The paper describes and compares three different approaches for grassland yield estimation. Given the limited availability of spatially explicit data on grassland yields and the pressure of global change on agricultural ecosystems, this paper addresses a relevant topic and emphasises the potential of fine-scale monitoring using high-resolution spatio-temporal earth observation and environmental data.

The main novelty of this paper seems to me to be the presentation of a novel remote sensing-based approach that combines modelled AGB time series and mowing events, all derived from free and globally available satellite data. This approach is compared with an established process-based model (LDNDC) and a simple rule-based approach. The paper is well written and clear.

I recommend it for publication but encourage the authors to make some improvements to clarify and increase its contribution to a larger audience. My main criticism is expressed in the first comment, which suggests strengthening the focus and improving the results section.

*Thank you for taking the time to review our manuscript as well as for this positive assessment and critique. It is correct that the remote sensing-approach is novel and we would like to follow your suggestion of putting more emphasize on the remote sensing results along with other important remarks made (more below). We appreciate the clear and constructive comments of the reviewer and think that the suggested changes will improve the manuscript. We answered all comments below in italic while the reviewer remarks are upright.*

**General comments**

1) I think the relevance of this paper could be increased with a stronger focus on the novel RS-based approach and its ability to capture small-scale spatial and temporal variations, with management information being a prerequisite, also for LDNDC. The importance of spatial/temporal variation is highlighted in the abstract and introduction, and its potential to be captured by the RS-based approach is also highlighted in the conclusion under L593/594, but in my opinion, this is not adequately represented in the results. I recommend a more detailed analysis of this.

*Thanks for this assessment. We agree with this and would like to follow your suggestions made below.*

A stronger focus on the RS-based approach would make it clearer what the novelty of this paper is, especially as the other two approaches are being currently under review in a separate paper, if I understood correctly. In this regard, it might be worth shortening the description of LDNDC in the methods and the detailed and rather descriptive comparison with the two other approaches (results and discussion section, e.g. L376-L407), which I think are interchangeable to some extent. This could be complemented with some in-depth analysis of the novel RS-based approach. For instance, I think the currently made statements such as “All three approaches reach plausible results of annual yields of around 4-9 t/ha and show overlapping as well as diverging spatial patterns” are not very meaningful. I'm also not sure how relevant this is for people who don't work with LDNDC, or if it allows general conclusions to be drawn for process-based models?

*Thank you for these suggestions. We agree with the shift of the focus towards the RS approach. Therefore, we will expand the RS-based results and shorten other parts of the manuscript, in particular the method description of LDNDC and some rather descriptive comparisons. The rule-based approach is just a small part of a larger agent-based modelling of ecosystem services in the region. In the meantime, the paper to this ecosystem service modeling approach was published (*https://doi.org/10.1016/j.eja.2025.127539*). However, the focus of this paper is not the yield modelling but the nitrogen fertilizer input, while the modeled yield is rather an intermediate product. The LDNDC yield modelling paper is probably resubmitted soon and not published yet. However, here the focus lies on the effect of drought conditions and a comparison of multiple years, also looking at the results with a different scale. Therefore, we think it is still necessary to present the results of the other two approaches as well. However, we will shorten the mentioned parts of the manuscript to reduce the repetitiveness. We assume that the results of LDNDC can be transferred to other process-based models (e.g. Daycent, APSIM). We will add this to the discussion.*

Can the authors show (map, stats) that small-scale variations are better captured with the RS-based approach (e.g. looking at adjacent in-situ measurements / within and across parcel variations)? Same for the temporal variation, e.g. which approach better captures AGB of the first cut (see statement L26/L470; it's not clear to me, was it compared with in-situ measurements?)? I think these points are rather well covered in the discussion but could be improved in the results section. Also consider changing the title to give more weight to the RS-based approach and clearly mention the focus in the introduction.

*Thank you for this important remark. We highly favor the suggestions made to highlight small-scale variations captured by the RS-based approach. It is indeed an advantage and not presented well enough in the current version of the manuscript. All three approaches result in yields per parcel. However, the RS-based approach considers within-parcel variations as biomasses per pixel are calculated. Therefore, we would like to add a map showing the estimated biomass for several time steps to highlight this. We will add the LDNDC biomass in a numeric way as the data is only available per parcel and we are not allowed to show parcel boundaries due to data restrictions. In addition, we will add the adjacent in-situ measurements in Figure 3 to enable a temporal comparison of results as well. We are open to change the title of the manuscript and suggest "Grassland yield estimations – potentials and limitations of remote sensing in comparison to process-based modelling and field measurements".*

If the authors choose to keep their focus on comparing the three approaches, more in-depth analysis is needed, e.g. to clearly understand why one model over/underestimates yield and under what conditions. This should go beyond "simply" describing differences at hexagon-level and frequency distributions in the result section.

*We decided to follow the suggestion of moving the focus more towards the RS-based results and agree that otherwise a more in-depth analysis of the comparison would be needed.*

In the first chapter of the results, please provide more details about the RS-based model, e.g. prediction vs. observed plot.

*Thank you for this suggestion. As suggested we will extend the first chapter of the results with more detailed results of the RS-based approach. A prediction vs. observed plot is a well-suited option to do so. We will include it in the revised version of the manuscript as well as*

*further more detailed results, such as a map of temporal biomass estimates showing within-parcel variability as described above.*

2) More information regarding the in-situ measurements is needed (section biomass field data). In my opinion it is not sufficient to simply provide a reference, at least an overview of the sampling design is needed, addressing the nested design, the spatial and temporal distribution, and the number of samples per mowing frequency.

*Thank you for this remark as well as the specific and helpful suggestions. We agree that the description of the in-situ measurements was relatively short. In an updated version of the manuscript we will extend the description, including the mentioned information.*

3) The results focus on showing the differences between the three approaches (map, histogram). However, these are difficult to interpret in terms of which approach performs better. A more detailed comparison/analysis with in-situ measurements would give more substance to this paper. However, I'm not sure if the in-situ data allows such further analysis? If so, this could help to address questions raised under comment #1, is linked to specific comments bellow and the potentially missing in-situ symbols in figure 3.

*Thank you for this comment. It is correct that the in-situ data is limited and might not allow for a reliable comparative analysis. However, the in-situ measurement symbols in Figure 3 were indeed missing, we apologize for this mistake. We hope that adding the symbols there will improve the interpretability of the plausibility of the approaches and their differences.*

4) The comparison of the RS-based model with LDNDC is a bit difficult to interpret, as both approaches rely on Sentinel-2-based mowing events. Please clarify what this means for the interpretation, e.g. what would be the performance of LDNDC without spatially explicit mowing information (or is there another data source for this)?

*Thank you for this remark. It is correct that all three of the approaches rely on the same mowing dates dataset. This might make it harder to differentiate the results but also provides the same baseline for the three approaches. All three approaches rely on the mowing dates and don't work without. For the LDNDC model, different solutions without a RS-based mowing dates dataset exist, for example using simulated mowing dates depending on standing biomass as described in Krischan et al. (2021) (*https://doi.org/10.1016/j.eja.2021.126306*). However, the missing mowing dates were usually seen as a disadvantage of models such as LDNDC in the past. We will add this aspect in the discussion to make the reader more aware of the circumstances.*

As both models are highly dependent on the RS-based mowing dates, the finding that the spatial patterns or box plots (Figure 8) between the two approaches are similar and strongly related to the number of mowing events is somewhat redundant/obvious (e.g. L418, L467). I don't know how this could be resolved, perhaps by running separate models with/without mowing events?

*Thank you for this comment. We agree that it seems straightforward that the RS and LDNDC approaches are strongly related to the mowing events. However, with Figure 8 we would like to show how the influence of the mowing frequency might differ between all three yield estimation approaches. For example, we found that the RS-based approach shows lower yields for low intensity grasslands and higher values for high intensity ones. We will review the text to make sure that the results are formulated in a clear manner. The difficulty with*

*running separate models, either with or without mowing data is, that the models need the mowing dates as input as there are no yields without harvesting. It is therefore not possible to use these approaches without mowing information.*

5) Please review figure and table captions to be more self-explaining and consistent (e.g. Figure 4 vs. Figure 5 vs. Figure 6; rather mention the different modelling approach like in Fig. 4) and use consistent terms in the whole document (e.g. LandscapeDNDC vs. LDNDC).

*Thank you for these remarks. We will adjust the figure and table captions to make them more informative and consistent and will revise the manuscript to use consistent terms in the entire document.*

**Specific comments**

L21: Change "present" with "compare", or clarify which approach is new (RS-based) and which already exist / will be published separately?

*Thanks, we will clarify that.*

L140 /Figure 1: Change to "CLCplus Backbone 2021", consider citation, remove not-shown legend items. Make clear that the study area is the Ammer catchment and add catchment border to sub-figures.

*We will follow the suggestions and make the according changes.*

L155: Was the validation done separately for this study or is the F1-score derived from Reinermann et al.? More details or citation is needed. Please check that the uncertainty of mowing detection is adequately addressed in the discussion section, as this is the most important parameter.

*Thank you for highlighting this. It is the same validation approach and based on the same data but in fact a separate validation as only the subset of data from the Ammer region was used. We will add this information in the methods and check that the uncertainty of the mowing detection is adequately discussed.*

L166: Add plot locations to the map in figure 1 (maybe top-left)?

*That is a good idea. We will do so.*

L170/171: More information regarding the sampling design is needed, I cannot quite follow how the n=111 samples are reached

*Thanks for the remark. We will revise the description of the in-situ data.*

L178 (Figure 2): Add validation procedure?

*We will think about that; however, we are not sure if it fits well.*

L180: Title not clear, consider "Remote sensing-based approach"

*Thanks for pointing us to that. We will revise the titles of the sub-headlines.*

L195: Please check EVI formula for correctness (position of factor 2.5 and apostrophe after 1)

*Thank you for highlighting this. It is correct, that the factor 2.5 should be positioned in front of the fraction line. The apostrophe is a comma after the formula. We will try make sure during the editing that this doesn't look confusing.*

L205: Please cite the HR-VPP (https://doi.org/10.2909/c1c46cb2-b02b-4013-aae5-a54a8c018b1e)

*Thank you for the source. We will cite it.*

L220: It would be helpful to get some information about the distribution of these data pairs, e.g. a figure with y=field plots, x=dates S2&field, colour=S2/field/pairs(train/test), table (supplementary) or at least providing some more information in the text (how many field plots etc.)?

*Thank you for the comment. We will add information to the data pairs, either as a table or as additional text.*

L244: Introduce r2 instead of L247; use PRMSE like in L248

*We will revise the sentences and introduce r2 when it is mentioned first and correct the abbreviations which seem to be confused here.*

L282: It sounds like the RVA is based on an official table, but then it gets more complicated with the tables and the link to Kaim et al. (under review). It is a bit confusing, please check/revise.

*We will revise the RVA description in the revised manuscript.*

L286: Do you mean "Table A1"?

*Yes, thanks for pointing to this. We will change "Appendix A1" to "Table A1".*

L311: Improve caption. Maybe move table to appendix, to make clear that it belongs to Table A1.

*We will advise the caption and think about moving it to the Appendix as you are right that it belongs to Table A1 and might be difficult to interpret without Table A1 anyways.*

L335 (Figure 3): I cannot find the in-situ measurements! Provide annual values (which are a bit hidden) as stacked-bar adding AGB from the individual events?

*Thank you for this important remark. The in-situ measurement points are missing from the Figure. We apologize for this mistake. We will add them in the revised version and will try out to add stacked bars of the event AGBs and see if the figure remains clear.*

L342: 0.97 seems to indicate overfitting and might need to be discussed later?

*It could indeed indicate overfitting. We will add some discussion to this.*

L389 (Figure 6): I suggest adding r2 and variance to the plots. Same for Figure 7.

*Thanks for the suggestion, we will add this to the figures.*

L433: Remove one bracket

*Thanks, we will do so.*

L518: You could mention the potential to tackle cloud coverage with SAR

*Yes, that is right. We will mention this here in the revised manuscript.*

L550: In this study LDNDC also depends on cloud-free observations for the mowing dates

*That is correct. We will mention this as well.*