Abstract.

Line 25: Is cover cropping only common in humid and sub-humid regions? Perhaps it would be more informative to rephrase to something similar to: ".. which is an agricultural management technique commonly used in the regions evaluated in this study." Alternatively, you can say that it is a technique growing in popularity to improve soil health and carbon storage.

Thank you very much for your suggestion, which we will gladly incorporate.

Line 26: Are you referring to the parameterization of new CFTs? Please clarify.

Yes, we will clarify this by using the new wording as suggested below.

Line 27: Please move the reference of RSME for LH and SH to just after the energy fluxes rather than after NEE.

Thanks, we will incorporate this suggestion.

Line 31: When you refer to the "LAI curve", is this the same as the season cycle of LAI? If so, please modify the wording to reflect this.

We will rephrase this into "terms of LAI magnitudes and seasonal cycle of LAI".

Lines 31-33: It would be more impactful if you strengthened the last sentence in the abstract. Here is one suggestion: "Our modifications significantly improve model simulations and should therefore be used in future simulations to better understand large-scale impacts of agricultural management on carbon, water, and energy fluxes."

Thanks, we agree and will rephrase accordingly.

Introduction: Overall, the introduction needs some reorganization. You need to more clearly highlight the role of management (make this a separate paragraph, include cover crops but also other types of management). The new representation of cover crops is a primary contribution to this paper and is barely mentioned here. The introduction also needs a broader overview of crops in LSMs (it currently only focuses on AgroIBIS and CLM). Last, most of the introduction emphasizes the global nature of models and that the variation in soils, plants, climate is important. When the reader finally gets to the end of the introduction, which highlights that you focus on a few sites in Europe (which some may argue has narrower variation in soils, plants, climate than if you were to compare to locations from other continents), it makes this study seem limited. It might help to instead describe that models are still limited by their ability to represent many crop types and important management practices, emphasizing the importance of your work adding these new capabilities, and also to highlight that Europe is a major agricultural hub for global food production.

Thank you for your constructive suggestions. We agree and will re-evaluate and reorganize our introduction for the revised manuscript.

Lines 44-49: The mention of cover crops here seems a bit out of place. The earlier part of this paragraph and the start of the next paragraph is focused on adaptation to climate change, whereas the description of cover crops here focuses on soil benefits and climate mitigation. I suggest reorganizing, moving the cover crop description to later in the introduction.

Thanks for your suggestions, we agree and will restructure the text.

Lines 68-70: I'm not sure I entirely understand the point of this sentence. Is this just to highlight the evaluation of crops in CLM4.5?

Here we give a short overview of improvements in earlier versions of CLM. We agree that it might not be necessary here and will restructure the text accordingly.

Lines 73-76: You should also reference the CLM5 crop overview paper here, which evaluates global crop yields: Lombardozzi, D. L., Y. Lu, P. J. Lawrence, D. M. Lawrence, S. Swenson,

K. W. Oleson, W. R. Wieder, and E. A. Ainsworth (2020), Simulating Agriculture in the Community Land Model Version 5, J Geophys Res-Biogeo, 125(8), 927–19, doi:10.1029/2019JG005529.

Thanks for pointing this out, we will add this to our references.

Lines 77-84: This paragraph seems too detailed for the introduction. I suggest summarizing and merging with the previous paragraph. For example: "The few studies that have evaluated CLM5 suggest inaccurate phenology and overestimated crop yields (Sheng et al. 2018)." However, you'll probably want to change/update this to also incorporate results from the Lombardozzi et al. CLM5 paper mentioned above.

Thanks a lot for your suggestions on the introduction. We will restructure the introduction taking your comments into account and agree that Lombardozzi et al. (2020) should be added and cited here and will do so in the revised manuscript.

Methods Overall: The methods section needs to be tidied up. There are redundancies in the first section, and a lack of detail in the cover crop description. Please pay careful attention to providing enough detail that the reader isn't left wondering how something was done, but keep the text succinct. You reference Lawrence et al. 2018 in several places throughout the text. However, I believe this paper was published in 2019 (not 2018). Please double-check.

Thanks for your suggestions. We will re-evaluate and restructure several parts of our methods section in the revised manuscript as discussed below. We agree that Lawrence et al. (2019) should be added and will do so in the revised manuscript.

Section 2.1:
When describing the crop model, please also cite Lombardozzi et al. (2020), as this has much more detailed information about the crop model updates than Lawrence et al. 2019.
The methods should be streamlined to avoid repetition. For example, allocation is mentioned in lines 134-138, and then again in the paragraph starting at line 153. When referring to C allocation, you state that it varies throughout the growing season (e.g., line 156), whereas the reference to N allocation states that it uses two different C/N ratios (lines 161-162). However, these are treated the same way in the model. Please update for consistency. I suggest switching the order of Eq. 1 and Eq. 2.

Thanks, we will incorporate this suggestion.

Line 114: Please define "CFT" the first time you use this term.

We will do this as suggested by the reviewer.

Line 115: land units are not separated by fertilizer, only by irrigation. Please update.

We agree that the phrasing is misleading and will update it accordingly.

Lines 204-206: This is a bit confusing and could use clarification. Does the vernalization factor always range from 0-1? Is it applied to GDD for air and soil temperatures (e.g., does it affect all phenological phases)? If it is only applied to grain C allocation, where does the remaining C get allocated?

Yes, the vernalization factor ranges from 0 to 1 (fully vernalized) and affects the GDD in the phenology phase after planting (vernalization starts after leaf emergence and ends before flowering). This leads to a reduced growth when the plant is not fully vernalized and the vf is smaller than 1:

For vf <1

$GDD_o * vf = GDD_n$ with $GDD_n < GDD_o$

where the subscripts $_o$ and $_n$ stand for original and updated GDD.

We will better clarify the methodology of this routine by giving more details in section 2.2.1 or in the appendix (please see our replies below to section 2.2.1).

Section 2.2.1
It would be helpful to start with an overview of how winter cereal representation differs from other crops. I suggest a high level overview of why it's important to include both vernalization and cold tolerance before diving into the details of each.

Thanks for your suggestions. We will add more detail to this part in the revised manuscript.

Equation 4: You specify that Tcrown is slightly higher than the freezing temperature when covered by snow. I see that snow height is used in the calculation, but where is the plant height? Without including the plant height, how do you know whether the plant is covered by snow?

In CLM, the crown temperature is the crown depth soil temperature calculated as a function of 2m air temperature and snow depth. The crown is the connecting tissue between the roots and the shoots at the base of the plant. For winter wheat, the crown node is located at about 3 – 5 cm soil depth (Aase and Siddoway, 1979). We will add this explanation and reference in the revised manuscript.

Line 213: The text describes what the accumulative parameters are, but what about the previous time step is used? It would also be useful to include a brief description of how some of the accumulative parameters accumulate (e.g., are these all based on some aspect of accumulated temperature?)

We will include a more detailed description and add some equations from the source literature Lu et al. (2017) in order to clarify this.

Equation 6: Please define the "alpha-surv" and the "t" variables in this equation.

Thanks, we will add this.

Equation 7: I am confused by this, partly because it's not clear what the equation is taking the max of. Also, can Tcrown be negative? That seems to be the only way the solution to this equation isn't 0. Please update to clarify. Also, I think 'fsurf' should in fact be 'fsurv'.

Thanks, f_{surv} is correct.

Paragraph starting at line 227: I find the description here a little confusing. Can you revise this to more clearly articulate the difference between survival probability and WDD? Is survival probability just a step function, where any value <1 causes the same amount of damage (simulated as part of the C and N pools being transferred to litter)? Should I be thinking of survival probability as the proportion of the plant that survives, or the probability that the whole plant survives? Also, part of my confusion is that this is the first place that a frost damage function is mentioned.

Thanks for your suggestions on this section. We are currently evaluating whether to move some of the very detailed description to the appendix as it is not new to our study (as suggested by the other referee). We hope to clarify several of your specific comments above by adding more detail to these descriptions in the revised version of the manuscript. In section 2.2.1 of the paper we would then give a slightly broader overview on winter cereal representation following your suggestions.

Section 2.2.2:
Since you use a pre-existing winter wheat parameterization, it would be helpful to include some information about what you changed in the parameterization and why.

For the CFTs sugar beet and potatoes, the same parameters as for spring wheat are used on the default parameter set due to the lack of parameters specifically calibrated for these crops. For winter wheat, there are pre-existing crop-specific parameters available on the default parameter set. However, this default parameterization of winter wheat performed poorly in representing the crop phenology for the evaluated study sites in this study. This was also reported in an earlier study by Lu et al. (2017). Thus, crop specific parameters were also added for the winter wheat CFT.

We will include an additional text in the revised manuscript to clarify this.

Table 1: This is a useful summary, but I'm not sure it adds much information to the main text.

We believe that this overview table is very helpful information for the reader. Thus, we would like to keep it.

Section 2.2.3: How do you determine when the cover crops (or rotations) are planted and the subsequent phenology phases? Is it based on GDD? Did you have to modify GDD parameters or add new ones? Did you add new CFTs to accomplish this? How is allocation determined? This section needs more detail about how modifications were made, as it is the bulk of the development work in this paper.

The rotation schemes are hardcoded in the new cover cropping subroutine. Basically, in the new routine, the phenology algorithm is reset and restarted after harvest of any crop that is assigned with the cover crop flag. We are currently working on a version where the rotation is more flexible and user-friendly defined by a control file.

Lines 267-270: It's great to hear that you introduced a flag to use the cover crop option, but I'm not sure you need to include that description here.

We will add more detail to the description of the new routine and will keep this information in the revised manuscript.

Lines 276-277: How are you predefining a rotation scheme?

At this stage this is hardcoded in the new cover cropping subroutine. However, we are currently working on a version where the rotation is more flexible and user-friendly defined by a control file

Line 283: "catch crop" this is the first time you mention it. Are you using this interchangeably with cover crop (which is how you described this in the previous sentence), or are you using a new phrase to distinguish this from cover crop? Please be clear and consistent with word choices.

Here the terms "cover crop" and "catch crop" were used synonymously. We will correct the wording to be more consistent.

Line 283: You mention plowing the crop into the soil. However, CLM does not represent plowing. How did you accomplish this. Do you assume that the plant biomass is transferred to the litter pool? Also, how did you decide when this happens?

We agree that this could be misleading. The plant biomass is moved to the litter pool instead of to the grain product pool. This is done upon harvest of the crop. We will clarify this in the revised manuscript.

Section 2.3:
I think it would help to describe the sites before the validation data, and/or mention whether you run CLM simulations at these sites. This section starts by describing validation data, but does not mention what is being validated.

We will restructure the text accordingly.

Table 2: Useful information about the sites, but I think the map describes the locations quite well, and most of the other information included in the table is not used in the simulation. Therefore, I'm not sure that this table is necessary in the main text.

We think this table gives the reader a nice overview without having to read this section in detail and therefore would like to keep it in the main text. We will include an additional table with textural fractions at the study sites in the appendix of the revised paper as requested by RC1.

Lines 318-319: You mention winter wheat twice here.

Thanks, we will correct this.

Lines 341-342: CLM's default time step is 30 minutes.

Here we mean the customized time step of input forcing data, which was set to hourly. Not all meteorological input data was available half-hourly, thus an hourly temporal resolution was used. The internal model time step remains at 30 minutes.

Section 3.1: Throughout this section, the differences in model version versus parameter set seem to be conflated. Please make this much clearer throughout, explaining what each of the model versions includes and what the default versus modified parameter sets include.

Thanks for your suggestions on this section. We will change wording accordingly, as discussed above, in the revised manuscript.

Table 3: Which simulations include the potato and sugar beet parameterization? It looks like it's the CLM_WW simulation, but this needs to be explicitly mentioned in the table description.

We agree that more clarification is needed in this section. This table was meant to give an overview of all the simulation runs that were compared in this study. We will change the wording and include clearer descriptions in the text.

Lines 364-366: This text is confusing: It is not clear what the difference is between the default model and the modified model. I assumed the "default" model did not include winter wheat, but this text suggests that it does. How, then, is the default model run with the modified winter wheat parameters different from the winter wheat model with the modified parameters?

The CFT of winter wheat is included in the default model but its specific parameter set yielded very poor representation of simulated winter wheat phenology at our sites and also in previous studies. Thus, next to the implementation of vernalization and cold tolerance representation in the model code, new crop specific parameters were supplied in order to optimize the model performance. Please see also our replies to RC1, where we supplied an additional plot showing the differences in simulation results using both the default model with and without new parameterization and the extended model with and without new parameterization.

We will reorganize the text and table to increase the comprehensibility.

Lines 369-370: What are the default parameterizations of sugar been and potatoes? These aren't included in CLM, so is there a "default"?

Sugar beet and potatoes are included in the structure of the CLM5 crop module and are amongst the 64 CFTs. The CFTs sugar beet and potatoes do not have assigned parameters specifically calibrated for these crops, instead the same parameters as for spring wheat are set as default for these CFTs. We will rephrase the text using the term 'control' instead of 'CLM_D'.

Section 4

In general, I find the use of CLM_D, CLM_WW, and CLM_WW_CC to be confusing, as the changes included in each are not clearly described. Additionally, it seems that sugar beet and potato parameterizations are added to CLM_WW. It might be more helpful to instead refer to CLM_D as "control" or "default" and then refer to updated parameterization (e.g., "improvements to winter wheat" rather than "CLM_WW" in Section 4.1 and "new potato" or "new sugar beet" parameterization in Section 4.2). Additionally, throughout this section, figures should include estimates of uncertainty.

> Thanks for your suggestions. We agree that the different model versions and parameterization approaches should be described in more detail. We will improve the descriptions of the model versions and the table in section 3.1 accordingly.
>
> Due to the small number of compared years (2 to max. 6 years), uncertainty estimates might not add much value to the plots. We may present the uncertainties in additional plots in the supplement.

Section 4.1: Throughout this section, the text could be streamlined to avoid repeating the description of trends for each site (see note below about Figs. 2-5). Additionally, the trends in energy fluxes are barely mentioned, leaving the reader wondering why you show these in Fig. 2-5, particularly since their mention focuses on cumulated monthly sums (which aren't shown). Also, yields are discussed frequently throughout the text in this section. Is it worth making a bar chart of yields to more clearly illustrate their evaluation? I realize that a bar chart may look busy, but perhaps averaging across years for the sites with multiple years and including standard deviations will work. Related, how are you calculating yields from CLM simulations? It's important to use the peak daily grain carbon value for the entire growth cycle rather than averaging this over some period of time.

I suggest reorganizing the text (and figures) have 4 paragraphs, focusing on the descriptions of: 1) LAI ; 2) yields; 3) NEE; and 4) energy fluxes. Highlight differences among sites within each paragraph. You can also include an opening paragraph that mentions that CLM_WW improves trends for nearly all variables compared to CLM_D, so the remainder of the discussion focuses on the evaluation of CLM_WW.

> Thanks for your suggestions. We will consider rearranging the text and figures accordingly. Also, we will consider adding a plot similar to Figure R1 below to the main text or the appendix in the revised manuscript.
>
> The simulated crop yield was calculated from the peak value of daily grain carbon.

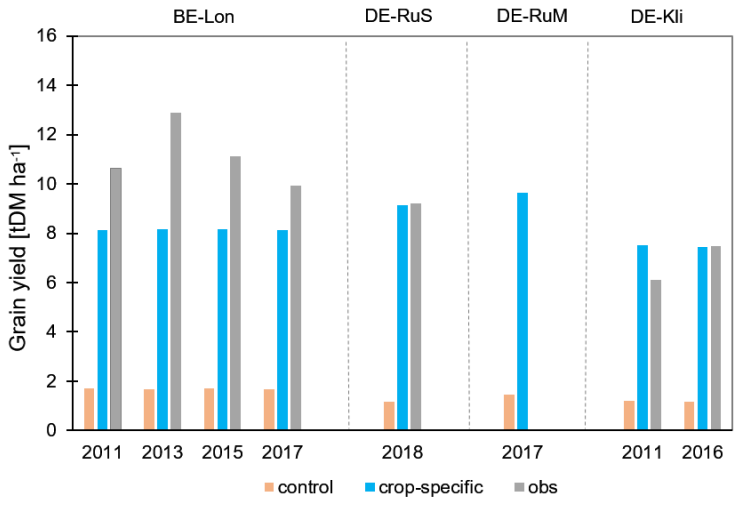

Figure R1: Simulated annual grain yield [tDM/ha] using the default model (orange) and the modified model and crop-specific parameterization (blue), compared to recorded harvest dates (grey) for all simulated winter wheat years at the sites BE-Lon, DE-RuS, DE-RuM and DE-Kli.

Figures 2-5: Is it possible to compile these into a single, multi-panel figure? Given that they all show the same variables for different sites, a single panel would allow the reader to compare across sites more easily. Another, possibly better, alternative is to combine all sites and separate the figures into LAI (Fig. 2) and energy fluxes (LH, SH in Fig. 3). It would also allow you to streamline the description of trends throughout Section 4.1. â˘A˘ TIf I understand the legends correctly, simulations and observations in Figs. 2 and 5 are averaged over multiple years. Can you add uncertainty estimates to these plots? If you plot all individual years (it looks like you possibly do that for observations, but not model), it might be easier to plot averages across years and then plot the uncertainty range associated with interannual variability.

Thanks for your suggestion. We will re-evaluate the plots and try out other options, such as the suggested multi-panel plot. Originally, these results were plotted individually for each site to be able to look at the data in more detail. In addition, we will add estimates of uncertainty to the figures when appropriate.

Fig. 2 states that the observations are GLAI, whereas Figs. 3-5 state that the observations are LAI. Are the observations LAI, GLAI, or does this vary by location? If it is different by location and both LAI and GLAI are used, how might this change the ability to evaluate CLM?

For winter wheat, the green leaf area index (one-sided green leaf area per unit ground surface area) is measured in the field and compared to CLM simulated LAI (defined as one-sided leaf area index, no burying by snow). We will add this in the revised manuscript.

Fig. 5: There aren't any LAI observations plotted in panel a, yet the figure legend suggests that there should be site observation data for LAI.

Unfortunately, there are no LAI observations available for that year. The legend also refers to the two plots below and therefore shows observations, too. We will consider this when rearranging the plots as suggested above.

Lines 394-5: As you state, it looks like the LAI peak is indeed too early. However, even more noticeable (and not mentioned) is the fact that the LAI peak looks to be dramatically underestimated.

Thanks, we will add this comment in the revised manuscript.

Lines 413-4: Table 4 suggests that crops are only harvested _ 1 month too early, but there are higher observed LH fluxes later in the season than just one month. Is this due to cover cropping, which is not included here?

Yes, the later peak in LH fluxes is related to cover cropping. At the Be-Lon site, mustard is often planted after winter wheat which is not represented in this simulations. This issue is discussed in a later section but we will include a short mention in this section as well.

Line 420: I think the phrasing "overestimated early growing season LAI" is potentially misleading. While it is technically correct, the simulated peak LAI values are actually similar to observed peak LAI values, but happen earlier in the year. I think it might be more informative to state that the peak magnitudes are similar, but that the peaks happen too early in the year.

We agree and will rephrase.

Lines 422-3: What does "growing cycle" refer to hear? As you mentioned earlier, LAI peaks too early and planting and harvest start early, suggesting that phenology is not accurate. Therefore, it is unclear what you mean by "generally good correspondence in growing cycle and LAI".

What is meant is the reasonable agreement of modelled and observed LAI magnitudes throughout the growing cycle, which is also reflected in the simulated grain yield, despite shifted planting and harvest dates. We will reformulate this statement to make it more comprehensible.

Lines 437-8: How can you say that CLM_WW resulted in more realistic magnitudes when you stated in the previous sentence that observations aren't available?

Although there are no observations available for the LAI at this site, we can still see that CLM_WW simulated more meaningful magnitudes of LAI compared to default simulations.

Lines 438-9: This is confusing. Does it refer to only the simulations, or also reference the observations? I get the sense that you are conflating simulated peak LAI with simulated and observed crop yields. It implies that lower LAI causes the lower crop yields, although I don't think you can say that for sure.

Thanks for pointing this out. For the phenology algorithm within CLM5, a lower LAI can indicate lower grain yield due to less growth. For observed crop yields, there is a multitude of factors that influences the grain yield, which is not necessarily reflected in the LAI. Here, we see that the lower simulated LAI corresponds to a lower simulated grain yield in comparison to the other sites. However, for the BE-Lon site for example simulated LAI is lower but yield is higher compared to DE-RuS site. This is explainable by an exposure of low temperatures at DE-RuS (temperatures are generally slightly higher at BE-Lon early in the year) that affected the grain carbon due to the frost damage routine that was implemented (section 2.2.1). Damage due to low temperatures affects grain yield more than growth and is therefore not represented in the LAI. We agree that a short discussion on this is missing in the manuscript and we will extend the revised version accordingly.

Line 440: I think this may be backwards. Table 4 suggests that yields are overestimated in 2011 and match really well for 2016.

Thanks, we will correct this.

Lines 453-4: Are all the subsequent mentions (including the metrics in Table 5) calculated using the cumulative monthly sums?

The metrics were calculated using simulation output and observation data at daily time step. We will add this information in the text.

Line 459: You just stated that the BE-Lon sites high some of the highest correlations in the previous sentence, and here single out this site as having high RMSE and biases with low correlations.

Thanks for pointing this out. We also calculated metrics for the whole year instead of only the time between planting and harvest, where low correlations were found for BE-Lon. We do not show these metrics and will consider adding them in the Annex of the revised manuscript. Although energy fluxes match reasonable well during the growing cycle of the crop, the annual correlation is relatively poor due to the influence of cover cropping and poor representation of post-harvest field conditions. We will correct this in the revised manuscript.

Lines 460-461: This sentence should be moved to above, where you briefly mention the mismatch in late-season LH. Also, how does this affect the metrics in Table 5 (see above comments as well).

Thanks for pointing this out. The metrics in Table 5 are calculated for the time between recorded planting and harvest of the crop and thus not affected by this. Please also see our reply to the comment above.

Line 464: Are you referring to CLM_WW? I suggest clarifying here, as you do include simulations that represent cover crops.

Up to this point, all simulations were run with either the default model or the model including the new winter cereal representation. The cover cropping approach is then introduced below. We see that throughout the text the usage of model versions can be confusing and will try to amend this by reorganizing wording and structure as mentioned above.

Lines 471-2: It is not quite accurate to say that NEE observations match better due to improved LAI. Consider changing to: "in part due to the better representation of LAI".

We agree and will change the text accordingly.

Line 473: Are you actually using cumulative monthly values? Fig. 6 show NEE in unites of umol CO2/mˆ2/sˆ1.

We are using average NEE rates in umol $CO_2/m^{-2}/s^{-1}$ for the respective month. We will rephrase in the figures and text for better clarity.

Line 475: Both sites? You mentioned three in the previous sentence. If only referring to two sites, please specify which ones.

We were referring to BE-Lon and DE-RuS.

Section 4.2:
Perhaps this should be titled "New Parameterizations" or "Sugar beet and Potato Parameterizations" to distinguish from the modified winter wheat parameterization in Section 4.1 The evaluation of corn here seems a bit out of place since this section focuses primarily on the new parameterizations. I'm not sure where it goes (perhaps in supplemental?), though. Perhaps this section could be refocused as "Evaluation of other crop types", which includes corn and also the new crop types.

We will think about splitting this into two separate sections (i.e. Evaluation of CLM5 default crop types and Evaluation of new crop-specific parameterization for sugar beet and potatoes), and maybe moving parts of it to the appendix.

Lines 489-91: I suggest rephrasing to add some detail: "The modifications to winter wheat in CLM_WW do not affect other crop types. Therefore, we add new parameterizations for sugar beet and potatoes to this code."

We agree and will change the text accordingly.

Lines 502-4: Is this parameter set modified, or new? What is it strongly improved compared to, if these didn't exist in CLM? I assume it was compared to the default CLM crop model (where the crop might be represented by another type of crop), and it would help to know for sure.

We used the term 'modified' as there is already a 'placeholder' for both sugar beet and potatoes available in the default CLM structure (with parameters adapted from the spring wheat CFT) and we did not have to change or add much to the structure of the model but rather replace single parameters on the input plant parameter file. As mentioned above, we will change wording to improve the clarity of the text.

Line 507: You reference spring wheat here. Is this the crop type that default CLM uses for these sites? If so, you might want to make this clearer (and mention it earlier). For example:

“The default parameterization in CLM uses spring wheat for these crop types and effectively reproduced the growth cycle and seasonal LAI, simulations using the potato and sugar beet parameterizations better captured harvest date and growth cycle. ”

Thanks for your suggestion. We will also clarify this according to your comments above at an earlier stage in the text.

Line 509: As in previous comments, I don’t think “modified” is the best way to describe this. I suggest using “crop-specific parameters” or “parameterizations for new crop types” or similar. As far as I understand, parameters for new crop types were added, not modified.

We agree to use better wording such as ‘crop-specific parameters’.

Lines 510-2: It looks like the latent heat flux is very similar for the other site, which might be worth mentioning.

Thanks for pointing this out. We will mention this fact in the revised manuscript.

Lines 528-30: Performed better for NEE? Please clarify.

Simulations of the NEE using the crop specific parameter set yielded a slightly better correlation of 0.58, compared to the control simulation that resulted in a correlation of 0.43 (Table 6). We will clarify this in the revised manuscript

Figures 8 & 9: I suggest updating the use of “default” and “modified” here based on above comments â˘A˘T Please specify that the LAI results are daily (if they indeed are).
In previous figures, NEE is described as “cumulative monthly”, but here is described as “monthly averaged”. Can these be calculated and referred to in the same way for consistency?

Thanks for pointing this out. The NEE is the average daily NEE per month, not cumulative monthly sums. We will rephrase this in the text and figures accordingly. LAI observations are single field measurements (point observations).

Section 4.3 It seems that this section focuses on crop rotation as much as cover cropping. I suggest updating the heading to “Cover cropping and crop rotation” or similar to reflect this.

Thanks for the suggestion. We agree and will change this in the revised manuscript.

Lines 553-4: Is the simulation of a second crop growth onset for the same crop or for the cover crop? The current wording suggests that a second onset is for the same crop within one year AND for the cover crop. If this isn’t intended, perhaps change to “simulation of a cover crop as a second crop growth onset within a single year”

The focus is set on the second onset within a single year. Both a second onset of the cash crop, as well as the onset of a cover crop are possible. We will rephrase the text for more clarity.

Line 556: “Greening mix” â˘A˘Tis this the same as cover crop, catch crop? Please be consistent in your terminology throughout.

Greening mix here is simply an example for a cover crop that was planted that year at DE-RuS. We used the terms ‘catch crop’ and ‘cover crop’ synonymously and will update this in the text and figures.

Lines 556-557: Perhaps it would be more accurate to say “the cash crop rotation of barley (simulated using the spring wheat CFT)”.

Thanks, we will rephrase this.

Line 557: Spring wheat in CLM is not considered a perennial. It can simulate multiple years of spring wheat in a row, but that doesn’t make it perennial.

Thank you for pointing this out. We did not mean to say that spring wheat is a perennial crop but that it is simulated repeatedly over multiple years. We agree that the use of 'perennial' is wrong in this context and will correct this.

Lines 559-561: Can the effects of planting cover crops and the crop rotation be isolated?

Here, we wanted to show that not only an easier crop rotation is possible (especially from summer to winter crop) but also the simulation of a crop that is not considered a cash crop. Technically, this follows the same schematic.

Line 563: Please change "plantation" to "planting"

Thanks, we will correct this.

Line 576: Similar to above, spring wheat is not a perennial crop in CLM, as it's planted every growing season.

Thank you for pointing this out. Please see our reply to line comment 557 above.

Figures 10-11: It looks like these are for the same site and continuous. Why not plot the full time series on the same panel, adding lines or shading to show the transitions and associated crop type labels. Also, do you not have observational data for LH for 2017-2019 (Fig. 11)?

We will take your suggestion into account and re-evaluate these two plots. Also, we will add a plot for latent heat flux to Figure 11.

Section 5

In addition to the benefits and challenges of the new model developments that you include, I was hoping to see further big-picture discussion, for example about how these new developments might improve future large-scale simulations, possible interactions with climate, etc. Consider adding a paragraph to highlight how your improvements can improve our understanding of larger-scale processes. Also, NEE isn't mentioned at all. Why do you think that NEE didn't improve as much as energy fluxes?

Thanks for your suggestions, we will add more detail to our discussion in the revised manuscript.

Field observations indicate that heterotrophic respiration from soil organic matter and litter acts as a carbon source, which is not simulated well in CLM. This is one of the reasons why the correlation of the NEE is relatively low.

Lines 597-8: As mentioned in a previous comment, higher LAI does not mean higher grain yield. There are many factors that affect yield, including photosynthetic rate, nutrient availability, etc. Also, the results presented in this sentence further support that LAI does not directly correspond to yield: grain yield was higher at BE-Lon (which had lower LAI) than DE-Rus.

Thanks for pointing this out. We agree that LAI is not the only factor governing yield and will change the text accordingly. Please see also our reply to line comment 438-9 above.

Line 603: CLM may not represent different varieties, but the parameters could be changed (as you did here) to represent different varieties, especially when simulated at single points.

In this study, different cultivars were not considered as CLM5 offers only one CFT for winter wheat representing all varieties. The list of CTFs could be extended with suitable plant parameterizations, but this information is not available due to missing measurements and due to the complexity of the phenology algorithm and parameter scheme. The introduction of a phenology scheme based on plant physiological trait information could be a major improvement in this field (e.g. Fisher et al., 2019), as plant trait information becomes more readily available. Whether including different varieties

and cultivars of one crop would have a significant impact on the performance of regional or global scale simulations remains to be evaluated. We will rephrase this accordingly in the revised version of the manuscript.

Line 607: It might be clearer to say "The early leaf onset and harvest for winter wheat simulated by CLM: : :"

We agree, thanks.

Lines 619-22: Can this be more specific? How would discretizing plant hydraulic properties improve yield prediction? Also, why does the reference include "Daniel"? How could the properties (parameters?) be estimated by inverse modeling or data assimilation?

There are many variables that influence CLM simulated yield, e.g. LAI cycle and peak, length of the leaf emergence phase, harvest date and soil water available for root water uptake. Except for soil moisture, these variables are strongly determined by the GDD scheme, which suggests that the simulated crop yield strongly depends on the GDD. Also, the carbon allocation of crops is strongly limited by soil water available to the plant. An improved soil hydrology (i.e. the inter-annual variability of soil moisture) and representation of plant hydrology could help to improve future yield predictions with CLM5. Crop properties could be estimated by assimilating measurement data like NEE, LAI, soil moisture and/or energy fluxes using for example an augmented state-vector approach.

We will add more detail to this part of the discussion in the revised manuscript.

Lines 629-31: Why isn't it applicable to regional simulations? If a simulation is set up to use land use change, the distributions of vegetation, including crop types, will change, even on a point scale, and can be customized by the user if desired.

This tool is useful to study general land use changes by changing the land cover of individual land units, e.g. from naturally vegetated to cropland or urban, from forest to cropland, from C3 to C4 cropping etc. However, it lacks flexibility in accounting for changes within land units of the same land cover and does not account for all 64 CFTs. Also, changes always happen on the 1st of January of the given year. We agree that it is generally applicable to regional scale and will rephrase the text for better comprehensibility.

Line 634: Do you mean before fall of 2018? Fall of 2017 would be the same year.

Yes, we meant before fall of 2018, thanks for pointing this out.

Line 635: I don't see Figure 12.

It should be Figure 11, we will correct this.

Section 6

Line 665: Is higher flexibility for crop rotations possible beyond your study and beyond single point simulations? Because it isn't clear how cover cropping was incorporated in the methods, the applicability of this beyond your study or single point sites isn't clear.

In the revised manuscript we will discuss how large scale simulations could be used to test 'conceptual' cover cropping schemes. For example, the effect of an overall coverage of greening mix during winter months on all crop land units where summer cash crops are planted and that would otherwise be fallow by default during winter. This could also be tested for specific cash crops only. In addition, it would be possible to simulate cover crop plantations based on harvest date thresholds. Here, a defined maximum harvest date for any specific cash crop could define whether a cover crop such as winter wheat would

be planted or not. For example, for all sugar beet land units with harvest dates before a certain threshold (e.g. day 290 of any given year) winter wheat could be planted as a cover crop during winter. If this harvest threshold would not be reached and the summer crop is harvested late in the year, no cover crop would be planted. Alternatively, these harvest thresholds could define the type of cover crop, e.g. early harvest - winter wheat, late harvest – simple greening mix, etc.

Lines 675-8: I appreciate that there are numerous improvements that will improve CLM. However, none of these seem strongly related to the work presented here. For example, there is no evidence that lack of management or incorrect plant hydraulic properties are contributing to model biases.

We agree. Before addressing general points of improvement for the crop module, we will discuss how the modification presented in this study could be further developed and evaluated in order to increase their applicability.