Review Wijmer et al.

General Comments

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!

If we take a closer look on 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.

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

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, R^2 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. 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.

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. 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.

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. 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.

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.

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. 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. 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.

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-series from 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.)

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. 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).

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.
(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.
(iii) to rethink the connection between the introduction that focusses on carbon farming and the actual content of the paper, which is more on yield modelling than on carbon sequestration.
(iv) to think about, if comparing different pixel resolutions would be more relevant than comparing field scale to pixel scale resolution.
(v) to analyze the impact of observations in specific phenological stages instead of taking only the absolute number of observations into account.
(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.
(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.

Yours sincerely

Specific Comments

Abstract

Line 1: Phrasing. I am not sure if "*mitigation solution*" is the right word here and would prefer "*mitigation strategy*" instead.

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.

Line 5: Phrasing. Which kind of resolution is referred to in this sentence (spatial, temporal, radiometric…)?

Line 7: Phrasing. Please be consistent. Either use "*assimilate in*" or "*assimilate into*". I'd prefer the latter.

Line 8: Typo. "*transfert*" → "*transfer*"

Lines 10/11: Grammar. "*The chain considers as input a land cover maps, multi-spectral reflectance maps from the Sentinel-2 and Landsat-8 satellites, and daily weather forcing.*" → "*The chain considers land cover maps, multi-spectral reflectance images from the Sentinel-2 and Landsat-8 satellites, and daily weather forcing as input.*"

Line 11: Terminology. "*inversed*" → "*inverted*"

Line 14: Typo. "*agaisnt*" → "*against*"

Line 15: Question. I fear I don't understand what "$R^2 = 0.88 - 0.88$" means.

Line 16: Phrasing. "*We quantify the difference between pixel and field and pixel scale simulations…*" → *"We quantify the difference between pixel and field scale simulations…"*

1 Introduction

Line 24: Grammar. *"Agriculture and land use changes accounts for 15%..."* → *"Agriculture and land use changes account for 15%..."*

Line 39/40: Grammar. *"The need to monitor soil carbon at Farm level to inform individual farmers guide policies and development of carbon markets led…"* → *"The need to monitor soil carbon at farm level to inform individual farmers together with the development of guide policies and carbon markets led…"*

Line 42: Grammar. *"They rely on a soil centered quantification approaches which has limitations in terms of accuracy and reliability of the soil and biomass input data and a field scale resolution that often does not match the spatial resolution of in-situ soil and plant growth variability"* →*"They rely on soil centered quantification approaches which have limitations in terms of accuracy and reliability of the soil and biomass input data. Also, their spatial aggregation level on the field scale often does not match the spatial resolution of the soil and plant growth variability observed in-situ"*

Lines 46/47: Phrasing. *"These models include the main components of the cropland's carbon budget, plants photosynthesis, and respiration, emission due to soil organic matter mineralisation. These models can also account…"* → *"These models include the main components of the cropland carbon budget, photosynthesis and respiration, and emission due to soil organic matter mineralization. They can also account…"*

Line 52: Grammar. *"autotrophic respiration - Ra, heterotroph respiration - Rh"* → *"autotrophic respiration - Ra, heterotrophic respiration - Rh"*

Line 54: Grammar. *"This can results in high…"* → *"This can result in high…"*

Line 57: Phrasing. *"Getting hold on this information…"* → *"Getting hold of this information…"*

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.

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.

Line 64: Phrasing. Again, I am not so happy with the term *"prohibitive"*. I'd recommend changing *"often computationally prohibitive"* to *„computationally demanding"*.

Line 67: Typo. *"shemes"* → "*schemes*"

Line 72: Grammar. *"…radiative transfer model to Obtain GLAI at 10 m resolution, GLAI that is thereafter assimilated…"* → *"..radiative transfer model to obtain GLAI at 10 m resolution that is thereafter assimilated…"*

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?

2 Methods

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

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

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

Line 99: Typo. *"Ligth Use Efficiency(LUE)"* → *"Light Use Efficiency (LUE)"*

Line 103: Typo. *"corpping year"* → *"cropping year"*

Line 105: Interpunctuation. *"…requirement of 5 GB per process, for the satellite images…"* → *"…requirement of 5 GB per process for the satellite images…"*

Line 119: Comment. I think it would be worth mentioning that parceled land use data are 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.

Figure 1: Typo. *"construt"* → *"construct"*

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.

Line 137: Layout. Please avoid line-breaks between numbers and corresponding physical units.

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.

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

Line 152: Terminology. LAI is widely classified as a biophysical rather than a geophysical variable throughout the remote sensing community.

Line 155: Structure. Abbreviations such as SAIL should be explained when first mentioned in the text.

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.

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.

Eqs. general: Structure. I am missing direct references from the text to the Equations. You have included these cross—references for some, but not for all Equations. I suggest going through the entire manuscript and adding direct references for every single Equation in the paper to avoid potential misunderstandings. Thereby, you should aim for a uniform layout and decide whether you'd like to write "Eq. (x)" or "(Eq. x)" (I'd prefer the latter).

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?

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?

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.

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

Eq. 10: Typo. *"$T_s$oil"* → *"T_{soil}"*

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?

Line 199. Typo. *"…function and RSM1 The relative soil moisture…"* → *"…function and RSM1 the relative soil moisture…"*

Line 203. Typo. *"Biomass"* → *"biomass"*

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?

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.

Line 231: Typo. *"…physiology,heterotroph activity…"* → *"…physiology, heterotrophic activity…"*

Line 232: Question. What is a *"Look out table"*?

Line 238: Grammar. *"…of the model parameters knowing that the observations x…"* → *"…of the model parameters knowing the observations x…"*

Line 257: Phrasing. *"…but not to the assimilation of satellite imaging."* → *"…but not to the assimilation of satellite images."* or *"but not to the assimilation of satellite imagery."*

Line 263: Typo. *"(e.N >> n)"* → *"(e >> n)"*

Line 269: Typo. *"i.e"* → *"i.e."*

Line 274: Grammar. *"…this expression leads to the manipulations of…"* → *"…this expression leads to manipulations of…"*

Line 276: Grammar. *"…vanilla matrix product."* → *"…vanilla matrix products."*

Line 278: Format. *"…re-scaled by their maximum. Eq.(23)."* → *"…re-scaled by their maximum (Eq. 23)."*

Line 299: Typo. *"…Equations (22,23 and 24)…"* → *"…Equations (22, 23 and 24)…"*

Line 303: Typo. "*ug* m^{-2}" → "*µg* m^{-2}"

Line 307: Comment. Please always include physical units. *"…that is considered constant at 0.02"* → *"…that is considered constant at 0.02 g* cm^{-2}*"*

Line 313: Typo. *"…Equations (22,23 and 24)…"* → *"…Equations (22, 23 and 24)…"*

Eq. 27: Format. I think it's nice that you are using overleaf as it provides a good platform for manuscript editing by many authors. However, the reference to Eq. 27 in line 321 seems to be messed up…

Table 2: Grammar. *"…is not allocated to the Leafs"* → *"…is not allocated to the leaves"*

3. Application

Line 325: Phrasing. I think that simply referring to *"the chain"* might not sound straightforward to many readers. I'd suggest referring to *"the model chain"* or *"the assimilation chain"* instead.

Line 325. Grammar. *"…is applied over a for winter wheat in years 2017, 2018, and 2019"* → *"…is applied for winter wheat in the years 2017, 2018, and 2019"*

Line 327. Phrasing. *"Several assimilation experiments were conducted to answer the specific objectives of the paper, they are summarised in Table 3. They alternate the use of…"* → *"Several assimilation experiments were conducted to answer the specific objectives of the paper. They are summarized in Table 3 and alternate the use of…"*

Table 3: Grammar. The *Objectives* are not complete sentences. The full stop-symbols therefore should be removed.

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

Figure 2 caption: Typo. *"…tile 31TCJ limts…"* → *"…tile 31TCJ limits…"*

Figure 3 caption: Phrasing. *"The bars plots represent the percentage…"* → *"The length of the bars represents the percentage…"*

Line 335. Style. Please decide to either use spaces between numbers and their respective physical units or not and apply uniformly throughout the manuscript (also see line 359).

Line 338. Interpunctuation. "…while year 2019, had a mild winter…" → "…while year 2019 had a mild winter…"

Line 340/341: Phrasing. *"It is mainly occupied by agricultural fields that cover about 90 % of the area, among which a majority of seasonal crops."* → *"It is mainly occupied by agricultural fields that cover about 90 % of the area and are predominantly cultivated with seasonal crops."*

Line 356. Typo. *"…from the database of the Environmental Information System the laboratory and the Regional Spatial Observatory (RSO).his Information systhem centralizes…"* → *"…from the database of the Environmental Information System, from the laboratory and from the Regional Spatial Observatory (RSO). This Information system centralizes…"*

Line 375. Grammar. "…each sample corresponds to a one linear meter of the crop row." → "…each sample corresponds to one linear meter of the crop row."

Line 376. Question. 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"*?

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 focusses 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.

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?

4. Validation

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

Line 393: Phrasing. *"…fitting statistics computed over the growing season show a good fit (R^2 = 0.93) in with a lower fit for the growing season in 2019"* → *"…fitting statistics computed over the growing season show a good fit in 2017 (R^2 = 0.93) with a lower fit for the growing season in 2019"*

Line 395: Phrasing. *"The GLAI for year 2019 senescence period is under-fitted while the…"* → *"The GLAI for the senescence period of 2019 is under-fitted while the…"*

Line 397: Grammar. "…with a R^2 of 0.88, 0.91, and 0.62..." → "…with R^2 0.88, 0.91, and 0.62..."

Line 397: Structure. A reference to Table 5, where all the error indicators are listed, is missing here.

Line 403: Typo. "(RMSE = 1.43-1.90 gCm^{-2},Pique et al. (2020b))" → "(RMSE = 1.43-1.90 gCm^{-2}, Pique et al. 2020b)"

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 modelled and observed variables indeed are higher for 2018/19. Would you think that the neglection of water stress dynamics contributed to these deviations?

Figures 5 and 6: Typo. In the y-axis label, please either write ($g * m^{-2}$) or better ($g\ m^{-2}$), but avoid ($g.\ m^{-2}$).

Lines 412/413: Grammar. *"The comparison shows a good fit when considering together all DAM measurements with a R^2 of 0.90, a RMSE of 250 gm^2 and a slight negative bias 52 gm^{-2}."* → *"The comparison shows a good fit when considering all DAM measurements together with R^2 0.90, RMSE 250 gm^{-2} and slight negative bias of 52 gm^{-2}."*

Line 417: Grammar. *"…better fit that year with an R^2 = 0.94…"* → *"…better fit that year with R^2 = 0.94…"*

Table 6: Comment. 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.

Figure 7: Style. Please aim for uniform labels throughput the manuscript. The text and all other figures print the units as "g m^{-2}". In Figure 7 it is "g/m2" [sic].

Figure 7: Comment. I understand that you cannot show a comparison for simulated and observed growth for the ESU-fields 2018, as there are not 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.

Line 423: Style. The reference to Grisso et al. 2022 is not according to format. This also accounts for the respective entry in the list of references.

Line 424: Phrasing. Incomplete sentence? *"ACEO-S2L8-Pixel by multiplying the final DAM by the Harvest Index (HI)."*

Line 425: Grammar. *"These maps shows the comparison…"* → *""These maps show the comparison…"*

Line 431: Comment. The listing of the different performance measures is confusing. I suggest including all these numbers into Figure 8.

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.

Figure 8 Caption: "…for the 2017 and 2019." → "…for the 2017 and 2019 growing period."

5. Large Scale

Line 444: Phrasing. *"…considering 5000 LUT size."* → *"…considering a LUT size of 5000."*

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?

Line 460: Typo. *"…the maps in Figure 9 are a presented in…"* → *"…the maps in Figure 9 are presented in…"*

Figure 9: Style. Coordinates are missing. The units should be g m^{-2} and not $g.m^{-2}$. In the overview map, the extent indicators of the zoomed maps at the bottom are too tiny to be discernible. The zoom maps are missing a scale bar. The color bars seem to have an inadequate color stretch. The map

for NEP-Mean is scaled to show high positive values in blueish colors. However, only very few fields with a blue hue are visible in the map when zooming in to a maximum. Maybe the color stretch should be applied more aggressively to reveal the spatial patterns in the negative (red) value range. This would also benefit the zoom maps. The same applies to the NEP-std map, where I was not able to find a single red pixel, even when zooming in to the maximum.

Line 466. Typo. *"…agricultural practices (ex. early vs…"* → *"…agricultural practices (early vs…"*

Line 468. Phrasing. I'm not sure if *"presents"* is the right word here. Maybe better *"shows"* or *"reveals"*?

Figure 10. Style. A Legend explaining the colors is missing. The labels partly overlap (e.g. SENb vs. SLAb). The cropping of the decimals places of the labels seems arbitrary (Do three decimal places make sense in this case: 19471,605?).

Line 470. Phrasing. I'd suggest changing *"milder"* for *"less pronounced"*.

Figure 11: Style. The units should be $m^2\ m^{-2}$ and not $m^2.m^{-2}$. A legend explaining the colors of the lines is missing. Interpretation is not very intuitive, if the readers have to look up all the color codes in the caption. The date labels in figure (d) and (e) are cropped. The numbers indicating the ranges of the parameters in the radar plots are too tiny and partly overlap. There seem to be errors in the name labels of the parameters with exception of figure (a).

Figure 11: Caption. *"…in red the SAFYE-CO2 simulations), and"* → *"…in red the SAFYE-CO2 simulations), and"*

Figure 11: Caption. *"…where a cloud date is not filtered in (d)."* → *"…where a cloud date is not filtered."*

Figure 12: Style. The units should be $g\ m^{-2}$ and not $g.\ m^{-2}$

Figure 13: Style. The units should be $g\ m^{-2}$ and not $g.m^{-2}$

Line 515: Typo. *"…10 m resolution ans smoothed images…"* → *"…10 m resolution and smoothed images…"*

Line 516/517: Phrasing. "Figure 14 (b) is produced from the input land cover maps and shows the density of winter wheat fields over the region where the two main winter wheat regions: they correspond to the hilly areas located South-East of Toulouse and to the Gers department (West of Toulouse)." → *"Figure 14 (b) is produced from the input land cover maps. The density of winter wheat fields indicates the two main wheat cultivation regions in the hilly areas South-East of Toulouse and in the Gers department (West of Toulouse)."*

Line 523: Typo. *"paterns"* → *„patterns"*

Figure 14: Style. Coordinates, North arrows and scale bars are missing for all five maps. The legends should show the respective physical units.

Line 528/529: Phrasing. *"A higher number of winter wheat pixels at high altitude tend to emerge earlier than pixels at lower altitudes. This can be seen intuitively given the altitudinal temperature gradient, however this difference 530 may be caused by hill shading effects…"* → *"Winter wheat pixels at high altitudes tend to emerge earlier than pixels at lower altitudes according to the altitudinal temperature gradient. However, this difference may also be caused by hill shading effects…*

Line 532/533: Phrasing. *"This later observation can be explained…"* → *"The latter can be explained…"*

Line 536: Typo. The major axes are N, W, S, E. BTW, the correct technical term is *"aspect"* and not *"exposition"*.

Line 544: Phrasing: *"(at a lower degree days values)"* → *"(at a lower degree days threshold)"*

Figure 15: Style. The physical units should be provided for all parameters next to the color bars and diagram axes. The frames of the legends in Figure (c) are cropped at the top. *"Exposition"* should be replaced by *"aspect"* in the plots and in the figure caption.

6. Discussion

Line 558: Grammar. *"…because the uncertainty on the EO derived GLAI are accounted for…"* → *"…because the uncertainties in the EO derived GLAI are accounted for…"*

Line 598: Typo. *"…appropriate solutions(Figure10),(Wang et al., 2022)."* → *"…appropriate solutions (Figure10; Wang et al., 2022)."*

Line 601: Typo. *"Note that ixed prior parameters…"* → *"Note that fixed prior parameters…"*

Line 604: Structure. Please explain abbreviations when first mentioned in the text and only use the abbreviation afterwards (*"ML"*).

Line 616: Typo. *"term ,large scale assessment…"* → *"term ,large scale assessment…"*

Line 620: Grammar. *"…requires in addition to the new soil module parameters, input dataset on initial soil…"* → *"…requires in addition to the new soil module parameters, input datasets on initial soil…"*

Line 644: Typo. *"Co_2"* → *"CO_2"*

Line 658: Typo. *"forsoil"* → *"for soil"*