Discussion reply on Anonymous Referee #2 comments on submitted paper :

AgriCarbon-EO: v1.0.1: Large Scale and High Resolution Simulation of Carbon Fluxes by Assimilation of Sentinel-2 and Landsat-8 Reflectances using a Bayesian approach

Taeken Wijmer, Ahmad Al Bitar, Ludovic Arnaud, Rémy Fieuzal, and Eric Ceschia

General comments

Perception of the manuscript

This preprint describes a newly developed method of estimating the soil carbon budget of major crops at high spatial resolution. It addresses a “hot” and very relevant topic in climate research, specifically the field of estimating soil carbon fluxes. The method embodies a novel combination of both high spatial resolution (the intra-field scale) with large spatial coverage (up to 100x100 km) of carbon flux-related variables and their variability. It assimilates high-resolution remote sensing data into an agronomic model and accounts for uncertainties across the processing chain. The paper deals with evaluating the model accuracy at multiple scales, assessing the impact of spatial and temporal scale in the remote sensing-based input data, and integration of multiple models/dataset for a specific crop type and study site (winter wheat in SW France).Although the topic of quantifying soil carbon fluxes holds high significance, the way the proposed methodology is presented in the manuscript offers only little knowledge gain and shows only limited potential to be used for further studies of the soil carbon budget.

We agree with the Referee #2 on the overall view of the scope that we wanted to give to the manuscript. Note however that our objective is not only to assess soil carbon fluxes (as written in the last sentence above), but also to determin the main carbon budget components (biomass, C exported at harvest, GPP, plant respiration and soil respiration). We agree that a better link between the assessment of carbon fluxes (vegetation + soil) and soil carbon budget needs to be presented. We addressed this issue in the following answers by presenting the Net Ecosystem Carbon Budget (NECB) equation and producing additional results that will be added to the manuscript.

Comment I:

First, although the paper starts with the need for better quantifications of soil organic carbon, it is too hung up on general carbon flux, data/model integration and comparison to make interesting statements about the variability and accuracy of soil carbon estimation. To me, rather than addressing the estimation of soil carbon fluxes, it seems that the proposed approach is a crop biomass model based on correlations from remote sensing-derived GLAI.

As a matter of fact, the annual carbon budget for croplands that represents the amount of organic C gain or loss depends on the annual CO_2 fluxes (with GPP-Rauto conditioning the biomass production) and the lateral fluxes of Carbon as organic amendments (Cimports) and exported at harvest (Cexports) as in the equation 1 (Ceshia et al. 2010, Woodwell and

Whittaker et al., 1968, Chapin et al., 2006) . Except for carbon imports, those variables are estimated by the SAFYE-CO2 model in Agricarbon-EO. We agree that this link is not presented clearly enough in the paper. In order to clarify the relation between soil organic carbon stock changes, CO_2 fluxes and biomass, we will add the equation 1 , that connects the different components, into the introduction section of the manuscript with an associated explanation:

$$NECB = \overbrace{\overbrace{GPP - Rauto}^{NPP} - Rh}^{NEE} + C_{imports} - C_{exports} \quad \text{with } Reco = \underbrace{Rauto - Rh}_{Reco}$$ (eq.1)

ΔNECB is the Net Ecosystem Carbon Budget. It can be divided into two components. First the carbon fluxes as CO_2 induced by the biological processes represented by the Gross primary production (GPP) resulting from photosynthesis, autotrophic respiration (Rauto i.e. plant respiration) and heterotrophic respiration (Rh, i.e. soil respiration). The add-up of those fluxes is the Net Ecosystem Exchange (NEE). When the NEE is integrated over a cropping year, it is referred to as the Net Ecosystem Productivity (NEP). The net flux for the plant (GPP-Rauto) is referred to as the Net Primary production (NPP) which represents the amount of biomass produced. The sum of respiration fluxes (Rauto + Rh) is the ecosystem respiration (Reco) . The Net Ecosystem Carbon Budget also depends carbon fluxes resulting from farming practices, namely the Cimports representing the amount of C brought as organic amendments (e.g. manure, compost) and the Cexports that represent how much C has been exported at harvest (e.g. grain, grain+straw, tubers).

When the crop is harvested, the unharvested plant biomass (litter and roots) are incorporated into the soil, which means ΔNECB = ΔSOC at the end of the cropping year. In the revised version of the manuscript, we will present a ΔSOC variation map to complement the NEP maps over the wheat growing period by considering the following imports and exports conditions:

- In the region of interest, carbon imports are negligible as the fraction of farms practicing animal husbandry is very low. Furthermore the mass of seed carbon for wheat is about 6 to 10 g/m^2.
- Cexports are the parts of the plant that are harvested or removed. For wheat, grains are usually the only part of the plant that is exported in the region of interest.

$$C_{exports} = \overbrace{DAM \times HI}^{\text{dry Yield}} \times Cfrac$$ (eq.2)

HI is the harvest index and Cfrac is the fraction of carbon per unit biomass.

As NEE and NEP (annually cumulated NEE) are computed by Agricarbon-EO, by relying on eq1 and eq2, ΔSOC maps can be computed and added after figure 9 in the manuscript considering the rules mentioned above (no organic amendments, only grain is harvested):

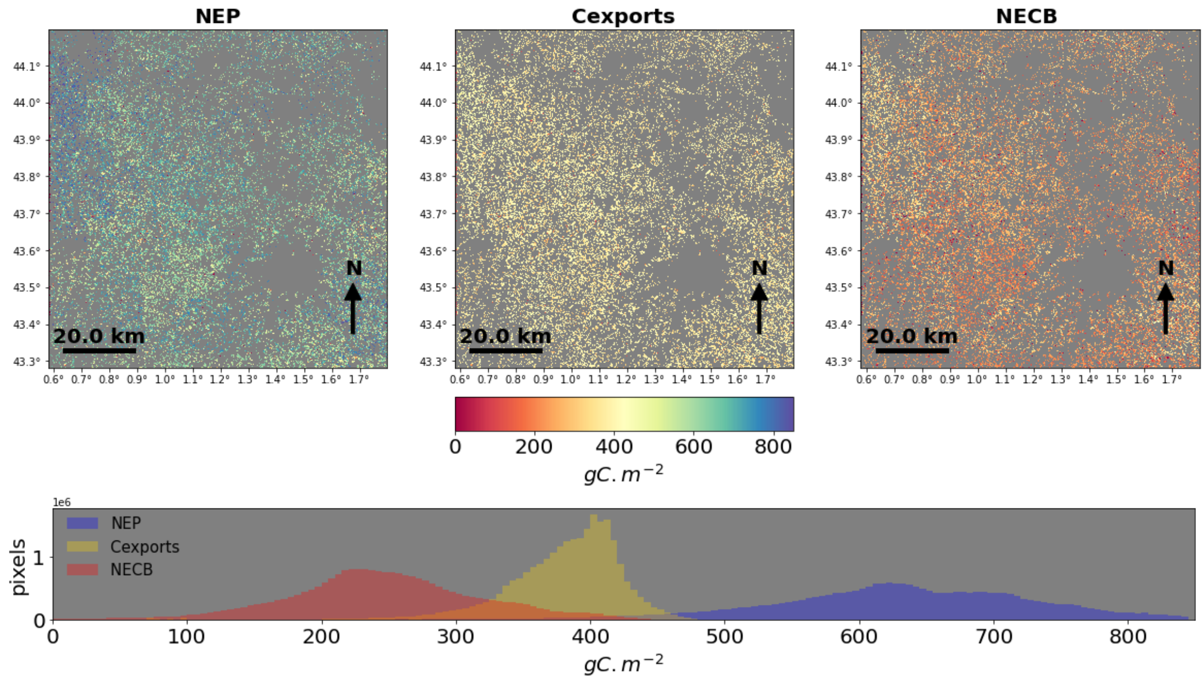

Figure R1: Net Ecosystem Productivity, Carbon exports based on yield and Net Ecosystem Carbon Budget maps between 20161001 and 20171001 over the T31TCJ SENTINEL2 tile and the distribution of those variables.

By highlighting and adding those elements, we hope that the utility of this processing chain for studies regarding soil carbon is now clarified. In the discussion section the limitations related to the soil module are already mentioned but will be also amended. More precisely, alternative solutions include adding more environmental constraints to the Rh terms in SAFYE-CO2, or estimating at high-resolution root and aboveground biomass inputs to the soil with AgriCarbon-EO that will be used as inputs in soil models such as AMG,Daycent or RothC.

The addition of these informations will not lengthen the paper as we are suggesting in the latter answers to remove and reduce parts of the manuscript. The **section 5.3 "Regional scale analysis"** containing **Figure14** and **Figure15** will be removed and the **Figure10** regarding posterior parameter distributions and the analysis related to this figure will be moved to supplementary material.

Comment II:

Second, it remains unclear why exactly the study area was chosen and to what extent it has greater relevance for other potential applications in different regions of the world with different crop types, different climatic and soil characteristics.

The study area is used as an application area for the presented method. It was chosen for several reasons that are mentioned across the section "3.1 Study area". We will summarize them in the updated manuscript at the beginning of section “3 Application in South-West France over Wheat”. The reasons are the following:

I. The extensive dataset available from the Space Regional Observatory (OSR - Observatoire Spatial Regional) and the ICOS measurement site of Auradé (FR-Aur). This large evaluation dataset contains multiple variables of interest for carbon stocks, as well as for the verification of the performance and soundness of an agronomic modeling approach. The amount and nature of this dataset corresponds to what is needed to characterize the accuracy/uncertainty of Monitoring Review and Verification(MRV) tools as described in Smith et al. (2010).

II. The spatial variability of pedoclimatic conditions resulting from the slope, aspect, soil properties, and historical land management "Remembrement" policy. This results in high crop growth variability at intra and inter-field scales. This is essential to assess the impact of using high-resolution modeling and assimilation schemes in quantifying the carbon budget components (e.g. biomass, CO_2 fluxes), and to assess the ability of the processing chain to reproduce those variations.

III. The area is a dense crop production zone. This is especially true for wheat production (up to 40% of the surface in the simulated area), with a large economic interest (Soft Wheat represents 75% of the Nation soft wheat exports).

These characterics of the study area make the arguments for a benchmark area for both high resolution agronomic modeling exercises and MRV methods. We would also like to recall that the application to study area is to “validate and demonstrate the capabilities of AgriCarbon-EO” (section 3), and that the focus of the paper is to present the innovative method as mentioned by the referee in his general comment.

Comment III:

Furthermore, it is not clear which possible users the presented processing chain has and which questions - apart from comparing/evaluating data and model parameters - it can answer.

The development of the AgriCarbon-EO modeling chain was done based on the needs expressed by the soil carbon communities in several research papers (Smith et al. 2020, Paustian et. al. 2019) or actions like the CIRCASA (Coordination of International Research Cooperation on Soil Carbon Sequestration in Agriculture), the Horizon ORCASA project, and the Eu Commission Expert Group on the implementation of the EU Soil Strategy for 2030…

The current presented approach (AgriCarbon-EO) is a hybrid solution that combines remote sensing, agronomic modeling, and assimilation strategy. It provides a scalable solution that enables the estimation of their main soil carbon budget components as presented and clarified in the first comment (eq.1). It answers a question related to: How to provide soil carbon budget at regional scale for cropland knowing that farming practices vary between the fields, and that intra-field heterogeneity is present?

Beyond this question the tool can be used to answer more or less specific questions, that were not addressed in this paper, like:

- What trade-off between water consumption and carbon storage is related to the deployment of cover crops?
- How to better define a soil sampling strategy based on the spatial heterogeneity of the carbon budget components?

The current study has limitations presented in the discussion section that open to other questions on how to enhance the approach:

- How to enhance the soil carbon budget by coupling SAFYE-CO2 to process based soil models (AMG, RothC, DayCent) ?
- What type of accuracy can be expected for crops in different regions with different crop rotations and pedoclimatic conditions ?
- …

The questions above address many profiles of research users and communities. Clearly there is a multidisciplinary aspect to this tool that can be of interest to the soil, crop modeling and climate impact communities.

We can include those elements in the outlook to illustrate the type of questions that can be answered by Agricarbon-EO for study areas that range from individual sample points to intra field to regional analysis, and may be of interest to agronomic modeling, soil carbon, water management and precision agriculture communities.

Comment IV:

Lastly, the paper touches on too many aspects – the discussion of scale (spatial and temporal), inter-comparison of different data/models, different crop-related parameters –so that it is rather difficult to read. The focus on relevant questions within the carbon modeling community and clear answers to those got lost.

We agree that the paper touches many elements. While some analysis are needed to justify driving elements of the method like scalability, accuracy, and high-resolution (figures 11 and 12), we reckon that reducing some elements would enhance the paper. We suggest the following major modifications to the paper to improve the readability:

I. **Clarifying the link between soil carbon, CO2/carbon fluxes and biomass:**
    A. We will add the NECB equation in the introduction, we will comment on it and make the link with the variables simulated by AgriCarbon-EO (see **comment I**).
    B. We will add the Figure R1 here above that presents the SOC stock change map using this same equation and the hypothesis explained here above.
II. **Reducing and simplifying parts of the paper:**
    A. Section 2.4 **"Bayesian normalized importance SAmpling using Look out Table - BASALT"** will be fused with section 2.4.2 **"Log-likelihood computations"**.
    B. Section 2.4.1 **"Normalised Importance Sampling and Look-up table"** will be moved to supplementary material.
    C. **Figure 6** and **figure 5** will be fused together as they treat the same years and field.

D. **Figure10** regarding posterior parameter distributions and the analysis related to this figure will be moved to supplementary material.
E. **section 5.3 "Regional scale analysis"** containing **Figure14** and **Figure15** will be removed.

III. Enhancements:

A. The "Discussion" section will be amended based on feedbacks from the discussion process.
B. Time series corresponding to the pixel wise validations over the ESU biomass dataset will be added to the supplemental material.
C. We leave it to the editor to decide whether or not part of or all of the equations of the Agronomic model should be put in the supplemental.

IV. **Grammatical and typo:**

The paper will be spell checked by an external service and several native speakers with knowledge in the domain.

We will also integrate the eventual comments provided by the Scientific Editor.

General conclusion:

Overall, the paper holds certain potential for wide scientific interest, but needs to be substantially revised for better informative value on soil carbon fluxes and scalability of the method for other regions. I recommend performing major revisions before considering for publication in GMD.

I suggest to address the following major points:

1. the many spelling and grammatical errors in the manuscript (I can't list all of these)

As mentioned above, the paper will be spell checked by an external service and several native speakers with knowledge in the domain.

2. missed topic: the discrepancy between the introduction of the topic of soil carbon fluxes and the presentation of the approach to a crop model that maps biomass and NPP instead of soil carbon.

As mentioned In **comment I.** we will add the NECB equation (eq1) in the introduction and we will use it as a common thread through the manuscript to highlight the relevance of each specific element of analysis to the carbon budget assessment and we will provide a SOC stock change with the NEP and Yield maps that allowed to compute it in section 5.1 **"Large scale simulations".** We will also discuss in more detail how besides providing evaluations of the SOC stock changes, above and below ground biomass maps from ACEO can be used to enhance the spatial representativity of existing soil models.

Clearer designation and focussed answering of relevant research questions/objectives

The main objectives that are:

I. To present a method to represent high resolution crop variability at large scale and its impact on the carbon cycle.

II. To demonstrate its validity through different evaluation exercises for the different components of the carbon budget.

III. To showcase its spatial capabilities through an application over winter wheat to estimate crop characteristics and the resulting SOC stock changes over a cropping year.

Those objectives will be put forward more clearly.

3. Justification for the selection of the study area in terms of what we can learn from it for other regions

We provided an extensive answer to this remark in **comment II.**

Specific comments

Lines73ff (objectives)

Please be more specific about the objectives here. What are specifically "the outputs" you aim to check for accuracy and coherence?

The regional analysis in the paper section 5.3 corresponds to checking if outputs present a statistical behavior with respect to altitude, aspect (exposition), and soil properties that is consistent with our expectations.

This section (5.3) will be removed in the updated manuscript to shorten the paper. So the following comment along with the mentions to it in other sections will be removed.

Instead of throwing in buzz words like "multi-scale validation", name the research questions precisely, e.g. what is the impact of (a) spatial resolution and (b) temporal resolution of the input data?

We agree that we can better show the research question. Still our intention was not to use the "multi-scale validation" because of its popularity. It was used because we argue that a proper presentation and assessment of a method that combines intra-field to regional scales requires validation and comparison at several resolutions.

"verifying the coherency of the outputs through intra-field as well as regional analysis" sounds spongy too me. What exactly is meant by "coherency", what is "intra-field and regional analysis"?

The intra-field analysis designs the evaluation of the models against spatialised aboveground biomass data and high resolution combine harvester yield maps, especially the ability of the model to retrieve intra-field resolution yield patterns. The regional analysis (that we propose to remove from the manuscript) aims at verifying that the model is able to represent the expected patterns when confronted to covariables with known effects on retrieved patterns. For example, we expect plants to grow faster and earlier on south-facing slopes in our region of interest. We have shown that this is the case in the simulations, lending additional credibility to the overall recovery of vegetation dynamics.

Line 112/land cover map

So, only the border extents of the shape file are used here to download the remote sensing and weather forcing data? Please describe how the land cover information is used, if at all, and how the land cover categories should look like.

Yes, in the first step. The extents of the shapefile are used to prepare the remote sensing data and the weather forcing (either they are in the database or they are downloaded).

Then the content of the shape file (land cover map) is used to run the modeling chain over all the pixels covered by the crop of interest. In the current application these are wheat fields.

The modeling chain will use the same vegetation prior parameters file for wheat for each of the pixels.

If the user aims at making estimates for maize, the user will need to provide a shapefile with identified maize fields and the prior parameters of maize.

We discuss limitations related to the land cover map in the Discussion section 6.2 of the paper.

The shapefile used for the simulation is provided in 'ACEO/EXPERIMENT/template/sim_shape/' folder in the zenodo code archive.

Figure 4

What do the different colours mean? Why not use different colours for each dataset groups: input, validation and regional datasets?

Thank you for the suggestion, the figure will be updated. Also the dataset used in section 5.3 will be removed from the figure.

Line 571f: "So high resolution allows more accurate estimates of the mean DAM values at field scale which enables more accurate field scale estimates of SOC changes by soil models in the perspective of monitoring SOC stock changes"

SOC was not really analyzed and assessed here, so how would you know the effect of high-spatial resolution input data on soil organic carbon modeling? I think that this conclusion is simply too far-fetched here.

We consider that more accurate representation of DAM and yield spatial variabilities allow more accurate estimates of the biomass that returns to the soil and of its spatial variability. As the amount of biomass that is returned to the soil is a key input for soil models, reducing its uncertainty or improving its spatial variability will improve the results of soil models.

Line 665: "an intra-field scale quantification of the DAM and of the biomass that returns to the soil is needed for accurate monitoring of soil organic carbon stock changes"

Again, this conclusion seems far-fetched here. It is not clear to me how you assessed soil organic carbon with the proposed modeling chain, as it was never computed or derived from the output variables.

As we mentioned above the NECB equation and harvest equations, the carbon from the unharvested dry biomass (NPP or (DAM + DBM) * cfrac) * (1 - HI) at the time of harvest is the main carbon input to agricultural soils (Soussana et al., 2019, Minasny et al.,2022). As such we argue that it is important to take into account the spatial variability of the biomass that returns to the soil to provide a spatially explicit representation of the evolution of soil organic carbon stocks.

Technical corrections (selection, please perform comprehensive spelling correction!)

- Line 8: misspelling – "radiative transfer"
- Line 10/11: misspelling - "a land cover maps"
- Line 14: misspelling - "against"
- Line 103: misspelling – "corpping"
- Line 111: misspelling - "Landcover"
- Line 165: misspelling – "In Contrast"
- Line 268 (and every other occurrence): misspelling "lykelyhoods"
- Line 304: misspelling- "dry Biomass"
- Line 322: "in the field 27" What is 27?
- Line 325: misspelling – "is applied over a for winter wheat"
- Table 3: "Quantify spatial and variability" What variability?
- Line 353: "Dry Aboveground biomass, DAM measurements" Please present abbreviations in a uniform manner.
- Line 453: misspelling - "respiration"
- etc.

Will be corrected.

Bibliography:

E. Ceschia, P. B eziat, J. F. Dejoux, M. Aubinet, C. Bernhofer, B. Bodson, N. Buchmann, A. Carrara, P. Cellier, P. Di Tommasi, J. A. Elbers, W. Eugster, T. Gr ̈unwald, C. M. J. Jacobs, W. W. P. Jans, M. Jones, W. Kutsch, G. Lanigan, E. Magliulo, O. Marloie, E. J. Moors, C. Moureaux, A. Olioso, B. Osborne, M. J. Sanz, M. Saunders, P. Smith, H. Soegaard, M. Wattenbach, Management effects on net ecosystem carbon and GHG budgets at European crop sites, Agriculture, Ecosystems & Environment 139 (3) (2010) 363–383. doi:10.1016/j.agee.2010.09.020. URL https://www.sciencedirect.com/science/article/pii/S0167880910002537

F. S. Chapin, G. M. Woodwell, J. T. Randerson, E. B. Rastetter, G. M. Lovett, D. D. Baldocchi, D. A. Clark, M. E. Harmon, D. S. Schimel, R. Valentini, C. Wirth, J. D. Aber, J. J. Cole, M. L. Goulden, J. W. Harden, M. Heimann, R. W. Howarth, P. A. Matson, A. D. McGuire, J. M. Melillo, H. A. Mooney, J. C. Neff, R. A. Houghton, M. L. Pace, M. G. Ryan, S. W. Running, O. E. Sala, W. H. Schlesinger, E.-D. Schulze, Reconciling Carbon-cycle Concepts, Terminology, and Methods, Ecosystems 9 (7) (2006) 1041–1050. doi:10.1007/s10021-005-0105-URL https://doi.org/10.1007/s10021-005-0105-7

G. M. Woodwell, R. H. Whittaker, Primary Production in Ter-restrial Ecosystems, American Zoologist 8 (1) (1968) 19–30.24 doi:10.1093/icb/8.1.19. URL https://academic.oup.com/icb/article-lookup/doi/10.1093/icb/8.1.19

J.-F. Soussana, S. Lutfalla, F. Ehrhardt, T. Rosenstock, C. Lamanna, P. Havl ́ık, M. Richards, E. L. Wollenberg, J.-L. Chotte, E. Torquebiau, P. Ciais, P. Smith, R. Lal, Matching policy and science: Rationale for the '4 per 1000 - soils for food security and climate' initiative, Soil and Tillage Research 188 (2019) 3–15. doi:10.1016/j.still.2017.12.002. URL https://www.sciencedirect.com/science/article/pii/S0167198717302271

B. Minasny, D. Arrouays, R. Cardinael, A. Chabbi, M. Farrell, B. Henry, L.-S. Koutika, J. K. Ladha, A. B. McBratney, J. Padarian, M. Rom ́an Dobarco, C. Rumpel, P. Smith, J.-F. Soussana, Current NPP cannot predict future soil organic carbon sequestration potential. Comment on "Photosynthetic limits on carbon sequestration in croplands", Geoderma 424 (2022) 115975. doi:10.1016/j.geoderma.2022.115975. URL https://linkinghub.elsevier.com/retrieve/pii/S0016706122002828