Discussion reply on Anonymous Referee #1 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:

Comment:
The five authors have submitted a rather long manuscript (approx. 13.500 words excluding references), in which they advertise an approach for improved carbon balance mapping in agricultural soils. Long-term storage of carbon in agriculturally used soils is a very hotly discussed topic in the frame of carbon farming, which again is part of the EU green deal. Finding new ways to quantify soil carbon fluxes in agricultural systems therefore is a very interesting and promising research topic!

Answer:
The referee mentions that the paper is long. We agree with him about reducing the size of the manuscript while maintaining the needed sections to answer the aim of the paper which is to provide the community with a methodology that enables realistic estimation of crop carbon fluxes and production at large scale (regional) and high resolution (intra-plots).

Comment:
If we take a closer look at the manuscript, however, it turns out that the proposed approach is more a method for producing yield maps for winter wheat from correlations with green LAI observations from Earth Observation time-series. As, by using a simple LUE-based growth model and neglecting water-stress effects, the interrelations of the carbon and water cycle are explicitly not considered in this study, so that mass and energy balance may not necessarily be closed and no direct link between atmospheric carbon dioxide concentration and carbon sequestration in agriculturally used soils is established, I have the feeling that what we as a modelling community can learn from the presented study for carbon-farming related questions unfortunately is limited.

Answer:
We don't agree with the statement above and we bring the following precisions:

- The main outputs of the approach are the carbon fluxes: Net Ecosystem Exchange (Equation 9), Gross Primary Production (Equation 1), heterotrophic respiration (Rh) (Equation 10), autotrophic respiration (Equation 5), Ecosystem respiration (Reco) as well as biomass (Equation 13), and yield. NEE, Reco, GPP are validated in figure, NEE is showcased in figure (9). We also propose to add the sum of NEE over the growing season to figures (15). Probably some confusions originate from the differences between the SAFY (Duchemin et al. 2008) and the SAFYE-CO2 (Pique et al., 2020a,b) models that share a limited number of equations.

- The input remote sensing information of the modelling chain is top of canopy (TOC) reflectances and not green leaf area index (GLAI). GLAI is an intermediate biophysical variable that is first retrieved with its associated uncertainty via the PROSAIL model and then assimilated into the SAFYE-CO2 crop model. As described in section (2.1). This ensures a propagation of uncertainties across the modelling chain.

- SAFYE-CO2 doesn't rely on correlation with GLAI but on a parsimonious modelling approach (Pique et al. 2020b) unless the referee means here "relation" and not "correlation". The model formalisms that are presented have been widely used in the community. They are less detailed than other radiative transfer and agronomic models but offer the advantage of limited amount of data inputs which enables large scale high resolution spatialisation.

- The impact of water stress is considered indirectly for the case of Gross primary production (GPP) through the impact of in situ water stress on the observed GLAI that is assimilated into SAFYE-CO2. On the other hand, the impact of water-stress on soil respiration and evapotranspiration is considered via the available water content in the soil. The question of irrigation remains to be answered, we provide some elements later in this comment.

Overall, the very rich comments and open questions raised by the referee are of interest for the modelling community. From our point of view while AgriCarbon-EO includes a parsimonious modelling approach it provides answers to questions related to the crop carbon fluxes at never applied levels of high spatial resolutions and regional scales with a crop model.

Comment:
The paper consists of at least five major parts/questions, whereby each of the topics potentially would provide substance for individual articles.

Answer
Concerning the paper structure: Again, the main question behind the aim of the paper is:
“ Is it possible to Provide reliable estimates of CO_2 fluxes (e.g. NEE) and other key carbon budget components (biomass, yield) at intra field resolution and large scale”.

In order to answer this question we need to provide a adapted methodology (presented in section 2) as well as to provide answers to the following specific questions (section 4 and 5):
- What is the accuracy and limits of the approach based on existing validation datasets (carbon fluxes, biomass and yield) ?
- What is the impact of using Sentinel-2 and Landsat8 data at intra-field compared to the more commonly applied spatial unit (field scale)?
- Are the aggregated regional scale estimates coherent ?

To answer the main question we firstly present Agricarbon EO in detail for a generic application (section 2) e.g. methodology for inversion, modelling tools and input datasets). We describe the choices of input data for the paper, and then we present the application over South-West France (section 3). In this case study, we provide a multiscale validation/verification that covers the different scales covered by the AgriCarbon-EO tool (see figure R1 / Figure 4 in the manuscript), we also provide in Table 3, the name of the different simulation and which sub-objective they answer.

We agree that we can better explain the sub objectives and the structure at the end of the introduction, by providing a more concise paper and enhanced transitions we will make it clearer for the referee and the reader.

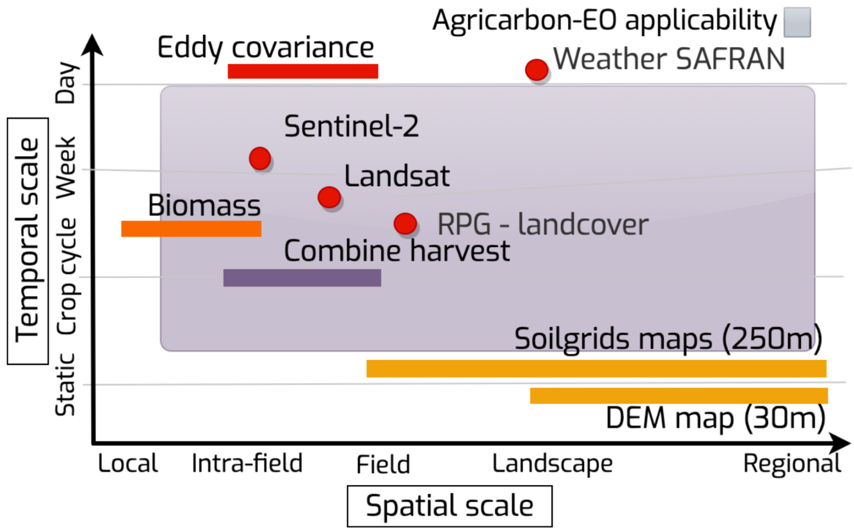

Figure R1: Temporal and spatial coverage of the input dataset (S2,L8, SAFRAN, and RPG), the validation dataset (Eddy covariance, Biomass, and Combine harvest), the regional dataset (SOILGRIDS and DEM), and AgriCarbon-EO outputs (blue zone).

Comment:
First, the model and the assimilation approach are introduced and the system is applied to field scale simulations, which are validated against destructively measured biomass and yield data. Thereby, ESU samples and combine harvester data are used. In figure 6 it can be seen that the ESUs were sampled on different dates throughout the growing period of 2017, while in 2019 only one date was sampled. Unfortunately, this is not clearly described in the text (e.g. sampling dates are not explicitly mentioned/listed) and confuses the validation of the temporal and the spatial performance of the algorithm. For example, in Figure 7 it remains largely unclear if the good performance of the year 2018 can be traced to the fact that the algorithm well follows the temporal dynamics of biomass accumulation (as biomass is suspected to continuously increase over the growing season, R2 correlation will necessarily be high...), or if the good performance is due to a good spatial mapping of heterogeneities of yield. Especially for 2019, where only one in-situ sampling date was available (as I deduce from Figure 6), the correlation is extremely poor, as the algorithm returns constant values of 2000 g/m2, while the in-situ data show a wide range of values. The way that I read the validation figures, this leads to the conclusion that the spatial heterogeneities of yield cannot be reasonably mapped with the proposed approach.

Answer:
First, it is incorrect to state that we deal with biomass and yield only. Indeed, before the section relative to the spatial biomass validation, we provide a validation of carbon fluxes (NEE, GPP, Reco) over the ICOS FR-Aur site (Section 4.1, Figure 5) at different periods (Table 5).

Second, concerning the biomass validation (figure 6 and figure 7). We mentioned in Table 4 of the manuscript, that ESU are sampled on 1 to 4 dates without providing the exact dates. **Table R1** shows the date and the field location of each of the measurements (it will be also added to the supplementary materials).

Table R1 : List of biomass records used in the validation for figures 6 and 7 of the manuscript

Data set	Lat	Lon	Name	Date
ESU_2018	1,2507	43,4999	Plot 1	20180702
ESU_2018	1,2507	43,4999	Plot 1	20180528
ESU_2018	1,2507	43,4999	Plot 1	20180504
ESU_2018	1,2507	43,4999	Plot 1	20180416
ESU_2018	1,2507	43,4999	Plot 1	20180406
ESU_2018	1,5095	43,5261	Plot 2	20180703
ESU_2018	1,5095	43,5261	Plot 2	20180524
ESU_2018	1,5095	43,5261	Plot 2	20180502
ESU_2018	1,5095	43,5261	Plot 2	20180416
ESU_2018	1,5095	43,5261	Plot 2	20180403
ESU_2018	1,5038	43,5259	Plot 3	20180703
ESU_2018	1,5038	43,5259	Plot 3	20180524
ESU_2018	1,5038	43,5259	Plot 3	20180502
ESU_2018	1,5038	43,5259	Plot 3	20180403
ESU_2018	1,2029	43,4663	Plot 4	20180528
ESU_2018	1,2029	43,4663	Plot 4	20180502
ESU_2018	1,2029	43,4663	Plot 4	20180416
ESU_2018	1,2029	43,4663	Plot 4	20180406
ESU_2018	1,2355	43,4916	Plot 5	20180702
ESU_2018	1,2355	43,4916	Plot 5	20180502
ESU_2018	1,2355	43,4916	Plot 5	20180416
ESU_2018	1,2355	43,4916	Plot 5	20180406
ESU_2018	1,2290	43,4870	plot 6	20180702
ESU_2018	1,2290	43,4870	plot 6	20180528
ESU_2018	1,2290	43,4870	plot 6	20180502
ESU_2018	1,2290	43,4870	plot 6	20180406
ESU_2018	1,1559	43,4919	plot 7	20180528
ESU_2018	1,1559	43,4919	plot 7	20180502
ESU_2018	1,1559	43,4919	plot 8	20180702

AUR_2019	1,1052	43,5497	AUR	20190701
AUR_2019	1,1053	43,5499	AUR	20190701
AUR_2019	1,1066	43,5494	AUR	20190701
AUR_2019	1,1065	43,5493	AUR	20190701
AUR_2019	1,1068	43,5494	AUR	20190701
AUR_2019	1,1065	43,5496	AUR	20190701
AUR_2019	1,1062	43,5496	AUR	20190701
AUR_2019	1,1057	43,5497	AUR	20190701
AUR_2019	1,1060	43,5498	AUR	20190701
AUR_2017	1,1069	43,5497	AUR	20170704
AUR_2017	1,1058	43,5497	AUR	20170704
AUR_2017	1,1065	43,5497	AUR	20170704
AUR_2017	1,1054	43,5496	AUR	20170704
AUR_2017	1,1070	43,5496	AUR	20170522
AUR_2017	1,1065	43,5496	AUR	20170522
AUR_2017	1,1067	43,5498	AUR	20170522
AUR_2017	1,1059	43,5496	AUR	20170522
AUR_2017	1,1063	43,5497	AUR	20170522
AUR_2017	1,1054	43,5496	AUR	20170522
AUR_2017	1,1057	43,5497	AUR	20170522
AUR_2017	1,1052	43,5497	AUR	20170522
AUR_2017	1,1054	43,5496	AUR	20170330
AUR_2017	1,1057	43,5497	AUR	20170330
AUR_2017	1,1070	43,5496	AUR	20170330

Here, we will detail information about each dataset, to bring very precise elements about the results.

ESU 2018 dataset: The data is sampled from 8 plots in 2018 (1 to 4 ESU measurements per plot). We extracted the time series for the ESU-2018 samples locations. Figure R2 shows the agreement between the observed ESU 2018 data and simulated biomass from AgriCarbon-EO at different stages of growth. *Figure R2* shows that the estimates are in good agreement across all dates and over all plots.

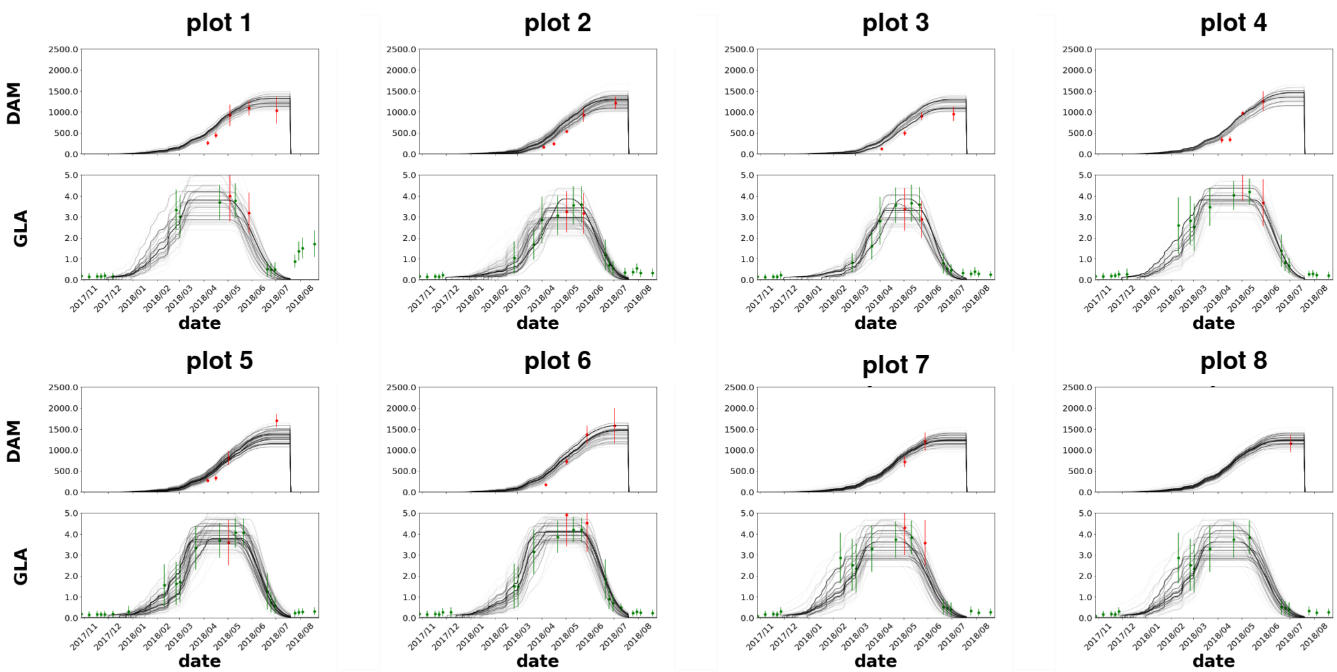

Figure R2: LAI and biomass timeseries for the 8 ESU sites. Simulated ensembles are in black with transparency proportional to their relative likelihood, field observations are in red, and satellite observations are in green.

FR-AUR 2017 and 2019 : The sampled locations for FR-AUR (Auradé site) are shown in Figure R3. The plot as a whole can exhibit strong heterogeneities. However the biomass measurements are done on a homogeneous part of the plot, close to the ICOS Eddy covariance flux tower. In fact, the flux tower location was intentionally installed in a relatively homogeneous and flat area according to ICOS site protocol. This is confirmed by the relatively homogeneous GLAI value next to the tower and in the sampling area (see Figure R3). This is why we associate the variability in the measurements to measurement uncertainties (std =300-400 g/m2). Year 2019, was a highly productive year, the model shows a very high estimate within the observation uncertainty (Figure 6) but also a flat estimate of biomass that the referee has commented on. We associate it to the saturation effect of optical images, but we still confirm that the measurement uncertainty was high considering the low heterogeneity of the sampled region.

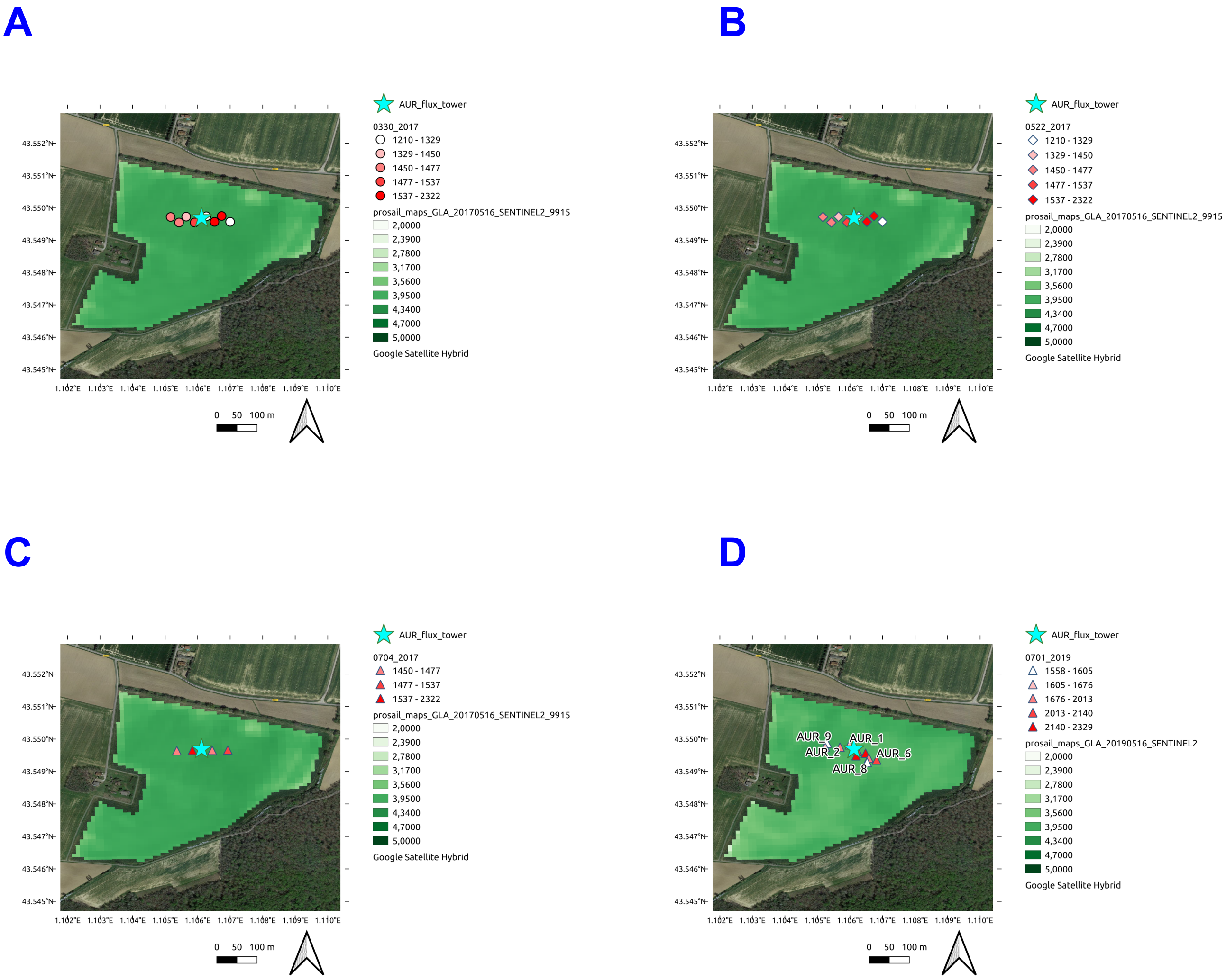

Figure R3: Location of the FR-AUR biomass measurement on 320170330(A), 20170522(B), 20170704(C) and 20190701 (D). The mean GLAI maps show the LAI on the 16th of may for both 2017 and 2019 years. This date corresponds roughly to the maximum GLAI.

In order to better represent the results, Figure 7 in the manuscript will be modified to exhibit the differences between end of cycle measurements and measurements during growth.

Comment:

This is confirmed then by Figure 8 also for the intra-field scale, where very poor correlations between the spatial combine harvester measurements and the satellite-based yield product become evident. For a satellite-based approach that explicitly targets the monitoring of intra-field variability in the context of precision farming measures, as claimed in the introduction, this is rather poor.

Answer:

Regarding the Spatial correlation of the yield maps, we agree that the R² is low compared to other studies focussing on yield predictions such as (i.e Hao 2021). In Hao (2021) we can see that state of the art agronomic models such as Apsim expect an RMSE of 1/t/ha which is higher than the RMSE obtained in our study. Furthermore it is important to keep in mind that this type of model is crop specific and most of the cases in the meta-analysis benefit from site specific calibrations in opposition to Agricarbon EO that only relies on weather data and satellite acquisitions.

In this context of application the small R^2 can however be explained by the range of variation of wheat yield that is smaller at intra field scale than regional or worldwide scale. As an illustration of this effect we realised a simple synthetic experiment. In this analysis we took the yield maps produced by the combine harvester as ground truth, added a normal simulated measurement noise with sigma = 1t/ha. The combine-harvester maps are compared to the maps with simulated observation noise to compute R^2 . This exercise was repeated 1000 times to be able to represent the expected distributions of R^2 given this measurement noise (**Fig R4**:). This figure we can see that agricarbon EO the obtained R^2 is not out of line with the expected R2 for applications of state of the art Agronomic model for plot 3.

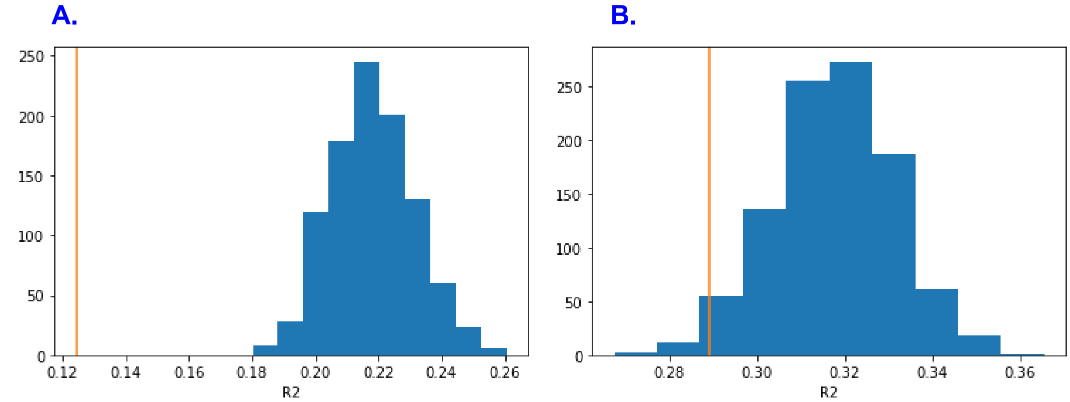

Fig R4: Expected R2 given simulated measurement error with sigma = 1t/ha for plots 6 (A) 3 (B). This value corresponds to the RMSE of SAFYE_CO2 at field scale and the expected RMSE for Apsim. The vertical line represents the value returned by Agricarbon-EO.

Furthermore when considering the correspondence between high and low Yield anomalies (i.e. |anomaly|>0.5) we got a good correspondence. This means that we can reliably identify high and low productivity areas in the field when there is significant variability in the yield to begin with as can be seen in **figure R5**.

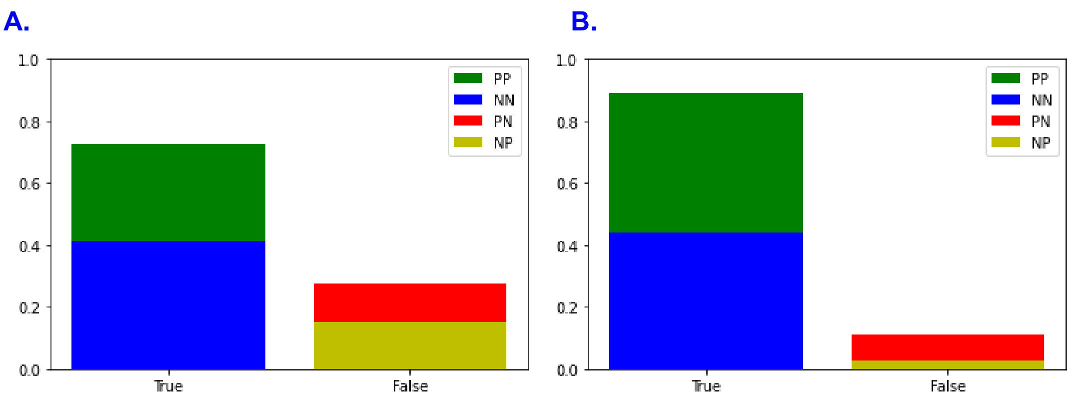

Figure R5: Correspondence between the sign of extreme anomalies inside the plot 6 (A) and plot 3 (B). PP represents true positives , NN true negatives, PN false positives and NP false negatives.

Finally, we wish to highlight that if we compare these simulations to standard field wise simulations (that explains no variability i.e.R^2=0) the explained spatial variance illustrated here is a net gain.

To clarify our position and the context in which the statistics are obtained we will amend the discussion with these elements and also strengthen the discussion about the uncertainties.

Comment:

Following the field-scale validation, the paper takes a sudden turn towards large-scale simulations and shows the application of the method for retrieving Net Ecosystem Production for a 110 x 110 km large scene for the growing season of 2017. Obviously, only the winter wheat pixels were investigated, although this is not clearly stated in the text.

Answer:

Large-scale is mentioned in the title of the paper. Section (5) content was announced in the abstract, at the end of introduction and in the application section (3). Yes, We use winter wheat as an application and transitioning at this stage of the paper to other available crop parameters like maize, sunflower would have been more confusing. Winter wheat is a major crop in South-west France as mentioned in the study area section 3.1, therefore this specific focus makes sense. However, because AEO is not "Winter wheat specific", the focus is not mentioned in the title.

Comment :
In this section, Figure 11, although suffering from some stylistic errors, makes the assimilation procedure and potential pitfalls of the algorithm transparent, so that the readers get a clear picture of how the Bayesian approach works. However, as the maps shown in Figure 9 due to the color stretch do not allow for the detection of large-scale patterns and also because these patterns are not discussed in the text, the readers wonder why this had not already been explained as part of the intra-field scale validation and why the jump to the large-scale actually was required for the purpose of this paper.

Answer:
We thank the reviewer for his positif feedback, we would add that it is important to note also that these represent a small percentage of the simulation. ("singular anomalies in the simulations that don't stand out in the regional statistics (line 473)".

Concerning Figure 9, we present here an updated formatting of the figure (**Figure R6**). It will also be updated in the manuscript. Note that in order to reduce the stretch positive values are omitted in this figure. These values correspond to abnormally low wheat development as well as places where no wheat grew contrary to the information provided by the RPG. We changed the zoomed maps to the limits of the NT-Plot6 field (for 2017). The regional patterns are more visible, but still the figure shows intra-field, inter-fields and regional heterogeneities at the same time.
We suggest moving Figure 10 to the supplementary material, and Figure 11 will be improved stylistically.
As explained above, large scale applicability is the core of the paper. It is important to address this scale if we aim at demonstrating the capabilities of the modelling chain. It suggests specific developments in order to address such scales. Removing the analysis at large scale, reduces the methodology to an application of the model to the field scale. As mentioned in section (2.4), demonstrating the capability of producing large scale and high spatial resolution estimates is one of the aims of Agri Carbon EO and this manuscript .

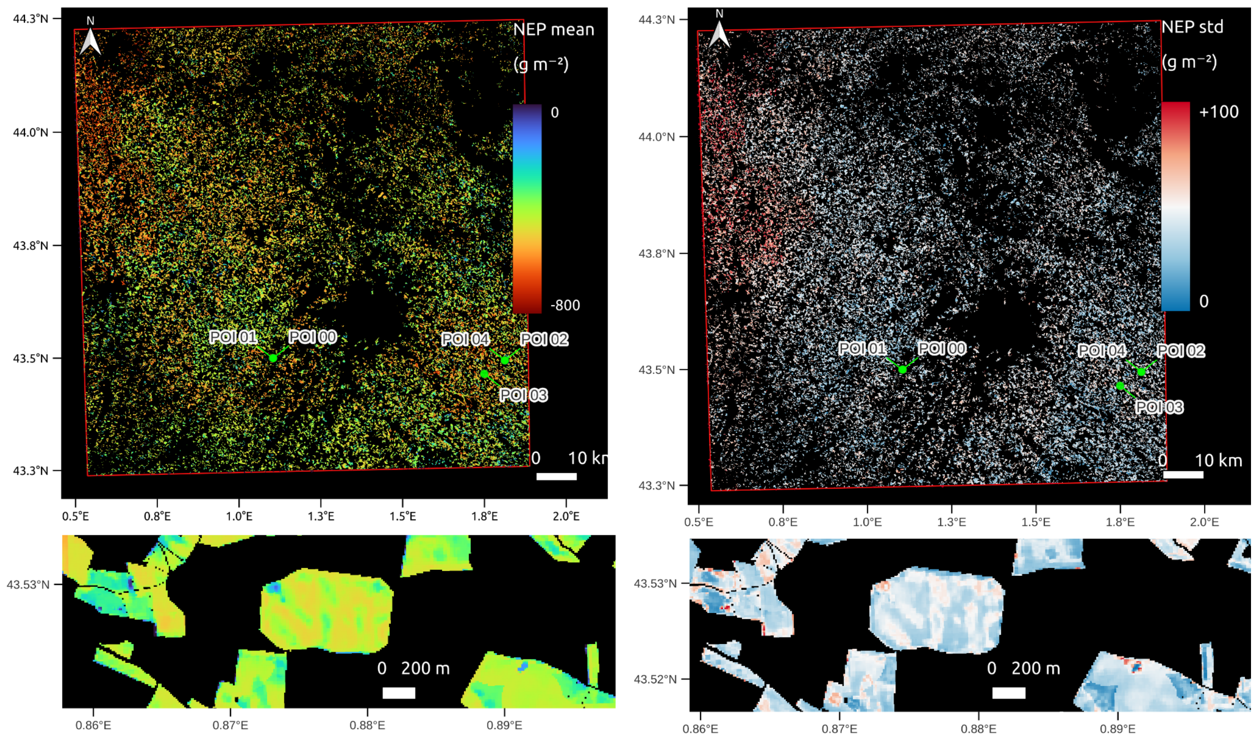

Figure R6: On the left, the NEP for winter wheat for the 2016-2017 cropping season. On the right, the corresponding uncertainties calculated with the AgriCarbon-EO processing by assimilating Sentinel-2 and Landsat-8 data (ACEO-S2L8-Pixel). The map also shows the position of selected Points Of Interest (POI) presented in Figure 11. The zoomed maps are on the NAT-Plot3.

Comment:
The next section opens two further side-questions about the impacts of spatial resolution and temporal sampling frequency (from my perspective these are questions three and four of the paper). The impact of spatial resolution is assessed by determining the bias between field averages and pixel-based values. I think that we no longer need to prove that intra-field variability indeed plays a major role. From my perspective it would have been more interesting to analyze, in a proper quantitative way, how much detail is lost when going from high-resolution of 10 m to a more modest resolution of 30 m, which widens the possibility for multi-mission observations. I think that in view of future Sentinel-Missions, which potentially will even provide higher resolutions of up to 5 m, this discussion is relevant. A similar drawback from my point of view is that the analysis of observation frequency impact on the DAM simulations is limited to the absolute number of satellite scenes and does not account for the differences in spatial resolution between Landsat and Sentinel-2.

Answer:
We explain hereby those choices:

- Concerning intra_field heterogeneities, in the discussion we mention the following: *"Intra-field heterogeneity is a well established issue in agricultural applications (Weiss et al., 2020; Blackmore et al., 2003; Grieve et al., 2019; Nowak, 2021), but it has not been thoroughly treated in terms of CO2 fluxes and uncertainties estimates. "* lines 566-567. So we address this point because if we are advocating for large scale intra-field estimation of carbon fluxes with its corresponding uncertainties, we should quantify the impact. Otherwise it is difficult to justify for data intensive methodologies that require higher computational needs.

- Concerning the choice of spatial resolution for the comparisons : From an agronomic modelling perspective which is central to this paper, the field scale and intra-field scale are well established concepts (Pasquel et al., 2022). Again, as we are advocating for intra-field scale we compare the highest resolution in our system (10m) to the field scale. It is important to bring the field scale to the picture. The only motivation to use 30m data is that it is the resolution of Landsat8 extracted bands, but actually the temporal frequency is 16 days which is not enough to constrain the crop model and anyway the Sentinel-2 is available at 10m. Still, we make a comparison in the paper between Sentinel-2 and Landsat8, when we compare the Sentinel-2 data against Sentinel-2 and Landsat8 combined which is most logical because Sentinel-2 is an operational mission with data available globally at 10m - 20m resolutions (figure 12 C).

- Concerning the context of Sentinel-2 Next Generation: Sentinel-2 NG at 5m (Löscher et al. 2020) is in study for an expected launch in 2034. To provide concrete elements in the discussion, we should consider existing similar or higher resolution data from drones, Planet Scope (Aragon et al., 2021), VENµS VM5 (Dick et al. 2022) (on daily revisit cycle). Such data is not operational or freely available globally. Also this will also have specific complexities (revisit, resolution, quality, spectral sampling) and is clearly out of scope of this paper's objectives, but we agree on the potential interest in the future.

- Finally on the DAM comparison in figure 13, the objective here is to show that increasing the absolute number of images reduces the uncertainties in the estimates but doesn't impact strongly the mean values which is important. Producing this figure for Landsat8 data only would not be feasible because the frequency of the Landsat8 data alone is not high enough to constrain SAFYE_CO2. As mentioned above the impact of adding Landsat8 specifically is shown in Figure 12 C.

Comment:
Also, the analysis does not consider the impact of observations that happen to cover specific growth stages. Satellite observations at specific growth stages might benefit the retrieval accuracy of certain parameters. By simply correlating the satellite images to the in-situ-sampled yield maps, it can be found that the pronunciation of intra-field patterns should high during BBCH 70-89 and reduced during the bolting phase. The occurrence of cloud cover during these phases may be highly specific for the region and may impact the application of the approach differently in different parts of the world.

Answer:

We agree with the Referee on this point. However doing such an analysis is not that straightforward for several reasons:

- First, in its current version our model does not estimate the timing of specific phenological phases as some classical agronomic models (e.g. STICS) apart from emergence, beginning and end of senescence. It makes difficult for instance to define when BBCH 70-89 will occur precisely on a given plot/pixel.
- Second, over our area of study, winter wheat varieties exist that have different sensitivities to degree days and we have no information on which winter wheat variety at 100x100 km. Also ,considering a fixed temporal window over the area of study in the attempt to catch accurately a precise phenological status is not an option as winter wheat pixels over our area of study may be characterised on a given day by rather different phenological status given the variability in sowing dates, climatic gradients, the differences in altitude and slope and aspects (even at intra field level). This is illustrated by the bivariate distribution of the phenological parameters in Figure 10 and by the spatial variability in the "*emerg"* and "*Sena"* parameters in Figure 14 (d) and (e).
- Third, the cloud coverage is not statistically stationary across the phenological stages over our area of study, therefore considering a fixed temporal window would bias the analysis.

Considering the above elements all together, it seems out of reach for this paper to do a meaningful analysis of the effect of the temporality of the gaps in EO observations on the performance of the method. Yet, we consider that a dedicated analysis on this topic would be very valuable.

Comment:

The fifth part of the paper then widens the scale even further up to the regional level by filtering the spatial results to correlation lengths of 2,5 km, thus entering a spatial scale beyond individual fields. The found large-scale patterns are explained by the soil characteristics in terms of water holding capacity and by the terrain situation (elevation, slope, aspect). While this undoubtedly explains the found spatial patterns, it remains unclear what the contribution of this section to the overarching subject of the study actually is.

Answer:

At this point in the manuscript we have provided performance metrics for several cases at pixel and plot scale applications for different variables. However this verification is operated on fractions of a percent of the total domain. It is thus possible to have doubt on the representativity of the simulations in other parts of the region with different pedoclimatic and growth conditions. In the absence of more validation data we deem important to verify if the simulations comply with the hypothesis about the effect of the pedoclimatic conditions in the region on crop growth. To perform this coherency analysis we compared the aspect, slope and soil texture to the phenological and growth outputs of Agricarbon-EO. However as the referee mentioned for Figure 9 regional tendencies can not be depicted clearly from the map. He is right on this and it is mainly because the map contains a mix of different variation scales. Those scales are the intra-field (soil texture, fertility depth, waterflow, slope, aspect), inter-field (soil texture, fertility, impact of choice of variety, farming practices, slope, aspect), and landscape (impact of, altitude, local weather conditions).These facts motivated the application of a gaussian smoothing at 2.5 km to retrieve landscape scale trends. By subtracting this landscape trend from the raw signal we can retrieve the anomaly of the landscape signal containing the information about plots and intraplot variability. After unmixing these scales of variation we can compare the phenological and growth to the environmental variables. As the reviewer noticed "The found large-scale patterns are explained by the soil characteristics in terms of water holding capacity and by the terrain situation" as expected if we respect our hypothesis regarding the effects of the environment on crop growth. This allows to improve credibility to the spatial variabilities at different scales that are retrieved by Agricarbon-EO from weather and EO data alone.

Concerning Figure 14: We suggest moving it to supplementary material as it is just an intermediate result that shows the smoothed maps and confuses the message of the section.

Concerning Figure 15: The emergence (*emerg*) and sum of temperature at senescence (*sena*) are in this regard interesting because they answer part of the comment raised by the referee about the analysis with regards to the phenological stages. Dry above ground biomass is also important as it determines the quantities of exported biomass and thus biomass incorporated in the soil.

Comment:

The last part of the paper is dedicated to the discussion, which, in comparison to the size of the rest of the paper, is rather short (three pages). Here, the authors - among other recommendations-suggest the use of SAR-data with their approach. As the proposed retrieval scheme which is based on the PROSAIL model does not apply to microwave data, it remains unclear how SAR data could successfully be integrated into the system.

Answer:
Naturally and considering the team's previous experiences with SAR data (Tomer et al. 2015, Tomer et al. 2016, Zribi et al. 2019, Fieuzal et al. 2017, Valero et al. 2021, Velozo et al. 2017, Baup et al. 2019), we consider assimilating SAR but we do not intend to assimilate SAR data into PROSAIL or to adapt PROSAIL for microwave data. As answered above in a previous question raised by the referee, the assimilation scheme which is based on a Bayesian approach can integrate additional information, but we would need to replace the PROSAIL model by an observation operator for SAR using a different model (Water cloud model WCM or other approaches). Two references are mentioned in the text (line 566), Veloso et al. (2017) who showed relation between SAR data GLAI and biomass, and Fieuzal et al. (2017) who derived GLAI from SAR and assimilated it into the SAFY model.

Comment:
In the second part and again in the fourth part of the discussion, the authors try to link their study, which is on winter wheat yield n South-West France, to soil carbon processes in general. As the connection is vague and indirect, because the relevant soil processes were neglected here, these attempts come across as rather endeavored. The further discussion treats well-known basic facts about remote sensing, e.g. the respective tradeoffs and advantages of physical modelling and machine learning.

Answer:
Again as explained above we re-emphasis: Our paper is not limited to winter wheat yield in South-West France. Actually only one figure over 15 concerns the yield ! We think that if we presented in the paper the methodology only without an application, it would have been impossible to show the advantages, drawbacks, and limitations of the approach. So alternatively producing simulations over a given area of interest of 100x100km shouldn't be an argument to reduce the paper to just this application and moreover reduce the application to one single variable.

Comment:
In general, I think that the Bayesian approach with the associated uncertainties makes a lot of sense and I would have been very curious about the explanations why the system performed so poorly with respect to spatial patterns. The discussion, however, only traces this major drawback to uncertainties in the in-situ data (e.g. in the combine harvester measurements), which I do not find very convincing.

Answer:
We agree that the discussion on the spatial patterns can be strengthened. We provided many elements in this comments that show that the results are not actually poor while we are aware of the limitations, including for a paragraph on the challenge to estimate very high aboveground biomass values (see previous answer relative to the 2019 biomass data) and more elements relative to the uncertainty on the yield in-situ data as mentioned. We didn't trace the results to the uncertainties in in-situ data only, here are excerpts from the text mentioning the model limitations:

- Line 429 : *"a clear saturation effect is observed in the simulations for the NAT-plt6 field"*
- Line 435 : *"The Low correlation as well as the difficulties in reproducing the range of yield observed variations of yield values may be caused by the simple representation of grain biomass allocation through the use of a HI which does not take into account potential variation of harvest index due to nutrient availability or crop cycle duration (Dai et al., 2016). "*

We also brought many elements in the previous responses that help clarify these points.

Comment:
A short conclusions and outlook section that mainly summarizes the main findings again closes the manuscript. As the authors state correctly at the end of the paper, the approach could potentially be used as "a coherent and multi-criteria full crop cycle agronomic diagnosis tool for production, carbon, phenology and water use". It is a pity that the presented paper does not advertise this potential of integrated remote sensing supported modeling approaches, but focuses on only one variable (aboveground biomass alias yield) and from my point of view misses to ask the relevant questions for time-seriesfrom EO data (e.g. What is the additional value of 10 m resolution compared to 30m? What

can be expected from future 5 m resolution data? What is the additional value of the spectral bands of Sentinel-2 compared to Landsat? What will be the impact of the new SWIR bands in the Sentinel-2 next generation for such integrated approaches? What does the interpretation of top of canopy spectral signals actually reveal about processes happening in the soil? etc.)

Answer:
First, we are really sorry to have to repeat the answer to the comment that we adresse "one variable" but we find it important to recall: we strongly disagree with the statement that the paper only focuses on yield and on the reasons why we limited our analysis to the carbon budget components (please see answers above).

Concerning the unraised questions that the referee would have liked the authors to ask:

- What is the additional value of 10 m resolution compared to 30m? We provided the reason why the 10 m to 30 m resolution analysis is not the most relevant (see answers above). We summarise this here again. 1) the only reason to go for specifically an arbitrary resolution of 30m is that the Landsat data is at 30m but this data is at 16 days revisit and is not sufficient to properly drive the SAFYE-CO2 parsimonious model. 2) The vast majority of crop modelling application at regional scale uses the field scale resolution, so we computed the impact of the use of field scale against 10m resolution 3) we also provided in the analysis the comparison to using S2 or S2+L8, as it has no sense to make an analysis by removing the S2 data which is freely available at 10m resolution globally.

  - What can be expected from future 5 m resolution data? This was also answered above. So in summary it is of interest to answer this question, but clearly out of scope of this paper. It can be partially investigated before the launch of Sentinel-2 Next Generation (2034) by using Planet Scope, Venus VM5. We have a research program on that, we hope studies from other groups could answer this question with AgriCarbon-EO.

  -What is the additional value of the spectral bands of Sentinel-2 compared to Landsat? We didn't make specific analysis by removing spectral bands for the PROSAIL retrievals, this has been already investigated in the literature (Dong et al. 2023).

- What will be the impact of the new SWIR bands in the Sentinel-2 next generation for such integrated approaches? It is an interesting question but again was the referee expecting us to add a section on the impact of potentially new SWIR bands in Sentinel-2 next generation expected to be launched not before 2034 ? Clearly this would require a specific analysis. One way to proceed would be to use a physically based RT model like DART model (Gastellu-Etchegorry et al. 2017) to generate synthetic scenes at different stages of development of the crops and to assimilate it into AgriCarbon-EO while updating for PROSAIL-Pro while coding PROSAIL-Pro in python. Referee can understand that this is out of scope of this paper.

- What does the interpretation of top of canopy spectral signals actually reveal about processes happening in the soil? Top of canopy/soil spectral signal can give access to different variables with varying levels of uncertainty. In a general case, optical remote sensing can provide information about the structure and chemical composition of the plant as well as the composition and moisture of the soil. The processes inside the soil are thus not observed directly but the variables mentioned before condition de biomass that is returned to the soil, the top soil composition and moisture that are all relevant elements that are needed to model biological activity in the soil. The assimilation of this information can thus help to constrain soil processes indirectly.

- We didn't discuss this point indeed. The model produces above ground biomass but also below ground biomass, and soil respiration constrained by the soil moisture, so the assimilation of optical remote sensing of TOC reflectances is enabling better estimation of the soil processes.

In summary we find it quite positive that the current paper opens so many questions. Clearly launching research programs in the community to answer them is of interest. We hope that AgriCarbon-EO could contribute to answering them.

Comment:
Overall, I think the fact that field scale temporal patterns, field scale spatial patterns, large-scale and regional scale modelling together with questions about spatial and temporal sampling density are all mixed in the paper, blurs the

structure of the presentation and makes the manuscript a rather demanding and rather exhausting read. I would highly recommend focusing on fewer aspects, e.g. maintaining sections 1-4 plus Figure 11 from section 5 and removing the rest. The link to the CarbonFarming buzzword seems artificial and should be removed or mitigated.

Answer:
Clarifications and argumentations about spatial and temporal scales and the need for the regional study have been provided in previous answers. We agree though that streamlining the paper by synthesising, and recontextualizing some details and removing some elements is needed to enhance the readability. We provide more details on this at the end of the general comments section.

Concerning the last comment on Carbon Farming, to make the link with carbon farming more explicitly, we propose to add at the end of line 72 the following sentence "One of the main motivation for developing AgriCarbon-EO is to answer the growing demand for a MRV tools allowing the production of high resolution maps of carbon budget components estimates and their uncertainties for different context of applications (e.g. carbon farming projects for the voluntary carbon market, Common Agricultural Policy, National Determined Contributions) compliant with the frameworks proposed by Smith et al. (2020) and Paustian et al. (2019). Those are relying among other things on the combined use of process based models, remote sensing data and a range of in-situ data (e.g. flux towers) for validation as in our approach." As a matter of fact, one of the main purpose for developing AgriCarbon-EO is to answer the demand from the scientific community working on soil carbon sequestration in Agriculture that span from several initiative like the CIRCASA project recommendations (https://www.circasa-project.eu/). The objective was thus to contribute in the development of tools for Monitoring Reporting and Verification (MRV) of soil carbon stock changes that would meet a set of criteria. Among other things it involves the use of process based models, assimilation of high spatial resolution remote sensing data, flux tower sites for validation of the models…to estimate the carbon budget components with associated uncertainties. This demand and context have also given rise to scientific projects that financed the development of AgriCarbon-EO mentioned in the acknowledgment (H2020 NIVA project, "Naturellement Popcorn", the Bag'ages (Agence de l'eau Adour Garonne) projects, the Horizon Europe ClieNfarm (n° 101036822) and ORCaSa (n° 101059863) projects, all of which are addressing carbon farming. Of course, to finalise the tools meeting all the criteria and comply with the frameworks proposed by Paustian et al. (2019) and Smith et al. (2020) some developments are still needed in AgriCarbon-EO (e.g. to validate the coupling of SAFYE-CO2 with several soil models) and some validations of the whole approach against in-situ data of soil organic carbon stock changes are foreseen, as mentioned in the last section of the discussion (section 6.4). More generally, our approach is meant to provide a solution for MRV of soil organic carbon stock changes following carbon farming practices compliant with the framework of Smith et al. (2020) and applicable to different contexts of MRV (voluntary carbon market, agrifood sector's insetting, national inventories, carbon indicators for the common agricultural policy as during the NIVA project…).

Comment:
Also, the large amount of typing, language and format errors is quite surprising, given the autocorrection capabilities of state-of-the-art text processing software (see specific comments below).

Answer:
We sincerely thank the referee for his extensive editing work ! We will approve all the suggested modifications for the upcoming review process. We will also have the manuscript verified by an editing service.

Comment:
Before the potential further processing of the manuscript I recommend:
(i) to resolve the very large number of formal and spelling errors that prevent the readers from focusing on the content.

Answer:
All the formal and spelling errors listed by the referee will be corrected and the paper will be processed by an editing service.

(ii) to decide, if indeed all the scales, the intra-field scale, the large scale and the regional scale, should be treated in a single paper.

Answer:

We commented on this point above, explaining the structure and the modifications we suggest to streamline the reading of the paper. Clearly, several modifications should be made to enhance the readability. We suggest to put the sub objectives at the end of the introduction, to move SAFYE-CO2 equations to supplementary materials; to reduce section 2.4.1 on LUT generation ; to enhance the objective for each results section, to enhance figure 6, 7 and 9 ; to remove figure 10 and figure 14 in the results; to enrich the discussion section based on the exchanges in this document. We would maintain the multiscale validation as augmented in our previous answers.

(iii) to rethink the connection between the introduction that focuses on carbon farming and the actual content of the paper, which is more on yield modeling than on carbon sequestration.

Answer:
We reiterate the objective of the paper is to present the carbon fluxes and biomass (yield is presented in one figure of the 15 figures). As mentioned above we will better explain how AgriCarbon-EO fits in a soil carbon sequestration approach.

(iv) to think about, comparing different pixel resolutions would be more relevant than comparing field scale to pixel scale resolution.

Answer:
We invite the referee to consider the elements in our previous replies. To summarise, common applications of crop models at regional scale are done over field scale, we are advocating for regional scale intra-field resolution, so considering the community the paper addresses it is most important to compare the field to intra-field scale in the context of carbon fluxes and biomass.

(v) to analyze the impact of observations in specific phenological stages instead of taking only the absolute number of observations into account.

Answer:
We agree that from an agronomic point of view this would be interesting, but to make such a proper analysis is really not straightforward as explained above. We consider that a dedicated study is needed on this.more areas should be considered, because cloud coverage is not statistically stationary across the phenological stages and would bias the analysis.

(vi) to rewrite the discussion so that not only very general aspects of remote sensing are discussed, but the approach is referenced against other studies/approaches in the same field and especially the poor performance with respect to spatial patterns at the intra-field scale is adequately explained.

Answer:
Many elements in this discussion will enrich the discussion.

(vii) to rephrase the rather general title so that the limitation of the study to certain crops (winter wheat), certain variables (biomass, yield) and the region (South-West of France) is reflected in the headline.

Answer:
As this is the first paper that presents the methodology behind AgriCarbon-EO. We proposed to replace the current title by :
"AgriCarbon-EO: v1.0.1: Large Scale and High Resolution Simulation of Crop Carbon Fluxes and production by Assimilation of Sentinel-2 and Landsat-8 Reflectances using a Bayesian approach".
As explained above we do not wish to focus on winter wheat in the title as the methodology can apply to different crops and locations, winter wheat was chosen as an example for the application of the method.

Specific Comments

Note:

All typo, grammatical and rephrasing comments will be answered (validated) in the discussion stage. They were not included in this clarification reply.

Abstract

Comment:
Line 2: Phrasing. I don't think that in-situ sampling is "prohibitive". It surely is extremely labor intensive and thus not feasible. But the main drawback in my opinion is that it will never be spatially continuous.

Answer:
We agree on the message to relay to the reader. We will replace the word "prohibitive" as it has a legal sense to it that is not intended here. We mention that making soil samples every 3-5 years at intra-field resolutions and at national scale is unrealizable.

1 Introduction

Comment:
Line 59: Comment. "...information about development dynamics..." I think it would be important to highlight here that GLAI incorporates both, information on environmental growth conditions as well as information on human (management) behavior.

Answer:
We agree on this, we will clarify this in section 2.3.2. For instance, by replacing the following sentence "The soil stress impact on the vegetation cycle scale is implicitly considered through the assimilation of GLAI." by "The effects on the vegetation cycle of environmental stress impacts (e.g. nutrient or soil water content availability) and management are implicitly considered through the assimilation of GLAI".

Comment:
Line 61: Phrasing. The term *"restituted to the soil"* is somewhat misleading in my opinion, because the main issue here is the long-term storage of atmospheric carbon in the topsoil, which to the largest part comes from geological reservoirs and not only from agricultural soils and thus is not strictly all given back to the soils. I'd thus phrase it a little more neutral and just speak of *"biomass and carbon storage"* in the soil.

Answer:
This is a misunderstanding. Here we are talking about the carbon contained in the biomass that is returned to the soil. The word "restituted" will be replaced by "returned" which is commonly used by the soil carbon scientific community.

Comment:
Line 79: Question. Again, as there are temporal (frequency of observation) as well as geometric questions (intra-field resolution) targeted in your study, what kind of resolution is focused here?

Answer:
Spatial resolution as mentioned above. Our study, target the production of outputs at daily timescale and 10m spatial resolution over large areas (100 x 100 km).

2 Methods

Comment:
Line 84: Question. Is a daily time-scale adequately suited to model crop growth and potential stressors?

Answer:
Yes for crop growth and carbon fluxes at crop cycle scale, provided that the model trajectory is corrected via assimilation of frequent observations, not surely for the instantaneous impact of stressors. Clearly the time-scale and spatial scale depends on the modelled phenomenons and processes. The vast majority of crop models have daily time steps (CropSyst (Stöckle, 2003), STICS (Brisson, et al. 2003), DNDC Gilhespy et al. 2014 , SAFY- Duchemin et

al. 2008, SAFYE-CO2 Pique et al. 2020a, Sunflo (Casadebaig et al. 2011). We consider that in the light of the announced objectives the daily time step is a good trade-off between precision and efficiency knowing that the objective of the paper is not to present a new crop model but the assimilation scheme and the overall performance of the approach to estimate carbon budget components for cropland.
Still, for more detailed modelling a higher time-scale would be needed (DSSAT runs at hourly time step, SCOPE also). If one wan't to aim at modelling for example the unsaturated flow and solute transport in the soil (Al Bitar, 2007), or turbulent flow above the canopy for energy balance, for specific stressors, a sub-second time step may be required not only to represent the process but also for numerical stability. In these cases a regional modelling at very high resolution would be extremely demanding, and an extensive amount of approximations need to be considered for the input parameters.

Comment:
Line 85: Question. Wouldn't the assimilation scheme not also work for microwave data?

Answer:
For passive microwave: for example if we are addressing soil moisture or vegetation optical depth (Al Bitar et al., 2017), a spatial downscaling would be needed considering the high spatial resolution gap with the model spatial resolution, also the representative depth should be considered (Lievens et al., 2016).

For active microwave: SAR data exist at 10m but it should be processed for speckle effect so resolution change should be considered. In summary specific processing should be done.

Nevertheless algorithmically, the Bayesian approach is a good framework for multi-sensor data assimilation and would work as long as an observation operator is provided to generate a variable that is modelled by SAFYE-CO2.

Comment:
Line 96: Question. Is analyzing each image independently making full use of time-series of satellite data?

Answer:
Yes, in the context of the defined objectives and trade-off between computational needs and precision. The other two options to consider spatio-temporal information would have been:

1) make an integrated retrieval of the PROSAIL (observation operator) and SAFYE-CO2 model in a spatio-temporal manner by assimilating the reflectances in the integrated system. For local application this would be possible but at large scale and 10 m resolutions it would require a large amount of computational resources which limits the scalability to 100 by 100 km scales.

2) We could get more information from the time series as there is temporal correlation inside of the LAI time series that could be used to constrain each image further. However, this correlation is variable in time and space depending on the different growth stages and thus difficult to characterise and implement.

Comment:
Line 119: Comment. I think it would be worth mentioning that parceled land use data is not available for many parts of the world. The requirement of parceled land use inputs limits the applicability of the AgriCarbon-EO approach to those areas where parcel information is available.

Answer:
Actually, AgriCarbon-EO can be applied if a pixel-based classification map is available without parcel's limits. The limiting factor would be a classification map which is much more easily accessible than parcel data. We can clarify this issue in the manuscript.

Comment:
Line 123: Question. I understand that a UTM map projection corresponds to the Sentinel-2 imagery that is used in the approach. I just wonder if staying with an equal-area projection would make sense here to facilitate the quantification of fluxes per area.

Answer:
A concept of abstract grid exists in AgriCarbon-EO. Actually, all data sets are projected to a common Discret Global Gridding (DGG) System. Users can define the DGG as the EASE Cylindrical grid (Brodzik et al., 2012) which is equal-area, but at 10m resolution, the scale of this exercice and considering the associated uncertainties we don't see the specific interest of using an equal-area projection .

Comment:
Section 2.2.3: Question. I am surprised that neither wind velocity nor atmospheric carbon dioxide concentration are required as meteorological input. For a model that is targeting carbon farming applications, I would have expected a direct link between the water and carbon cycle to be present in the algorithm.

Answer:
The effect CO2 concentration is implicitly accounted for through the LUE calibration approach. Concerning wind speed, previous work on SAFY, SAFYE or SAFYE-CO2 showed that biomass, yield or CO2 fluxes could be estimated with good accuracy without accounting for wind speed. Also, When the user activates water balance in SAFYE-CO2, the wind velocity is implicitly considered in the Potential evapotranspiration computation.

Comment:
Line 149: Question. How does 8 km resolution weather data correspond to the intra-field geometric detail that is targeted in this study?

Answer:
The 8 km weather data corresponds to a mean value for the zone. However, we agree that the climatic variables environmental stressors (temperature radiation and rain, wind, air humidity etc..) that constrain the plant development and soil processes are forced by the microclimate (especially in hilly landscapes such as in the application section). In our approach, this effect is compensated by the value of the variable parameters that adapt to local conditions through the assimilation of GLAI remote sensing datasets. The processing chain has been developed to take into account the state-of-the art global and open weather dataset (e.g. ERA5-land) which is an important criteria to provide future carbon farming MRV tools. Providing weather data with better resolution will surely have an impact on accuracy and computational performances. In the manuscript, we discussed this in Section 6.2 while providing the impact on performances (Equation 28) of future weather data.

Comment:
Line 159: Question. One is wondering why in a study about carbon storage instead of Prospect-5d not the most recent version PROSPECT-PRO is used, which specifically includes absorption coefficients for the carbon-based constituents of aboveground biomass.

Answer:
Prospect-Pro (Féret et al. 2021) is a recent evolution of the Prospect model. Féret et al. (2021) mentions the following: "Our results indicate the importance of narrow SWIR domains, which will remain to be important also at the canopy level. Current multispectral spaceborne data (e.g., Landsat 8/9 and Sentinel-2 images) do not comply with the narrowband SWIR spectral requirements that we identified, and further investigations are necessary to conclude on feasibility and limitations of its potential use for N mapping using PROSPECT-PRO." If this limitation remains applicable, we would potentially integrate Prospect-Pro when narrow SWIR is available (launch of Sentinel-2 NG in 2034).

Comment:
Eq.1: Comment. I think it's a pity that such a "simple" LUE-model is used to describe carbon fixation at the land surface. There are gas-exchange models available that create a direct link between carbon and water cycles. For a study in carbon farming, I would expect a more complex approach.

Answer:

The scope of this paper is to analyse carbon budget components only. Also, we chose a parsimonious approach because our objective is to simulate in a diagnostic mode only those components at large scale and high resolution benefiting from the assimilation of GLAI derived from remote sensing which implicitly accounts for some environmental stress including atmospheric CO2 fertilisation effect. Many other agronomic models rely on LUE approaches (e.g. STICS). Using more complex photosynthesis modelling approaches (e.g. based on the Farquahr model) may be justified for analysing infra daily photosynthesis process or for forecasting (e.g. in future climate). In our case it would require more parameters, more calibration processes, more computation time and it might result in larger uncertainties in the key outputs. Therefore, we consider that given our objectives using a "simplified LUE approach" is justified.

Comment:

Line 170/171: Question. What was the reason for ignoring water stress effects in the simulation? I think that very interesting findings can be made for example when the modelled biomass according to the natural water budget does not correspond to the biomass accumulation observed from satellites. Ignoring the water stress also means ignoring uncertainties in the soil parameterization. How does your model explain the differences found between simulation and satellite observation, if soil processes are ignored? If the model does not even try to explain them and simply accepts the observation and carries on, what can we learn about natural processes from such a model?

Answer:

- Question. What was the reason for ignoring water stress effects in the simulation?

We recall that the impact of water stress on production of biomass is implicitly accounted for in the assimilation of frequent remote sensing observations which are a representation of the in situ plant developpement. Impact of water availability on soil respiration and percolation is actually taken into account.

- I think that very interesting findings can be made for example when the modelled biomass according to the natural water budget does not correspond to the biomass accumulation observed from satellites. Ignoring the water stress also means ignoring uncertainties in the soil parameterization. How does your model explain the differences found between simulation and satellite observation, if soil processes are ignored?

It is not clear what the referee means with biomass accumulation from satellites. We are assimilating GLAI which is not a direct proxy of biomass. Still, we find the comment on discrepancies of interest. Actually, in such cases to identify the discrepancies, a crop model would be run without assimilation of remote sensing GLAI and forced by weather inputs only. The discrepancies between modelled GLAI and observed GLAI (or modelled soil moisture and remote sensing based soil moisture) can then be interpreted in an inversion scheme as the impact of unrepresented processes ie. soil properties, irrigation, water flow or pathogens depending on the model. These effects can all be present at the same time which makes them difficult to discriminate. These are all alternative research objectives in which water stress should be taken into account explicitly.

- If the model does not even try to explain them and simply accepts the observation and carries on, what can we learn about natural processes from such a model?

Again, concerning soil processes, they are not ignored as mentioned above. Also the model doesn't just accept the observations and "carries on" as we are not doing a direct insertion. As we are using a Bayesian approach for the assimilation, the assimilation scheme takes the model uncertainties and the observation uncertainties into account, which are also the basis for obtaining the output variables uncertainties. The assimilation of GLAI in the agronomical model is to rely on the observations which has many benefits such implicitly accounting for not only stress but also for pest and pathogens effects on the plant development that cannot be reliably simulated by most agrometeorological models at large-scale, while simulating several processes (e.g. photosynthesis, plant respiration, biomass allocation…). Additionally the model provides useful information on the effects of climatic variable on the carbon budget components that allow for instance to analyse the effect of straw export or return on the annual C budget (Pique et al. 2020a), the effect of cover crops or spontaneous regrowth on the carbon budget components (Al bitar et al. 2021), the water use efficiency (Pique et al. 2020b). But again, As mentioned above the objective of the paper is not to present the SAFYE-CO2 model, but to present AgriCarbon-EO and to evaluate its potential for estimating carbon budget components at high resolution and large scale, and we argue that the community can learn from AgriCarbon-EO on questions linked to the uncertainty of the estimations of the carbon budget components.

Comment:

Line 175: Question. Is a multiplicative factor well suited to describe the temporal dynamics of senescence? A multiplicative factor will result in a rapid decrease of "greenness" at the onset of senescence, while the increments of the greenness-decrease will become smaller as senescence progresses. From my experience, S-shaped sigmoid functions better correspond to the dynamics of senescence that are observed in the field. The model results shown in Fig. 5 also do not look like as if a constant senescence factor was applied. Could you please explain?

Answer :
The senescence is modelled in the study, as a function of the sum of temperature (SMT) so it can adapt to changes in weather conditions and changes in GLAI as implemented in SAFY (Duchemin et al. 2008), SAFYE (Battude et al. 2017), and SAFYE-CO2 (Pique et al. 2020b) models. Actually, the function used here is a discrete form of a sïgmoid function in SMT. It depends on sena: the sum of temperature at which senescence begins, and senb the parameters that control the slope of LAI decrease.

Comment:
Eqs. 6 and 7: Comment. From my perspective, the maintenance respiration should be connected to the tissue that already has been accumulated which requires “maintenance energy”. I don’t see a link to the accumulated biomass here.

Answer:
The maintenance respiration is linked to the NPP that integrates all the carbon that has been accumulated by the plant. It is equivalent to (DAM+DBM)/C content.

Comment:
Eq. 8: Question. What is Yg? Growth conversion efficiency? This could be included in line 189.

Answer:
Yes it is growth conversion efficiency. We will clarify this in the text.

Comment:
Eq. 11. Question. If the water stress response of the vegetation in the model indeed has been deactivated as stated in line 170, no realistic simulation of the soil moisture status is possible. Does it then make sense to use soil moisture as a proxy for Rh?

Answer:
The water stress corresponds only to the impact of water availability on the production of the vegetation (GPP and DAM), which is continuously updated by the assimilation of the remote sensing data. The available water on the other hand impacts the transpiration and soil respiration so there is a realistic modelling of the water budget. Actually, Pique et al. (2020b) showed for contrasted climatic years that water stress effect was already accounted for though GLAI assimilation in SAFYE-CO2 when estimating GPP and DAM for winter wheat (no improvement of GPP estimates when the stress function was activated) (again only for GPP).

Comment:
Eq. 15.: Question. I’m curious. SLA obviously is the key parameter for GLAI development in most models. I understand that in you study SLA (Cm) is constrained in the PROSPECT model inversions to the ranges given in Table 2. However, does the growth model, if running without satellite data to assimilate, consider changes of SLA over the course of the growing cycle?

Answer:
Yes, *SLA* can be configured to be either constant (Pique et al. 2020 a) or dynamic (Battude et al. 2017) over the course of the growing cycle. So yes theoretically, it is possible to consider changes of *SLA*. It is important to mention that SAFYE-CO2 has been developed specifically for spatialised simulations through GLAI assimilation. It has not been applied without remote sensing assimilation.

Comment:
Lines 224/225: Comment. I understand that decoupling the water and carbon cycle is convenient, because it alleviates the necessity to explain discrepancies between vegetation growth simulated according to the natural conditions and growth observed by the satellite. However, I see potential gaps evolving from that. E.g., your approach allows you to force the model into the reproduction of GLAI values that may be found in the satellite data, but cannot be explained by the meteorological budget (water, temperature) or the natural conditions (soil structure, nutrient supply etc.), as it

might for example be the case for irrigated areas. In this case, the mass and energy balance of your approach would not be maintained.

Answer:
As mentioned above the decoupling is only applied for the computation of the GPP by removing the impact of water stress on the photosynthesis. The water budget that is expressed via the percolation in the soil and the evapotranspiration is still dependent on the available soil moisture. For the natural conditions, it is also used like in the case of the soil respiration which depends on the soil moisture. So the water balance is maintained. For the energy balance we are only using the FOA56 method to compute the evapotranspiration based on the potential evapotranspiration provided by the weather dataset.
The case of irrigation doesn't concern this winter wheat application as winter wheat is not irrigated in our area of interest. On other crops such as corn the user will activate an automatic irrigation module that ensures a consistency between evaporation demand, and plant growth (Battude et al. 2017). Nevertheless, applying irrigation schemes in crop and land surface models is a hot subject and requires specific attention (Druel et al. 2022).
Referee has raised this point several times and we answered it. We find it is important to add this information to the manuscript to clarify it also for the readers.

3. Application

Comment:
Figure 2: Question. What is the data source of the DEM in the background?

Answer:
The map is ESRI World Topo Map
The map sources are: Esri, HERE, Garmin, Intermap, increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China (Hong Kong), (c) OpenStreetMap contributors, and the GIS User Community.

Question:
Line 376. I think the term "relative humidity" should be reserved for the meteorological variable. What was the reason for avoiding the commonly agreed term "canopy water content"?

Answer:
We agree that the term "canopy water content" is in fact more accurate. The term was taken from the sampling protocol technical documents.

Comment:
Line 378. Question. It is unclear what you mean by "Eight fields were also sampled using the ESU protocol in 2018"? As you didn't mention how many fields were sampled in 2017 and 2019. As your analysis focuses on the growing seasons of 2016/2017 and 2018/19, including the 2018 fields is somewhat confusing. Please provide a clear overview of how many fields were sampled according to which protocol in which year. This ideally should correspond to the points displayed in Figure 6.

Answer:
This was thoroughly answered earlier. We also provided the locations and dates of the 53 measurements used. We also provided the time series of simulated vs observed DAM for the year 2018 showing the fitting. For 2017-2018 over field FR_AUR the crop was not winter wheat. Table 4. in the manuscript shows all the list of simulations and the way they are used.

Comment:
Section 3.2.2. Question. It is known that yield data from combine harvesters is well-suited for describing relative spatial heterogeneities of yields, but may suffer from large errors concerning the absolute yield values. Were the CH measurements corrected, e.g. by determining the absolute weight of the harvest of the fields on a scale and applying the bias?

Answer:
Thanks for pointing this out. No, this correction wasn't applied. It may explain some mismatch between the results and the in-situ data and may be of interest. We will mention it in the discussion.

4. Validation

Comment:
Line 391: Question. I don't understand the reference to Equation 27, please explain.

Answer:
We referenced equation 27 to indicate the method used to obtain the plot scale time series from the original 10 m pixel scale simulation.

Comment:
Figure 5: Question. It appears that growth activity in terms of GPP was overestimated in the model compared to the observations in the months February to May 2019. The simulated GLAI development in March 2019, however, is underestimated compared to the observations. Could you please elaborate on that? Compared to 2017/18, the deviations between modeled and observed variables indeed are higher for 2018/19. Would you think that the neglection of water stress dynamics contributed to these deviations?

Answer:
It is true that modelled GPP is higher than the measured GPP during the growing phase in 2019, and the GLAI in March seems to be retrieved correctly but not at the end of the cycle. The performances in 2018-2019 are less accurate than in 2016-2017. We attribute this discrepency to three factors. First The quality of flux data seems to be better in 2016-2017 than in 2008-2019, we illustrate this in **Figure R7**. In this figure we overlay 2 flag information from the flux tower data over Figure 5 from the manuscript. The grey color corresponds to date where more than 50% of the eddy measurements - that are initially provided at 30 min intervals - are gap filled. This reveals that the majority of the information given during the month of february is produced by the gap-filling procedure. It is however known that gap-filling is less performant for large gaps (Moffat et al., 2007). The second flag in red represents days where there is doubts on the partitioning of Net Ecosystem exchange in Green primary production and Ecosystem respiration as the respiration values are incoherent for the end of march in the absence of freezing. It is notable that this period also presents a high number of gap filled information over a shorter time span. Finally 2018-2019 was an exceptional cropping season for wheat in or region of interest in terms of production. This is due to climatic as well as management factors through a high Nitrogen fertilisation on this specific plot. It is a well known fact that nitrogen fertilisation can influence Chlorophyll concentrations (Hinzman et al., 1986). This can lead to slight overestimations of LAI. This is due to the similar effect of LAI and chlorophyll content increase have on the reflectances simulated by PROSAIL. This can be mentioned in the manuscript.

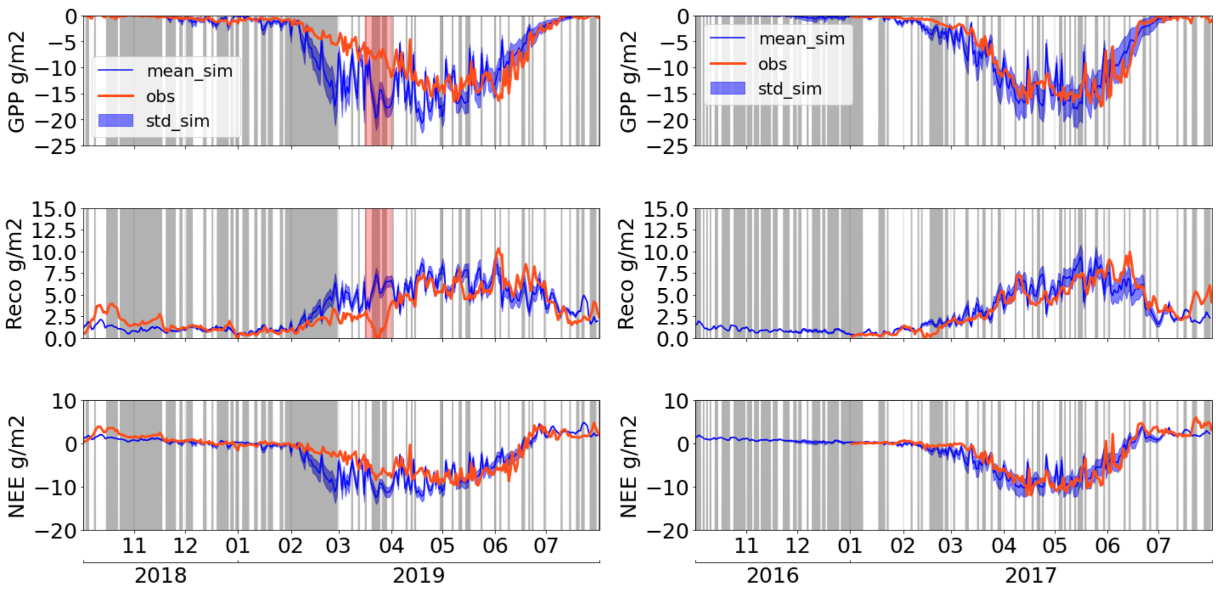

Figure R7: Time series of CO2 fluxes. In blue the a posteriori distribution and the standard deviation. In red the GLAI derived from the satellite observations and the NEE, Reco and GPP at the FR-AUR site for two cropping years (2017 and 2019).

Comment:

Table 6: I think it is important to highlight somewhere that the statistics given in Table 6 and in the text for the FR-AUR-Fields correspond to the agreement of the temporal biomass development and not to the agreement of the spatial yield patterns within or between the fields.

Answer:

It is of interest as pointed out to specify spatial and temporal variability and keep this figure to illustrate the representation of the temporal component.

Comment:

Figure 7: Comment. I understand that you cannot show a comparison for simulated and observed growth for the ESU-fields 2018, as there are no flux towers installed at these fields. However, first reading about the detailed model results for 2017 and 2019 and then seeing a validation including lots of points for 2018, is somewhat confusing. I think a well-structured overview about all the samples that are used is missing. To me, it is not clear to which ground samples the different data pairs in the scatter plot actually correspond. Obviously, the model returned constant values of 2000 g m-2 for 2019, while the in-situ data showed large variations. Are these data from different fields? Or are they from different ESUs in the same field? Or are they from different ESUs in the two combine harvester fields? Sorry, if I'm sounding confused here, but I think this must be made more clear.

Answer:

The questions around figure7 were answered above. And all the raised questions were answered. As the referee noticed (Figure R1) we did provide the observed and simulated growth for the ESU-2018 fields. They show a good agreement. What we don't have is flux towers on all these fields to show the carbon fluxes. As mentioned above figure 7 will be enhanced to show better these information.

Comment:

Figure 8: Comments/Questions. What do "plot3" and "plot6" mean? Why are the names of the fields as given in the caption not displayed here? Why are the respective harvest years not printed? Why are there no scale bar, no North-arrows and no coordinates? The units should be t ha-1 and not t.ha-1. The variable is "Yield" and not "Yiled". In the right part of the figure, there are small black dots between the fields. What do they represent? The agreement of the spatial patterns is surprisingly poor, given that the assimilation of GLAI should above all enable the simulation of intra-field heterogeneities.

Answer:

The figures will be enhanced but concerning the representation of intra-field heterogeneity we showed an analysis earlier in the answers that show that the outputs have a reasonable agreement regarding the range of variability seen in the field.

5. Large Scale

Comment:
Line 445/446: Question. While the scene is 110 x 100 km, the number of pixels with wheat fields is much lower as it can be seen in Figure 9. For which number of pixels do the given computing performances apply?

Answer:
In fact, the number of wheat pixels (20 M pixels line 329) which corresponds to about 20 000 000 /(11000*11000)*100=16.5% of the scene. To be clear we simulate all the 20 000 000 wheat pixels given by the RPG in the 110*110 km window.

References :

Al Bitar, Modélisation des écoulements en milieu poreux hétérogènes 2D/3D, avec couplages surface/souterrain et densitaires, Thése de Doctorat, SDUS2E, 2007.

Al Bitar, A., Mialon, A., Kerr, Y. H., Cabot, F., Richaume, P., Jacquette, E., ... & Wigneron, J. P. (2017). The global SMOS Level 3 daily soil moisture and brightness temperature maps. *Earth System Science Data*, *9*(1), 293-315.

Al Bitar, A., Wijmer, T., Arnaud, L., Fieuzal, R., Soussana, J. F., Gibrin, H., ... & Ceschia, E. (2022, July). Quantification of the impact of cover crops on Net Ecosystem Exchange using AgriCarbon-EOv0. 1. In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5781-5784). IEEE.

Aragon, B., Ziliani, M. G., Houborg, R., Franz, T. E., & McCabe, M. F. (2021). CubeSats deliver new insights into agricultural water use at daily and 3 m resolutions. Scientific reports, 11(1), 1-12.

Broberg, M. C., Högy, P., Feng, Z., & Pleijel, H. (2019). Effects of elevated CO2 on wheat yield: non-linear response and relation to site productivity. *Agronomy*, *9*(5), 243.

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., & Savoie, M. H. (2012). EASE-Grid 2.0: Incremental but significant improvements for Earth-gridded data sets. *ISPRS International Journal of Geo-Information*, *1*(1), 32-45.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P. and Bussière, F., 2003. An overview of the crop model STICS. *European Journal of agronomy*, *18*(3-4), pp.309-332.

Casadebaig, Pierre, et al. "SUNFLO, a model to simulate genotype-specific performance of the sunflower crop in contrasting environments." *Agricultural and forest meteorology* 151.2 (2011): 163-178.

Deininger, K., Monchuk, D., Nagarajan, H. K., and Singh, S. K.: Does Land Fragmentation Increase the Cost of Cultivation? Evidence from India, The Journal of Development Studies, 53, 82–98, https://doi.org/10.1080/00220388.2016.1166210, https://doi.org/10.108000220388.2016.1166210, publisher: Routledge _eprint: https://doi.org/10.1080/00220388.2016.1166210, 2017.

Dick, A., Dedieu, G., Hagolle, O., Raynaud, J. L., Pelou, S., Farges, M., & Peschoud, C. VENµS Mission Evolutions and Radiometric Performances During VM5 In-orbit Test Phase. Conference on Characterization and Radiometric Calibration for Remote Sensing (CALCON) (2022)..

Dong, T., Liu, J., Liu, J., He, L., Wang, R., Qian, B., ... & Shang, J. (2023). Assessing the consistency of crop leaf area index derived from seasonal Sentinel-2 and Landsat 8 imagery over Manitoba, Canada. *Agricultural and Forest Meteorology*, *332*, 109357

Druel, A., Munier, S., Mucia, A., Albergel, C., & Calvet, J. C. (2022). Implementation of a new crop phenology and irrigation scheme in the ISBA land surface model using SURFEX_v8. 1. *Geoscientific Model Development*, *15*(22), 8453-8471.

Duchemin, B., Maisongrande, P., Boulet, G., and Benhadj, I.: A simple algorithm for yield estimates: Evaluation for semi-arid irrigated winter wheat monitored with green leaf area index, Environmental Modelling & Software, 23, 876–892, https://doi.org/10.1016/j.envsoft.2007.10.003,
https://www.sciencedirect.com/science/article/pii/S1364815207002010, 2008.

Féret, J. B., Berger, K., De Boissieu, F., & Malenovský, Z. (2021). PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents. *Remote Sensing of Environment*, *252*, 112173.

Gastellu-Etchegorry, J. P., Lauret, N., Yin, T., Landier, L., Kallel, A., Malenovský, Z., ... & Mitraka, Z. (2017). DART: recent advances in remote sensing data modeling with atmosphere, polarization, and chlorophyll fluorescence. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(6), 2640-2649.

Gilhespy, S. L., Anthony, S., Cardenas, L., Chadwick, D., del Prado, A., Li, C., ... & Yeluripati, J. B. (2014). First 20 years of DNDC (DeNitrification DeComposition): model evolution. *Ecological modelling*, *292*, 51-62.

Hao, S., Ryu, D., Western, A., Perry, E., Bogena, H., & Franssen, H. J. H. (2021). Performance of a wheat yield prediction model and factors influencing the performance: A review and meta-analysis. *Agricultural Systems*, *194*, 103278.

Hinzman, L. D., Bauer, M. E., & Daughtry, C. S. T. (1986). Effects of nitrogen fertilization on growth and reflectance characteristics of winter wheat. *Remote sensing of environment*, *19*(1), 47-61.

Pasquel, D., Roux, S., Richetti, J. *et al.* A review of methods to evaluate crop model performance at multiple and changing spatial scales. *Precision Agric* 23, 1489–1513 (2022). https://doi-org.insu.bib.cnrs.fr/10.1007/s11119-022-09885-4

Paustian, K., Larson, E., Kent, J., Marx, E., and Swan, A.: Soil C Sequestration as a Biological Negative Emission Strategy, Frontiers in Climate, 0, https://doi.org/10.3389/fclim.2019.00008, https://www.frontiersin.org/articles/10.3389/fclim.2019.00008/full, publisher: Frontiers, 2019.

Pique, G., Fieuzal, R., Al Bitar, A., Veloso, A., Tallec, T., Brut, A., Ferlicoq, M., Zawilski, B., Dejoux, J.-F., Gibrin, H., and Ceschia, E.: Estimation of daily CO2 fluxes and of the components of the carbon budget for winter wheat by the assimilation of Sentinel 2-like remote sensing data into a crop model, Geoderma, 376, 114 428, https://doi.org/10.1016/j.geoderma.2020.114428, https://linkinghub.elsevier. com/retrieve/pii/S0016706119321998, 2020a.

Pique, G., Fieuzal, R., Debaeke, P., Al Bitar, A., Tallec, T., and Ceschia, E.: Combining High-Resolution Remote Sensing Prod- ucts with a Crop Model to Estimate Carbon and Water Budget Components: Application to Sunflower, Remote Sensing, 12, 2967, https://doi.org/10.3390/rs12182967, https://www.mdpi.com/2072-4292/12/18/2967, number: 18 Publisher: Multidisciplinary Digital Pub- lishing Institute, 2020b.

Löscher, A., Martimort, P., Jutz, S., Gascon, F., Donlon, C., Manolis, I., and Del Bello, U.: The ESA Sentinel Next-Generation Land & Ocean Optical Imaging Architectural Study, an Overview, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-6675, https://doi.org/10.5194/egusphere-egu2020-6675, 2020

Lievens, H., De Lannoy, G. J. M., Al Bitar, A., Drusch, M., Dumedah, G., Franssen, H. J. H., ... & Pauwels, V. R. N. (2016). Assimilation of SMOS soil moisture and brightness temperature products into a land surface model. Remote sensing of environment, 180, 292-304.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., ... & Stauch, V. J. (2007). Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and Forest Meteorology*, *147*(3-4), 209-232.

Smith, P., Soussana, J.-F., Angers, D., Schipper, L., Chenu, C., Rasse, D. P., Batjes, N. H., van Egmond, F., McNeill, S., Kuhnert, M., Arias-Navarro, C., Olesen, J. E., Chirinda, N., Fornara, D., Wollenberg, E., Álvaro Fuentes, J., Sanz-Cobena, A., and Klumpp, K.: How 920 to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal, Global Change Biology, 26, 219–241, https://doi.org/10.1111/gcb.14815, https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb. 14815, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.14815, 2020.

Stöckle, C. O., Donatelli, M., & Nelson, R. (2003). CropSyst, a cropping systems simulation model. *European journal of agronomy*, *18*(3-4), 289-307.

Tomer, S. K., Al Bitar, A., Sekhar, M., Zribi, M., Bandyopadhyay, S., Sreelash, K., ... & Kerr, Y. (2015). Retrieval and multi-scale validation of soil moisture from multi-temporal SAR data in a semi-arid tropical region. Remote Sensing, 7(6), 8128-8153.

Tomer, S. K., Al Bitar, A., Sekhar, M., Zribi, M., Bandyopadhyay, S., & Kerr, Y. (2016). MAPSM: A spatio-temporal algorithm for merging soil moisture from active and passive microwave remote sensing. Remote Sensing, 8(12), 990.

Valero, S.; Arnaud, L.; Planells, M.; Ceschia, E. Synergy of Sentinel-1 and Sentinel-2 Imagery for Early Seasonal Agricultural Crop Mapping. *Remote Sens.* **2021**, *13*, 4891. https://doi.org/10.3390/rs13234891

Zribi, M., Muddu, S., Bousbih, S., Al Bitar, A., Tomer, S. K., Baghdadi, N., & Bandyopadhyay, S. (2019). Analysis of L-band SAR data for soil moisture estimations over agricultural areas in the tropics. *Remote Sensing*, *11*(9), 1122.