Disentangling land model uncertainty via Matrix-based Ensemble Model Inter-comparison Platform (MEMIP)

Cuijuan Liao, Yizhao Chen, Jingmeng Wang, Yishuang Liang, Yansong Huang, Zhongyi Lin, Xingjie Lu, Yuanyuan Huang, Feng Tao, Danica Lombardozzi, et al.

To cite this version:

Cuijuan Liao, Yizhao Chen, Jingmeng Wang, Yishuang Liang, Yansong Huang, et al.. Disentangling land model uncertainty via Matrix-based Ensemble Model Inter-comparison Platform (MEMIP). Ecological Processes, 2022, 11 (1), pp.14. 10.1186/s13717-021-00356-8. hal-03566553

HAL Id: hal-03566553
https://hal.archives-ouvertes.fr/hal-03566553
Submitted on 16 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Disentangling land model uncertainty via Matrix-based Ensemble Model Inter-comparison Platform (MEMIP)

Cuijuan Liao¹, Yizhao Chen²,³*, Jingmeng Wang¹, Yishuang Liang¹, Yansong Huang⁴, Zhongyi Lin⁵, Xingjie Lu⁵, Yanjuan Huang⁶, Feng Tao⁵, Danica Lombardozzi⁷, Almut Arneth⁸, Daniel S. Goll⁹, Atul Jain¹⁰, Stephen Sitch¹¹, Yanluan Lin¹, Wei Xue¹², Xiaomeng Huang¹ and Yiqi Luo¹³*

Abstract

Background: Large uncertainty in modeling land carbon (C) uptake heavily impedes the accurate prediction of the global C budget. Identifying the uncertainty sources among models is crucial for model improvement yet has been difficult due to multiple feedbacks within Earth System Models (ESMs). Here we present a Matrix-based Ensemble Model Inter-comparison Platform (MEMIP) under a unified model traceability framework to evaluate multiple soil organic carbon (SOC) models. Using the MEMIP, we analyzed how the vertically resolved soil biogeochemistry structure influences SOC prediction in two soil organic matter (SOM) models. By comparing the model outputs from the C-only and CN modes, the SOC differences contributed by individual processes and N feedback between vegetation and soil were explicitly disentangled.

Results: Results showed that the multi-layer models with a vertically resolved structure predicted significantly higher SOC than the single layer models over the historical simulation (1900–2000). The SOC difference between the multi-layer models was remarkably higher than between the single-layer models. Traceability analysis indicated that over 80% of the SOC increase in the multi-layer models was contributed by the incorporation of depth-related processes, while SOC differences were similarly contributed by the processes and N feedback between models with the same soil depth representation.

Conclusions: The output suggested that feedback is a non-negligible contributor to the inter-model difference of SOC prediction, especially between models with similar process representation. Further analysis with TRENDY v7 and more extensive MEMIP outputs illustrated the potential important role of multi-layer structure to enlarge the current ensemble spread and the necessity of more detail model decomposition to fully disentangle inter-model differences. We stressed the importance of analyzing ensemble outputs from the fundamental model structures, and holding a holistic view in understanding the ensemble uncertainty.

Keywords: Soil organic carbon, Inter-model comparison, Uncertainty analysis, Carbon–nitrogen coupling, Vertical resolved soil biogeochemistry structure

Introduction

Whether or not the current land carbon (C) sink will persist into the future is a major source of uncertainty in assessing the global C budget (Friedlingstein et al. 2019; Piao et al. 2020). Proper understanding and identification of the sources of uncertainty is a critical step to improve
the prediction of C dynamics in a fast-changing world (Hoffman et al. 2017).

Outputs of Model Inter-comparison Projects (MIPs) provide information of global land C storage and budget for single model and ensembled projections. Meanwhile, uncertainty widely exists over the various components of the ensemble outputs (Friedlingstein et al. 2006). Model analysis indicated that various potential uncertainty sources, e.g., climate forcing (Ahlström et al. 2017; Bonan et al. 2019; Wu et al. 2018), process representation (Koven et al. 2015; Wang and Houlton 2009; Wieder et al. 2018, 2013; Zaehle et al. 2015) and parameterization (Keenan et al. 2012; Luo and Schuur 2020; Shi et al. 2018), can cause substantial differences in model outputs. However, it is difficult to fully distinguish the contribution of each factor, with little information being extracted from the model structure. One important reason for the situation is that the earth system itself is holistic, with multiple feedbacks among different components (Heinze et al. 2019). Every single model advance in one component probably causes changes in the related components to varying degrees. The complexity then dramatically increases in the ensemble simulations, with models sharing similar general structures but are divergent in representing specific processes (Rogers et al. 2017). Until now, little quantitative information has been attained on the uncertainty contributions from the process itself and the interactions between different processes.

Soil organic carbon (SOC) is the largest C reservoir in the biosphere, and its interactions with vegetation and atmosphere substantially influence the variability of global C cycling and C-climate feedback (Cox et al. 2000; Friedlingstein et al. 2006; Jones et al. 2003; Koven et al. 2011). Despite its importance, great uncertainty exists in estimating SOC stocks and fluxes and the corresponding responses to key environmental variables. Ensemble outputs from Coupled Model Intercomparison Project 5 (CMIP5) suggested that the estimated global SOC varied from 510 to more than 3000 Pg C and the spatial distributions are poorly correlated with observation-based data (Anav et al. 2013; Luo et al. 2015a; Todd-Brown et al. 2014). Outputs from the Multi-Scale Synthesis and Terrestrial Model Inter-comparison (MSTMIP) showed a similar spread of global estimation of historical SOC storage to that from CMIP5 (Tian et al. 2015). Such large inter-model differences seriously affect the credibility of our prediction on future global land C budget.

A major recent advance in SOC modelling in earth system models (ESMs) is the explicit representation of the vertically resolved structure of SOC exchange and mixing (Burke et al. 2017; Camino-Serrano et al. 2018; Koven et al. 2013). This helps ESMs to make better SOC prediction over the permafrost areas, which contain large quantities of soil organic material (SOM) in deep frozen soils and is sensitive to future warming (Schuur et al. 2015). With the quantification of deeper soils, a significant increase of the global estimation of SOC storage was found in land models (Shi et al. 2018). In addition, the feedback of C and N cycling to climate warming was predicted to be significantly modified once the deeper SOM was explicitly considered, especially through the release of C from deeper soils over permafrost ecosystems (Koven et al. 2015; McGuire et al. 2018). These model advancements have potentially improved the estimation of current and future amount of SOC stocks and fluxes. However, it remains largely unknown that if incorporating such a new process in specific models would cause larger uncertainty of SOC predictions in ensemble simulations.

A matrix-based framework has been developed to decompose various land C processes into unified components (Xia et al. 2013). Using a unified framework, crucial elements of the land C cycling can become traceable in a matrix form. Then the inter-model differences can be explicitly analyzed and quantified using the traceability analysis. This framework offers effective diagnosis of model uncertainty through a three-dimensional model output space: C residence time, C flux, and C storage potential (Luo et al. 2017), which has been widely applied in multiple state-of-art land models, such as in the community land model (CLM) (Huang et al. 2018a; Lu et al. 2019), the community atmosphere biosphere land exchange model (CABLE) (Xia et al. 2013), the organizing carbon and hydrology in dynamic ecosystems model (ORCHIDEE) (Huang et al. 2018b) and the boreal ecosystem productivity simulator (BEPS) (Chen et al. 2015). Following the framework, the inter-model differences can be quantitatively attributed to the sources.

We developed a Matrix-based Ensemble Model Intercomparison Platform (MEMIP, ver. 1.0) targeting to fully disentangle the uncertainty sources in inter-model comparisons. We incorporated SOC modules from different land models as ensemble members, and a vertically resolved biochemistry structure into an ESM, CIESM (Fig. 1). Each ensemble member was converted to a unified matrix form. Following the framework from Luo et al. (2017), we then decomposed and compared the SOC components in single-layer (SL) and multi-layer (ML) models using members in MEMIP. Through this study, we investigate how the ML structure modifies the SOC prediction and the corresponding traceable components, the explicit contribution of the process (PC), and the N feedback results from the competition between vegetation and soil (NFB) to the inter-model difference, and illustrate the potential implications to future MIP analysis.
Methods
Development of the MEMIP
SOC modeling schemes and the multi-layer structure
The land biosphere model used in this study is the land component of community integrated earth system model (CIESM), CIESM-LAND. The model is based on the framework of the Community Land Model version 4 (CLM 4.0). The CN biogeochemistry module (CLM-CN) is from Biome-BGC, which is a first-order kinetic scheme of 7 pools (3 litter and 4 soil pools) specified with different decomposition rates (Oleson et al. 2010). The soil temperature (ξ_T) and water scalars (ξ_W) were considered to limit SOC decomposition. In the single layer version of the Biome-BGC model (BGC-SL), the weighted means of surface soil temperature and water potential from the top 5 soil layers were used to quantify ξ_T and ξ_W, respectively. The organic N cycling follows the same pool structure with the C cycling. N mineralization and immobilization occur along with the C decomposition. The N scalar (ξ_N) is the ratio of actual and potential N immobilization, which is determined by the N competition between vegetation and soil (Additional file 1: Text S1).
CIESM-LAND introduced following modifications: a newly released global soil property database and an improved thermal roughness parameterization. The International Geosphere–Biosphere Programme
(IGBP) soil data sets were replaced by the Global Soil Data set for use in ESMs (GSDE, Shangguan et al. 2014). GSDE provides soil texture information of eight layers extending from the surface to the deep soil which will impact the land surface heat and water flux, as well as soil respiration. Meanwhile, a new thermal roughness length scheme developed by Yang et al. (2008, 2002) was adopted to improve the sensible heat flux and land surface energy budget. A test of the scheme in the Noah land surface model suggested that simulated land surface temperature was improved relative to the original scheme (Chen et al. 2011). Full introduction about CIESM-LAND can be found in Lin et al. (2020).

Four additional SOC models from CENTURY, CABLE, the Lund–Potsdam–Jena model (LPJ), and the Joint UK Land Environment Simulator (JULES) were explicitly implemented as ensemble members of MEMIP (Fig. 1). The models share similar multi-pool structure with first-order decay equations, but with different designs of SOM pools, transferring scheme and environmental scalars. The CENTURY scheme has a 6-pool structure of SOC, with 3 litter and 3 soil pools. Equations for $\xi_T$ and $\xi_U$ are the same as the BGC scheme. The CABLE scheme has a 5-pool SOC structure, with 2 litter pools (structure and metabolic litter pools) and 3 soil pools (microbial, slow and passive pools). The LPJ scheme is a 4-pool model, with C transitions from the above- and below-ground litter pools to the fast soil pool with a fixed proportion and the rest to the slow soil pool. The JULES scheme follows a Roth-C framework, with plant materials input to decomposable (DPM) and resistant (RPM) litter pools. C goes into the two soil pools [microbial biomass (BIO) and long-lived humified (HUM) pools]. In the CABLE, LPJ and JULES models, the surface soil temperature and water content, i.e., the average condition of soil water content from top 5 layers, are used to quantify $\xi_T$, $\xi_W$, respectively. Detail equations for each model can be found in Additional file 1: Table S1.

The vertically resolved biogeochemistry model, i.e., the multiple layer (ML) structure, was further incorporated following Koven et al. (2013). In the model, SOM dynamics of each soil layer and the vertical exchanges were explicitly considered. More detailed information can be found in Additional file 1: Table S1 and Text S2.

At the current stage, two members, Biome-BGC (BGC) and CENTURY (CEN), were coupled with the ML structure. Therefore, seven ensemble members were created in total, i.e., BGC-SL, BGC-ML, CEN-SL, CEN-ML, CABLE-SL, LPJ-SL and JULES-SL (Fig. 1). The same CN coupling model, i.e., CLM-CN, was used for all members.

**Matrix representation of the models**

All ensemble members were organized in one matrix form that captures the entire SOC dynamics. The overall control equation is

$$\frac{dX(t)}{dt} = U(t) - \xi(t)AX(t) - V(t)X(t) \quad (1)$$

The left side of the equation represents the net C pool changes per unit time.

The first term on the right side of the equation, $U(t)$, is the vector representing the C input into the system at time $t$. It is the product of the C input rate at time $t$ ($C_{\text{in}}(t)$) and the vector representing the C input allocation into different pools ($B$):

$$U(t) = C_{\text{in}}(t)B \quad (2)$$

The second term of Eq. (1), $\xi(t)A X(t)$, represents the C transfers and losses over different C pools at the same soil layer. $\xi(t)$ is the matrix for environmental scalar at time $t$:

$$\xi(t) = \xi_T(t)\xi_W(t)\xi_N(t)\xi_D(t) \quad (3)$$

where $\xi_T(t)$, $\xi_W(t)$, $\xi_N(t)$ and $\xi_D(t)$ are the soil temperature, water, $N$ and depth scalars at time $t$. In the SL models, $\xi_T$, $\xi_W$, and $\xi_N$ are considered as the average condition of surface layers; in the ML models, $\xi_T$, $\xi_W$ and $\xi_N$ are calculated for every single layer and an extra depth scalar ($\xi_D$) was considered following the scheme of CLM 4.5 (Oleson et al. 2013).

Matrix $A$ is the C transfer matrix and Matrix $K$ is a diagonal matrix for baseline decomposition rates of each C pool.

The third term, $V(t)X(t)$ represents the vertical exchange between different SOC layers per unit time. This term is only for the ML models. N cycling mainly influences SOC prediction by modifying $U(t)$ through limiting vegetation C assimilation and modifying SOC decomposition through $\xi_N$ (see detail in Additional file 1: Text S1).

To verify the validity of the matrix models, both original and matrix models were created for each member in the MEMIP. We conducted 100-year global simulations for both original and matrix models. The results proved that the matrix models can exactly reproduce the outputs from the original models (Additional file 1: Fig. S1).

**Model simulations**

We conducted spin-ups for each ensemble member to get equilibrium conditions using the climate forcing of 1900–1920, and pre-industrial atmospheric CO$_2$ concentration and N deposition. Then we ran transient simulations of CIESM-LAND for the historical
period (1900–2000) using each ensemble member, which represent different SOC calculations, in parallel. All ensemble members were run using the same climate forcing, CRU-NCEP 4.0 (Viovy 2012), transient historical atmospheric CO₂ concentration, and N deposition dynamics (Lamarque 2005). We ran each ensemble member under both C-only and CN modes.

**Traceability analysis**

To explicitly trace and disentangle the sources of model difference, we applied the traceability analysis framework following Luo et al. (2017):

\[
X(t) = X_e(t) + X_p(t)
\]

(4)

where \(X(t), X_e\) and \(X_p\) are the C storage, C storage capacity and the C storage potential at time \(t\), respectively. \(X_e(t)\) represents the C storage at the equilibrium condition, which is the product of C residence time (\(\tau_E(t)\)) and \(C_{in}(t)\):

\[
X_e = \tau_E(t)C_{in}(t)
\]

(5)

\(\tau_E(t)\) can be further decomposed into baseline residence time (\(\tau_E^b\)) and \(\xi(t)\):

\[
\tau_E(t) = \tau_E^b \xi(t)^{-1}
\]

(6)

where \(\tau_E^b\) is determined by baseline decomposition rate and C transfer schemes in the SL models:

\[
\tau_E^b = (KA)^{-1}B
\]

(7)

In the ML models, an additional depth-related factor was added to modify \(\tau_E^b\) (Additional file 1: Text S2). \(\xi(t)\) is calculated as the product of multiple soil environmental variables including \(\xi_T(t), \xi_M(t)\) and \(\xi_D(t)\) (the depth scalar is only for ML members).

\(X_p(t)\) defines the ecosystem potential to store or lose C in addition to C storage capacity. It is the product of C storage change rate (\(X'(t)\)) and the chasing time (\(\tau_{ch}(t)\)):

\[
X_p(t) = X'(t)\tau_{ch}(t)
\]

(8)

where \(\tau_{ch}\) is defined as the common part for \(X_e\) and \(X_p\), which represents the expected residence time for C transferring among different SOC pools (Luo et al. 2017):

\[
\tau_{ch}(t) = (KA\xi(t))^{-1}
\]

(9)

In this study, we focused on \(X_e\) analysis as the contribution of \(X_p\) to the total amount of SOC is very small (~10 Pg C for \(X_p\) versus 600–1000 Pg C for \(X_e\), Additional file 1: Fig. S2).

**Quantifying the contribution of modeled SOC differences**

Following the model decomposition scheme introduced above, the SOC differences between ensemble members can be represented as following:

\[
R_{SOC,MOD12,t} = R_{Cin,MOD12,t}, R_{\tau_E^b,MOD12,t}, R_{\xi,MOD12,t}, R_{\tau_{ch},MOD12,t}
\]

(10)

where \(R_{SOC,MOD12,t}\) are the ratios between SOC, \(C_{in}\), baseline residence time, and environmental scalars from one model to another, respectively (i.e., \(R_{SOC,MOD12,t} = \frac{SOC_{MOD12,t}}{SOC_{MOD2,t}} = \frac{C_{in,MOD12,t}}{C_{in,MOD2,t}} R_{\tau_E^b,MOD12,t} R_{\xi,MOD12,t} R_{\tau_{ch},MOD12,t}\)).

Equation 10 was then log-transformed to quantify the additive contribution of different components:

\[
\log(R_{SOC,MOD12,t}) = \log(R_{Cin,MOD12,t}) + \log(R_{\tau_E^b,MOD12,t}) + \log(R_{\xi,MOD12,t}) + \log(R_{\tau_{ch},MOD12,t})
\]

(11)

Thereafter, the following approach was used to calculate the relative changes of each component against the SOC change:

\[
f_i = \sum_t \frac{\log(R_{i,MOD12,t})}{\log(R_{SOC,MOD12,t})}
\]

(12)

where \(f_i\) is the multi-year mean of the contribution of component \(i\) (i.e., \(C_{in}, \tau_E^b\) and \(\xi\)) relative to the SOC difference. Its positive or negative value represents if the change direction of component \(i\) is the same or the opposite to SOC change direction.

We then conducted a pairwise comparison to illustrate the sources of SOC difference with two SOC decomposition structures (CEN versus BGC) and two SOC layer representations (SL versus ML). The specific comparisons are BGC-SL versus BGC-ML (Layer_BGC comparison), CEN-SL versus CEN-ML (Layer_CEN comparison), BGC-SL versus CEN-SL (Model_SL comparison), and BGC-ML versus CEN-ML (Model_ML comparison).

When included in CIESM-LAND, MEMIP can explicitly attribute inter-model differences to processes and feedbacks. In this study, the relative contributions from processes (PC, i.e., different SOC decomposition structures and layer representations) and NFB (i.e., N competition between vegetation and soil) were explicitly disentangled. To quantify the relative contributions of PC and NFB, the C-only outputs of the 4 ensemble members
were compared to determine the relative changes of the corresponding component induced by PC. Thereafter the outputs were compared against the results from CN mode to get the contributions derived from NFB.

In this case, \( f_{\tau E}' \) is mainly from PC, i.e., different SOC decomposition scheme and depth representation. NFB also contributes to \( f_{\tau E}' \) in ML models with the incorporation of the vertical mixture term through modifying canopy condition and then the soil moisture and temperature (Additional file 1: Text S2). Model differences in \( f_{C_{\text{in}}} \) is from NFB through soil-vegetation N competition. \( f_{\xi T}^{-1} \) and \( f_{\xi M}^{-1} \) are contributed by both PC and NFB. \( f_{\xi T} \) and \( f_{\xi M} \) are from the different uses of soil temperature and soil moisture layers as the PC term, and soil temperature and moisture modifications due to vegetation structure changes with the NFB term. \( f_{\xi N}^{-1} \) differences are from different N pool structures and corresponding parameterization as the PC term, and the modification through C input to the soil as the NFB term (See detailed explanations in Additional file 1: Text S3). The magnitude and direction of \( \xi_N \) contribution depend on the relative magnitude of changes in PC and NFB (Additional file 1: Text S1). Briefly, higher \( \tau E \) increases (i.e., change from PC) the demand of soil N immobilization and, therefore, decreases \( \xi_N \), which in turn, promotes SOC accumulation, while lower \( C_{\text{in}} \) (i.e., change from NFB) decreases both total mineral N and vegetation N demand. The former (i.e., lower total mineral N) decreases \( \xi_N \) (thereby increasing SOC) and the latter (i.e., lower vegetation N demand) increases \( \xi_T \) (thereby decreasing SOC). Therefore, the final contribution of NFB depends on the relative importance of the two components.

Due to the same biophysical model being used in all simulations, the contributions of NFB to \( f_{\tau E}' \), \( f_{\xi T}^{-1} \) and \( f_{\xi M}^{-1} \) were very small in this study (Additional file 1: Fig. S3). Hence, we assumed \( f_{\tau E}' \), \( f_{\xi T}^{-1} \) and \( f_{\xi M}^{-1} \) were all from PC.

**Results**

**Comparison between the single and multiple layer models**

By explicitly representing the vertically resolved soil biogeochemistry structure, the two ML models predicted higher SOC amounts across large regions (Fig. 2). Areas with increasing SOC over the high-latitudes of the Northern hemisphere were mainly produced through the prolonging of \( \tau E \), while the other regions with increasing SOC had co-contributions by the changes in both \( C_{\text{in}} \) and \( \tau E \) (Fig. 2). For example, SOC increase in major parts of China was the result of \( C_{\text{in}} \) increase and \( \tau E \) decrease in the Layer_BGC comparison and the increase of both \( C_{\text{in}} \) and \( \tau E \) in the Layer_CEN comparison, respectively. Areas with decreasing SOC were mainly located in central Siberia, eastern Europe and part of Canada, mainly due to \( C_{\text{in}} \) decreases in the ML models. Further decomposing \( \tau E \) into separate components, the ML structure caused similar changes of spatial patterns in various \( \tau E \) components for the two models compared to the SL structure (Fig. 3). \( \xi_T \) only increased in areas with latitudes approximately higher than 40°N and the
Tibetan plateau, while it showed major decreases for the remaining land areas (Fig. 3a). $\xi_M$ showed an overall increase in the ML models (Fig. 3b). The spatial pattern of $\xi_N$ was similar for the two comparisons. Larger $\xi_N$ decrease was found for the Layer_CEN comparison, especially for Eastern Europe, Canada, South Africa and Australia (Fig. 3c). $r'_E$ increase was also found for major land areas, except Western North America, Central Asia, Southern Africa, and Australia (Fig. 3d).
Layer_CEN comparison exhibited a larger magnitude of changes of $\tau_E$ (Fig. 4b).

Globally, the ML models showed higher $\tau_E$ but lower $C_{in}$ (Fig. 4a). The combination led to higher predictions of global SOC from 664 and 662.3 Pg C to 881.8 Pg C and 1003.4 Pg C for the BGC and CEN models, respectively. Changes in $\tau_E$ resulted from both $\tau_E^f$ and $\xi$ (Fig. 4b). Extra consideration of the depth scalar ($\xi_D$) was the main source of $\xi$ changes, accounting for 14.5% (from 0.144 to 0.123) and 14% (from 0.152 to 0.131) of $\tau_E^f$ changes for the BGC and CEN models, respectively (Fig. 4b).

**Comparison between the BGC and CEN models**
SOC divergences between the two model schemes were larger with the ML structure (i.e., the Model_ML comparison) compared to that with the SL structure (i.e., the Model_SL comparison). Differences of SOC in the Model_ML comparison greatly increased over 10 times for most of the land pixels compared to those from the Model_SL comparison (Fig. 2a). Spatially, the SOC difference between BGC_SL and CEN_SL was relatively small in magnitude with no consistent spatial patterns. Predicted SOC of BGC-ML exceeded CEN-ML mainly over the permafrost areas, and the pattern reversed in the rest part of global land (Fig. 2a). In the Model_ML comparison, the SOC difference was primarily from the $\tau_E$ difference with very similar spatial patterns for the BGC and CEN models (Fig. 2). $C_{in}$ difference influenced the pattern of predicted SOC over particular areas, e.g., Northeastern North America and Western Europe (Fig. 2). Differences in $\xi_T$ and $\xi_M$ presented small magnitude of changes in both comparisons (Fig. 3a and b). The spatial pattern of $\xi_N$ was similar for the two comparisons. Main differences are from Northwestern North America, Siberia, South Africa and Australia, where $\xi_N$ from the BGC model was higher in the Model_ML comparison but was lower in the SL comparison (Fig. 3c). $\tau_E^f$ displayed an overall larger discrepancy in the Model_ML comparison and the spatial distribution was similar to that from $\tau_E$. Taken together, SOC difference between models in the Model_ML comparison was 71.5 times of that in the Model_SL comparison (Fig. 4a).

**Vertical distribution of soil organic carbon and environmental scalars**
Vertical distribution of SOC from the ML models showed a contrasting pattern between the high latitude areas of North Hemisphere (> 60° N) and the rest part of global land (Fig. 5a–c). The mean SOC density was much higher over the high latitude areas than the other land, with a large proportion of SOC accumulated over the deep layers. The SOC density increased from top layers to layer 9 (~2.3 m deep) and then decreased to layer 10 (~3.8 m deep). Following the distribution of SOC, soil organic nitrogen (SON) presented a very similar vertical distribution pattern globally (Fig. 5d–f).

Regarding soil moisture and temperature scalars, both values were much lower for the high latitude areas (Fig. 6), probably due to the harsh environment, i.e., low soil temperature and moisture condition (Additional file 1: Fig. S4). The combination led to a much lower decomposition rate over those areas than the other land areas, and thus highly limited soil decomposition (Fig. 3). In return, $\tau_E$ was much longer over high latitude areas than the rest part of the global land (Fig. 2).

**Traceability analysis and difference contributions**
The contribution of SOC variation from different sources was quantified in each pairwise comparison of ensemble members (Fig. 7). In the Layer_BGC and Layer_CEN comparisons (Fig. 7a and b), the SOC increases in the ML models were mainly from $f_{\tau_E^f}$, $f_{\xi_D}$, and $f_{\xi_T^1}$, and were buffered by $f_{C_{in}}$ and $f_{\xi_M^1}$. $f_{\xi_N^1}$ represented a small fraction of differences in the Layer_BGC and Layer_CEN comparisons. Little $f_{\xi_N^1}$ was from PC, indicating that the ML structure had a limited impact on the structural aspect of N cycling (Fig. 7a, b). Overall, PC-induced model differences dominated in the two comparisons, with contributions of 82.45% and 85.26% for the Layer_BGC and Layer_CEN comparisons, respectively (Fig. 7a, b).

The contribution from NFB became much more important in explaining the model differences in the Model_SL and Model_ML comparisons (Fig. 7c, d). In comparing the two SL models, the relative change rate from every element was much higher than the SOC difference, due to the similar SOC predictions from the two models (Fig. 4a). Higher SOC prediction in BGC_SL compared to CEN_SL was mainly from $f_{\xi_{N,ML}}$ and was buffered by $f_{C_{in}}$ and $f_{\xi_E^f}$ (Fig. 7c). $f_{\xi_{N,ML}}$ from NFB exceeds that from PC, suggesting that the different models influenced SOC through modifying vegetation C uptake rather than regulating N cycling. Similar pattern was found in comparing the ML representations (Fig. 7d). However, $f_{\xi_{N,ML}}$ is smaller with increasing $f_{\xi_E^f}$ in this comparison, leading to a higher SOC prediction from CEN_ML (Figs. 4b and 7d). Taken together, the contribution of NFB to SOC difference was comparable to PC in these two comparisons. NFB contributions were 53.02% and 46.64% for the Model_SL and Model_ML comparisons, respectively.

**SOC outputs from TRENDY v7 and MEMIP members**
Predicted SOC from TRENDY v7 (TRENDY) and more extensive outputs from MEMIP are shown in Fig. 8. TRENDY output presented a spread of SOC predictions
with a range of 540–3355 Pg C. Among the models, the result from CLM5.0, one includes an explicit ML structure, predicts a much higher SOC than the other models. The ML models from MEMIP, using a similar ML model in CLM5.0, also exhibits obvious higher SOC prediction than the corresponding the SL models. However, the soil depth considered in CLM5.0 is much deeper than the one in CIESM-LAND (8.6 m versus 3.8 m). With the consideration of N limitation, SOC predictions from CN mode are lower than C-only mode in MEMIP outputs. Correspondingly, the C-only models (1340–3011 Pg C) showed much larger range of spread than the CN models.

Fig. 5 Vertical distribution of soil organic carbon (SOC) for a global, b high latitude land (latitude > 60°) and c the rest land, and soil organic nitrogen (SON) for for d global, e high latitude land (latitude > 60°) and f the rest part of global land, from CEN-ML and BGC-ML. Each horizontal line represents a soil layer.
Fig. 6 Vertical distribution of soil temperature scalar for a global, b high latitude land (latitude > 60°) and c the rest part of global land, and soil temperature scalar for d global, e high latitude land (latitude > 60°) and f the rest part of global land, from CEN-ML and BGC-ML. Each horizontal line represents a soil layer.

(See figure on next page.)

Fig. 7 Relative contributions of different elements to the final SOC differences in pairwise comparisons. Pink and light blue bars represent the feedback and process contribution, respectively. Purple bars represent the combined effect from feedback and process from a single element. Model names with bold font and an asterisk indicate the model with higher SOC prediction. Positive values of the x-axis indicate that the direction of the element is the same with models of higher SOC prediction and vice versa. $C_{in}$: C input rate, $\tau_E$: baseline residence time, $\xi_T$: soil temperature scalar, $\xi_W$: soil water scalar, $\xi_D$: depth scalar, $\xi_N$: N scalar.
Fig. 7 (See legend on previous page.)
(666–1481 Pg C), with only changes on the SOM models. The situation is much more complex in real ensemble output. Contrary to the MEMIP output, the CN models showed larger spread than the C-only models (802–1601 Pg C for C-only models, 540–3355 Pg C for CN models with CLM5.0 and 540–1885 Pg C for CN models without CLM5.0, respectively). With various SOM models used in different biosphere models, it is difficult to tell the sources of uncertainty without looking into specific model structures.

**Discussion**

Additional processes characterizing deeper soil dynamics showed a significant impact on the global SOC storage, especially for the permafrost region. As presented from the MEMIP output, SOC outputs from the ML models were higher than those of the SL models. Similar patterns were also confirmed in the SOC output from TRENDY, in which CLM5.0 with the ML structure showed a much higher SOC prediction (Fig. 8). Consistent conclusions were found in previous studies through comparing the SL and ML models (Burke et al. 2017; Camino-Serrano et al. 2018; Koven et al. 2013). Although further analysis is needed to confirm the effect in the ensemble outputs, the current results suggest that the ML structure potentially enlarges the spread of current ensemble SOC prediction. Within the MEMIP, we further investigated why the final SOC was different by analyzing its elements.

SOC difference between the SL and ML models was mainly from $\tau_E$, which can be further attributed to the differences from both $\xi$ and $\tau'_E$. The effect from $\xi$ is straightforward due to its direct modifications to $\tau'_E$. Extra consideration of $\xi_D$ in the ML models imposed a strong effect on the C decomposition with its decreasing tendency to the deeper soil layers (Additional file 1: Text S2). With SOC gradually diffused into the deeper layers, where decomposition rates were much lower than the top layers, $\xi_D$ thus caused higher $\tau_E$ for the ML models (Fig. 2). Since the SL models used average condition of top 5 layers of soil water potential and temperature to quantify the corresponding scalars, $\xi_T$ and $\xi_W$ decreased and increased with the incorporation of the ML structure, respectively (Fig. 6). Because the same biogeochemical model and $\xi$ functions are used for the BGC and CEN models, the contributions from the two scalars were relatively small comparing to $\xi_D$. However, with more diverse equations incorporated from different models, extra uncertainties from both process and feedback would be generated to the outputs (Chen et al. 2015; Falloon et al. 2011), and are likely a contribution to the differences in the MEMIP and TRENDY results.

More complicated interactions were generated in $\tau'_E$, which includes the major model structural elements. In the SL members, $\tau'_E$ can be easily quantified and compared with time-invariant inputs. As presented in Fig. 2, $\tau'_E$ is higher from the CEN model than that from the BGC model. With the incorporation of the vertical mixture term, $\tau'_E$ was different for each layer (Additional file 1: Text S2). Due to the low decomposition rate over the deep layers, our results of $\tau'_E$ from the ML models clearly elongated and thus strongly contributed to the higher SOC predictions from the ML models than the SL models (Fig. 7). The exchange rates to the upper and lower layers depend on the depth and environmental factors and, therefore, create a more complex interaction between vertical (layer to layer) and horizontal (pool to pool) C exchanges. Further studies are needed to understand the internal and external controls in this component. A thorough parameter sensitivity analysis may be a good start point (Huang et al. 2018b).

The ML structure further modified the vegetation–soil interactions through N cycling. In the CIESM, stronger plant N competition between vegetation and soil happened in the ML models due to $\tau_E$ elongation. On one hand, the vegetation production was reduced because of the lower availability of soil mineral N that stems from relatively higher N demand for decomposition (Additional file 1: Text S2).
Lower vegetation production due to N limitation is a feedback that counteracts the SOC increase by reducing vegetative C input to the soil. On the other hand, the $\xi_N$ to soil C transition and decomposition was also changed. With the increasing demand for N immobilization but a decrease of both total mineral N and plant demand, $\xi_N$ exhibited different responses in different models, which promoted or buffered the SOC increase based on different model schemes (Fig. 7 and Additional file 1: Text S1). Nevertheless, the magnitude of the effect of $\xi_N$ on soil decomposition is much smaller than the effect on vegetation C uptake (Fig. 5a, b). Taken together, the total NFB effect buffered the SOC increase in the BGC and CEN comparisons. It should be mentioned that here, SOC increment from the SL to ML models is mainly an effect of additional soil layers leading to deeper SOM. The warming effect on soil respiration would be amplified with this deeper SOM, especially for the permafrost regions (Schuur et al. 2015). It can then be inferred that a decrease in soil mineral N availability may reduce the CO$_2$ fertilization effect on plants; and deeper SOM associated with additional soil layers might increase respiration compared to the CN models with SL schemes (Sokolov et al. 2008; Thornton et al. 2009; Zaehle et al. 2015), both of which would lead to decreasing estimations of future land C uptake. Future studies based on MEMIP can be implemented to explicitly quantify the potential consequences using various model schemes.

Difference between the SOM schemes was significantly amplified due to the ML structure for both spatial pattern and global mean, suggesting that the ML structure brings more complex interactions than simply adding deeper soils. For example, the ML structure caused different distributions of SOC to different depths, generating a contrasting response in permafrost areas and the rest of the land (Figs. 5 and 6). NFB showed a similar contribution as PC to the SOC difference in the SL and ML comparisons, indicating that the holistic response from the system may play an equal role as PC in the inter-model comparison. Meanwhile, the effect of NFB may buffer or expand the inter-model differences depending on the model structure and the corresponding parameterization (Fig. 7 and Additional file 1: Text S1), acting as a potential unpredictable part in the inter-model comparisons. Without explicit comparison and quantification, its contribution can hardly be elucidated. With increasing feedbacks represented in land models, it also implies that comparisons only showing the relative change from one specific process would miss more important systematic variations in real ensemble simulations and comparisons.

The CN coupling scheme, i.e., CLM-CN, used in this study offers insights into how CN coupling influences C storage, while further improvements in this process representation are necessary (Wieder et al. 2019). The potential biases in N representation seem not change the importance of NFB found from this study, as similar global response patterns to increasing N and atmospheric CO$_2$ were found in the CN coupling comparisons (Davies-Barnard et al. 2020). However, the importance of N varies among models (Davies-Barnard et al. 2020). Meanwhile, the CN coupling is one major nutrient feedback that has recently been incorporated in the land component of many ESMs. With further incorporations of other nutrient feedbacks, e.g., phosphorus, the actual ecosystem feedback to the ML structure should be stronger than the current prediction (Wieder et al. 2015).

**Conclusions**

Great efforts and resources have been invested to predict global C cycling using multiple models, but there is still wide variation in terrestrial C uptake across models. Compared to the ensemble outputs from CMIP5 and CMIP6, similar spreads were found in both land C uptake and feedback against climate change (Arora et al. 2020). Does this mean that incorporating new processes in ESMs does not improve the performance? To answer the question, we first suggest grouping and analyzing models based on major model advances. The bottom line is that the C centric land models include similar overarching biogeochemical processes, although representation of the processes can vary widely. In real applications, an improvement on a single process would fundamentally change the model projections (Wieder et al. 2013, 2015). Different models fit to the same observations, which then creates relatively good agreements on historical conditions but greater divergence in future predictions (Luo et al. 2015b). Once grouping models together, it then creates a situation that is difficult to analyze and understand. As illustrated in this study, switching one single process can only explain the magnitude of model change from that, but the situation in the real ensemble is much more complex (Fig. 6). One possible solution is to learn, cluster, and analyze models at the process level. As shown by Luo et al. (2016), the common baseline structural elements of land C cycling can be extracted from different models. Following the traceability framework, it is possible to explicitly separate and compare the common structural components and the unique elements in each model. The model differences can be traced back to its origin step by step as shown in this study. Moreover, with new observations and tools emerge, models can then be compared and benchmarked more properly and efficiently (Basile et al. 2020; Han et al. 2019; Minasny et al. 2017; Shi et al. 2020; Thum et al. 2020). Stemming from the model fundamental structure and properties, more meaningful and usable information can be obtained from the ensemble outputs.
Our findings further suggest the important role of feedbacks, i.e., the interactions between different components of an ecosystem, to contribute to the inter-model differences. With increasing feedbacks being represented in land models, it would reflect more complex trade-offs between biochemical and biophysical processes in nature, and potentially place extra emphasis on ensemble outputs. Given the current inclinations on process-based uncertainty analysis, the contribution of feedbacks is not yet well understood. Comprehensive analyses with full consideration of both feedback and process responses will pave the way to the holistic dis-entangling of ensemble uncertainty in land C cycling.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13717-021-00356-8.

Acknowledgements

This study is supported by the funding from the National Key Research and Development Program of China under Grant 2017YFA0604600. YC was supported by the National Science Foundation (NSF) under Cooperative Agreement 1852977. DL’s computing and data storage resources, including the Cheyenne supercomputer (https://doi.org/10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. DSG receives support from the ANR CLAND Convergence Institute.

Availability of data and materials

The MEMIP data that support the findings of this study are available from the corresponding authors upon reasonable request. The TRENDY v7 data are available from the TRENDY project coordinators upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing, China. 2Joint Innovation Center for Modern Forestry Studies, College of Biology and the Environment, Nanjing Forestry University, Nanjing, China. 3Department of Geography, University of Cambridge, Cambridge, UK. 4School of Life Science, Nanjing University, Nanjing, China. 5Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-Sen University, Guangzhou, China. 6CSIRO Oceans and Atmosphere, Aspendale 3195, Australia. 7Terrestrial Sciences Section, Climate and Global Dynamics, National Center for Atmospheric Research, Boulder, CO 80305, USA. 8Karlruhe Institute of Technology, Institute of Meteorology and Climate Research/Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany. 9Université Paris Saclay, CEA-CNRS-UVSQ, LSCE/IPSL, Gif sur Yvette, France. 10Department of Atmospheric Sciences, University of Illinois, Urbana, IL, USA. 11College of Life and Environmental Sciences, University of Exeter, Exeter, UK. 12Department of Computer Science and Technology, Tsinghua University, Beijing, China. 13Center for Ecosystem Science and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, USA.

Received: 10 March 2021 Accepted: 28 December 2021

Published online: 08 February 2022

References

Ahlstrom A, Schwurs G, Smith B (2017) The large influence of climate model bias on terrestrial carbon cycle simulations. Environ Res Lett 12:014004

Anav A, Friedlingstein P, Kidston M, Bopp L, Ciais P, Cox P, Jones C, Jung M, Myneni R, Zhu Z (2013) Evaluating the land and ocean components of the global carbon cycle in the CMIP5 earth system models. J Clim 26:6801–6843

Arora VK, Katavouta A, Williams RG, Brovkin V, Friedlingstein P, Schwinger J, Bopp L, Boucher O, Cadule P, Chamberlain MA, Christian JR, Deleire C, Fisher RA, Hajima T, Ilyina T, Joetzjer E, Kawamiya M, Koven C, Kasting J, Law RM, Lawrence DM, Lentot A, Lindsay K, Pongratz J, Raddatz T, Séférian R, Tachiiri K, Tjiputra JF, Wu T, Ziehn T (2020) Carbon-concentration and carbon-climate feedbacks in CMIP6 models, and their comparison to CMIP5 models. Biogeosciences 17:4173–4222

Basile SJ, Lin X, Wieder WR, Hartman MD, Keppel-Aleks G (2020) Leveraging the signature of heterotrophic respiration on atmospheric CO2 for model benchmarking. Biogeosciences 17:1293–1308

Bonan GB, Lombardozzi DL, Wieder WR, Oleksen KW, Lawrence DM, Hoffman FM, Collier N (2019) Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). Glob Biogeoch Cycles 33:1310–1326

Additional file 1: Table S1. Model information and set-ups in MEMIP. Figure S1. Matrix model validations for each branch: (a) spatial pattern of the model difference between the regular and matrix form, and (b) the pixel level correlations. Figure S2. The C storage potential (Xp) for the 4 branches (i.e., BGC_Sg, CEN_Sg, BGC_ml and CEN_ml) from 1900 to 2000. Figure S3. Correlations of soil temperature and soil water potential of C-only and CN models. Values are global mean annual outputs of top 5 layers from 1900 to 2000. Figure S4. Vertical distribution of soil temperature and soil water potential for global, high latitude land (latitude > 60°) and the rest land from CEN-ML and BGC-ML. The upper panel represents soil temperature and the lower panel represents soil water potential. Figure S5. Model outputs of (a) Mineral N, (b) Soil mineral N to plant, (c) N limitation to GPP and (d) LAI for different model members. Text S1. The soil-vegetation N competition. Text S2. Structural elements in soil baseline residence time. Text S3. Distinguish the contributions of model difference from process and feedback.
Schuur EAG, McGuire AD, Schädel C, Grosse G, Harden JW, Hayes DJ, Hugelius G, Koven CD, Kuhry P, Lawrence DM, Natali SM, Olefeldt D, Romanovsky VE, Schaefer K, Turetsky MR, Treat CC, Vonk JE (2015) Climate change and the permafrost carbon feedback. Nature 520:171–179. https://doi.org/10.1038/nature14338

Shangguan W, Dai Y, Duan Q, Liu B, Yuan H (2014) A global soil data set for earth system modeling. J Adv Model Earth Syst 6:249–263

Shi Z, Crowell S, Luo Y, Moore BK, Mahowald N, Randerson JT, Fung I, Lamarque J-F, Feddema JJ, Lee Y-H (2009) Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model. Biogeosciences 6:2099–2120. https://doi.org/10.5194/bg-6-2099-2009

Shi Z, Allison SD, He Y, Levine PA, Hoyt AM, Beem-Miller J, Zhu Q, Wieder WR, Thurman S, Randerson JT (2020) The age distribution of global soil carbon inferred from radiocarbon measurements. Nat Geosci 13:555–559

Sokolov AP, Kicklighter DW, Meillió JM, Felzer BS, Schlosser CA, Cronin TW (2008) Consequences of considering carbon-nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle. J Clim 21:3776–3796. https://doi.org/10.1175/2008JCLI2038.1

Thornton PE, Doney SC, Lindsay K, Moore JK, Mahowald N, Randerson JT, Fung I, Lamarque J-F, Feddema JJ, Lee Y-H (2009) Carbon-nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere-ocean general circulation model. Biogeosciences 6:2099–2120. https://doi.org/10.5194/bg-6-2099-2009

Thum T, Nabel JE, Tsuruta A, Aalto T, Dlugokencky EJ, Liski J, Luijkhx IT, Markkanen T, Pongratz J, Yoshida Y (2020) Evaluating two soil carbon models within a global land surface model using surface and spaceborne observations of atmospheric CO2 mole fractions. Biogeosciences 12:1–34

Tian H, Lu C, Yang J, Banger K, Huntzinger DN, Schwalm CR, Michalak AM, Cook R, Ciais P, Hayes D, Huang M, Ito A, Jain AK, Lei H, Mao J, Pan S, Post WM, Peng S, Poulette B, Ren W, Ricciuto D, Schaefer K, Shi X, Tao B, Wang W, Wei Y, Yang Q, Zhang B, Zeng N (2015) Global patterns and controls of soil organic carbon dynamics as simulated by multiple terrestrial biosphere models: Current status and future directions. Glob Biogeochem Cycles 29:775–792. https://doi.org/10.1002/2014GB005021

Todd-Brown KEO, Randerson JT, Hopkins F, Arora V, Hajima T, Jones C, Shevliakova E, Tjiputra J, Volodin E, Wu T, Zhang Q, Allison SD (2014) Changes in soil organic carbon storage predicted by Earth system models during the 21st century. Biogeosciences 11:2341–2356. https://doi.org/10.5194/bg-11-2341-2014

Viovy N (2012) CRU-NCEP Version 4. Available Href Httpdods Extra Cea Frdata-p529viowyoc2012. Accessed 11 Oct 2012

Wang Y, Houlton BZ (2009) Nitrogen constraints on terrestrial carbon uptake: implications for the global carbon-climate feedback. Geophys Res Lett 36:L24403

Wieder WR, Bonan GB, Allison SD (2013) Global soil carbon projections are improved by modelling microbial processes. Nat Clim Change 3:909–912. https://doi.org/10.1038/nclimate1951

Wieder WR, Cleveland CC, Smith WK, Todd-Brown K (2015) Future productivity and carbon storage limited by terrestrial nutrient availability. Nat Geosci 8:441–444. https://doi.org/10.1038/ngeo2413

Wieder WR, Hartman MD, Sultan BN, Wang Y, Koven CD, Bonan GB (2018) Carbon cycle confidence and uncertainty: exploring variation among soil biogeochemical models. Glob Change Biol 24:1563–1579

Wieder WR, Lawrence DM, Fisher BA, Bonan GB, Cheng SJ, Goodale CL, Grandy AS, Koven CD, Lombardozzi DL, Olesen KW, Thomas RQ (2019) Beyond static benchmarking: using experimental manipulations to evaluate land model assumptions. Glob Biogeochem Cycles 33:1289–1309. https://doi.org/10.1029/2018GB006141

Wu Z, Boke-Olen N, Fensholt R, Arndt J, Ekulundh L, Lehnert V (2018) Effect of climate dataset selection on simulations of terrestrial GPP: highest uncertainty for tropical regions. PLoS ONE 13:e0199383

Xia J, Luo Y, Wang Y-P, Harukaki Y (2013) Traceable components of terrestrial carbon storage capacity in biogeochemical models. Glob Change Biol 19:2104–2116. https://doi.org/10.1111/gcb.12172

Yang K, Koike T, Fuji H, Tamagawa K, Hirose N (2002) Improvement of surface flux parametrizations with a turbulence-related length. Q J R Meteorol Soc 128:2073–2087

Yang K, Koike T, Ishikawa H, Kim J, Li X, Liu H, Liu S, Ma Y, Wang J (2008) Turbulent flux transfer over bare-soil surfaces: characteristics and parameterization. J Appl Meteorol Climatol 47:276–290

Zaehe S, Jones CD, Houlton B, Lamarque J-F, Robertson E (2015) Nitrogen availability reduces CMIP5 projections of twenty-first-century land carbon uptake. J Clim 28:2494–2511. https://doi.org/10.1175/JCLI-D-13-00776.1

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.