

Authors' response to the second anonymous referee

We thank the anonymous referee for his/her review and all the relevant remarks and questions that will help us to further improve this manuscript.

(1) Lines 40-45: while I have experience is 'classical' machine learning (Random forest) or data assimilation, deep learning is something new for me. Please state briefly (and explicitly) what the difference is between deep learning and more classical machine learning approaches.

RESPONSE 2.01: Thank you for pointing out this lack of description. The term 'deep learning' refers to the fact that we are using a neural network with several layers. Training over a large database can be either unsupervised or semi-supervised. Here, we use supervised training, since we feed the neural network labelled data.

(2) Line 46: if near-real time is the concern, you could have used near-daily SIF data. What is the advantage of using the 8-day TROPOMI? Please clarify.

RESPONSE 2.02: In this study, we do not use the 8-day TROPOMI data, we use daily TROPOMI SIF data instead.

(3) A thought: the TROPOMI SIF data have never been properly calibrated/validated. They also contain quite some negative values, so it is likely that some of its values are biased. The traditional answer here is that the variability in SIF contains most of the information, as opposed to the absolute SIF value.

RESPONSE 2.03: Thank you for sharing this with us. The negative values are definitely an issue. We chose to remove them before creating the database and assimilating the observations. This also biases the product to some extent, since it was initially statistically calibrated to be unbiased over bare soil areas. We have also decided to keep this strictly positive rule for SIF in the observation operator design. As you said, we believe that enough information in the SIF is still important to be used anyway. Another approach would be to use the product with negative values and then train the neural network to match the zero bias over bare soil (allowing it to take on negative values).

(4) Line 97-100: the airborne campaign is not mentioned in the introduction. How important is it for the research question? The 'why an airborne campaign' needs more context.

RESPONSE 2.04: The HyPlant airborne campaign is here to provide context for the discussion of the limitations of observation operators. This part will be moved to a supplement.

(5) Line 106: the introduction (line 46) mentions an 8-day SIF product, while this line mentions a daily SIF product. Please clarify.

RESPONSE 2.05: In this study, we do not use the 8-day TROPOMI data, we use daily TROPOMI SIF data instead.

(6) General thought: section 2.2 within methodology would be easier to understand if there were some kind of visual. Figure 3 is helping a bit, but it focusses on the neural network itself, rather than on the assimilation.*

RESPONSE 2.06: Thank you for your comment. We will include a figure presenting the assimilation.

(7) Line 193: why having two LAI datasets? LAI and LAIv1?

RESPONSE 2.07: One LAI is an observation product (LAI-V1) and the other is the LAI produced by ISBA model. This part will be moved to a supplement.

(8) Line 195: when doing machine learning for SIF, I often note that not DOY, but $\sin(\text{doy}/365*2\pi)$ and $\cos(\text{doy}/365*2\pi)$ are more potent predictors compared to doy itself. What's the reasoning for not using these?

RESPONSE 2.08: Cycling the DOY metadata predictor could actually improve the NN training. More research is needed to address this question.

(9) Line 208: year missing in reference.

RESPONSE 2.09: Thank you for pointing this out, we will correct as: Lundberg and Lee (2017).

(10) Line 228: leave out 'as mentioned earlier in this study'. Be either specific on where it was mentioned or leave it out altogether (I recommend the latter)

RESPONSE 2.10: Yes. We propose to replace:

"As mention earlier in this study, this area comprises heavily irrigated croplands which are usually not accurately simulated by the land surface model"

by

"This area comprises heavily irrigated croplands which are usually not accurately simulated by the land surface model".

(11) Section 2.3.3 the term 'learning process' does not properly cover the content of this paragraph. Consider something like 'feature selection' or 'training variable identification'. My main point is, 2.3.3. does not describe the neural network itself, but rather the selection of the variables and the train-test split.

RESPONSE 2.11: Yes. We will replace the title of Section 2.3.3 (Learning process and data split) by "Predictor identification".

(12) Section 2.4: role of airplane data in the assimilation is not clear here. The methodology just mentions the existence of the dataset, but not its purpose within the study.

RESPONSE 2.12: The description of the airborne campaign was included here to bring together all the SIF datasets used in the same section. However, it is not referenced until the discussion section. We will move this description to a supplement.

(13) Line 236: a detail, but I am struck the use of the future tense here. Why would you 'will run'. Haven't you done it? I presume you mean it's a continuous (and therefore unfinished, but already initiated) process. The future tense made me think the process wasn't even started yet.

RESPONSE 2.13: We will remove the future tense when it is irrelevant.

(14) Line 239: repeat what precisely you mean by the ‘baseline experiment’.

RESPONSE 2.14: Thank you for pointing out this missing, we will repeat that the baseline experiment corresponds to “LAI-V1” (Table 1).

(15-16) Line 243-244: again a future tense. Please don’t

RESPONSE 2.15: We will change the tense in both case, thank you for pointing this.

(17) Line 256: I agree on the good agreement, but the word ‘good agreement’ looks sloppy. Just write a metric here.

RESPONSE 2.16: Yes. We propose to replace:

“As shown in Table 2, the values based on the training and test datasets are in good agreement.”
by

“As can be seen in Table 2, the skill score values obtained for the training and test datasets are similar.”

Also, we will improve the readability of Table 2.

(18) Line 260: I agree with this interpretation but Huber loss is a bit of an esoteric term and I want to encourage you to make this paper readable for a larger audience than the people that are familiar with both SIF and Huber loss. Explain it a bit more in conceptual terms.

RESPONSE 2.17: Yes, we will define the Huber loss in the Methods section.

(19) Line 263: how do you determine a bias? What is the reference here? The difference between the TROPOSIF measurement and the modelled SIF? Or difference with the airplane?

RESPONSE 2.18: As written in Table 2, the bias we are referring to is the mean difference between NN estimates of SIF and observed TROPOMI SIF. We will include it in the main body of the article. Thank you for pointing out its absence.

(20) Figure 5: the non-existent correlation above the 60th parallel is striking. Why would that be? Presence of snow?

RESPONSE 2.19: It is mainly due to the lack of observations during the winter at these latitudes.

(21) Figure 5: I noticed you cut out mountainous regions. Why?

RESPONSE 2.20: To avoid any frozen soil where the assimilation cannot be done.

(22) Conclusion: really strong, but it ends on a false note. What’s stopping you from adding additional variables (you have e.g., temperature maps to your disposal) or you can tune the hyperparameters.

RESPONSE 2.21: Yes; We propose replacing:

“This study assesses the benefits of incorporating TROPOMI SIF products into a land data assimilation system via a deep learning operator, which opens up new possibilities for systematically assimilating such products. As the architecture and training of the neural network

operator are not overly specific to the TROPOMI SIF product, it would be interesting to adapt it for the assimilation of other SIF products. Better optimisation of the hyperparameters or the use of additional variables could enhance the neural network's accuracy.”

by

“We demonstrated the advantages of integrating TROPOMI SIF products into a land data assimilation system using a deep learning operator. This opens up new possibilities for the assimilation of new SIF products. Although the neural network operator was designed for TROPOMI SIF, it can be adapted for other SIF products. We also addressed the co-assimilation of SIF with LAI, which could be extended to other satellite products, such as land surface temperature.”