## REVIEWER 2

**General comments: This manuscript presents a neural-network-based observation operator to enable the direct assimilation of passive microwave brightness temperatures from SMAP (L-band) and AMSR2 (X-band) within LDAS-Monde, with the aim of improving soil moisture and leaf area index (LAI) estimates. The approach links model states and auxiliary predictors to observed TB, and evaluates the impact of assimilating each sensor (separately) over cropland-dominated regions at the global scale. Results show reductions in TB departures and some improvements in LAI—more consistent at the global scale than regionally—while the impact on soil moisture and the physical consistency of the framework are less clear. In addition, I found it difficult to reconcile the relatively weak sensitivity of the NN to LAI (compared to SWI) with the magnitude of the resulting LAI increments, which raises further questions about how information is being propagated through the system. I have some reservations about the physical consistency of the observation operator as currently presented. The use of observed LAI, along with static predictors (e.g. latitude/longitude) and even forcing variables within the NN, makes it less clear that this is a clean mapping from model state to observation. In particular, it raises questions about how the operator remains dynamically consistent with the model state, and what exactly the NN is learning (physical relationships vs. geographical or climatological patterns). Clarifying the role of these predictors, and why they are required, would strengthen confidence in the approach. More generally, I found parts of the methodology difficult to follow, with some key elements either not clearly described or introduced relatively late. In particular, it is not always clear whether observed or modeled LAI is used within the operator, how anomalies are defined, what exactly is included in the control vector, and how the assimilation timing is implemented. These are all important for interpretation and reproducibility, and would benefit from a clearer and more explicit description earlier in the manuscript. Finally, I am not fully convinced by the experimental design and its interpretation. The rationale for assimilating the different sensors separately is not entirely clear, and it would be useful to understand whether their information content is complementary when used together. Similarly, the role of static and forcing predictors, and the apparent mismatch between NN sensitivities and the resulting analysis increments, are not fully explained. Addressing these points would help clarify how the system is behaving and how robust the conclusions are. Overall, I find the topic timely and the approach promising, and I appreciate the effort involved in developing and implementing the NN-based operator within LDAS-Monde. However, the concerns outlined above—particularly regarding the physical consistency of the observation operator, the clarity of the methodological description, and aspects of the experimental design—make it difficult to fully interpret and trust the results as currently presented. Addressing these points would significantly strengthen the manuscript.**

### RESPONSE 2.0:

Thank you for the review and all your suggestions. We propose to expand the section about the predictors, as requested by both reviews. Observation operators can generally be designed using other than model variables as input. Since the goal is to match the observations, potentially biased model variables (like LAI in our case) should not be used in the training or be replaced if possible. Our LDAS is not fully coupled with the atmosphere and the forcing offline. The DA experiments show that the addition of atmospheric forcing variables actually improves the system. The DA experiment without atmospheric forcing variables is also presented. In the revised version of the paper, we will elaborate more on this to make that clear. After implementing all the suggestions, the text should be easier to follow. Instead of model LAI, we use observed LAI for training the observation operator in a first phase. In the data assimilation phase, model LAI is used as input to the observation operator.

We fully agree that it could be interesting to assimilate both AMSR2 and SMAP sensors simultaneously, however, to understand their impact on the DA system, it is essential to study them separately before putting together two not well understood systems.
The apparent mismatch between the sensitivities of the NN to the different predictors and the resulting analysis increments is indeed something to discuss more. In DA, large increments can cause the model to rebound in the subsequent time step. In any case, the physics must be consistent. This is not necessarily always the case after the update step.

**COMMENT 2.1: L-1 "A neural-network based forward operator for the assimilation of microwave satellite observations with LDAS-Monde" The title is somewhat generic and does not fully reflect the scope of the study. In particular, the focus on LAI and soil moisture is not evident, despite being central to the results. It would be helpful to revise the title to better highlight the scientific objectives and outcomes of the work, rather than focusing primarily on the methodological aspect.**

**RESPONSE 2.1:**

We agree, we propose: “Analysis of global cropland leaf area index and soil moisture using a neural network-based forward operator for microwave satellite observations.” for the title.

**COMMENT 2.2: L1 Abstract. The abstract would benefit from a more balanced description of the results and limitations. In particular, the improvements appear more consistent for LAI than for soil moisture, and vary depending on spatial scale, which is not clearly reflected. It would also be helpful to more explicitly acknowledge key aspects of the experimental setup (e.g. cropland focus), as these define the scope of the conclusions.**

**RESPONSE 2.2:**

We propose to change L. 3 to:
---
[…] we assimilate the observations of the satellite microwave sensors SMAP and AMSR2 with our land-data assimilation system (LDAS) on global croplands.

---
and add the following in L. 13:
---
In agricultural regions of the world that are particularly susceptible to drought, AMSR2 experiments outperform SMAP experiments in terms of LAI, in most cases. We also investigated the impact on GPP and SWI anomalies and found relatively little impact with mixed results.
---

**COMMENT 2.3: L23 “observing in near-real time and providing gridded data” Not sure what “gridded” adds here — this could be clarified or removed.**

**RESPONSE 2.3:**

Gridded data is more straightforward to work on in general, which can have some importance for users. Since this is not the place to elaborate on that, we will remove this word.

**COMMENT 2.4: L53 “assimilating land-surface variables using SMAP and AMSR2**

**observations” maybe “assimilating land-surface variables using both SMAP and AMSR2 observations”?**

RESPONSE 2.4:

We agree, we propose the following:
---
To our knowledge, this is the first time that an observation operator based on a neural network has been used to assimilate SMAP or AMSR2 observations over land.
---

**COMMENT 2.5: L67 “interpolated linearly onto the model grid” This seems a little basic — should this not be part of the observation operator (ie its model states that are interpolated, not observations), rather than just preprocessing? Could be clarified.**

RESPONSE 2.5:

This could be part of an observation operator indeed. The observational data that we interpolate here is additionally needed to train the neural network on. Since downscaling is not in the scope of our training strategy, we need input and output data to be at the same resolution. We will add a sentence to clarify this:
---
The data is linearly interpolated onto the 0.25° model grid to provide a common basis for training the NN and for assimilation.
---

**COMMENT 2.6: L67 What is the resolution of the model grid?**

RESPONSE 2.6:

Indeed, this information appears only later in the manuscript, we will add this information, see Response 2.5.

**COMMENT 2.7: L95 “converted into anomalies by subtracting the mean from the actual value for each data set and scaling with the standard deviation.” Mean over what exactly (grid cell? time period?) — this should be specified.**

RESPONSE 2.7:

We agree, we propose the following:
---
the data are converted into anomalies by subtracting the temporal mean of the SWI from the actual value for each grid point and data set, and then scaling by the standard deviation over the entire verification period (2021).
---

**COMMENT 2.8: L106 “The LAI is computed once per day.” So LAI is a fully prognostic variable?**

RESPONSE 2.8:

Yes, we will add this:

---
The prognostic LAI is calculated daily.
---

**COMMENT 2.9: L106 “interaction between CO2 and photosynthesis with regard to stomatal respiration and plant growth” “stomatal respiration” seems like an unusual description — could this be clarified?**

**RESPONSE 2.9:**

We agree, we propose the following:
---
[...] with regard to stomatal conductance and plant growth.
---

**COMMENT 2.10: L122 “control variable xj (here LAI and soil moistures of layers 2-7, i.e., 0.01- 1 m depth)” It would help to state more explicitly which model variables are included in the control vector, and why they were chosen. Should surface/soil temperature also be included?**

**RESPONSE 2.10:**

We propose to adapt the beginning of the paragraph to the following:
---
In previous studies (e.g. Albergel et al., 2017), the control vector was optimised in our LDAS and contains LAI and soil moisture data from seven layers between 0.01 and 1 m in depth (SM2-8). The control vector variables evolve over a 24-hour assimilation window.
---

**COMMENT 2.11: L133 “croplands (Corchia et al., 2023; Shan et al., 2024), we use exclusively grid cells that are dominated by crops.” The focus on croplands should be introduced earlier, as it has important implications for the interpretation of the results.**

**RESPONSE 2.11:**

We agree, we propose to move that sentence to the introduction (L.50):
---
As previous studies have shown that the main impact of assimilating vegetation-sensitive observations is on croplands (Corchia et al., 2023; Shan et al., 2024), we only use grid cells dominated by crops.
---
and replace it in L.133 with:
---
As previously mentioned, our focus is on croplands due to the higher sensitivity of microwave observation assimilation found in previous studies.
---

**COMMENT 2.12: L160 and Table 1 “variable model fields as predictors” It would be helpful to clarify how this list of predictors was selected.**

**RESPONSE 2.12:**

We agree. To also address the remark of the other referee, we propose the following addition:
---
The selection of input parameters was not only motivated by potentially correlated variables but also by what is easily accessible for the training/testing period. We tested variables of deeper soil layers as well in the beginning but since they did not seem to have an impact, we omitted them for the systematic tests. We will elaborate on this.
We propose the following addition:
---
For the input fields, we employ observed LAI (Sect. 2.3) as well as different static (e.g., coordinates, physiographic data) and variable model fields as predictors and analyse their contribution to the final training performance. Since the ISBA LAI is known to be biased, we expect more accurate information from observed LAI.
As model fields, we select soil moisture and temperature of the soil layer (0.01-0.04 m), which corresponds to the second layer in the ISBA model. Deeper layers can be used as well in principle, but tests did not show any benefits, which is likely due to the strong cross-correlation of the different layers, i.e. no additional information is brought into the system. The top most layer (0-0.01 m) is subject to rapid changes determined by atmospheric forcing and therefore, generally considered challenging in the context of assimilation (Draper et al. 2009, Vural et al. 2021). Unlike X-band measurements, such as those taken by AMSR2, which have a shallower penetration depth, observations at longer wavelengths, such as the L-band (SMAP), are more sensitive to the subsurface soil layer (Escorihuela et al. 2010; Enthekabi et al. 2010, Yee at al. 2017). Passive observations are highly dependent on surface altitude and surface roughness at different scales (Mätzler and Standley 2000, Njoku et al. 2006, Monerris et al. 2008). Consequently, terrain altitude and topographic complexity are expected to be important factors in the accurate simulation of Tb. Depending on the wavelength, microwave observations are more or less affected by the atmosphere. Thus, atmospheric effects cannot be completely neglected when retrieving surface variables from microwave emissions, especially at the X band. To account for this, we also examine the effect of using relevant atmospheric forcing variables (see Sect. 3.1) as predictors, such as air temperature and humidity (Tair, Qair), long-wave downwelling radiation (LWdown) and shortwave downwelling direct and scattered radiation (SWdir, SWsca). We are aware that other variables might contribute as well to the Tb. At one point, too many predictors might increase the probability of overfitting though. As a consequence, we try to keep the number of predictors small and, additionally, test the systematic omission of static predictors.
---

**COMMENT 2.13: L172 "In addition to the predictors that have already been employed in similar studies, we achieved a significant improvement of the model equivalent when using forcing data." Clarify exactly what you mean by "forcing data" here.**

**RESPONSE 2.13:**

We propose to change this sentence to the following:
---
[…] we achieved a significant improvement to the model equivalent by using the variables employed for the atmospheric forcing within LDAS-Monde (cf. Sect. 3.1).

---

**COMMENT 2.14: L175 I'm not sure about having latitude and longitude as static predictors. What do they really represent?**

**RESPONSE 2.14:**

To elaborate further on this, we propose to add in L.178:
---
We suspect that the coordinates, especially latitude, provide information on fine-scale regional patterns that the other predictors do not capture sufficiently.
---

**COMMENT 2.15: L185 "strong benefit seen from adding the forcing variables" I assume this refers to meteorological forcing used to drive the model? These seem unusual to include in an observation operator, as they break the mapping from model state to observation, and would benefit from further justification.**

**RESPONSE 2.15:**

Yes, meteorological forcing as introduced in 3.1, we already added a bit of clarification as answer to Comment 2.13. Observation operators can be constructed using other than the mere model variables (see Garrigues et al. 2026). Especially here, the actual observation operator is just the application of the weights found for the NN.

**COMMENT 2.16: L. 213 "observation operator described in Sect. 4" The observation operator is not described clearly enough here — in particular, it is not clear whether observed or modeled LAI is used. It seems the NN is trained with observed LAI?**

**RESPONSE 2.16:**

We use observed LAI for training the observation operator in a first phase. In the data assimilation phase, model LAI is used as input to the observation operator.

**COMMENT 2.17: L218 "LDAS-Monde is set up to work on a 24 h assimilation window and the observations are assimilated at a given hour. To account for the diurnal variability of the observations, we ran four separate assimilation cycles using 6-hour time bins around each of 0300UTC, 0900UTC, 1500UTC, and 2100UTC, which correspond to the model output times in our setup." Don't really understand what you have done here – clarify.**

**RESPONSE 2.17:**

We propose to rephrase as follows:
---
LDAS-Monde can only run on a 24-hour assimilation window, with observations assimilated at a specific time each day. This means that several assimilation cycles must be run to cover the full day if variables with a diurnal cycle are to be fully exploited. To account for this, we ran four separate assimilation cycles at 03:00 UTC, 09:00 UTC, 15:00 UTC and 21:00 UTC. The assimilated observations were binned into six-hour time bins around these times. This strategy strikes a balance between the time efficiency of the assimilation experiments and the smoothing of the signal over one day due to averaging.
---

**COMMENT 2.18: L263 Figure 8 is difficult to read in its current form — in particular, SWI2 is indecipherable.**

**RESPONSE 2.18:**

We agree, it’s quite difficult to make the overlapping lines be distinguishable, especially with the required colorblind-friendly colormap. Stretching the figure helps a bit. We also implement the proposition from the other referee to split into Northern and Southern hemisphere.

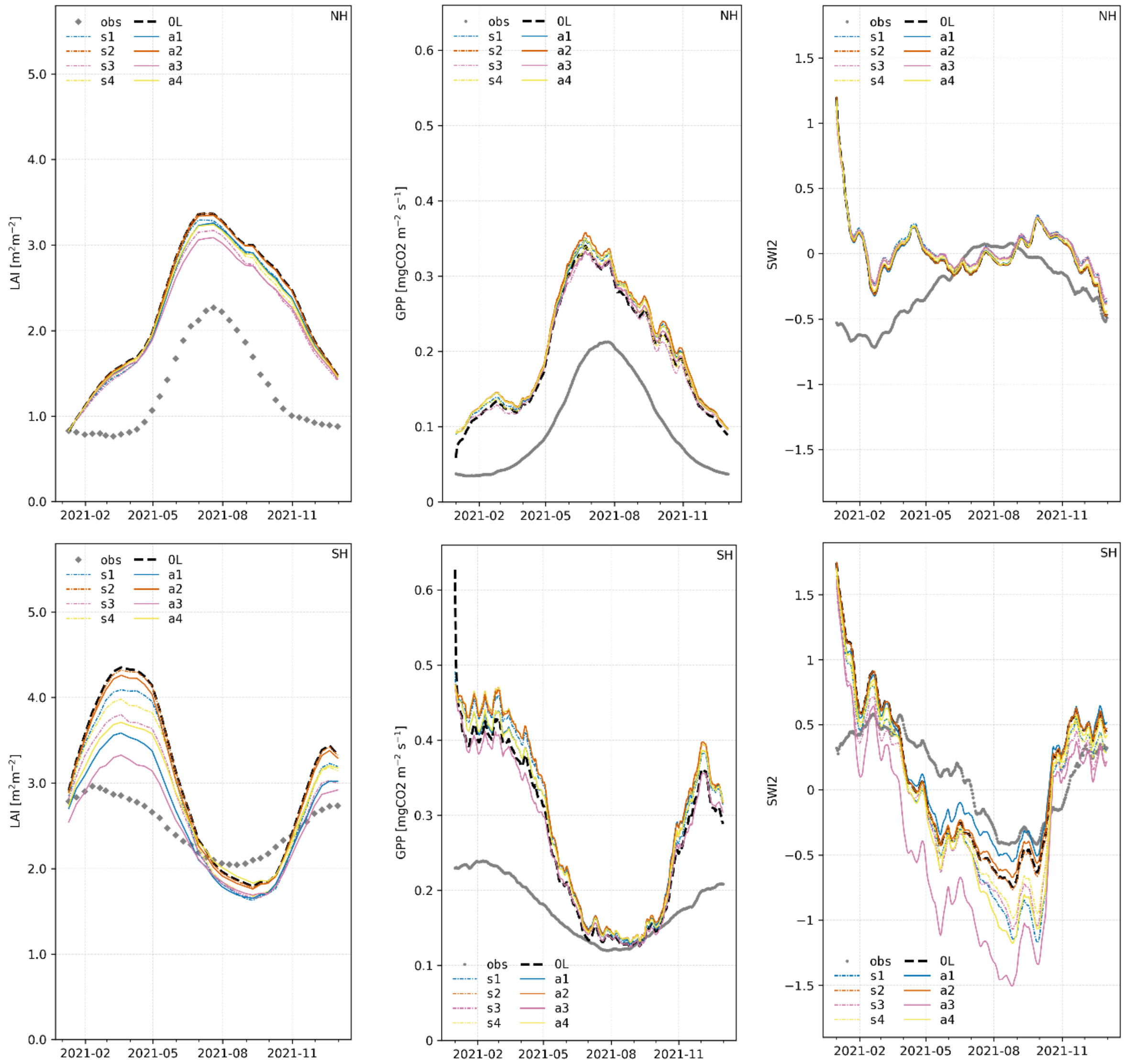


**Figure R2.1:** Comparison of (from left to right) time series of LAI, GPP, and scaled SWI anomalies averaged over global croplands for northern (NH; top row) and southern hemisphere (SH; bottom row). Observations (grey points), OL (black dashed line), and DA experiments (SMAP: dash-dotted lines, AMSR2: solid lines) are shown in comparison.

**COMMENT 2.19: L315 “static predictors are essential for our data sets” This would benefit from further elaboration — in particular, why a coordinate-independent NN is not sufficient in this case. Why is a coordinate-independent NN seemingly not possible with this approach?**

**RESPONSE 2.19:**

We propose to add this:

---

In Section 4.2, we demonstrated the importance of static predictors for our data sets, particularly for AMSR2 compared to SMAP (see Fig. 4). As mentioned, static predictors may help to classify information obtained from variable predictors more effectively depending on the exact location.

---

**COMMENT 2.20: The manuscript would benefit from some minor improvements in terminology consistency (e.g. “forward operator” vs. “observation operator”, TB vs brightness temperature), clarification of certain terms (e.g. “model equivalent”), and overall readability (sentence length, figure referencing)**

**RESPONSE 2.20:**

Where appropriate, we will replace 'forward operator' with 'observation operator' and 'brightness temperature' with 'Tb'. 'Model equivalent' (L. 30) corresponds to the model variables that are transformed in the observation space. We will strive to improve the text's overall readability.