yang et al., 2007; Yu et al., 2007, 2010). The differences in these estimates at different scales may result from distinct soil profile sizes, locations, methods, and periods of sampling. The area-weighted average method usually shows a poor spatial representation and coarse resolution (Stockmann et al., 2013). The second widely used method to investigate the SOC dynamics is the biogeochemical modeling built on prior knowledge (Allison et al., 2010; Franko, 1996; Li et al., 1992; Liu et al., 2003; Parton et al., 1998; Wieder et al., 2013). Estimates from different models usually present substantial divergences (both in magnitudes and trends) mainly caused by different model mechanisms (or structures), incomplete specification of processes, and key inputs and parameters (Zhang et al., 2017). Model calibration that relies on long-term observations plays a key role in influencing the capability of these process models and is the prerequisite for accurate simulations of SOC stocks and dynamics (Hui et al., 2020; Stockmann et al., 2013).

In recent years, growing studies using state-of-the-art data-driven methods (or machine learning approaches) to estimate the spatiotemporal patterns of SOC have emerged and showed promising results (Padarian et al., 2020; Xiao et al., 2019). In terms of the spatial patterns of SOC, for instance, the global SOC density was mapped at a high spatial resolution (250 m × 250 m) (SoilGrids) using a random forest (RF) model (Hengl et al., 2017). The Food and Agriculture Organization (FAO) of the United Nations generated the Global Soil Organic Carbon Map (GSOCmap) using several data-driven approaches (FAO & ITPS, 2018). There are also a large number of studies that utilized these state-of-the-art methods to map SOC densities at landscape or regional scales (de Anta et al., 2020; Gomes et al., 2019; Zhou et al., 2019b, 2020). As for the dynamics of SOC at large spatial domain, the research community usually use independent data-driven models trained by samples in different periods to predict the SOC maps in the corresponding period and evaluate the changes between these maps (Ding et al., 2016; Liu et al., 2018; Yang et al., 2014; Zhou et al., 2019b). However, these models lack the temporal expansion ability, and a major reason is the lack of detailed time information with the measurements used in the model training. Over the past decades, substantial efforts have been made to survey and collect soil profile information across the country. This provides us a great opportunity to investigate the spatiotemporal dynamics of SOC using a data-driven model trained by time intensive samples.

Deciphering the impacts of different biotic, abiotic, and anthropogenic factors on SOC dynamics are critical to understand how SOC will respond to future climate change, environmental variation, and anthropogenic disturbances (Rossel et al., 2019). There is a broad consensus that the dynamics of SOC are largely controlled by the carbon input and carbon loss processes, which are closely associated with biotic, abiotic, and anthropogenic drivers (Jackson et al., 2017). Soil physicochemical properties are strongly related

to SOC stocks at different scales and climatic regimes (Wiesmeier et al., 2019), due to their decisive function on the diversity and abundance of soil microorganisms that can notably influence the decomposition rate of organic matter. Biogeochemical processes related to SOC dynamics are substantially affected by climate change. Taking temperature as an example, warming can increase soil carbon input via promoting vegetation productivity, but it can also promote microbial activity and accelerate carbon decomposition (Prietzel et al., 2016; Wang et al., 2020b). Hydrological condition is another important factor controlling the SOC dynamics (Suseela et al., 2012; Wang et al., 2019; Zhou et al., 2019a). Vegetation growth directly controls the carbon source of both soil surface through litter production and deep soil layer via root growth, exudation, and turnover, hence playing an essential role in controlling SOC distribution and dynamics (Plante & Conant, 2014). Environmental variations such as enriched CO_2 and nitrogen deposition may also exert substantial impacts on SOC dynamics (Cheng et al., 2012; van Groenigen et al., 2014; Liu et al., 2005). Anthropogenic activities can bring about great disturbance on SOC balance, and among all these kinds of drivers, land-use change may be the most significant factor of SOC destabilization (Poeplau et al., 2011; Wiesmeier et al., 2019). Agricultural management practices such as nitrogenous (N) fertilizer addition are also important for the balance of SOC. China has experienced significant variation ($p < .01$) in precipitation, temperature, and vegetation growth during the last few decades (Li et al., 2021; Liang et al., 2015; Yao et al., 2018). Previous assessments based on site-scale measurements usually revealed the correlations between the spatial pattern of SOC and some influencing drivers. However, our knowledge about the long-term SOC dynamics in different soil depths at the national scale under comprehensive influences of these factors is uncertain, not to mention the controlling drivers and their contributions.

To understand the dynamics of SOC pool in China and the impacts of the key drivers, we developed SOC maps in the topsoil (0–20 cm) and the first meter soil at 1km by 1km spatial resolution over China for the periods of 1980s, 2000s, and 2010s based on a machine learning approach and extensive measurements. We analyzed the spatiotemporal dynamics of SOC in the two soil layers and further quantified the contributions of climate change, vegetation growth, and environmental variation to SOC dynamics based on factorial analysis experiments. More importantly, we presented the spatially explicit estimates of the dominant drivers for SOC dynamics in both soil layers.

2 | MATERIALS AND METHODS

2.1 | Soil organic carbon samples

Soil profile investigations over China during the early 1980s (1979~1985) were collected from the Second National Soil Survey of the country (National Soil Survey Office, 1993, 1994, 1995, 1996). This nationwide soil survey provided the most comprehensive soil property information of the nation. For each sampling site, the SOC density (kg C m^{-2}) was calculated by the following equation:

$$SOC = \sum_{i=1}^{n} D_i \times B_i \times SOCC_i \times (1 - C_i/100)/100 \quad (1)$$

where D_i, B_i, and C_i are soil depth (cm), bulk density (g cm^{-3}), and volume fraction (%) of the rock fragments with a particle size>2 mm in the soil horizon for each soil layer (i), respectively. SOCC is soil organic carbon content (g kg^{-1}), which is calculated by multiplying soil organic matter (SOM) by 0.58. A total number of 2203 soil samples in the top 20 cm soil (topsoil, SOC_{20}) and 2436 soil samples in the top 100 cm (SOC_{100}) soil layers were obtained from this soil survey. The SOC measurements in the topsoil and the top 100 cm soil layers during the period of 2000 to 2014 were obtained from an open access database (Xu et al., 2018). This publicly available SOC data set contains 4515 soil samples for the topsoil and 3026 soil samples for the top 100 cm soil layer across China during this period, and those samples were collected from *in-situ* investigations of 1036 publicly published papers and database. Most soil samples (>90% of topsoil samples and >81% of samples in the top 100 cm soil) contain the specific sampling time and some of them were annually-continuously monitored (>3 years). All collected samples were widely distributed in China and their locations encompassed the five main ecosystems in China, with over 35% of samples in cropland, followed by forest (~28%) and grassland (~20%) (Figure 1a,b).

2.2 | Soil organic carbon estimation based on the random forest model

2.2.1 | Splitting the training and testing time-series samples

In this study, the RF model was selected to estimate the spatiotemporal dynamics of SOC_{20} and SOC_{100}. RF is an ensemble of decision tree predictors, which uses the bootstrap resampling method to build decision tree for each sampling (Breiman, 2001). For the construction of each tree, samples are independently selected; however, the distributions for all trees in the forest are the same, which guarantees the robustness of the model. A specific number of trees (here 1200) were constructed during the model training. Two RF models were built to simulate SOC in the topsoil and the top 100 cm soil, respectively. To validate the temporal prediction capability of a RF model, we split the time-series SOC samples into three different time periods—observations during the early 1980s (1979–1985), the 2000s (2000–2009), and the early 2010s (2010–2014). Observations during the first and last periods were used to train the two RF models, and the rest (those during the 2000s) were for validation. Under such a sample segmentation scheme, we can test whether the RF models can predict the temporal variation of SOC during these periods.

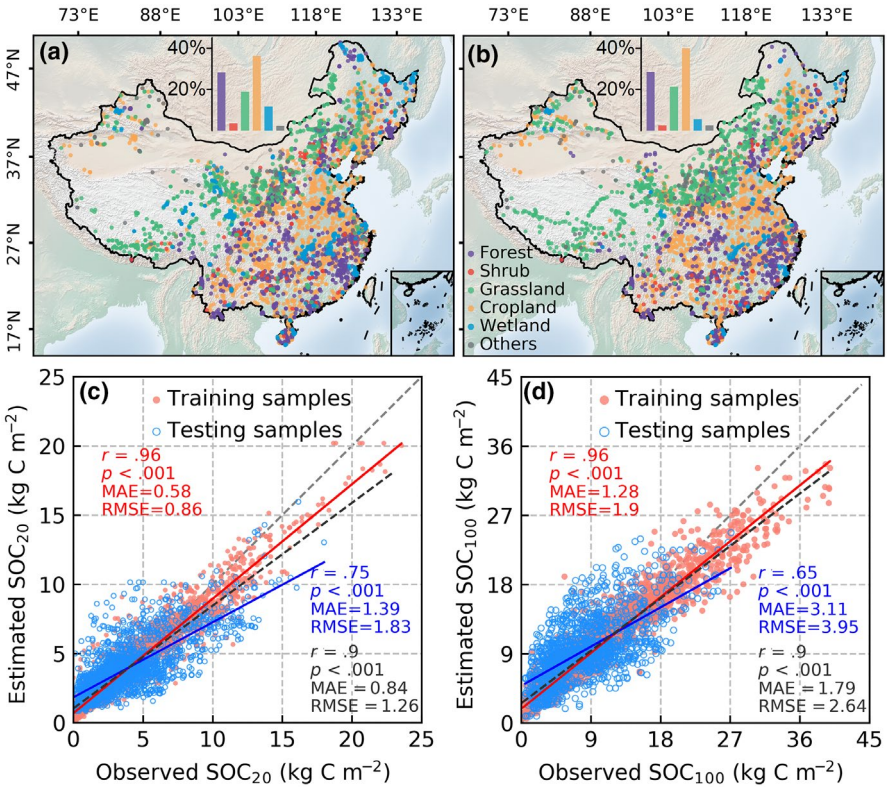

FIGURE 1 Locations of soil organic carbon (SOC) samples used in the study and comparison between predicted and observed SOC at the training and testing phases in the two soil layers. (a) and (b) are samples in the top 20 cm and top 100 cm soil layers in different ecosystems. Insets in (a) and (b) indicate the relative amount ratios of samples in different ecosystems. Measurements during the early 1980s (1979–1985) and 2010s (2010–2014) were used to train the random forest models and the measurements during the 2000s (2000–2009) were used to validate the spatiotemporal prediction ability of the models. The dashed gray line shows the 1:1 line. The red line shows the performance of the model during the training phase and the blue one indicates that during the testing phase. The black dashed line shows the prediction accuracy of the model compared to all measurements, including training and testing samples. *r*, Pearson correlation coefficient; MAE, mean absolute error; RMSE, root mean square error [Colour figure can be viewed at wileyonlinelibrary.com]

2.2.2 | Model driver selection

We conducted the collinearity and independence detections to determine the model drivers. Among the list of candidate drivers, one is opted out if its Pearson correlation coefficient with any other drivers was larger than 0.7 (Liu et al., 2022). Finally, 16 driving factors were determined to estimate the SOC over China (Table S1). These drivers involved five categories, including soil property, terrain feature, climate change, vegetation growth, and environmental variation. The data sources and detailed information about these drivers were described in the Supplementary Materials. We extracted the predictor variable (e.g., climate and land-use type) values linked to each sample, and the corresponding sampling time (a specific year or a period) was adopted as the temporal information. For the period 2000–2014, the exact sampling years for most SOC measurements were available. However, SOC measurements during the 1980s did not contain a specific sampling time, then we used the mean values of the time-variant drivers during this period (1979–1985). The land use and land cover (LULC) map during this period was replaced by that in 1980. The Net Primary Productivity (NPP) during this period was substituted by the average during 1982 and 1985. Similarly, for those samples with no exact sampling year (2000s and 2010s), we used the mean values of the time-variant drivers during that period (2000s or 2010s). When we extract the feature properties linked to SOC samples beside a water body, we built a buffer ring with a radius of 5km that covered the sample and then set each feature value as the average of all gird cell values or the most frequent value for soil types and LULC in the buffer. For extracting soil feature properties of different soil layers, the mean values of those in the corresponding soil depth were calculated. For the estimation of SOC_{100}, apart from the above drivers, the estimated SOC_{20} maps were also considered as a new driver to build the model (Figure S1a,b).

2.2.3 | Model construction scheme

When training the RF model, feature numbers between two and the maximum (i.e., 16) were randomly selected to find out the best split at each node, and the mean absolute error (MAE) of about one-third of the samples (out-of-bag, OOB) that were left out of the bootstrap sampling and decision tree construction was used to measure the

quality of the splits. After building all trees, the final predictions were determined by the weighted average outputs of all ensemble trees. The RF model generally shows a high ability in reducing the experimental noise and can enhance the prediction accuracy by aggregating the predictions of multiple trees, so that is why the RF model has been successfully used in mapping the SOC in many regions (Gomes et al., 2019; Hengl et al., 2017; Li et al., 2021; Wang et al., 2020b). As mentioned by those past studies, a RF model without a validation strategy for the time-varying prediction may only be able to simulate the spatial variability of SOC while hard to predict the temporal variation of SOC. In this study, our validation strategy of splitting the time-series samples into training (1980s and 2010s) and testing (2000s) parts can test the model ability in predicting SOC variations between time periods. The Python module "scikit-learn" (Pedregosa et al., 2011) was used to build the RF regression models for SOC simulations in the two soil layers.

2.2.4 | Model validation and the uncertainty detection

To validate the spatiotemporal prediction ability (or robustness) of the RF model, we adopted the nine-fold cross-validation method, which means the initial training samples during the 1980s and 2010s were equally divided into 9 parts. For each split, 8/9 of the initial training samples were used to build a new model and the rest 1/9 samples and those during the 2000s were used to validate the model. Then, we can have 9 groups of training and testing samples. It means the cross-validation method can detect not only the model performance in predicting the spatial variability of SOC but also the temporal variability of SOC across China, at least between time periods. Statistical indices of the Pearson correlation coefficient (r), coefficient of determination (R^2), the MAE, and the root mean square error (RMSE) were used to quantify the bias of the RF models. For the unbiasedness of model sampling, we detected the spatial distributions of the training and testing samples, as well as the density distributions of the corresponding SOC observations during the model construction. Combining the validation strategy for the time-varying prediction of the RF models and the nine-fold cross-validation, we can fully detect the spatiotemporal prediction ability of the RF model for SOC across China.

The verified RF models were then used to estimate SOC_{20} and SOC_{100} during the early 1980s, 2000s, and the early 2010s. Specifically, the RF models used the average maps of the time-variant variables during the 1979–1985 (except for the LULC and NPP, see section "Model driver selection") to simulate the SOC maps in the two soil layers during the 1980s. For the period of 2000–2014, firstly, the time-variant driver maps of each year were used to run the RF models, and the SOC maps in each year during this period were predicted. Then, the average SOC maps of the two soil layers during 2000–2009 and those during 2010–2014 were calculated as the final SOC maps for the periods of 2000s and early 2010s, respectively.

The uncertainties of our estimates during different periods were quantified using the standard deviations of the nine-fold cross-validation simulations. It is worth noting that the uncertainties here exhibit the robustness of the model during the spatial prediction, indicating the uncertainties caused by the random sampling. Specifically, for the periods of 2000s and 2010s, we first used the nine-fold cross-validation models to simulate the SOC maps in each year during each period. Second, the average of SOC maps in both layers estimated by all validation models during each period were used as the final SOC maps. Third, the estimation uncertainties during the 2000s and 2010s for each soil layer were quantified using the standard deviations of the nine simulations by the cross-validation models during these two periods. The uncertainties of the estimates during the 1980s were determined as the standard deviations of the nine-fold cross-validation model simulations. Apart from examining the prediction ability of the two RF models, we adopted several approaches to validate the accuracy of our estimates. First, we validated the spatial patterns of our estimates by comparing them with other widely used SOC maps (SoilGrids and GSOCmap). Second, we compared the magnitudes of our estimated SOC densities and stocks in different ecosystems with previous investigations. Third, variation trends of our predictions were also compared with previous investigations.

2.3 | Experimental design for driving factor analysis

2.3.1 | The contributions of time-variant driver categories to soil organic carbon dynamics

The dynamics of SOC among different periods were controlled by the three time-variant driver categories (i.e., climate change, vegetation growth, and environmental variation). Following a previous study (Yuan et al., 2019), we designed factorial experiments to isolate the impacts of a specific time-variant driver category on SOC in the two soil layers for the period of 2000–2014 (Figure S1c,d). In the experiment, we just kept one specific driver category (e.g., climate) constant at their initial states (data at the first year, here 2000) in a single experiment, while allowing the other driver categories vary with time as they are (i.e., the time-series in the real world) for each soil layer. For instance, the $SOC_{20\text{-}Cli}$ (or $SOC_{100\text{-}Cli}$) means that all drivers in the climate category remained constant at values in year 2000 but with real (time-variant) features of drivers in vegetation and environment categories. Similarly, $SOC_{20\text{-}Veg}$ (or $SOC_{100\text{-}Veg}$) estimation experiment just kept vegetation drivers constant at values in year 2000, and $SOC_{20\text{-}Env}$ (or $SOC_{100\text{-}Env}$) kept environmental drivers constant at values in year 2000.

By comparing the simulations between experiments, we could isolate the impact of a target driver category and quantify its contribution. To implement the comparison, we defined the differences between simulation driven by all variables in soil layer s and period y ($SOC_{s\text{-}All,y}$) and that driven by partial time-variant driver categories

($SOC_{s\text{-}Vi,y}$) as the actual contribution of the target variable *i* (*Vi*) in soil layer *s* and period *y* ($ActCon_{s\text{-}Vi,y}$):

$$ActCon_{s-Vi,y} = SOC_{s-All,y} - SOC_{s-Vi,y} \quad (2)$$

The relative contribution (%) of variable *i* in soil layer *s* and period *y* ($RelCon_{s\text{-}Vi,y}$) was defined as the following equation:

$$RelCon_{s-Vi,y} = \frac{\left|ActCon_{s-Vi,y}\right|}{\sum_{i}^{n}\left|ActCon_{s-Vi,y}\right|} \times 100\,\% \quad (3)$$

The grid annual mean relative contributions of the three driver categories in the two soil layers were used to derive the RGB (Red-Green-Blue) combinations to visualize the spatial patterns of the influential intensity of different driver categories. We defined the single dominant driver category of each pixel as the variable with a relative contribution larger than 60%. If the maximum relative contribution of any driver category is below this threshold, the dominant driver category of that pixel is defined as the combination of the two driver categories with larger relative contributions among the three driver categories. When we carried out the factorial analysis model experiments for SOC_{100}, the initial driver SOC_{20} was replaced by the new estimated $SOC_{20\text{-}Vi,y}$ (Figure S1d). Due to negligible SOC changes in desert area, the controlling driver analyses in this region are excluded.

2.3.2 | The impacts of key drivers on soil organic carbon dynamics

We adopted three ways to detect the impacts of the key drivers (i.e., warming, precipitation variation, vegetation growth, and fertilization) on SOC dynamics. For warming and precipitation variation, we conducted a sensitivity-analysis experiment to quantify the impacts of these two drivers on SOC dynamics through changing the magnitude of each driver during 2000–2014. We designed three scenarios: 20% more precipitation (P+20%), 20% less precipitation (P-20%), and a raise of 2°C in air temperature (T+2°C). The differences between the estimated SOC between actual situation and designed scenarios indicate the impacts of warming and precipitation variation on SOC dynamics. All estimates of the nine-fold cross-validation models were used to quantify the uncertainties as described in section "Model validation and the uncertainty detection". For the impact of vegetation growth, we used the forest canopy height map with a resolution of 1 km × 1 km (Simard et al., 2011) to investigate the relationship between forest growth and SOC dynamics in the two soil layers. In addition, China has experienced a rapid growing application of N fertilizer since 1979 (Figure S2). However, because the N fertilizer consumption during 2000–2014 was at a relatively stable level (22.85 million tons of N yr^{-1}) compared with the early 1980s (10.47 Mt N yr^{-1}), we compared the differences between cropland SOC_{20} during the 1980s and 2000–2014 to detect the impact of N fertilizer on cropland SOC across China.

3 | RESULTS

3.1 | Evaluation of the random forest models

The estimated and measured SOC densities during the model training phase (using samples during the early 1980s and 2010s) were significantly correlated, with a correlation coefficient (*r*) of 0.96 ($p < .001$) for both soil layers (Figure 1c,d). Positive correlations were also observed between the predicted and measured SOC densities during the model testing phase (using samples during the 2000s), with *r* being 0.74 and 0.65 for SOC_{20} and SOC_{100} ($p < .001$), respectively. Take all measurements as a whole, the predicted SOC densities were closely associated with the measurements for both soil layers, with a *r* of 0.9 ($p < .001$). The predictions and measurements of all samples were also distributed nearby the 1:1 lines, with the RMSE equaling to 1.26 and 2.64 kg C m^{-2} for SOC_{20} and SOC_{100}, respectively. These validations suggested that the trained RF models are capable to predict the spatiotemporal variability of SOC during the three periods (i.e., early 1980s, 2000s, and 2010s) over China.

The wide distributions of both sampling locations and SOC-density magnitudes (training and testing samples) (Figure S3) jointly guaranteed the capabilities of the RF models for estimating SOC for the two soil layers across the country. Based on the nine-fold cross-validation, we found that all robustness-detection models explained more than 75% of the inter-site variabilities of all SOC measurements across the country (Figure S4 and S5). The estimation error level is small (less than 1.38 and 2.9 kg C m^{-2} of RMSE for SOC_{20} and SOC_{100}, respectively) (Figure S4 and S5), indicating that the RF models were able to generate reliable predictions of SOC in both soil layers. The robustness of the models mainly benefited from the unbiasedness in sampling which was reflected in the widely distributed samples for both soil layers and SOC density distributions (Figure S6 and S7). With the validated RF models, we produced the SOC maps in the topsoil and the top 100 cm soil over China during the periods of early 1980s, 2000s, and 2010s. The uncertainties (standard deviations of the cross-validation model simulations) of the SOC estimates (Figure S8) showed that the high uncertainties were mainly occurred in the Qinghai-Tibetan Plateau for both layers where the samples were relatively sparse. Overall, the national average uncertainties of SOC_{20} and SOC_{100} are in the range of 0.16–0.18 kg C m^{-2} and 0.38–0.49 kg C m^{-2} during different periods, respectively (Table 1).

3.2 | Spatiotemporal patterns of soil organic carbon

The estimated SOC in the two soil layers exhibited similar spatial patterns during the study periods (Figure 2a,b; Figure S9), with

TABLE 1 Synthesis of soil organic carbon (SOC) densities and stocks in the two soil layers in China from this study and previous investigations

Soil layer (cm)	Period	Area (10^6 ha)	SOC stock (Pg C)	SOC density (kg C m^{-2}) [a]	Estimation methods[b]	References
0–20	1979–1992	901.63	33.74 ± 5.93	3.74 ± 0.66	Estimated based on 2473 soil profiles during China's second national soil survey	(Wang et al., 2004)
	1990s	881.81	32.54	3.69	Estimate based on published data from 34,411 soil profiles investigated during the China's second national soil survey	(Wu et al., 2003)
	2004–2014	925.64	34.32 ± 3.37	3.71 ± 0.36	Synthesis analysis of field surveys from multiple sources	(Xu et al., 2018)
	1980s	918.93	29.96 ± 1.65	3.26 ± 0.18	Random forest model estimates based on spatiotemporal *in-situ* observations in the topsoil during the 1980s (2203) and 2000–2014 (4515) from the Second National Soil Survey of China and a publicly available database	This study
	2000s		33.91 ± 1.47	3.69 ± 0.16		
	2010s		35.29 ± 1.56	3.84 ± 0.17		
	2000–2014		34.37 ± 1.47	3.74 ± 0.16		
0–30	1979–1985	880.37	32.93	3.7	Investigation based on 2473 soil profiles from China's second national soil survey and 810 soil profiles from 270 field observation sites	(Yang et al., 2007)
0–100	1970-1980s	945	185.7	19.7	Investigation by 745 soil profiles	(Fang et al., 1996)
	1970-1980s	959.63	119.76	12.48	Estimate based on soil carbon densities and area of different vegetation types	(Ni, 2001)
	1979–1984	877.63	92.42	10.53	Investigation based on 2473 soil profiles from China's second national soil survey	(Wang et al., 2000)
	1979–1985	880.37	69.1	7.8		(Yang et al., 2007)
	1979–1992	901.63	82.48 ± 19.46	9.15 ± 2.16		(Wang et al., 2004)
	1980s	925.45	100.18	10.81	Investigation based on 236 representative soil profiles from China's first national soil survey	(Shaoqiang Wang & Zhou, 1999)
	1980s	870.94	89.61	10.29	Investigation based on 2473 soil profiles from China's second national soil survey and publicly available literatures	(Xie et al., 2007)
	1980s	928	89.14	9.6	SOC was calculated based on gravel content, soil bulk density, and organic carbon content from 7292 soil profiles	(Yu et al., 2007)
	1980s–2000s	928.1	87.78 ± 9.98	9.46 ± 1.08	Based on soil map and SOC density from publicly available literatures	(Yu et al., 2010)
	1990s	881.81	70.31	7.97		(Wu et al., 2003)
	2004–2014	925.64	84.55 ± 8.09	9.13 ± 0.87		(Xu et al., 2018)
	2010–2015	946.4	74.98 ± 1.28	7.92 ± 0.14	Investigation based on an intensive field campaign involving 14,371 field plots over the country	(Tang et al., 2018)
	1980s	918.93	73.61 ± 4.5	8.01 ± 0.49	Random forest model estimates based on spatiotemporal *in-situ* observations in the top 100 cm soil during the 1980s (2436) and 2000–2014 (3026) from the Second National Soil Survey of China and a publicly available database	This study
	2000s		79.86 ± 3.58	8.69 ± 0.39		
	2010s		82.52 ± 3.77	8.98 ± 0.41		
	2000–2014		80.68 ± 3.49	8.78 ± 0.38		

[a]Estimation uncertainties of this study in different soil layers were determined as the standard deviations of the nine cross-validation model simulations. It is worth noting that the uncertainties here exhibit the robustness of the model during the spatial prediction, indicating the uncertainties caused by the random sampling.

[b]Estimation methods were presented where they first appeared.

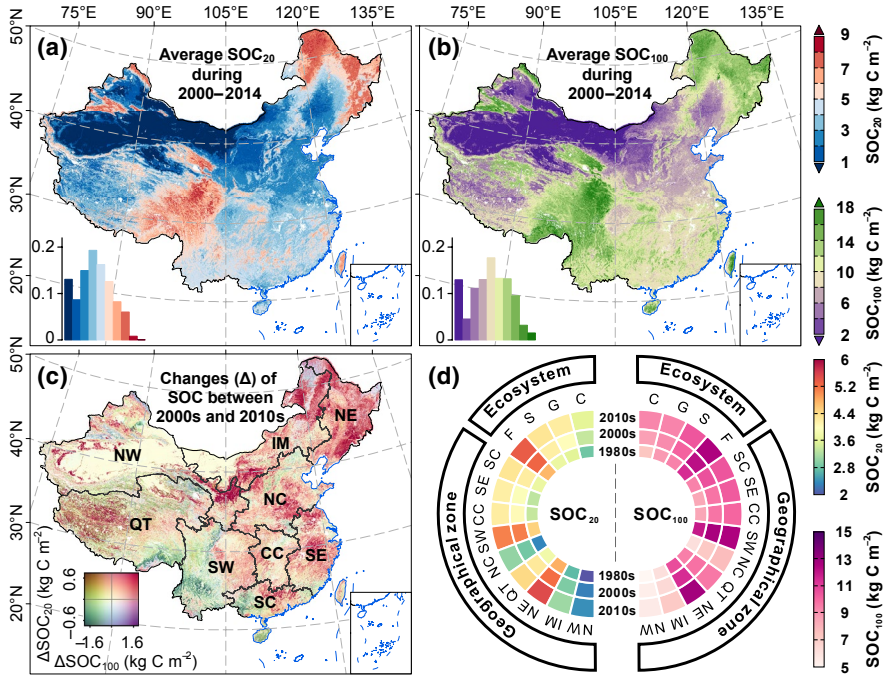

FIGURE 2 Spatial patterns and the changes of SOC in the two soil layers. (a) and (b) are the average SOC density maps in the top 20 cm (SOC_{20}) and top 100 cm soil (SOC_{100}) during the period 2000–2014. The spatial patterns of SOC during different periods are shown in Figure S7. (c) illustrates the spatial changes of SOC_{20} and SOC_{100} during the 2000s (2000–2009) and 2010s (2010–2014). The spatial changes of SOC during different periods are shown in Figure S8. (d) presents the average SOC_{20} and SOC_{100} in different ecosystems and geographical zones during the 1980s, 2000s, and 2010s. The left insets in (a) and (b) indicate the area ratios of different SOC levels. NW, Northwest China; IM, Inner Mongolia; NE, Northeast China; QT, Qinghai-Tibetan Plateau; NC, North China; SW, Southwest China; CC, Central China; SE, Southeast China; SC, South China; F, forest; S, shrub; G, grassland, and C, cropland [Colour figure can be viewed at wileyonlinelibrary.com]

high values in the eastern Qinghai-Tibetan Plateau (Mounts Qilian and Bayan Har in Qinghai), Mount Xing'an in Northeast China, and Mounts Tianshan and Alta in northern Xinjiang. The low SOC levels were found on the northwestern China, Loess Plateau, Inner Mongolia, the Northeast China plain, the North China Plain, and the western Qinghai-Tibetan Plateau.

The average densities of SOC_{20} and SOC_{100} over China during 2000–2014 were 3.74 ± 0.16 and 8.78 ± 0.38 kg C m^{-2}, respectively, suggesting an increase of 14.72% and 9.61% when compared to the 1980s (3.26 ± 0.18 kg C m^{-2} for SOC_{20} and 8.01 ± 0.49 kg C m^{-2} for SOC_{100}). The total SOC_{100} storage of China during 2000–2014 was estimated to be 80.68 ± 3.49 Pg C with 42.6% (34.37 ± 1.47 Pg C) stocking in the top 20 cm soil. Taking the middle year of 1982 (the mean of 1979–1985) and 2007 (the average year of 2000–2014) to represent these two periods, the average accumulation rates of SOC_{20} and SOC_{100} were 19.2 ± 3.3 and 30.8 ± 12.37 g C m^{-2} yr^{-1}, respectively. Over the recent 15 years (2000–2014), SOC_{20} increased from 3.69 ± 0.16 kg C m^{-2} (33.91 ± 1.47 Pg C as storage) during the 2000s (2000–2009) to 3.84 ± 0.17 kg C m^{-2} (35.29 ± 1.56 Pg C as storage) during the early 2010s (2010–2014), an increase of 4.07%. Similarly, the overall SOC_{100} in China also presented an increasing pattern from 8.69 ± 0.39 kg C m^{-2} (79.86 ± 3.58 Pg C as storage) during the 2000s to 8.98 ± 0.41 kg C m^{-2} (82.52 ± 3.77 Pg C as storage) during the early 2010s, an increase of 3.34%. Taking middle year 2004 and 2012 to represent the 2000s and early 2010s, the overall average SOC accumulation rates in China were 18.75 ± 7.18 g C m^{-2} yr^{-1} and 36.25 ± 15.81 g C m^{-2} yr^{-1} for SOC_{20} and SOC_{100}, respectively.

Spatially, compared with SOC during the 1980s, SOC in most regions showed an increasing trend for both soil layers during the 2000s and 2010s (Figure S10). However, the variation of SOC over China from the 2000s to 2010s exhibited substantial spatial heterogeneity across the country (Figure 2c). Considering this spatial heterogeneity of SOC change, our following analysis mainly focused on the periods of 2000s and 2010s. SOC in 70.87% of the topsoil showed an increasing trend from the 2000s to 2010s; meanwhile, SOC in 64.56% of the top 100 cm soil presented a positive change. Regions with a substantial increase mainly occurred in Northeast China, Inner Mongolia, North China, Southeast China, the Loess Plateau, and the western Qinghai-Tibetan Plateau (Figure 2c), whereas Southwest China and South China were the main regions that suffered from decreasing trends from the 2000s to 2010s.

The average SOC_{20} and SOC_{100} in the four major ecosystems (forest, shrubland, grassland, and cropland) and nine geographical zones across China exhibited increasing trends from the 1980s to 2000s and 2010s (Figure 2d; Table S2). For the period of 2000–2014, the average SOC density of the forest ecosystem was the highest for both topsoil (5.28 ± 0.16 kg C m^{-2}) and top 100 cm soil (12.24 ± 0.37 kg C m^{-2}), whereas the minimum level of SOC densities in the topsoil (3.46 ± 0.12 kg C m^{-2}) and top 100 cm soil (8.9 ± 0.29 kg C m^{-2}) were

both for cropland. For carbon stock in the top 100 cm soil, grassland and forest had the two largest amount of carbon, accounting for 31.88% and 31.68% of the national SOC_{100} stock, respectively, followed by agricultural land (19.65%) and shrubland (8.99%).

3.3 | The contributions of time-variant driver categories

The relative contribution (RelCon) maps of the three time-variant drivers (i.e., climate change, vegetation growth, and environmental variation) for the two soil layers during 2000–2014 are shown in Figure 3a,b. At the national scale, the average RelCons of climate, vegetation, and environmental drivers to SOC_{20} were 48.5%, 17.32%, and 34.18%, respectively (inset in Figure 3a), indicating that climate change was the most important driver category controlling the SOC_{20} dynamics at the national scale during this period. It is clear that the RelCon of climatic driver category was still the most significant one for SOC_{100} but decreased to 37.14%; whereas the RelCon of vegetation driver category increased to 30.49% (inset in Figure 3b), suggesting strengthened impact of vegetation on SOC along depth. Meanwhile, the RelCon of environment driver category showed a very slight decrease along the soil depth (32.37% for SOC_{100}).

The spatial distributions of the driver RelCon showed heterogeneous geographical patterns for the two soil layers (Figure 3a,b). The dominant driver maps also illustrated the spatial differences of the driver impacts (Figure 3c,d). The climate-dominant regions in SOC_{20} had the largest area, which were widely distributed across China, especially in Southeast China, Northeast China, South China, and the Qinghai Tibetan Plateau (Figure 3a,c). Environment controlling regions were mainly located in the eastern China and western Qinghai-Tibetan Plateau. The vegetation-dominant regions were mainly distributed in North China and Southwest China. Different from the topsoil, the controlling area of vegetation on SOC dynamics in subsoil showed a higher percentage, especially in the eastern China (Figure 3b,d). Interestingly, a single factor category (climate, vegetation, or environment) has limited impact (quantified by the dominant area of the dominant driver) on controlling the spatiotemporal dynamics of SOC (insets in Figure 3c,d), except for the climatic drivers for SOC_{20}, which controlled the SOC_{20} dynamics of about 24.6% of

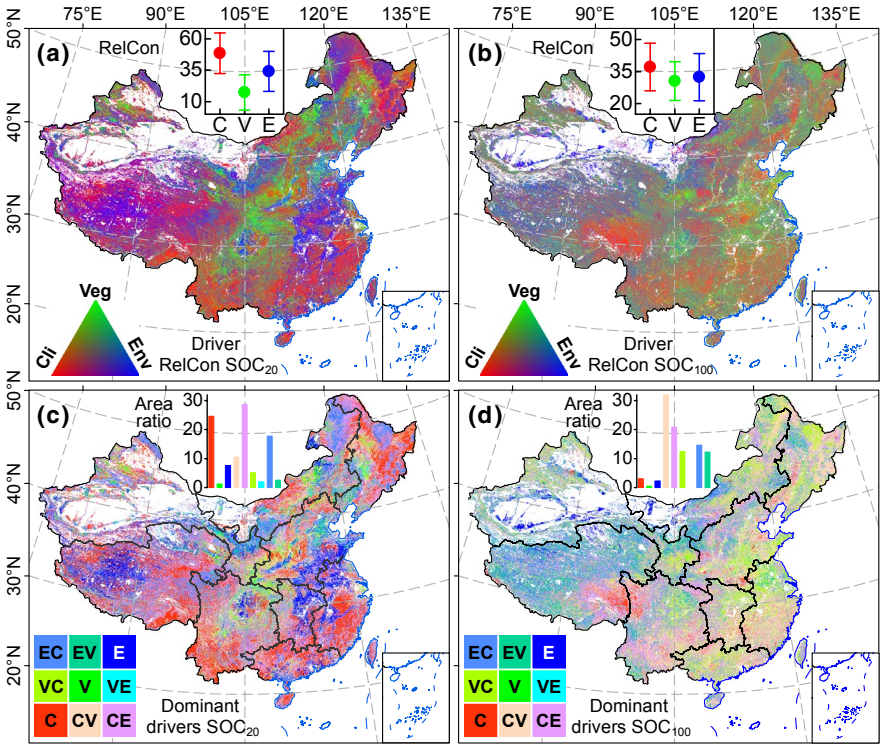

FIGURE 3 The relative contributions (RelCon) of different drivers and the dominant drivers for soil organic carbon (SOC) dynamics in the two soil layers. (a) and (b) are the spatial patterns of the relative contributions of climate change (Cli), vegetation growth (Veg), and environmental variation (Env) to SOC dynamics in the two soil layers. The top insets in (a) and (b) indicate the average relative contributions (%) of different drivers over China with whiskers indicating the standard deviations of all grid cell values. The top insets in (c) and (d) indicate the controlling area ratios of the dominant drivers over China. C, V, and E in (c) and (d) represent climate, vegetation, and environmental drivers, respectively. The combinations of C, V, and E indicate the coupled impacts of different drivers dominate the SOC dynamics. For instance, CV (VC) indicates the coupled impacts of climate and vegetation drivers and the front driver has a larger relative contribution. CE (EC): the combined influences of climate and environmental drivers, VE (EV): the combined impacts of vegetation and environmental drivers [Colour figure can be viewed at wileyonlinelibrary.com]

the topsoil at the national scale. Along with the single impact of climate change, the coupled influences of climate change and environmental variation (CE and EC in Figure 3c) are also important drivers for SOC_{20}, which together dominated 46.4% of the SOC dynamics in the topsoil (Figure 4j). Although the combined impacts of climate change and environmental variation are still important drivers for SOC_{100}, dominating 35.54% of the dynamics of SOC_{100} (Figure 3d), the combined impacts of climate change and vegetation growth (CV and VC), controlling about 46.51% of the variation inSOC_{100}, became the most important drivers for subsoil (Figure 4j). In addition, the combined influences of environmental variation and vegetation growth (EV) also accounted for 12.21% of the dominant area in SOC_{100}.

We further quantified the drivers' dominant areas in the nine geographical sub-regions (Figure 4) and found that, similar to those at the national scale, a single driver showed limited impacts on the SOC dynamics in almost all sub-regions, except for the climate change. Actually, CE, climate change, and CV are the three most important drivers for SOC_{20} in almost all sub-regions. Among these three drivers, CE controlled the largest are a proportion (>~40%) of SOC_{20} in the nine sub-regions, whereas CV and CE played the dominant role in controlling the subsoil organic carbon in all sub-regions. In addition, the dominant area of climate change and CE in most sub-regions showed a decreasing pattern with soil depth, while that of CV exhibited an increasing trend with soil depth. Considering the decreasing relative contribution of climate change to SOC_{100} with soil depth, this increasing impact of CV was largely contributed by the enhanced influence of vegetation growth on SOC_{100}. The increased impact of vegetation growth on SOC_{100} with soil depth are also reflected by the raised dominant area of VE with soil depth in all sub-regions.

3.4 | The impacts of the key drivers

The sensitivity-analysis experiments by changing the magnitudes of air temperature and precipitation indicated that a 2°C increase in temperature may significantly ($p < .05$) reduce China's SOC stocks by 0.37 ± 0.18 and 0.55 ± 0.33 Pg C for the topsoil and top 100 cm soil, respectively (Figure 5a,b). A 20% increase in precipitation can significantly ($p < .05$) raise China's topsoil organic carbon storage by 0.39 ± 0.18 Pg C, whereas the decreased precipitation by 20% did not significantly reduce the SOC stocks in the two layers. Overall, we can observe that warming mainly presents negative correlation with SOC dynamics while increase in precipitation may exhibit positive impact on SOC at the national scale.

Although the mean SOC significantly ($p < .05$) correlated with the forest canopy height in the two soil layers at the national scale, the relationships between SOC and forest canopy height in these two soil layers exhibited different patterns (Figure 5c,d). For the topsoil, SOC_{20} showed the largest accumulation rate (0.12 kg C m^{-2} m^{-1}) at the rapid growing phase of forest (forest shorter than the average canopy height ~19 m), while its increasing rate substantially decreased (0.02 kg C m^{-2} m^{-1}) when forest getting mature (forest taller than the average canopy height). Interestingly, SOC_{100} basically presented a positive linear correlation with canopy height during both the rapid growing phase and mature stage, but the increasing rate at the mature phase (0.16 kg C m^{-2} m^{-1}) was smaller than that (0.25 kg C m^{-2} m^{-1}) during the prior phase. This phenomenon suggested the important potential of forest deep soil in sequestrating carbon at the national scale.

Our estimates suggested that SOC_{20} in most croplands (94.02%) across China increased through the 1980s to 2000–2014 and the decreasing regions mainly located in Northeast China (Figure 5e). The national average SOC_{20} in croplands increased from 2.87 ± 0.15 kg C m^{-2}

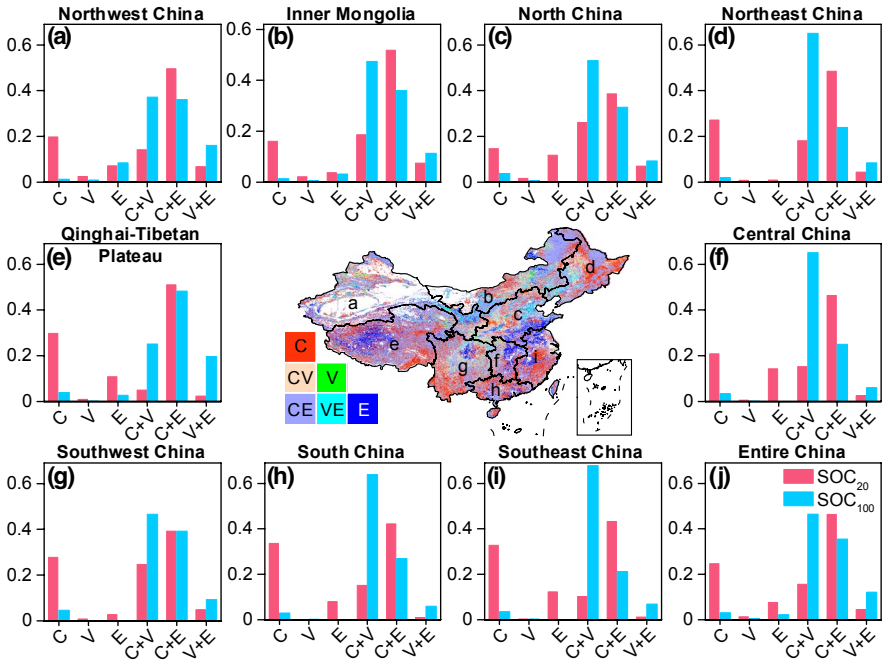

FIGURE 4 The dominant drivers of soil organic carbon (SOC) dynamics in the nine sub-regions and entire China. The dominant areas of the controlling drivers to SOC dynamics in both layers in the nine sub-regions and entire China are shown in (a–j). Map shows the spatial distribution of the dominant drivers to topsoil SOC over China. C, V, and E represent climate, vegetation, and environmental drivers, respectively. The combinations of C, V, and E indicate the coupled impacts of different drivers dominate the SOC dynamics [Colour figure can be viewed at wileyonlinelibrary.com]

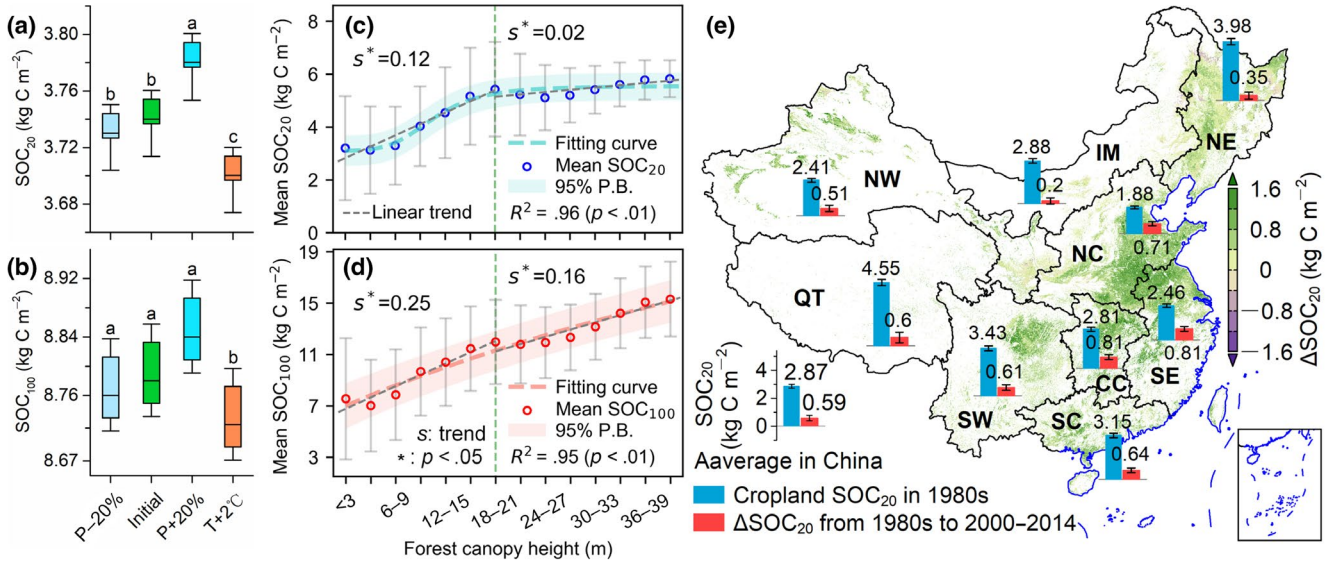

FIGURE 5 Impacts of the key drivers on soil organic carbon (SOC) dynamics and the changes in cropland SOC in the topsoil from the 1980s to 2000–2014. (a) and (b) illustrate the predicted average SOC under different precipitation variation and warming scenarios in the two soil layers. The letters upon the boxes indicate the significant differences ($p < .05$) between average SOC over China under different scenario simulations. The range of each box shows the difference among the national averages of the nine cross-validation model estimates. (c) and (d) present the distributions of SOC among different forest canopy height levels in the two soil layers. The fitting curves were fitted using logistic regression and "95% P.B." represents the 95% confidence interval of the prediction band. The vertical green dashed lines separate the rapid growing and mature phases of forests in China, which was determined as the average canopy height of forests (~19 m) in China. The error bars indicate the standard deviation of SOC in the corresponding canopy height range. (e) shows the spatial changes of the topsoil SOC in China's nine agronomic regions (NW, Northwest China; IM, Inner Mongolia; NE, Northeast China; QT, Qinghai-Tibetan Plateau; NC, North China; SW, Southwest China; CC, Central China; SE, Southeast China; SC, South China) from the 1980s to 2000–2014 [Colour figure can be viewed at wileyonlinelibrary.com]

during the 1980s to 3.46 ± 0.12 kg C m^{-2} during 2000–2014 with a net increase of 0.59 ± 0.19 kg C m^{-2} from the 1980s to 2000–2014 (or 23.6 ± 7.6 g C m^{-2} yr^{-1}, calculated based on the middle year of the two periods—1982 and 2007). Cropland SOC_{20} in the nine geographical regions exhibited different increasing trends with the largest one (0.81 ± 0.17 kg C m^{-2}) occurring in Southeast and Central China, followed by North China (0.71 ± 0.15 kg C m^{-2}). Overall, cropland soils in China played an important role in sequestrating carbon since the 1980s, and the widely applied N fertilizer (Figure S2) can be an important stimulus, especially for the agricultural lands in North China, Southeast China, Central China, and Southwest China (Figure 3c).

4 | DISCUSSION

4.1 | Comparison-based evaluation of model performance

We adopted several approaches to validate the accuracy of our estimates. First, our sample splitting strategy and the validation confirmed the capability of the RF model in predicting the spatiotemporal dynamics of SOC across China. Second, the nine-fold cross-validation exhibited the robustness of the RF model in predicting the spatiotemporal variations of SOC. Third, there is no systematic deviation in the magnitude of our estimates compared to previous investigations (Table 1; Table S2), even though, the estimated shrubland SOC density in the top 100 cm soil was a little higher than some previous investigations (Table S2). This may be caused by the classification differences of shrubland in different land cover maps because shrubs usually mix with herbaceous and woody vegetation, leading to difficulties in separating them from other land cover classes in remote sensing images (Herold et al., 2008). In this section, we further evaluated the spatial patterns and variation trends of our predictions with other widely-used SOC maps and previous investigations.

As for the spatial patterns, our estimated SOC density maps in the topsoil is consistent with that of the SoilGrids ($R^2 = .73$, $p < .001$) and GSOC map ($R^2 = .57$, $p < .001$) SOC maps (Figure S11 and S12). The estimated SOC_{100} showed an overall similar spatial pattern with the SoilGrids SOC_{100} (Figure S13a,b) and other studies (Shangguan et al., 2014; Tang et al., 2018). Notably, the SoilGrids SOC stock in the first meter soil layer is ~144.58 Pg C with a density of 15.53 kg C m^{-2} for an area of 930.97 km^2, which is much larger than our estimate and most previous studies (Table 1). The overestimation of the SoilGrids SOC maps in China and other regions were also reported by previous studies (Chen et al., 2020; Liang et al., 2019; Tifafi et al., 2018). We checked the global soil profiles that were used to produce the subsoil SOC maps (under 15 cm) in SoilGrids (Figure S12c) and found that although there are 2886 sites distributed over China, they only accounted for 1.52% of all global subsoil sites. SOC maps in the SoilGrids were produced based on a global scale model

trained by these observations around the world (Hengl et al., 2017). However, calculation based on these observations showed that the average subsoil organic carbon content in China (7.7 g/kg) was much less than that of the rest region (14.97 g/kg). Thus, the limited sample amount in China and the relatively lower organic carbon content may lead to the overestimation of the SoilGrids subsoil SOC in China.

As for the variation trends of our predictions, our estimated SOC_{20} variation trend in forest (23.48 ± 4.46 g C m^{-2} yr^{-1}) from the 1980s to 2000s is basically consistent with a previous investigation (20 ± 5.5 g C m^{-2} yr^{-1}) (Yang et al., 2014). The average trend (29.43 ± 13.31 g C m^{-2} yr^{-1}) of our estimated SOC_{20} in the Tibetan Plateau grassland from the 2000s to 2010s is also close to a previous investigation (28.0 g C m^{-2} yr^{-1} for the topsoil) (Ding et al., 2017). Our estimated increasing magnitude of cropland SOC_{20} (0.59 ± 0.19 kg C m^{-2}) from the 1980s to 2000–2014 is comparable with a previous investigation (0.43 ± 0.09kg C m^{-2}) from 1980 to 2011 (Zhao et al., 2018). Above comparisons suggest that our estimates are relatively reasonable and satisfactory in both magnitude and spatiotemporal patters, illustrating the capability of the RF model in predicting the spatiotemporal dynamics of SOC across China.

It is noted that although our simulations are comparable with previous studies as described above, there are three distinct improvements. First, our model has been tested to have the spatiotemporal expansion capability, which provided new insights for the temporal variation of SOC. Second, we provided the maps of long-term SOC dynamics in different soil depths with full evaluation, and can be important and attractive supplements to the existing static SOC maps or site-scale measurements. Last but not least, we spatially quantified the contributions of the driving forces to SOC dynamics.

4.2 | Divergent soil organic carbon responses in different soil depths

Our simulations revealed that climate change exerted the largest relative contribution compared to environmental variation and vegetation growth at the national scale, especially for topsoil. This is because the chemical and biochemical reactions related to the soil carbon input and loss are temperature-dependent and also subject to moisture condition (Armas-Herrera et al., 2012; Davidson & Janssens, 2006; Meyer et al., 2018). Our sensitivity-analysis experiments based on air temperature and precipitation further deciphered how climate change affect the SOC dynamics and demonstrated that both warming and drying can reduce the SOC storage while a wetter condition showed positive impact on it at the national scale (Figure 5a,b). Our results not only confirmed the previous conclusion based on correlation analysis (Jobbágy & Jackson, 2000; Luo et al., 2020; Schindlbacher et al., 2009; Suseela et al., 2012; Tang et al., 2018) but also provided the quantified contributions of these factors, providing a paradigm for evaluating the impacts of different influcing drivers. As for the weakened impact of climate change on SOC with soil depth, this is mainly due to that shallow soil organic matter was more vulnerable to warming than that of deep soil organic matter (Jobbágy & Jackson, 2000; Walz et al., 2017). The labile carbon formed by plant residues and topsoil root exudates dominate the turnover of organic matter in shallow soil with the highest microbial abundance and diversity (Liebner et al., 2008). Simultaneously, most biological mediated processes take place in the shallow soil, which could quickly respond to slight changes of climate (Knapp et al., 2008).

Plant biomass allocation between aboveground and belowground and the root depth control the relative distribution of organic carbon in soil profile (Jobbágy & Jackson, 2000), and root is the main source of organic matter in deep soil, dominating the dynamics of deep SOC (Nepstad et al., 1994). In addition, roots are generally harder to be decomposed than topsoil litter (Rasse et al., 2005), resulting in a relatively greater impact of vegetation growth on SOC dynamics in deep soil. The relationships between forest canopy height and SOC in the two soil layers further illustrated these different impacts of vegetation growth on SOC dynamics in different soil depths (Figure 5c,d). Although the topsoil carbon accumulation rate substantially decreased when the forest canopy height exceeds the average level across China (i.e., getting mature), subsoil in mature forest still showed a comparable carbon sequestration rate as the rapid growing phase. Our results are consistent with previous investigations that revealed the old-growth forests can accumulate atmospheric carbon at a high rate (Tang et al., 2011; Zhou et al., 2006). It can be attributed to the impact of litter quality (i.e., litter carbon-to-nitrogen ratio) on the SOC accumulation (Sumiyoshi et al., 2017; Xiong et al., 2020; Zhou et al., 2019a), because a lower litter C/N ratio in the mature forest can increase litter turnover and the carbon-use efficiency of microorganisms (Huang et al., 2011; Zhou et al., 2019a). China have experienced a rapid expansion of afforestation over the past decades, and these rapid-growing plantation forests have substantially promoted the carbon sink capacity of ecosystems through photosynthesis (Li et al., 2021; Wang et al., 2020a). The ongoing nationwide ecological protection programs such as the Grain-for-Green and the Natural Forest Protection projects have greatly improved the soil environment, causing the increase of SOC stock, as previously confirmed (Tang et al., 2018). Our results highlighted that, with the rapid growth of the large proportion of young forests in China, soils in these new planted forests may also exhibit high potential in sequestrating carbon.

The agricultural development in China largely depended on N fertilizer, and its amount increased rapidly from1979 to 2000–2014 to meet the rising demand for food (Figure S2). Our estimates revealed that the increasing trend of cropland SOC_{20} throughout the 1980s to 2000–2014 (Figure 5e), and fertilization showed the dominant impact on controlling the SOC_{20} dynamics in these agricultural lands (Figure 3c). The increased N fertilizer over this period played a key role in promoting the soil carbon sequestration in the croplands through the enhanced biomass production (Snyder et al., 2009; Xu et al., 2021; Zhao et al., 2018). Nonetheless, it should be noted that fertilization is not absolutely beneficial to soil carbon sequestration (Zhao et al., 2018) because excessive use of N fertilizers can result in soil acidification, which can deteriorate the stability of SOC

(Kuzyakov et al., 2021). Chinese government has issued a series of eco-agricultural policies to optimize the intensive agricultural development. It is predictable that the optimized fertilization and continuously enhanced straw/stover return to soils (Zhao et al., 2018) can further benefit the SOC stocks in croplands.

4.3 | Uncertainties and limitations

Although the visual comparisons and statistical evaluations of all model validations showed that our estimates were reasonable and acceptable, there still exist uncertainties and limitations in the study. A potential limitation is the different spatial resolutions of the model inputs. We used several grid datasets to predict the SOC distributions and dynamics, and some of them were extracted from the global-scale datasets. Take the WoSIS SoilGrids soil parameters as an example, these datasets were constructed based on extensive field investigations (~150,000 soil profiles) across the world and owned a higher spatial resolution compared with other global soil property data sets (Hengl et al., 2017). Although the producers used a few approaches to control the accuracy of the data sets (Batjes et al., 2020; Hengl et al., 2017), the performance and accuracy of such a high resolution global data set may cause some biases and bring uncertainties in our study. Therefore, extensive evaluation of these data sets locally in different regions may deserve further work in future by the scientific community. Another limitation is that some factors that may induce SOC variations were not introduced into our model simulation due to the data unavailability at large scale. For example, litter quality (i.e., litter C/N ratio) may be a key factor controlling the SOC accumulation (Sumiyoshi et al., 2017; Xiong et al., 2020; Zhou et al., 2019a). The nationwide ecological protection programs, grazing activity, and agricultural practices may also affect the SOC dynamics locally or regionally (Deng et al., 2014; Lal, 2004). Developing reliable data of these programs/practices and investigating their impacts on SOC dynamics at the national scale would be a promising subject in future study.

5 | CONCLUSIONS

This study provided the investigated SOC dynamics for different soil layers over China using a machine learning approach and extensive measurements during the early 1980s, 2000s, and early 2010s. Our spatiotemporal estimates of SOC confirmed soils in China functioned as carbon sink at the nation scale, with an accumulation rate of 30.8 ± 12.37g C m^{-2} yr^{-1} in the first meter soil since the 1980s. Using factorial analysis experiments, we further quantified the impacts of different time variant factors on SOC dynamics and found that climate change exerted the largest relative contributions to SOC dynamics in the two soil layers. However, the climate change-induced influence on SOC weakened with soil depth, while vegetation growth exhibited a strengthened impact with soil depth. Relationships between SOC and forest canopy height confirmed this strengthened impact of vegetation with soil depth and implied that forest soil in China may exhibit a high potential in sequestrating atmospheric CO_2 with the rapid growth of young forests owing to the nationwide expansion of afforestation over the past decades. Using the sensitivity-analysis experiment, we concluded that warming and drying can both negatively affect the SOC stocks, while the increase in precipitation showed a positive contribution. Our simulations further suggested that SOC dynamics is not simply controlled by a single driver. In fact, climate change and its coupled impact with environmental variation (CE) were the two key drivers for SOC_{20}, controlling the SOC dynamics in 71% of topsoil. Meanwhile, CE and the combined influence of climate change and vegetation growth were the most important drivers for SOC_{100}, dominating the SOC dynamics in 82.05% of the first meter soil. Additionally, we have spatially quantified the changes in cropland SOC_{20} across China and found an increasing rate of 23.6 ± 7.6 g C m^{-2} yr^{-1} through the 1980s to 2000–2014. The widely applied N fertilizer since 1979 can be an important stimulus to the increase in cropland SOC. Overall, our spatiotemporal estimates can be important and attractive supplements to the existing static SOC measurements, and emphasized the combined impacts of different factors on SOC dynamics.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation of China (31961143011, U20A2089), the National Key Research and Development Program of China (2019YFC0507403), the Strategic Priority Research Program of Chinese Academy of Sciences (XDB40020205), Shaanxi Key Research and Development Program (2022ZDLSF06-04), the Fundamental Research Funds for the Central Universities (xzy022020008), Technology Innovation Center for Land Engineering and Human Settlements, Shaanxi Land Engineering Construction Group Co. Ltd and Xi'an Jiaotong University (201912131-B2), Shaanxi Major Theoretical and Practical Program (20ST-106), the Key Laboratory of Eco-Environment and Meteorology for The Qinling Mountains and Loess Plateau, Shaanxi Meteorological Bureau (2021K-6), the Innovation Team of Shaanxi Province (2021TD-52). We also thank the HPCC Platform at Xi'an Jiaotong University for the computing equipment and computer maintenance.

CONFLICT OF INTEREST

All other authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

Y.W. and H.L. conceived the project. H.L. developed and conducted the model simulations and analysis with support from Y.W., Y.W. and H.L. drafted the manuscript with contributions from all co-authors. Y.W., S.L., J.X., and G.A. participated in the discussion and analysis of the results, and edited the manuscript. H.L. and Y.C. drew the figures. All authors (H.L., Y.W., S.L., J.X., W.Z., J.C., G.A., and Y.C.) reviewed the results, revised, and approved the manuscript.

DATA AVAILABILITY STATEMENT

The soil property maps from the World Soil Information Service and the Resource and Environment Data Cloud Platform can be obtained at https://www.isric.org and www.resdc.cn, respectively. The terrain maps from the Global Land One-km Base Elevation Project (GLOBE) can be acquired at https://www.ngdc.noaa.gov/mgg/topo/globe.html. The air temperature and precipitation gridded maps are available at https://data.tpdc.ac.cn. The solar radiation maps were gathered from the TerraClimate at http://www.climatologylab.org/terraclimate.html. The potential evapotranspiration maps were extracted from the Famine Early Warning Systems Network (FEWS NET) Land Data Assimilation System (FLDAS, https://ldas.gsfc.nasa.gov/FLDAS). The European Space Agency Climate Change Initiative Land Cover (ESA CCI LC) maps can be download at https://www.esa-landcover-cci.org/. All other primary data that support the findings of this study are openly available in Data Dryad at https://doi.org/10.5061/dryad.0cfxpnw4m.

ORCID

Huiwen Li https://orcid.org/0000-0001-9719-2237
Yiping Wu https://orcid.org/0000-0002-5163-0884
Jingfeng Xiao https://orcid.org/0000-0002-0622-6903

REFERENCES

Allison, S. D., Wallenstein, M. D., & Bradford, M. A. (2010). Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, *3*(5), 336–340. https://doi.org/10.1038/Ngeo846

Armas-Herrera, C. M., Mora, J. L., Arbelo, C. D., & Rodríguez-Rodríguez, A. (2012). Interannual variations of soil organic carbon fractions in unmanaged volcanic soils (Canary Islands, Spain). *Ecology and Evolution*, *2*(10), 2374–2386. https://doi.org/10.1002/ece3.355

Batjes, N. H. (1996). Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, *47*(2), 151–163. https://doi.org/10.1111/j.1365-2389.1996.tb01386.x

Batjes, N. H., Ribeiro, E., & van Oostrum, A. (2020). Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). *Earth System Science Data*, *12*(1), 299–320. https://doi.org/10.5194/essd-12-299-2020

Bradford, M. A., Wieder, W. R., Bonan, G. B., Fierer, N., Raymond, P. A., & Crowther, T. W. (2016). Managing uncertainty in soil carbon feedbacks to climate change. *Nature Climate Change*, *6*(8), 751–758. https://doi.org/10.1038/Nclimate3071

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Calvo de Anta, R., Luis, E., Febrero-Bande, M., Galinanes, J., Macias, F., Ortiz, R., & Casas, F. (2020). Soil organic carbon in peninsular Spain: Influence of environmental factors and spatial distribution. *Geoderma*, *370*, 114365. https://doi.org/10.1016/j.geoderma.2020.114365

Chen, S., Mulder, V. L., Heuvelink, G. B. M., Poggio, L., Caubet, M., Román Dobarco, M., Walter, C., & Arrouays, D. (2020). Model averaging for mapping topsoil organic carbon in France. *Geoderma*, *366*, 114237. https://doi.org/10.1016/j.geoderma.2020.114237

Cheng, L., Booker, F. L., Tu, C., Burkey, K. O., Zhou, L. S., Shew, H. D., Rufty, T. W., & Hu, S. J. (2012). Arbuscular mycorrhizal fungi increase organic carbon decomposition under elevated CO2. *Science*, *337*(6098), 1084–1087. https://doi.org/10.1126/science.1224304

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, *440*(7081), 165–173. https://doi.org/10.1038/nature04514

Deng, L., Liu, G. B., & Shangguan, Z. P. (2014). Land-use conversion and changing soil carbon stocks in China's 'Grain-for-Green' Program: a synthesis. *Global Change Biology*, *20*(11), 3544–3556. https://doi.org/10.1111/gcb.12508

Ding, J., Chen, L., Ji, C., Hugelius, G., Li, Y., Liu, L., Qin, S., Zhang, B., Yang, G., Li, F., Fang, K., Chen, Y., Peng, Y., Zhao, X., He, H., Smith, P., Fang, J., & Yang, Y. (2017). Decadal soil carbon accumulation across Tibetan permafrost regions. *Nature Geoscience*, *10*(6), 420–424. https://doi.org/10.1038/ngeo2945

Ding, J., Li, F., Yang, G., Chen, L., Zhang, B., Liu, L., Fang, K., Qin, S., Chen, Y., Peng, Y., Ji, C., He, H., Smith, P., & Yang, Y. (2016). The permafrost carbon inventory on the Tibetan Plateau: A new evaluation using deep sediment cores. *Global Change Biology*, *22*(8), 2688–2701. https://doi.org/10.1111/gcb.13257

Fang, J., Liu, G., & Xu, S. (1996). Soil carbon pool in China and its global significance. *Journal of Environmental Science*, *8*(2), 249–254.

FAO & ITPS (2018). Global soil organic carbon map (GSOCmap) Technical Report. Retrieved from Rome.

Franko, U. (1996). *Modelling approaches of soil organic matter turnover within the CANDY system*. Berlin, Heidelberg.

Gaitan, J. J., Maestre, F. T., Bran, D. E., Buono, G. G., Dougill, A. J., García Martinez, G., Ferrante, D., Guuroh, R. T., Linstadter, A., Massara, V., Thomas, A. D., & Oliva, G. E. (2019). Biotic and abiotic drivers of topsoil organic carbon concentration in drylands have similar effects at regional and global scales. *Ecosystems*, *22*(7), 1445–1456. https://doi.org/10.1007/s10021-019-00348-y

Gomes, L. C., Faria, R. M., de Souza, E., Veloso, G. V., Schaefer, C. E. G. R., & Fernandes, E. I. F. (2019). Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, *340*, 337–350. https://doi.org/10.1016/j.geoderma.2019.01.007

Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotic, A., Shangguan, W., Wright, M. N., Geng, X. Y., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment*, *112*(5), 2538–2556. https://doi.org/10.1016/j.rse.2007.11.013

Huang, Y.-H., Li, Y.-L., Xiao, Y., Wenigmann, K. O., Zhou, G.-Y., Zhang, D.-Q., Wenigmann, M., Tang, X.-L., & Liu, J.-X. (2011). Controls of litter quality on the carbon sink in soils through partitioning the products of decomposing litter in a forest succession series in South China. *Forest Ecology and Management*, *261*(7), 1170–1177. https://doi.org/10.1016/j.foreco.2010.12.030

Hui, J., Wu, Y., Zhao, F., Lei, X., Sun, P., Singh, S. K., Liao, W., Qiu, L., & Li, J. (2020). Parameter optimization for uncertainty reduction and simulation improvement of hydrological modeling. *Remote Sensing*, *12*(24), 4069. https://doi.org/10.3390/rs12244069

Jackson, R. B., Lajtha, K., Crow, S. E., Hugelius, G., Kramer, M. G., & Piñeiro, G. (2017). The ecology of soil carbon: Pools, vulnerabilities, and biotic and abiotic controls. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 419–445. https://doi.org/10.1146/annurev-ecolsys-112414-054234

Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications*, *10*(2), 423–436. https://doi.org/10.2307/2641104

Knapp, A. K., Beier, C., Briske, D. D., Classen, A. T., Luo, Y., Reichstein, M., Smith, M. D., Smith, S. D., Bell, J. E., Fay, P. A., Heisler, J. L., Leavitt, S. W., Sherry, R., Smith, B., & Weng, E. (2008). Consequences of more extreme precipitation regimes for terrestrial ecosystems. *BioScience*, *58*(9), 811–821. https://doi.org/10.1641/B580908

Kuzyakov, Y., Kuzyakova, I., Raza, S., Zhou, J., & Zamanian, K. (2021). Letter-to-the-Editor: Does acidification really increase soil carbon in croplands? How statistical analyses of large datasets might mislead the conclusions. *Geoderma*, *384*, 114806. https://doi.org/10.1016/j.geoderma.2020.114806.

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Li, C. S., Frolking, S., & Frolking, T. A. (1992). A model of nitrous-oxide evolution from soil driven by rainfall events. 1. Model structure and sensitivity. *Journal of Geophysical Research-Atmospheres*, *97*(D9), 9759–9776. https://doi.org/10.1029/92jd00509

Li, H., Wu, Y., Chen, J., Zhao, F., Wang, F., Sun, Y., Zhang, G., & Qiu, L. (2021). Responses of soil organic carbon to climate change in the Qilian Mountains and its future projection. *Journal of Hydrology*, *596*, https://doi.org/10.1016/j.jhydrol.2021.126110. 126110.

Li, H., Wu, Y., Liu, S., & Xiao, J. (2021). Regional contributions to interannual variability of net primary production and climatic attributions. *Agricultural and Forest Meteorology*, *303*, https://doi.org/10.1016/j.agrformet.2021.108384

Liang, W., Yang, Y. T., Fan, D. M., Guan, H. D., Zhang, T., Long, D., Zhou, Y., & Bai, D. (2015). Analysis of spatial and temporal patterns of net primary production and their climate controls in China from 1982 to 2010. *Agricultural and Forest Meteorology*, *204*, 22–36. https://doi.org/10.1016/j.agrformet.2015.01.015

Liang, Z., Chen, S., Yang, Y., Zhao, R., Shi, Z., & Viscarra Rossel, R. A. (2019). National digital soil map of organic matter in topsoil and its associated uncertainty in 1980's China. *Geoderma*, *335*, 47–56. https://doi.org/10.1016/j.geoderma.2018.08.011

Liebner, S., Harder, J., & Wagner, D. (2008). Bacterial diversity and community structure in polygonal tundra soils from Samoylov Island, Lena Delta. *Siberia. International Microbiology*, *11*(3), 195–202. https://doi.org/10.2436/20.1501.01.60

Liu, F., Wu, H., Zhao, Y., Li, D., Yang, J.-L., Song, X., Shi, Z., Zhu, A. X., & Zhang, G.-L. (2022). Mapping high resolution National Soil Information Grids of China. *Science Bulletin*, *67*(3), 328–340. https://doi.org/10.1016/j.scib.2021.10.013

Liu, L. L., King, J. S., & Giardina, C. P. (2005). Effects of elevated concentrations of atmospheric CO2 and tropospheric O-3 on leaf litter production and chemistry in trembling aspen and paper birch communities. *Tree Physiology*, *25*(12), 1511–1522. https://doi.org/10.1093/treephys/25.12.1511

Liu, S. G., Bliss, N., Sundquist, E., & Huntington, T. G. (2003). Modeling carbon dynamics in vegetation and soil under the impact of soil erosion and deposition. *Global Biogeochemical Cycles*, *17*, 1074. https://doi.org/10.1029/2002gb002010

Liu, S. S., Yang, Y. H., Shen, H. H., Hu, H. F., Zhao, X., Li, H., Liu, T. Y., & Fang, J. Y. (2018). No significant changes in topsoil carbon in the grasslands of northern China between the 1980s and 2000s. *Science of The Total Environment*, *624*, 1478–1487. https://doi.org/10.1016/j.scitotenv.2017.12.254

Luo, Z., Luo, Y., Wang, G., Xia, J., & Peng, C. (2020). Warming-induced global soil carbon loss attenuated by downward carbon movement. *Global Change Biology*, *26*(12), 7242–7254. https://doi.org/10.1111/gcb.15370

Meyer, R. S., Cullen, B. R., Whetton, P. H., Robertson, F. A., & Eckard, R. J. (2018). Potential impacts of climate change on soil organic carbon and productivity in pastures of south eastern Australia. *Agricultural Systems*, *167*, 34–46. https://doi.org/10.1016/j.agsy.2018.08.010

Nepstad, D. C., Decarvalho, C. R., Davidson, E. A., Jipp, P. H., Lefebvre, P. A., Negreiros, G. H., Dasilva, E. D., Stone, T. A., Trumbore, S. E., & Vieira, S. (1994). The role of deep roots in the hydrological and carbon cycles of amazonian forests and pastures. *Nature*, *372*(6507), 666–669. https://doi.org/10.1038/372666a0

Ni, J. (2001). Carbon storage in terrestrial ecosystems of China: Estimates at different spatial resolutions and their responses to climate change. *Climatic Change*, *49*(3), 339–358. https://doi.org/10.1023/A:1010728609701

National Soil Survey Office (1993). *China Soil Series*, Vol. *I*. China Agricultural Press.

National Soil Survey Office (1994). *China Soil Series*, *Vols. II–III*. China Agricultural Press.

National Soil Survey Office (1995). *China Soil Series*, *Vols. IV–V*. China Agricultural Press.

National Soil Survey Office (1996). *China Soil Series*, *Vol. VI*. China Agricultural Press.

Padarian, J., Minasny, B., & Mcbratney, A. B. (2020). Machine learning and soil sciences: A review aided by machine learning tools. *Soil*, *6*(1), 35–52. https://doi.org/10.5194/soil-6-35-2020

Parton, W. J., Hartman, M., Ojima, D., & Schimel, D. (1998). DAYCENT and its land surface submodel: description and testing. *Global and Planetary Change*, *19*(1–4), 35–48. https://doi.org/10.1016/S0921-8181(98)00040-X

Paustian, K., Collier, S., Baldock, J., Burgess, R., Creque, J., DeLonge, M., Dungait, J., Ellert, B., Frank, S., Goddard, T., Govaerts, B., Grundy, M., Henning, M., Izaurralde, R. C., Madaras, M., McConkey, B., Porzig, E., Rice, C., Searle, R., … Jahn, M. (2019). Quantifying carbon for agricultural soil management: From the current status toward a global soil information system. *Carbon Management*, *10*(6), 567–587. https://doi.org/10.1080/17583004.2019.1633231

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Plante, A., & Conant, R. T. (2014). Soil Organic Matter Dynamics, Climate Change Effects. In: B. Freedman (Eds) *Global Environmental Change. Handbook of Global Environmental Pollution*, *vol 1* (Vol. 1). Springer.

Poeplau, C., Don, A., Vesterdal, L., Leifeld, J., Van Wesemael, B., Schumacher, J., & Gensior, A. (2011). Temporal dynamics of soil organic carbon after land-use change in the temperate zone – carbon response functions as a model approach. *Global Change Biology*, *17*(7), 2415–2427. https://doi.org/10.1111/j.1365-2486.2011.02408.x

Prietzel, J., Zimmermann, L., Schubert, A., & Christophel, D. (2016). Organic matter losses in German Alps forest soils since the 1970s most likely caused by warming. *Nature Geoscience*, *9*(7), 543–548. https://doi.org/10.1038/Ngeo2732

Ramesh, T., Bolan, N. S., Kirkham, M. B., Wijesekara, H., Kanchikerimath, M., Rao, C. S., Sandeep, S., Rinklebe, J., Ok, Y. S., Choudhury, B. U., Wang, H. L., Tang, C. X., Wang, X. J., Song, Z. L., & Freeman, O. W. (2019). Soil organic carbon dynamics: Impact of land use changes and management practices: A review. *Advances in Agronomy*, *156*(156), 1–107. https://doi.org/10.1016/bs.agron.2019.02.001

Rasse, D. P., Rumpel, C., & Dignac, M. F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil*, *269*(1–2), 341–356. https://doi.org/10.1007/s11104-004-0907-y

Rossel, R. A. V., Lee, J., Behrens, T., Luo, Z., Baldock, J., & Richards, A. (2019). Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls. *Nature Geoscience*, *12*, 547–552. https://doi.org/10.1038/s41561-019-0373-z

Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., & Kapos, V. (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, *5*(1), 81–91. https://doi.org/10.4155/Cmt.13.77

Schindlbacher, A., Zechmeister-Boltenstern, S., & Jandl, R. (2009). Carbon losses due to soil warming: Do autotrophic and heterotrophic soil respiration respond equally? *Global Change Biology*, *15*(4), 901–913. https://doi.org/10.1111/j.1365-2486.2008.01757.x

Shangguan, W., Dai, Y. J., Duan, Q. Y., Liu, B. Y., & Yuan, H. (2014). A global soil data set for earth system modeling. *Journal of Advances in Modeling Earth Systems*, *6*(1), 249–263. https://doi.org/10.1002/2013ms000293

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, *116*, G04021. https://doi.org/10.1029/2011jg001708

Snyder, C. S., Bruulsema, T. W., Jensen, T. L., & Fixen, P. E. (2009). Review of greenhouse gas emissions from crop production systems and fertilizer management effects. *Agriculture Ecosystems & Environment*, *133*(3–4), 247–266. https://doi.org/10.1016/j.agee.2009.04.021

Stockmann, U., Adams, M. A., Crawford, J. W., Field, D. J., Henakaarchchi, N., Jenkins, M., Minasny, B., McBratney, A. B., de Courcelles, V. D., Singh, K., Wheeler, I., Abbott, L., Angers, D. A., Baldock, J., Bird, M., Brookes, P. C., Chenu, C., Jastrow, J. D., Lal, R., … Zimmermann, M. (2013). The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture Ecosystems & Environment*, *164*, 80–99. https://doi.org/10.1016/j.agee.2012.10.001

Sumiyoshi, Y., Crow, S. E., Litton, C. M., Deenik, J. L., Taylor, A. D., Turano, B., & Ogoshi, R. (2017). Belowground impacts of perennial grass cultivation for sustainable biofuel feedstock production in the tropics. *GCB Bioenergy*, *9*(4), 694–709. https://doi.org/10.1111/gcbb.12379

Suseela, V., Conant, R. T., Wallenstein, M. D., & Dukes, J. S. (2012). Effects of soil moisture on the temperature sensitivity of heterotrophic respiration vary seasonally in an old-field climate change experiment. *Global Change Biology*, *18*(1), 336–348. https://doi.org/10.1111/j.1365-2486.2011.02516.x

Tang, X., Wang, Y.-P., Zhou, G., Zhang, D., Liu, S., Liu, S., Zhang, Q., Liu, J., & Yan, J. (2011). Different patterns of ecosystem carbon accumulation between a young and an old-growth subtropical forest in Southern China. *Plant Ecology*, *212*(8), 1385–1395. https://doi.org/10.1007/s11258-011-9914-2

Tang, X., Zhao, X., Bai, Y., Tang, Z., Wang, W., Zhao, Y., Wan, H., Xie, Z., Shi, X., Wu, B., Wang, G., Yan, J., Ma, K., Du, S., Li, S., Han, S., Ma, Y., Hu, H., He, N., … Zhou, G. (2018). Carbon pools in China's terrestrial ecosystems: New estimates based on an intensive field survey. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), 4021–4026. https://doi.org/10.1073/pnas.1700291115

Tifafi, M., Guenet, B., & Hatté, C. (2018). Large differences in global and regional total soil carbon stock estimates based on SoilGrids, HWSD, and NCSCD: Intercomparison and evaluation based on field data from USA, England, Wales, and France. *Global Biogeochemical Cycles*, *32*(1), 42–56. https://doi.org/10.1002/2017GB005678

van Groenigen, K. J., Qi, X., Osenberg, C. W., Luo, Y. Q., & Hungate, B. A. (2014). Faster decomposition under increased atmospheric CO2 limits soil carbon storage. *Science*, *344*(6183), 508–509. https://doi.org/10.1126/science.1249534

Walz, J., Knoblauch, C., Bohme, L., & Pfeiffer, E. M. (2017). Regulation of soil organic matter decomposition in permafrost-affected Siberian tundra soils - Impact of oxygen availability, freezing and thawing, temperature, and labile organic matter. *Soil Biology & Biochemistry*, *110*, 34–43. https://doi.org/10.1016/j.soilbio.2017.03.001

Wang, G., Huang, W., Mayes, M. A., Liu, X., Zhang, D., Zhang, Q., Han, T., & Zhou, G. (2019). Soil moisture drives microbial controls on carbon decomposition in two subtropical forests. *Soil Biology and Biochemistry*, *130*, 185–194. https://doi.org/10.1016/j.soilbio.2018.12.017

Wang, J., Feng, L., Palmer, P. I., Liu, Y., Fang, S., Bösch, H., O'Dell, C. W., Tang, X., Yang, D., Liu, L., & Xia, C. (2020). Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature*, *586*(7831), 720–723. https://doi.org/10.1038/s41586-020-2849-9

Wang, S., Huang, M., Shao, X., Mickler, R. A., Li, K., & Ji, J. (2004). Vertical distribution of soil organic carbon in China. *Environmental Management*, *33*(S1), S200–S209. https://doi.org/10.1007/s00267-003-9130-5

Wang, S., & Zhou, C. (1999). Estimating soil carbon reservior of terrestrial ecosystem in China. *Geographical Research*, *18*(4), 349–356. https://doi.org/10.11821/yj1999040003

Wang, S., Zhou, C., Li, K., Zhu, S., & Huang, F. (2000). Analysis on spatial distribution characteristics of soil organic carbon reservoir in China. *Acta Geographica Sinica*, *67*(5), 533–544. https://doi.org/10.11821/xb200005003

Wang, T. H., Yang, D. W., Yang, Y. T., Piao, S. L., Li, X., Cheng, G. D., & Fu, B. J. (2020). Permafrost thawing puts the frozen carbon at risk over the Tibetan Plateau. *Science Advances*, *6*, eaaz3513. https://doi.org/10.1126/sciadv.aaz3513

Wieder, W. R., Bonan, G. B., & Allison, S. D. (2013). Global soil carbon projections are improved by modelling microbial processes. *Nature Climate Change*, *3*(10), 909–912. https://doi.org/10.1038/Nclimate1951

Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lutzow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Liess, M., Garcia-Franco, N., Wollschlager, U., Vogel, H. J., & Kogel-Knabner, I. (2019). Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales. *Geoderma*, *333*, 149–162. https://doi.org/10.1016/j.geoderma.2018.07.026

Wu, H. B., Guo, Z. T., & Peng, C. H. (2003). Distribution and storage of soil organic carbon in China. *Global Biogeochemical Cycles*, *17*(2), 1048. https://doi.org/10.1029/2001gb001844

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., Ichii, K., Ni, W., Pang, Y., Rahman, A. F., Sun, G., Yuan, W., Zhang, L., & Zhang, X. (2019). Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment*, *233*, 111383. https://doi.org/10.1016/j.rse.2019.111383

Xie, Z. B., Zhu, J. G., Liu, G., Cadisch, G., Hasegawa, T., Chen, C. M., Sun, H. F., Tang, H. Y., & Zeng, Q. (2007). Soil organic carbon stocks in China and changes from 1980s to 2000s. *Global Change Biology*, *13*(9), 1989–2007. https://doi.org/10.1111/j.1365-2486.2007.01409.x

Xiong, X., Zhou, G., & Zhang, D. (2020). Soil organic carbon accumulation modes between pioneer and old-growth forest ecosystems. *Journal of Applied Ecology*, *57*(12), 2419–2428. https://doi.org/10.1111/1365-2664.13747

Xu, C., Xu, X., Ju, C., Chen, H. Y. H., Wilsey, B. J., Luo, Y., & Fan, W. (2021). Long-term, amplified responses of soil organic carbon to nitrogen addition worldwide. *Global Change Biology*, *27*, 1170–1180. https://doi.org/10.1111/gcb.15489

Xu, L., Yu, G., He, N., Wang, Q., Gao, Y., Wen, D., Li, S., Niu, S., & Ge, J. (2018). Carbon storage in China's terrestrial ecosystems: A synthesis. *Scientific Reports*, *8*(1), 2806. https://doi.org/10.1038/s41598-018-20764-9

Yang, Y., Li, P., Ding, J., Zhao, X., Ma, W., Ji, C., & Fang, J. (2014). Increased topsoil carbon stock across China's forests. *Global Change Biology*, *20*(8), 2687–2696. https://doi.org/10.1111/gcb.12536

Yang, Y., Mohammat, A., Feng, J., Zhou, R., & Fang, J. (2007). Storage, patterns and environmental controls of soil organic carbon in China. *Biogeochemistry*, *84*(2), 131–141. https://doi.org/10.1007/s10533-007-9109-z

Yao, Y., Wang, X., Li, Y., Wang, T., Shen, M., Du, M., He, H., Li, Y., Luo, W., Ma, M., Ma, Y., Tang, Y., Wang, H., Zhang, X., Zhang, Y., Zhao, L., Zhou, G., & Piao, S. (2018). Spatiotemporal pattern of gross primary productivity and its covariation with climate in China over the last thirty years. *Global Change Biology*, *24*(1), 184–196. https://doi.org/10.1111/gcb.13830

Yu, D. S., Shi, X. Z., Wang, H., Sun, W. X., Chen, J. M., Liu, Q. H., & Zhao, Y. C. (2007). Regional patterns of soil organic carbon stocks in China. *Journal of Environmental Management*, *85*(3), 680–689. https://doi.org/10.1016/j.jenvman.2006.09.020

Yu, G., Li, X., Wang, Q., & Li, S. (2010). Carbon storage and its spatial pattern of terrestrial ecosystem in China. *Journal of Resources and Ecology*, *1*(2), 97–109. https://doi.org/10.3969/j.issn.1674-764x.2010.02.001

Yuan, W. P., Zheng, Y., Piao, S. L., Ciais, P., Lombardozzi, D., Wang, Y. P., Ryu, Y., Chen, G. X., Dong, W. J., Hu, Z. M., Jain, A. K., Jiang, C. Y., Kato, E., Li, S. H., Lienert, S., Liu, S. G., Nabel, J. E. M. S., Qin, Z. C., Quine, T., … Yang, S. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science Advances*, *5*(8), eaax1396. https://doi.org/10.1126/sciadv.aax1396

Zhang, F., Wang, Z., Glidden, S., Wu, Y. P., Tang, L., Liu, Q. Y., Li, C. S., & Frolking, S. (2017). Changes in the soil organic carbon balance on China's cropland during the last two decades of the 20th century. *Scientific Reports*, *7*, 7144. https://doi.org/10.1038/s41598-017-07237-1

Zhao, Y. C., Wang, M. Y., Hu, S. J., Zhang, X. D., Ouyang, Z., Zhang, G. L., Huang, B. A., Zhao, S. W., Wu, J. S., Xie, D. T., Zhu, B., Yu, D. S., Pan, X. Z., Xu, S. X., & Shi, X. Z. (2018). Economics- and policy-driven organic carbon input enhancement dominates soil organic carbon accumulation in Chinese croplands. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), 4045–4050. https://doi.org/10.1073/pnas.1700292114

Zhou, G. Y., Liu, S. G., Li, Z., Zhang, D. Q., Tang, X. L., Zhou, C. Y., Yan, J. H., & Mo, J. M. (2006). Old-growth forests can accumulate carbon in soils. *Science*, *314*(5804), 1417. https://doi.org/10.1126/science.1130168

Zhou, G., Xu, S., Ciais, P., Manzoni, S., Fang, J., Yu, G., Tang, X., Zhou, P., Wang, W., Yan, J., Wang, G., Ma, K., Li, S., Du, S., Han, S., Ma, Y., Zhang, D., Liu, J., Liu, S., … Chen, X. (2019). Climate and litter C/N ratio constrain soil organic carbon accumulation. *National Science Review*, *6*(4), 746–757. https://doi.org/10.1093/nsr/nwz045

Zhou, T., Geng, Y. J., Chen, J., Pan, J. J., Haase, D., & Lausch, A. (2020). High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms. *Science of the Total Environment*, *729*, 138244. https://doi.org/10.1016/j.scitotenv.2020.138244

Zhou, Y., Hartemink, A. E., Shi, Z., Liang, Z. Z., & Lu, Y. L. (2019). Land use and climate change effects on soil organic carbon in North and Northeast China. *Science of the Total Environment*, *647*, 1230–1238. https://doi.org/10.1016/j.scitotenv.2018.08.016

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Li, H., Wu, Y., Liu, S., Xiao, J., Zhao, W., Chen, J., Alexandrov, G., & Cao, Y. (2022). Decipher soil organic carbon dynamics and driving forces across China using machine learning. *Global Change Biology*, 28, 3394–3410. https://doi.org/10.1111/gcb.16154