Reply to Comment by Anonymous Referee #1

April 3, 2023

We thank the anonymous reviewer for the very thorough and well put comment. We will start by addressing the general comments for the larger issues, and afterwards address the particular comments.

In the following, review comments will be styled in *blue italics*, our responses in black normal text, and proposed changes in green.

1 General comments

1.1 "Logic issue"

Logic: the Authors recommend dropping their most efficient bias-blind LAI DA option using specious theoretical arguments.

The main novelty of our study is the impact of bias-aware DA methods for LAI DA and an assessment under which circumstances this might be helpful. We therefore focused mainly on pointing these out in the conclusions.

However, based on this and the other review comment we see that this can be easily misunderstood. We therefore propose to include the following sentence in the conclusion after l.549:

The bias-blind DA is most effective at reducing the disagreement in modelled and observed LAI, and leads to the largest improvements in GPP and runoff. For many applications it is therefore a suitable option, especially if large bias reductions are intended, even if bias-blind Kalman filtering is suboptimal. A temporal interpolation of the observation data, or even a direct insertion approach, could be even more efficient in this case. However, this approach does not necessarily improve other variables, e.g. if the model simulates biased LAI in conjunction with unbiased soil moisture. As alternative, we recommend to use observation rescaling techniques for LAI DA with Noah-MP if there are strong biases and if

- ... (same as in original manuscript)

Additionally, we will reformulate the last paragraph of the conclusion as follows:

A drawback of the observation rescaling approaches is that they result in estimates in the model climatology. If the observation-forecast bias is due to erroneous precipitation forcing or missing irrigation input, joint updates of LAI and RZSM in a bias-blind system can be considered instead. This might lead to large bias corrections while still retaining a stable model state even after large

updates. However, if the bias is not only caused by bias in the precipitation/irrigation, this poses the risk of seriously degrading the soil moisture estimates.

Alternatively, updates to model parameters, either via joint parameter and state update DA, or via a priori model calibration can also lead to more stable and persistent updates and LAI estimates in the observational climatology. This is especially desirable for research on the carbon cycle, where absolute values of carbon fluxes are required. Parameters to consider for calibration are parameters related to model leaf growth, but potentially also photosynthesis or soil parameters.

To gain the most benefit from LAI data assimilation into Noah-MP, further research and model improvement of the coupling mechanisms between water and carbon cycle is necessary.

1.2 "Scope issue"

Scope: both Abstract and Conclusion sections present results that are specific to the Noah-MP model and DA framework as if they were valid for all models and DA tools. Much different conclusions could have been obtained using other tools, and probably more logical conclusions. The title is also too general.

We agree that different conclusions might have been obtained with different models or different DA systems. We mentioned this in the discussion section (l. 462-465 or the more detailed discussion of shortcomings of the Noah-MP vegetation model in subsection 4.4 of the manuscript). In a revised version, we will provide a more detailed discussion of the Noah-MP model (see below). This will make it easier to transfer our results to other models or datasets, because it makes the underlying reasons in the model more transparent.

Additionally, we will adapt the title and the abstract as follows:

New title: Bias-blind and bias-aware LAI data assimilation over Europe with Noah-MP

In the abstract, we will modify l.12-15 as follows:

Furthermore, the bias-blind LAI DA produces a pronounced sawtooth pattern due to model drift between update steps, because the update step pushes the Noah-MP leaf model to an unstable state. This model drift also propagates to short-term estimates of GPP and ET, and to internal DA diagnostics that indicate a suboptimal DA system performance.

1.3 Validation data quality

Basic hypotheses on the validation data used to perform the analysis of results are disputable, especially over semi-arid areas around the Mediterranean where the model tends to overestimate LAI. Over such areas, SIF is probably not proportional to GPP (the Authors assume that SIF is proportional to GPP) and the ESA-CCI soil moisture (SM) product has shortcomings (not mentioned in the current version of the paper).

We agree that SIF has some shortcomings as GPP proxy, especially under extremely dry conditions. However, according to the study linked by the reviewer, this occurs during extreme events, and not generally over semi-arid areas:

"While a linear GPP–SIF relationship is expected in most conditions and driven by NPQ changes at the seasonal scale, under extreme stress a shift in energy allocation can occur" (Martini et al. 2022)

We propose to change l. 220 as follows:

Under extreme conditions, the linear relationship of SIF and GPP can break down (Martini et al. 2022). Therefore, similarly to FluxSat GPP, evaluations against SIF should be analysed carefully. Since we do not explicitly model SIF but only use it as GPP proxy, we evaluated it only in terms of R and R_{anom}.

Regarding ESA-CCI soil moisture, we do not agree with the reviewer. It is true that *active* microwave-derived soil moisture has problems due to subsurface scattering in Spain, but this does not affect passive soil moisture retrievals. Since the ESA-CCI soil moisture uses a probabilistic merging approach, it takes the low reliability of the active SM retrievals in these areas into account, and assigns a relatively large weight on the passive data (see Fig. 4 in Gruber et al. 2019).

Evaluations of the ESA-CCI soil moisture product against ERA5-Land also indicates good performance in these regions: `https://qa4sm.eu/ui/validation-result/6f4b382a-60bc-4296-9b58-d8802a1b12dd`

1.4   Incomplete analysis

The main problem I see in the Noah-MP DA system is that a partial (incomplete) analysis of the state variables of the soil-plant system is made. Root-zone soil moisture (RZSM) is not analyzed from the assimilation of LAI. Like LAI, RZSM changes relatively slowly. This is why RZSM needs to be analyzed together with LAI. Ignoring this tends to weaken the theoretical arguments used to criticize the bias-blind approach. The "negative effects" of the bias-blind approach are due to the incomplete use of the assimilated LAI data. Rescaling observations to the model range of values is relevant when units are different or when background model-dependent parameters affect the range of values. When model errors, model forcing errors, and model process uncertainties are responsible for the bias, the incorporation of the "true" observed values should improve the model simulations. In this case, artificially removing the bias is counterproductive and reduces the information content of the observations. This is particularly true for LAI. The exact value of this variable governs the biological regulation of soil moisture. Model LAI and RZSM biases can be due to model parameterization errors but also to biases in precipitation for example. ERA5 can present marked seasonal precipitation biases. The same difficulty would occur in irrigated areas since the Noah-MP model version used by the Authors does not represent irrigation. How can DA compensate for the impact of these biases if their influence on LAI is artificially removed? A solution is to analyze RZSM through the assimilation of the original LAI values. In the model you use, does LAI depend on RZSM?

Thank you for this detailed and well thought out remark. We agree that a complete analysis of the soil-plant system by jointly updating LAI and RZSM might also be another worthwhile approach for handling some of the peculiarities of the current literature standard for LAI DA into Noah-MP, where normally only LAI is updated (Kumar et al. 2019; Mocko et al. 2021; Rahman et al. 2022a; Rahman et al. 2022b). The joint updates of LAI and RZSM would present an alternative solution to the problem that we tried to solve with rescaling, especially in cases where the large bias between OL and observations is caused by biases in the forcing data (e.g. seasonal precipitation biases or lack of irrigation information). We briefly mention this in lines 460-464 in our discussion:

"[...] in strongly irrigated areas the change in soil moisture climatology leads to a decrease in soil moisture, even though the ba model performance originates from an underestimation of soil moisture due to the lack of an irrigation process in the model. Joint updates of LAI and root zone soil moisture as done in LDAS-Monde (Albergel et al., 2017) could alleviate this problem

caused by "missing" water to some extent but requires a good estimation of the coupling strength of LAI and soil moisture."

We agree that this does not give the topic the attention it deserves, so we propose the following changes to the manuscript:

1. We will add an appendix section giving a more detailed overview of the Noah-MP vegetation model, and make the description of the Noah-MP vegetation model in the main text more detailed, with a focus on how parameters and RZSM can affect the equilibrium state.

2. We will add an analysis of equilibrium LAI for the Majadas pixel and the Nile delta pixel to demonstrate how the different DA setups affect the stability of the model between update steps.

In the following, we will give initial drafts for the proposed changes.

Firstly, we will add the following below line 101:

We give a more detailed overview of the dynamic leaf model in Noah-MP in Appendix B.

In the appendix, we will add the following short overview of the Noah-MP leaf model:

Appendix B: Noah-MP dynamic leaf model

This section gives a short overview of the Noah-MP vegetation model, focusing on the interaction of LAI and soil moisture. For a more detailed description we refer to Niu et al. (2011).

Noah-MP calculates LAI from a prognostic leaf biomass C_l and a vegetation-type specific leaf area per leaf mass (specific leaf area; SLA):

$$LAI = \text{SLA} \cdot C_l$$

Leaf biomass is updated in each step via a mass balance equation:

$$\frac{\mathrm{d}C_l}{\mathrm{d}t} = \frac{1}{\text{SLA}} \frac{\mathrm{d}LAI}{\mathrm{d}t} = (1 - \text{FRAGR}) \cdot [f_l(LAI) \cdot GPP(LAI, \beta, T_c, F) - R_m(LAI, \beta, T_c)]$$
$$- D_c(LAI, T_c) - D_d(LAI, \beta) - T_l(LAI) \tag{1}$$

with FRAGR the fraction of GPP minus maintenance respiration invested in growth respiration, R_m the maintenance respiration, D_c the death rate due to cold stress, D_d the death rate due to drought stress, and T_l the turnover rate, F the atmospheric forcings, and T_c the canopy temperature. f_l is the fraction leaf carbon allocation, which governs how much of the total assimilated carbon (GPP) is allocated to the leaf pool (the rest will be allocated to other carbon pools) and is given by

$$f_l(LAI) = \left(1 - \frac{LAI}{10}\right) \cdot \exp\left(0.01 \cdot LAI \left(1 - e^{0.75 \cdot LAI}\right)\right)$$

The dependence of vegetation growth on available soil moisture is controlled via the soil moisture factor β, which represents the relative amount of plant available water in the root zone:

$$\beta = \sum_i \frac{\Delta z_i}{z_{root}} \min\left(1, \frac{\theta_i - \theta_{wilt}}{\theta_{fc} - \theta_{wilt}}\right)$$

where θ_i is the volumetric soil moisture in layer i, θ_{fc} is the field capacity, and θ_{wilt} is the wilting point.

The sink terms of Equation 1 are calculated the following way:

$$R_m(LAI, \beta, T_c) = 12 \times 10^{-6} \cdot R_{25} \cdot \text{FNF} \cdot 2^{\frac{T_c - 298.16}{10}} \cdot \beta \cdot LAI \tag{2}$$

$$D_c(LAI, T_c) = 10^{-6} \cdot \frac{LAI^2}{120 \cdot \text{SLA}^2} \cdot c_c \exp\left(-0.3 \max(0, T_c - T_{c,min})\right) \tag{3}$$

$$D_d(LAI, \beta) = 10^{-6} \cdot \frac{LAI}{\text{SLA}} \cdot c_d \cdot \exp\left(-100\beta\right) \tag{4}$$

$$T_l(LAI) = 5 \times 10^{-7} \cdot c_t \cdot \frac{LAI}{\text{SLA}} \tag{5}$$

with land cover specific parameters R_{25}, FNF, $c_c, T_{c,min}, c_d, c_t$.

We will also add the following as subsection 2.6, moving the current subsection 2.6 to 2.7.

2.6   Analysis of Noah-MP equilibrium LAI

As we will see later, each update step in the bias-blind DA is followed by a strong drift of the model LAI towards the OL values. This indicates that there is a stable equilibrium LAI to which the model tries to return after being updated by the DA step. Calculating this equilibrium value exactly is difficult, because it depends on the current forcing conditions, model state, and model parameters, but it is possible to approximate it for typical conditions within a month.

Assuming that vegetation growth is not light limited (reasonable for the areas with large bias in the southern half of the domain), GPP is proportional to LAI and to the relative amount of plant available water β:

$$GPP(LAI) \approx \alpha(F) \cdot \beta \cdot LAI \tag{6}$$

The proportionality factor α depends on the set of forcings F, and therefore (i) varies in time and (ii) an estimation of GPP requires knowledge of all forcings. Since the forcings, and therefore also α have a strong seasonal component, we can simplify the analysis by making a climatological approximation, i.e., approximating $\alpha(F)$ with the typical conditions in a given calendar month m. This is done by relating the model-simulated GPP with model-simulated LAI and β for the calendar month m, and then solving for α by taking the mean of $GPP/(LAI \cdot \beta)$.

We can then set the change in LAI in Equation 1 to zero and solve for the equilibrium LAI for this month, $LAI_{eq,m}$, by solving the following equation:

$$\begin{aligned}
0 = &\beta_m \cdot (1 - \text{FRAGR}) \cdot [f_l(LAI_{eq,m}) \cdot \alpha_m \cdot LAI_{eq,m} - R_{m,w}(LAI_{eq,m}, T_{c,m})] \\
&- D_c(LAI_{eq,m}, T_{c,m}) - D_d(LAI_{eq,m}, \beta) - T_l(LAI_{eq,m})
\end{aligned} \tag{7}$$

where $T_{c,m}$ is the mean canopy temperature and β_m the mean plant available water for month m, respectively. The solution can be obtained numerically with common root finding algorithms.

We will use this approximation method also to estimate how the equilibrium LAI would change, if we would modify the model parameters or the RZSM in the model.

Based on this method section, we will also add a new figure to the results section, showing the results of the equilibrium analysis:
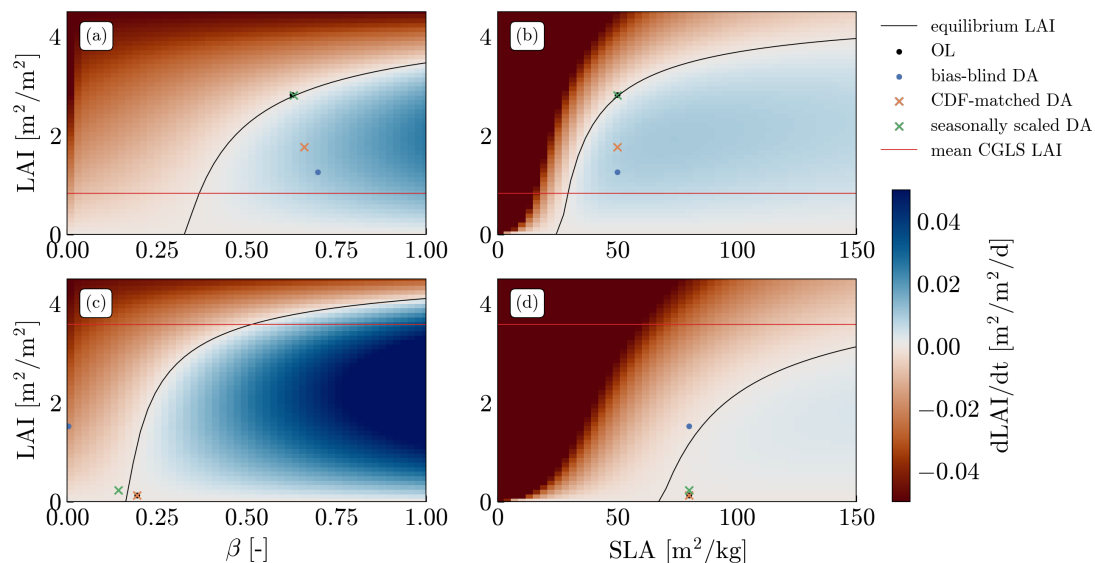
3.6 Analysis of equilibrium LAI

Figure 1: Estimates of the Noah-MP LAI change (dLAI/dt) and equilibrium LAI for the Majadas pixel in June (upper row), and the Nile delta pixel in July (lower row), as function of relative amount of plant available water β (left column) and specific leaf area (SLA, right column). Additionally, the mean conditions for the OL and the DA runs are shown as dots/crosses, and the mean CGLS LAI for the respective month as line.

We additionally assessed the dependency of the equilibrium LAI value on the model root-zone soil moisture (via the relative amount of plant available water β) and on model parameters (using the specific leaf area SLA as example parameter) for the two example sites discussed above.

Figure 1 shows how the equilibrium LAI would change if we would change β or SLA while keeping everything else constant. For both shown sites (Majadas, Nile delta) we chose the month where the mean difference between OL and observations is largest (June for Majadas, July for Nile delta). We approximated the GPP-LAI relationship for these sites and months based on Equation 6 (see Supplement for plots showing quality of fit).

For all 4 considered cases, the mean OL conditions are close to the estimated equilibrium LAI, giving confidence in the approximations performed in the derivation of the method.

For Majadas, the mean June conditions of the seasonally scaled DA are very close to the OL, while the bias-blind DA shows a strongly reduced LAI and an increased β, consistent with Figure 9. The CDF-matched DA is between OL and bias-blind DA. The latter are both further apart from the estimated equilibrium LAI, i.e. they are not in a stable state. To obtain a stable state but at the same time reduce the LAI towards the CGLS LAI observations, β would have to be reduced to about 0.3, or alternatively SLA would have to be reduced to 30.

For the Nile delta pixel, both bias-aware DA runs are very close to the OL, while the bias-blind DA shows a strongly reduced β, consistent with summer conditions in Figure 10. In this pixel, the equilibrium LAI shows a much higher sensitivity on β; a small increase in β would already lead to a large increase in LAI. Conversely, the sensitivity to SLA is low. To obtain a stable state close to the observed LAI, β would have to be increased to 0.5, while SLA would have to be increased to values larger than 150.

Based on these results, we propose to split 4.4 into 2 parts, one part discussing the potential of updating RZSM or model parameters, the other discussing potential model structural changes. Here, we will only show the new discussion section on updating RZSM or parameters, as the other section will stay mostly the same.

4.5 Alternatives to rescaling of observations

The bias-aware DA uses observation rescaling methods to reduce the effects of biased updates like drifts and sawtooth patterns, but leads to lower improvements than the bias-blind DA, because it leads to LAI estimates in the model climatology.

To obtain LAI estimates within the range of the observational LAI climatology while still keeping the model in a stable state, additional model state variables or parameters have to be updated or calibrated, or an observation or forecast bias could be estimated separately (and removed from the innovations) (G. J. M. De Lannoy et al. 2007b; G. J. M. De Lannoy et al. 2007a).

Calibration of model parameters is the best option if the bias is due to uncertainty in the model parameters. It has been successful at improving vegetation models in previous studies MacBean et al. (2015), MacBean et al. (2016), Scholze et al. (2019), Forkel et al. (2019), Kolassa et al. (2020), and Mahmud et al. (2021). Updates of specific leaf area together with LAI have already been used successfully by Xu et al. (2021) and He et al. (2022). Based on Equation 1, changes in the leaf turnover coefficient or the respiration coefficient might lead to similar results.

The calibration can either be done prior to the DA simulations or can be incorporated into a joint parameter and state update DA scheme. An EnKF (as used in this study) can in principle be used for the joint updates by augmenting the control vector to contain both state variables and parameters (Evensen 2009). If the model predictions' dependency on the parameters is highly nonlinear, particle methods might be more suitable (Frei et al. 2013; Leeuwen et al. 2019). Hybrid methods that combine the EnKF with particle methods could be used to obtain a DA system that performs well both for state updates and parameter updates (Frei et al. 2013; Leeuwen et al. 2019; G. J. De Lannoy et al. 2022).

Instead, if the (short-term) bias is caused by erroneous forcing data, e.g. a seasonal wet or dry bias in water-limited areas, joint updates of LAI and RZSM are better suited to improve the analysis. This is the approach chosen for LAI DA in LDAS-Monde (C. Albergel et al. 2017). With an EnKF, joint updates are achieved via error cross-correlations between LAI and RZSM

(obtained from the ensemble). This means that RZSM will only be updated in phases in which model LAI and RZSM show a strong coupling.

In the Nile delta, a known source of bias is the missing irrigation input in the model. Additionally, we found a very strong water limitation on vegetation growth, and a strong sensitivity of the equilibrium LAI on changes in RZSM, implying a strong model coupling of LAI and RZSM. Therefore, joint updates of LAI and RZSM are likely to improve the LAI DA results here, because RZSM is temporarily adjusted, but changes to soil parameters might also be necessary to sustain the increased moisture values.

At the Majadas site, the sensitivity of the equilibrium LAI to changes in RZSM is lower. Furthermore, Figure 9 indicates that in winter and spring, RZSM is largely dominated by soil parameters and precipitation input. In these periods, changes in LAI do not lead to changes in RZSM, and differences in RZSM between the OL and the bias-blind DA vanish quickly. Therefore, RZSM updates are unlikely to sustainably decrease RZSM during spring, which would be required to decrease LAI to a more stable state. This could make joint updates of LAI and RZSM even less efficient than in the Nile delta. Parameter updates are likely most useful for improving the LAI estimates at the Majadas site.

Is RZSM analyzed when you assimilate LAI? This is not clear in the present version of the manuscript and should be made clear.

RZSM is not analyzed jointly with LAI in our DA setup, following earlier research on LAI DA with Noah-MP, as detailed above. We apologize for not making this clear in the method section and propose to add the following sentence after l. 135

Following prior work on LAI DA with Noah-MP (Kumar et al. 2019; Mocko et al. 2021; Rahman et al. 2022a; Rahman et al. 2022b), we use the EnKF to update the model LAI, i.e., the state vector consists only of LAI.

3 Particular comments

3.1

L. 80 (Noah-MP): More should be said here on the representation of phenology and LAI in the version of Noah-MP used by the Authors. For example, is the day-to-day change in LAI impacted by RZSM? If yes, RZSM could be analyzed through the assimilation of LAI. Is this done? If not, conclusions are only valid for this model and DA system and have no universal significance.

As mentioned above, LAI is impacted by RZSM in Noah-MP. The detailed description is shown in the subsection 1.4.

3.2

L. 120 (CGLS LAI): this product has several versions/options. Which one is used in this study?

We used product version 2 (1 km resolution). We will clarify this in the revised version of the manuscript.

3.3

L. 130-135: "land surface state" is too vague. What are the analyzed variables? Is RZSM analyzed? Please list the analyzed variables.

Only LAI is analyzed. We will clarify this in the revised version, as proposed above.

3.4

L. 223: I guess that another reason for not using RMSD is that you do not simulate SIF. Correct? Please clarify.

Correct. We will clarify this in the revised version, as proposed above.

3.5

L. 231 (ESA-CCI SM): For which soil layer? Is it surface soil moisture? Please clarify.

We compare the ESA CCI SM product to the uppermost soil layer in Noah-MP (0-10cm).

3.6

L. 267 (temperature): Do you mean accumulated precipitation?

Yes, thanks you, this will be updated.

3.7

L. 283 (section 2.6): This is a bit obscure. Probably not that interesting for a majority of potential readers. I suggest moving this part and the corresponding results to a Supplement.

We agree with the reviewer that this section is difficult to understand, and that we might not have used the best way to explain and visualize our intention with this analysis.

The goal of this analysis is to show that the biased LAI updates, which are followed by an drift of the model LAI back towards the original equilibrium LAI, lead to similar drifts in flux estimates (GPP, ET). Furthermore, we want to demonstrate that these drifts are unphysical, i.e. they are caused by model instability instead of forcing input.

We will revise this section to make our intentions and the analysis more clear, and similarly adapt the related results section.

Section 2.6 will be revised as follows (as section 2.7)

2.7   Evaluation of short-term DA effects

To evaluate how biased updates affect the short-term model performance, we analyse day-to-day differences of model states and fluxes. In the OL, the day-to-day differences are driven by day-to-day variations in the forcing input. If averaged over larger areas or multiple years, this corresponds mainly to the day-to-day differences of the mean seasonal cycle. For LAI, GPP, and ET, which are high in summer and low in winter, we therefore expect positive day-to-day differences in spring, corresponding to leaf growth and increase in GPP and ET, and negative day-to-day differences in autumn, corresponding to leaf shedding and decrease in GPP and ET.

Large update steps in the bias-blind DA can induce model instabilities. In this case, the subsequent day-to-day differences are strongly impacted by the unstable artificial response to the update step, instead of reacting to the physical forcing input.

9

To detect if such model instabilities occur, and to what extent they propagate to flux estimates of the model, we evaluated day-to-day differences of 2 days minus 1 day after the DA update, as well as 1 day before minus 2 days before the DA update (the latter can also be interpreted as approximately 9 days minus 8 days after the DA update). A comparison of these also gives an indication of how long the DA induced effects remain.

For each pixel and month, we calculated the median of these day-to-day differences over the years 2003-2019 and normalised it with the monthly standard deviation of the variable values over the same multi-year time range (as a measure of the local within-month variation).

Additionally we will revise section 3.6 (new section 3.7) as follows
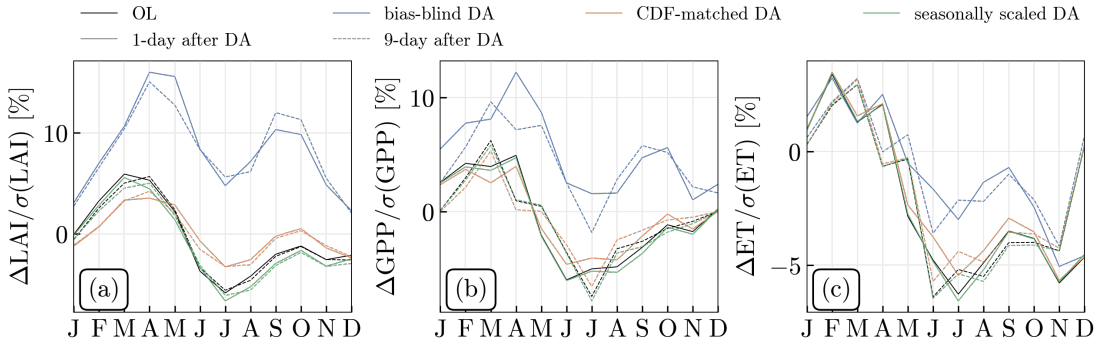
3.7 Evaluation of short-term DA effects

Figure 2: Normalised monthly median day-to-day forecast differences for (a) LAI, (b) GPP, and (c) ET. The differences are computed as the forecast value at 2-day after DA (not applied for OL) minus that of 1-day after DA (directly after DA, solid lines) and 1-day before DA minus 2-day before DA, corresponding to approximately 9-day after DA minus 8-day after DA (9-day after DA, dashed lines) for the OL (black), the bias-blind DA (blue), the CDF-matched DA (orange) and the seasonally scaled DA (green). The median was calculated from all grid cells at which the relative LAI difference between OL and bias blind DA (see Figure 1) is below -25%. For each grid cell and month, the median was normalised with the monthly standard deviation of the variable for this grid cell. The graph shows the median results across 17 years (2003-2019).

Figure 2 shows the monthly median day-to-day forecast differences for all performed simulation runs for LAI, GPP, and ET.

The OL shows a seasonal cycle with high values in spring and low values in summer, as expected (corresponds to the derivative of seasonal cycle of variable values). The bias-aware DA runs follow the OL seasonal cycle closely.

The bias-blind DA also shows the same seasonal cycle, but has an offset compared to the OL. For LAI, this offset is of the same size as the magnitude of the mean seasonal cycle, so that the mean day-to-day differences in the bias-blind DA in summer have the same magnitude as the day-to-day differences in the OL in spring, even though physically a decrease in LAI is expected. In fact, the day-to-day differences in LAI in the bias-blind DA are always positive, meaning that LAI is

expected to increase in all seasons. This is caused by the large DA update steps in the bias-blind DA, which pull the model to an unstable state. As a consequence, model instability instead of physical forcing input governs the short-term temporal evolution of LAI in the model in between update steps. Even 9 days after the DA update, right before the next update step, day-to-day differences do not significantly change, indicating that the model instability can persist for long time periods.

The instability effect also strongly affects GPP estimates throughout all seasons, and to a lesser extent ET estimates in summer.

3.8

L. 331 (Figure 4): I had a hard time understanding Fig. 4. Why are CDF and seasonal simulations missing? In the caption of Figure 4 I suggest replacing: "SIF and "scaled OL" have been rescaled to have the same maximum as "bias-blind DA"" by ""scaled SIF" and "scaled OL" correspond to rescaled SIF observations and OL simulations presenting the same maximum as "bias-blind DA", respectively".

Thank you for the suggestion of reformulating the captions, we will add it in a revised version. We did not include CDF and seasonally scaled simulations here, because we did not want to overload the plot with even more lines, since the climatologies of the CDF-scaled and seasonally scaled DA are much closer to the OL.

3.9

L. 332 (deterioration of the agreement of SIF and GPP in regions with a large bias): I disagree. Regions with large bias correspond to semi-arid areas commonly affected by droughts. SIF is not linearly correlated to GPP in all conditions. In very dry conditions, this correlation disappears. See Martini et al. (2021) for example https://doi.org/10.1111/nph.17920

See response in subsection 1.3.

3.10

L. 343 (sawtooth pattern): why should "sawtooth pattern" be considered as a problem? This is a sign that DA does its job of pulling the model closer to the observations, and that increasing the number of observations would improve the simulation.

We agree that the abrupt changes of LAI at the time of an DA update is not problematic, and simply a result of the filter algorithm doing its job. The problem is the large drift following an update, which indicates that the filter pulls the model to an unstable state and DA updates do not "stick".

As we tried to show with sections 2.6 and 3.6, this also leads to unphysical drifts/biases in flux estimates. The revised sections (now 2.7 and 3.7) are shown above.

3.11

L. 349 (GLEAM ET): Can GLEAM ET be considered as a reference dataset? Why should it be better than the simulations performed by the authors?

We included GLEAM ET for comparison due to the lack of in situ data in the areas most strongly affected by the bias-blind DA. We agree that it does not provide direct flux observations, but

(i) it is an ET product that is forced mostly with Earth observations of meteorology and land surface states and thus as close as possible to an observation-based product, (ii) it has been extensively evaluated against other products in various benchmarking activities (Greve et al. 2014; B. Martens et al. 2016; B. Martens et al. 2017; Brecht Martens et al. 2018), and (iii) it is also commonly used for assessing DA systems (e.g., Clément Albergel et al. 2019; Bonan et al. 2020; Kumar et al. 2019; Rahman et al. 2022a; Rahman et al. 2022b).

We will add the following in l. 230:

GLEAM has been evaluated against other products in various benchmarking activities (Greve et al. 2014; B. Martens et al. 2016; B. Martens et al. 2017; Brecht Martens et al. 2018), and it is commonly used for assessing DA systems (e.g., Clément Albergel et al. 2019; Bonan et al. 2020; Kumar et al. 2019; Rahman et al. 2022a; Rahman et al. 2022b). However, since it is a model-based product, comparisons against GLEAM should be analysed carefully, and GLEAM should not be treated as a ground truth validation data set.

3.12

L. 361-362: These regions are also those for which microwave derived SM is more uncertain because of subsurface scattering in dry conditions (see Wagner et al. 2022, https://doi.org/10.1016/j.rse.2022.113025)

See response in subsection 1.3.

3.13

L. 394 (Figure 9): CDF LAI is much larger than both OL and observations from January to April. Seasonally rescaled LAI is much larger than both OL and observations from April to July 2016. How is this possible? Rescaled LAI should be somewhere between the OL and the observations. Correct?

We are unsure how to interpret this comment: Does the reviewer (i) wonder why the seasonally rescaled LAI analysis (green line) is above both OL (black line) and original observations (blue dots), or (ii) wonder why the seasonally rescaled LAI observations (green dots) is above both OL (black line) and original observations (blue dots)?

In case (i), we want to point the reviewer to the seasonally rescaled observations (green dots), which are the assimilated observations. The seasonally rescaled analysis (green line) is between these observations and the OL, as expected. We will revise the figure caption to make the different dots and lines more clear.

If the author wonders about case (ii): The rescaling tries to transform the original observations in a way such that certain statistics of the OL are reproduced in the rescaled observations. In the case of the seasonal rescaling, we try to match the mean seasonal cycle and the overall standard deviation. It neither can nor intends to induce or conserve any ordering relationships with respect to the OL. Instead, it conserves the ordering relationships in the original observations. For example, if in the original observations April 2016 was the April with the highest LAI values within the baseline period (2002-2019), April 2016 will also be the April with the highest LAI values within this period in the rescaled observations, no matter what the OL values are.

3.14

L. 399: replace "suppressing" by "reducing".

Thank you, we will update the manuscript accordingly.

3.15

L. 401 (Figure 10): Why is the number of curves/dots in Fig. 10a different from Fig 9a? This not logical.

In the rescaled DA runs not much is happening in this example. We therefore excluded the lines to not overload the plot with irrelevant lines. We can also include the lines or add the complete plot in the supplements, if required.

3.16

L. 404 (strongly decreases SM2): is this because RZSM is not analyzed?

The reason for this is that the increase in LAI also strongly increases transpiration, which in turn decreases RZSM/SM2. Updating also RZSM in this case would mean adding water. If the system is properly tuned and cross-correlations are estimated well, this might counterbalance the increased water uptake. See subsection 1.4 (this document) for a more in depth discussion of the question of joint RZSM+LAI updates.

3.17

L. 409 (section 3.6): Move this section to a Supplement.

We refer to previous comments on this issue.

3.18

L. 440-442: ... and possible seasonal biases in ERA5 precipitation

Thank you, we will add this in a revised section of the manuscript.

3.19

L. 494: For the sake of clarity, it should be written here that rescaling LAI observations has a negative impact on DA efficiency.

We rephrased the conclusion to make this more clear.

3.20

L. 496: Are "standard assumptions" valid in a context where key variables (such as RZSM in this study) are not analyzed?

Even in such a scenario, the diagnostics based on these assumptions can be used to assess if the Kalman filter update sequence has the intended statistical properties, i.e. if the standard assumptions are valid.

3.21

L. 521: Do you mean that RZSM has no impact on the simulated LAI? This is far from the state of the art. Is there a more advanced version of Noah-MP able to correctly simulate LAI?

RZSM impacts LAI in Noah-MP, as elaborated in more detail above. Only the phenology index (growing season vs. no growing season) does not take into account water-driven phenology.

3.22

L. 550: I completely disagree with this recommendation. I would instead recommend paying attention to the consistency between LAI and RZSM in the LSM, and to the "fitness for purpose" of the Noah-MP LSM.

See comment in subsection 1.4.

References

Albergel, C. et al. (2017). "Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area". In: *Geoscientific Model Development* 10.10, pp. 3889–3912. DOI: 10.5194/gmd-10-3889-2017. URL: https://gmd.copernicus.org/articles/10/3889/2017/.

Albergel, Clément et al. (2019). "Monitoring and Forecasting the Impact of the 2018 Summer Heatwave on Vegetation". In: *Remote Sensing* 11.5. ISSN: 2072-4292. DOI: 10.3390/rs11050520. URL: https://www.mdpi.com/2072-4292/11/5/520.

Bonan, B. et al. (2020). "An ensemble square root filter for the joint assimilation of surface soil moisture and leaf area index within the Land Data Assimilation System LDAS-Monde: application over the Euro-Mediterranean region". In: *Hydrology and Earth System Sciences* 24.1, pp. 325–347. DOI: 10.5194/hess-24-325-2020. URL: https://hess.copernicus.org/articles/24/325/2020/.

De Lannoy, Gabriëlle J. M. et al. (2007a). "Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter". In: *Water Resources Research* 43.9. DOI: https://doi.org/10.1029/2006WR005449. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006WR005449.

De Lannoy, Gabriëlle J. M. et al. (2007b). "State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency". In: *Water Resources Research* 43.6. DOI: https://doi.org/10.1029/2006WR005100. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006WR005100.

De Lannoy, Gabrielle JM et al. (2022). "Perspective on Satellite-Based Land Data Assimilation to Estimate Water Cycle Components in an Era of Advanced Data Availability and Model Sophistication". In: *Frontiers in Water*, p. 156.

Evensen, Geir (2009). "The ensemble Kalman filter for combined state and parameter estimation". In: *IEEE Control Systems Magazine* 29.3, pp. 83–104. DOI: 10.1109/MCS.2009.932223.

Forkel, Matthias et al. (2019). "Constraining modelled global vegetation dynamics and carbon turnover using multiple satellite observations". In: *Scientific Reports* 9.1, pp. 1–12. DOI: 10.1038/s41598-019-55187-7.

Frei, M. and H. R. Künsch (July 2013). "Bridging the ensemble Kalman and particle filters". In: *Biometrika* 100.4, pp. 781–800. ISSN: 0006-3444. DOI: 10.1093/biomet/ast020. URL: https://doi.org/10.1093/biomet/ast020.

Greve, Peter et al. (2014). "Global assessment of trends in wetting and drying over land". In: *Nature geoscience* 7.10, pp. 716–721.

Gruber, Alexander et al. (2019). "Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology". In: *Earth System Science Data*, pp. 1–37.

He, Xinlei et al. (2022). "Improving predictions of evapotranspiration by integrating multi-source observations and land surface model". In: *Agricultural Water Management* 272, p. 107827. ISSN: 0378-3774. DOI: `https://doi.org/10.1016/j.agwat.2022.107827`. URL: `https://www.sciencedirect.com/science/article/pii/S0378377422003742`.

Kolassa, Jana et al. (2020). "An Observation-Driven Approach to Improve Vegetation Phenology in a Global Land Surface Model". In: *Journal of Advances in Modeling Earth Systems* 12.9. e2020MS002083 10.1029/2020MS002083, e2020MS002083. DOI: `https://doi.org/10.1029/2020MS002083`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002083`.

Kumar, Sujay V et al. (2019). "Assimilation of remotely sensed leaf area index into the Noah-MP land surface model: impacts on water and carbon fluxes and states over the continental United States". In: *Journal of Hydrometeorology* 20.7, pp. 1359–1377.

Leeuwen, Peter Jan van et al. (2019). "Particle filters for high-dimensional geoscience applications: A review". In: *Quarterly Journal of the Royal Meteorological Society* 145.723, pp. 2335–2365. DOI: `https://doi.org/10.1002/qj.3551`. URL: `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3551`.

MacBean, N. et al. (2015). "Using satellite data to improve the leaf phenology of a global terrestrial biosphere model". In: *Biogeosciences* 12.23, pp. 7185–7208. DOI: `10.5194/bg-12-7185-2015`. URL: `https://bg.copernicus.org/articles/12/7185/2015/`.

MacBean, N. et al. (2016). "Consistent assimilation of multiple data streams in a carbon cycle data assimilation system". In: *Geoscientific Model Development* 9.10, pp. 3569–3588. DOI: `10.5194/gmd-9-3569-2016`. URL: `https://gmd.copernicus.org/articles/9/3569/2016/`.

Mahmud, Kashif et al. (2021). "Optimizing Carbon Cycle Parameters Drastically Improves Terrestrial Biosphere Model Underestimates of Dryland Mean Net CO2 Flux and its Inter-Annual Variability". In: *Journal of Geophysical Research: Biogeosciences* 126.10. DOI: `https://doi.org/10.1029/2021JG006400`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021JG006400`.

Martens, B. et al. (2016). "Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture". In: *International Journal of Applied Earth Observation and Geoinformation* 48. Advances in the Validation and Application of Remotely Sensed Soil Moisture - Part 2, pp. 146–162. ISSN: 1569-8432. DOI: `https://doi.org/10.1016/j.jag.2015.09.012`. URL: `https://www.sciencedirect.com/science/article/pii/S0303243415300350`.

Martens, B. et al. (2017). "GLEAM v3: satellite-based land evaporation and root-zone soil moisture". In: *Geoscientific Model Development* 10.5, pp. 1903–1925. DOI: `10.5194/gmd-10-1903-2017`. URL: `https://gmd.copernicus.org/articles/10/1903/2017/`.

Martens, Brecht et al. (2018). "Towards Estimating Land Evaporation at Field Scales Using GLEAM". In: *Remote Sensing* 10.11. ISSN: 2072-4292. DOI: `10.3390/rs10111720`. URL: `https://www.mdpi.com/2072-4292/10/11/1720`.

Martini, David et al. (2022). "Heatwave breaks down the linearity between sun-induced fluorescence and gross primary production". In: *New Phytologist* 233.6, pp. 2415–2428. DOI: `https://doi.org/10.1111/nph.17920`. eprint: `https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.17920`. URL: `https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.17920`.

Mocko, David M. et al. (2021). "Assimilation of Vegetation Conditions Improves the Representation of Drought over Agricultural Areas". In: *Journal of Hydrometeorology* 22.5, pp. 1085–1098. DOI: `10.1175/JHM-D-20-0065.1`. URL: `https://journals.ametsoc.org/view/journals/hydr/22/5/JHM-D-20-0065.1.xml`.

Niu, Guo-Yue et al. (2011). "The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements". In: *Journal of Geophysical Research: Atmospheres* 116.D12. DOI: `https://doi.org/10.1029/2010JD015139`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015139`.

Rahman, Azbina et al. (2022a). "Global Assimilation of Remotely Sensed Leaf Area Index: The Impact of Updating More State Variables Within a Land Surface Model". In: *Frontiers in Water* 3. ISSN: 2624-9375. DOI: `10.3389/frwa.2021.789352`. URL: `https://www.frontiersin.org/article/10.3389/frwa.2021.789352`.

Rahman, Azbina et al. (2022b). "The Joint Assimilation of Remotely Sensed Leaf Area Index and Surface Soil Moisture into a Land Surface Model". In: *Remote Sensing* 14.3. ISSN: 2072-4292. DOI: `10.3390/rs14030437`. URL: `https://www.mdpi.com/2072-4292/14/3/437`.

Scholze, M. et al. (2019). "Mean European Carbon Sink Over 2010–2015 Estimated by Simultaneous Assimilation of Atmospheric CO2, Soil Moisture, and Vegetation Optical Depth". In: *Geophysical Research Letters* 46.23, pp. 13796–13803. DOI: `https://doi.org/10.1029/2019GL085725`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL085725`.

Xu, Tongren et al. (2021). "Improve the Performance of the Noah-MP-Crop Model by Jointly Assimilating Soil Moisture and Vegetation Phenology Data". In: *Journal of Advances in Modeling Earth Systems* 13.7. e2020MS002394 2020MS002394, e2020MS002394. DOI: `https://doi.org/10.1029/2020MS002394`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002394`.