Authors response

Title: Impacts of land-use change on biospheric carbon: an oriented benchmark using ORCHIDEE land surface model
Author(s): Thi Lan Anh Dinh, Daniel Goll, Philippe Ciais, and Ronny Lauerwald
MS No.: gmd-2024-42
MS type: Methods for assessment of models

Dear referees,

We thank two referees for their highly constructive comments and suggestions. Our responses (in blue) are listed below. The updated manuscript is also included underneath our responses.

We hope that our adjustments will make the manuscript clearer and more pleasing to the referees and other readers.

Thanks & best regards,
Lan Anh & co-authors.

Reviewer #1:

This study evaluates the performance of the ORCHIDEE dynamic global vegetation model (DGVM) in accurately depicting the effects of land-use change (LUC) on biospheric carbon stocks across Europe. Through a systematic evaluation, the authors compare ORCHIDEE's predictions with observation-based estimates of key variables such as net and gross primary productivity (NPP and GPP), biomass growth patterns, and soil organic carbon (SOC) stocks. Additionally, they employ interpretable machine learning techniques to uncover factors contributing to discrepancies between model simulations and observational data.

A notable aspect of this study is its comprehensive analysis of ORCHIDEE's ability to reproduce spatial patterns and temporal trends in biospheric carbon stocks. The results demonstrate encouraging agreement between model predictions and observed data, particularly in capturing biomass accumulation with age and the general trends in SOC responses to LUC. Moreover, the authors adeptly discuss both the strengths and weaknesses of ORCHIDEE, providing a balanced assessment of its performance.

Overall, this paper makes a significant contribution to the fields of biogeochemistry and dynamic vegetation model development. Its meticulous methodology, insightful findings, and implications for model refinement enhance our understanding of the complex interactions between land-use change and biospheric carbon dynamics. I commend the authors for their rigorous approach and eagerly anticipate further advancements in this important area of research.

Thank you very much for your appreciation.

Reviewer #2:

Summary

The authors perform a series of land model simulations for Europe and evaluate model performance against observations for key carbon variables, including soil carbon change due to land use change. They conclude that the model's performance is reasonable, given its limitations.

Overall comments

I appreciate this effort to validate key carbon outputs of ORCHIDEE. In particular, the approach for assessing soil carbon change due to land use change is clever. The paper is also well organized and relatively clear.

Thanks for your appreciation.

I do have the following main concerns, however:

1) Some of the methods are unclear; see below for details. In particular, it isn't clear how you obtained your forest age outputs from the listed simulations.

Sorry for that. We have improved the manuscript to clarify the method section by including more detailed information on biomass (forest age) simulation and the computation of model biases.

2) I think that the experimental design for land use change is not adequate. Steady-state assumptions and inappropriate transition years introduce error into the desired

relationships. You have information on the age of transitioned sites or the year of transition. And you have historical driving data for the model. You may even have site-specific meteorological driving data. While it may be difficult to simulate hundreds of individual sites, you can at least set up historical simulations that represent site-specific transition years, and don't assume steady-state values. You can use the transient land cover up to the specified year in a cell, then make the desired PFT transition and then keep it constant for the rest of the simulation. Since SOC is PFT specific, just set the other PFTs to something else and then hold them constant also. This way the climate data are also more aligned with your observations. I think this could improve your results.

We have added further information to clarify this approach. We fully agree that our method is generally "idealised" by fixing the 1950 model transition year. However, producing more realistic predictions is challenging given the available data and the European scope of this study (please see detailed discussion below). Despite this, we have followed your suggestions and conducted additional historical simulations to account for site-specific transition years. The new results, however, do not significantly differ from our initial findings. Therefore, we decided to keep our initial results in the main manuscript but included the additional simulations and their results in the Supplementary Material. Additional discussions and comments on this matter have also been added to the manuscript.

3) The conclusions regarding model performance are overstated. If these results are comparable to other models, this needs to be referenced.

We have rephrased the conclusions regarding model performance to avoid overstating our findings. We have also incorporated more references and comparisons with other models into the manuscript.

Please find below our responses to each of your specific comments.

Specific comments and suggestions

Abstract

Clarify that carbon stock change analysis is just for SOC.

We add in line 7: "Second, we evaluate the predicted response of soil carbon stocks to LUC …"

Materials and Methods

line 142: but you include data with the forest floor. so this statement about excluding this is not true at all sites.

This statement refers only to LUCAS data, while the forest floor refers to the meta-analysis introduced in the latter part of the sub-section.

We added in the text:

"The LUCAS sampling was conducted at different depths, … while excluding above-ground vegetation residues, grass, and litter. "

lines 219-227: are the subsequent sims (eg, FG2F) done without wood harvest?

No, they are done with wood harvest. We clarified this point here, as well as in the text in Table 3:

line 220: ".. (with historical meteorology, CO_2 concentrations, and wood harvest data) ..."

Tab. 3: in FG2F section: Historical simulation with transient climate, CO_2, and wood harvest …

line 220: this does not seem to be in steady state to me. climate is changing and did wood harvest stop only 50 year prior?

The equilibrium stage refers to the end of FG1F simulation, during which wood harvest was held constant at the values for the year 1900. For the following transient simulation from 1901-1950, we used the specific wood harvest values for each year as prescribed by the forcing data we used (see in Section 2.1).

We clarified this in the manuscript:

"... Here, we fixed land cover to 100 % TeBS. At this stage, the biomass and SOC stocks are in equilibrium. In the second step, we ran a historical simulation FG2F from 1901 to 1950 for this same PFT (with historical meteorology, CO_2 concentrations, and wood harvest data), restarting from the last year of the spin-up simulation FG1F. To perform the LUCs …"

line 255: what is the soil depth in the model? I don't see it in the methods.

We add this information in line 99 : "In ORCHIDEE, the module of soil has an assumed globally uniform depth of 2 m. Note however that soil carbon is not depth discretized and average values of soil temperature, moisture and clay content are used."

line 256: should there be different beta^30 and beta^d0?

Here, we used the same value for beta (= 0.9786), the difference is from d0 which varies depending on the considered sample (e.g. d0 = 200 for all model simulation samples).

We add:

"For instance, a simulated sample X_{ORC} at 2 m depth (d0 = 200) is converted into the topsoil sample using the equation $X_{30} = (1-0.9786^{30}) / (1-0.9786^{200}) \times X_{ORC} = 0.48 \times X_{ORC}$."

lines 284-290: unclear. how do you do model biases per site when you have an aggregate observed SOC change function? is the subset taken from model cells corresponding to the sites?

We add more information to this paragraph to make it clearer:

" The bias is calculated for each site-observation taken from the meta-analyses (Tab. 2). For this, we compared the observed SOC stock changes per site with corresponding simulated values from the corresponding ORCHIDEE grid cells. Then we analysed which predictor variables best explain the site-to-site variations in model bias for each LUC scenario ."

orchidee performance

lines 299-301: the mode alignment to observations is overstated. 40% is not a relatively small difference. figure 1 clearly shows that some model medians outside of the 25-25th percentile boxes.

We rephrase this comment into:

"Figure 1 compares the simulated NPP and GPP values with site observations (Sect. 2.2.1). Both simulations and observations exhibit comparable value ranges across various PFTs, notably showcasing good performance in temperate forests and temperate C3 grasslands, where the relative differences in the medians are around 10 %. However, the ORCHIDEE simulation results often present a narrower range than the observed site data."

lines 310-328: It isn't clear how you obtained the model data for the different age classes. your BM simulations appear to be spun up with the PFTS, and then continue with these PFTs. Where is the starting point for a PFT to determine its age? see lines 197-200.

Sorry for the confusion. In fact, there is a clear-cut simulation before the transient simulation, after which the biomass in the PFTs regrows from approximately zero. This information is now added in the model simulation section to clarify this point.

line 200: " In addition, a forest clear-cut simulation is performed before running the transient simulation, and during FG2 the biomass regrows from approximately zero. Thus, the simulated forest age was defined as time since the beginning of the FG2 simulation. "

lines 330-337, figure 4: it is difficult to see the differences in figure 4. i suggest replacing 4b with the difference plot: model - obs
We update the figure and adjusted the description and relevant text accordingly:
line 331: " The difference between our simulated SOC stocks from LUCAS data is presented in Fig. 4b."

In the Figure caption: "Maps showing the stocks after soil organic carbon (SOC in kg m^{-2}) based on the LUCAS topsoil database (a) and the deviation of simulated SOC stocks from these observation based estimates (b)."

lines 265-282: not sure how comparable the simulated and observed response functions are. the simulated one depends on particular climate forcing during the 20th century before and after transition. the observed one is based on contemporary measurements using space for time and transition years that do not correspond with your fixed 1950 model transition year, and there is a general and possibly invalid assumption that everything is in steady-state prior to transition. lines 373-381:I suspect these results are affected by a site-specific training to simulations that are not site specific.
We acknowledge that our simulations are an approximation of reality, as indicated in the manuscript. However, it is very challenging to produce more realistic predictions given the available data, due to the following issues:

- Many studies provide only the age of land conversion, not the exact transition year.
- Most experiments use paired plots (or chronosequences), comparing adjacent sites: one with original land cover and the other with new land cover after LUC. Consequently, even with the transition year included, there is uncertainty related to representativeness of present day SOC stock from the paired site for the SOC prior to the LUC.
- In addition, within one study (or one observation field), several sub-sites with different transition years (and/or different transition types) are selected to compare with the site of original land cover. It often occurs that several sites of field observation are often located within the same grid cell, as the resolution of the ORCHIDEE model grid is quite coarse (0.5°≈50 km) . This makes the attempt to

account for all observed samples challenging, as we would have to produce input fields for ORCHIDEE (e.g soil properties, climate etc) at a higher resolution and to set-up several simulations with different transition years for the grid cells of concern.

- Furthermore, it is not possible to reproduce the land use history of a given site based on the PFT maps, because they only give the shift in areal proportions of different PFTs within a grid cell of an already substantial spatial extent , but no information of the exact land use and land use change at the observation site. Please note that we already treat this issue in the discussion section (line 455-458).

Due to difficulties mentioned above, it is not possible to perform more realistic simulations, in particular given the European scope of this analysis.

Nevertheless, we followed your suggestions and conducted additional historical simulations to account for site-specific transition years. However, these new results are not significantly different from our initial results , while we end up with less data points for our analysis after discarding data points with no identifiable year of conversion (see Fig. S5 below). For that reason, we decided to keep our initial results, but provided discussions/comments on this matter, referring to these additional simulations and their results that we added to the Supplementary Material.

We added at the end of Sect. 2.3.2. Idealised LUC simulations:
"We acknowledge that defining 1950 as the same year of LUC in our simulations increases the uncertainties when comparing simulations to observations which relate to different years of LUC. Note that only a fraction of the studies from which we source the observed LUC impacts on SOC stocks specify the year of LUC. To explore this source of uncertainty, we thus conducted tailored simulations matching the individual years of LUC reported in these studies, and compared the results to simulations using 1950 as the year of LUC. Detailed information about these additional simulations is provided and discussed in Sect. S1 of the Supplement."

We added in the Discussion:
line. 429: "Additionally, our idealised assumption regarding the transition year in 1950 may introduce uncertainties to the model outputs. However, as shown in Sect. S1 of the Supplement, considering the actual transition year does not significantly enhance agreement with observations. This might be due to the limited number of available

samples. It is also possible that the impact of climate change on LUC effects over the past century is not substantial. If the latter is true, using an idealised transition year should not create significant issues."

We added in the Supplementary Material:

"S1. Idealised land-use change (LUC) simulations using realistic land-use transition matrices

We performed additional simulations to account for transition years specific to each site. These simulations are similar to the idealised LUC simulations presented in Sect. 2.3.2, involving two main processes: (1) the spin-up simulation FG1 to establish the equilibrium state for each land cover type, and (2) the historical simulation FG2. In the second step, we conducted FG2 using land-use transition matrices instead of a fixed 1950 transition year. The land-use transition matrix is constructed based on the land cover change and transition year information. For example, to perform the LUC from forest to crop (e.g. Temperate needleleaf evergreen forest (TeNE) to C3 crop (C3C)), the land cover is initially set to 100 % TeNE for all grid cells across Europe. Then, we adjusted this map to carry out the LUC by setting the land cover at the respective grid cell to 100 % C3C at the corresponding transition year. This grid cell is then kept constant for the remainder of the simulation time. The simulation stops at the observed experiment year.

The land-use transition matrices depend on the observed transition information. However, this information may not always be provided in the meta-analyses (Tab. 2). Only 53 of the 102 study sites have a specific LUC year reported. Therefore, sites without transition year information are excluded. Additionally, within one study or one observation field, observation sites are often very close to each other. On the other hand, the ORCHIDEE model has a spatial resolution of 0.5° × 0.5°. Therefore, if multiple observation sites are located in the same grid cell, only one sample is chosen. The final number of selected samples for each LUC transition case is shown in Fig. S5. Due to the limited number of samples, we do not conduct additional fitted carbon response functions (CRFs, see Sect. 2.4.2).

The comparisons of observed and simulated SOC changes for different LUC transitions are shown in Fig. S5. Here, we directly compared the observed absolute change in SOC with the corresponding simulated change for each selected site. The observed and simulated CRFs (see Tab. S6) that are based on available observation sites and the assumed LUC year 1950 are presented for comparison.

The simulated responses in both cases - with the individually reported transition years (Fig. S5) and the idealised transition year in 1950 (Fig. 6) - show similarities: the model aligns with the observed changes but underestimates the amount of carbon gained or lost. The improvement is not significant for several reasons. One major factor could be the absence of site historical information. Although the exact transition year is known, accurately reconstructing the land use history of a particular site is impossible. Additionally, most experiments utilise paired plots or chronosequences to compare two adjacent sites: one with the original land cover and the other with new land cover following LUC. On the other hand, our simulations consider only a single site within a grid cell, analysing its behaviour before and after the LUC. This approach is used to accommodate the current European scope of the analysis and to minimise computational costs. More realistic simulations incorporating detailed information on land use history at a site-specific scale would provide more precise and reliable results. ”

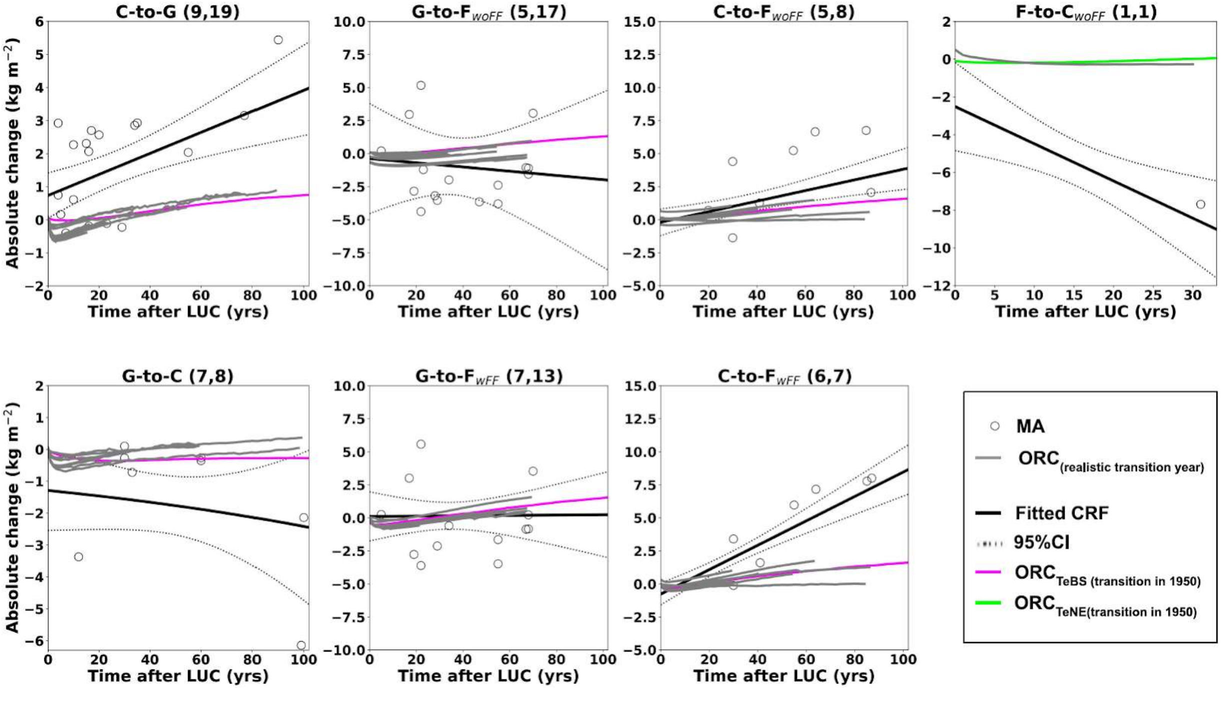

Figure S5. The absolute soil organic carbon changes (in kg m^{-2}) from site observations in meta-analyses (black circles) compared to corresponding ORCHIDEE simulated data (grey lines) for different land-use changes (LUCs: cropland-to-grassland (C − to − G), grassland-to-cropland (G−to−C), grassland-to-forest (without and with forest floor G−to−F$_{woFF}$, G−to−F$_{wFF}$), cropland-to-forest (C−to−F$_{woFF}$ and C−to−F$_{wFF}$), and forest-to-cropland (F −to−C$_{woFF}$). The first number in the parenthesis indicates the number of study sites, and the second is the number of samples in the meta-analyses. Here, temperate broadleaf summergreen (ORC_{TeBS}) is considered for the forest sites in all ORCHIDEE simulations, except for F−to−C_{woFF} in which temperate needleleaf evergreen (ORC_{TeNE}) is considered. The fitted carbon response functions (CRFs, black lines) ± 95 % confidence interval (black dotted lines) and simulated CRFs (magenta and green line) corresponding to all observation samples (Tab. S6, Fig. 6) are included here for comparison.

Discussion

line 394: this does not seem like a positive 'noteworthy' here. you focused on sites that matched the model PFTs, and while several processes may not be represented, the correlations are not very good.

If this is comparable to other models, you need to reference their evaluations.

We add here a comparison to other studies:

"For SOC stock simulation, a Pearson's correlation of 0.4 between observed and simulated SOC values (Fig. 4) is significant, given the absence of certain controlling factors and processes in the model version used. This score is similar to those in other DGVM models (Wu et al., 2019; Seiler et al., 2022). For example, Wu et al. (2019) demonstrated a correlation coefficient of approximately 0.45 between LPJ-GUESS (a global dynamic ecosystem model) and SoilGrids (an observation-driven global soil dataset) on a global scale, and lower correlation scores among different land cover classes. In this study, … "

References:

Wu, Z., Hugelius, G., Luo, Y., Smith, B., Xia, J., Fensholt, R., Lehsten, V., and Ahlström, A.: Approaching the potential of model-data comparisons of global land carbon storage, Scientific Reports, 9, 3367, https://doi.org/10.1038/s41598-019-38976-y, 2019.

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., Anthoni, P., Goll, D., Jain, A. K., Joetzjer, E., Lienert, S., Lombardozzi, D., Luyssaert, S., Nabel, J. E. M. S., Tian, H., Vuichard, N., Walker, A. P., Yuan, W., and Zaehle, S.: Are Terrestrial Biosphere Models Fit for Simulating the Global Land Carbon Sink?, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 946, https://doi.org/https://doi.org/10.1029/2021MS002946, e2021MS002946 2021MS002946, 2022.

lines 399-400: this is actually a feature of averaging, not climate.

Yes, it is possible that this could be due to the smoothing effect. However, climate could also contribute to this improvement since climate factors operate at larger scales. When data are aggregated, the influence of these large-scale climatic factors might become more apparent, improving correlation.

We also ran a simple simulation (not shown in the manuscript) to strengthen this hypothesis: We performed a random forest regression between soil organic carbon stock (observed and modelled) and a set of explanatory climate variables (i.e. air temperature, rain, and solar net radiation). Then, we performed spatial aggregation comparisons across different grid scales (0.5° x 0.5°, 1° x 1°, 2° x 2°, and 3° x 3°) to identify if climate is one of the main drivers for large-scale patterns.
The results indicate a stronger coefficient of determination (R^2 score) at a coarser resolution for LUCAS data (e.g. $R^2 = 0.22$ at 0.5° grid-scale and $R^2 = 0.51$ at 3° grid scale over all grassland sites), whereas this is not observed for ORCHIDEE outputs (e.g. $R^2 = 0.86$ at 0.5° grid-scale and $R^2 = 0.83$ at 3° grid scale over all grassland sites). As expected, since the ORCHIDEE model is based on climate data at 0.5° grid-scale, increasing the size of climate patterns does not impact the model outputs. On the other hand, a higher score obtained at a 3° grid scale signifies that climate is one of the main drivers for large-scale patterns of the observed SOC stock.

Finally, we rephrase the sentence and add a reference to it:
"... , we find that larger grid scales demonstrate a stronger correlation, which may be driven by climate patterns (Wang et al., 2023). …"

Reference:
Wang, X., Lin, D., Zhao, L., and Michalet, R.: The Relative Importance of Coarse-Scale Climate and Fine-Scale Nitrogen Availability Contrasts in Driving Home-Field Advantage Effects in Litter Decomposition, Ecosystems, 26, 1456–1467, https://doi.org/10.1007/s10021- 023-00844-2, 2023

Conclusions
lines 447-449: this agreement is overstated.
We rephrase this paragraph to:
"Our research investigated the ability of the DGVM ORCHIDEE model to reproduce what is known from experimental studies about LUC impacts on biospheric carbon. We performed various comparisons between simulations and experimental data, including on-site measurements and data from meta-analyses."