Reviewer #1:

The paper illustrates how satellite-based vegetation data assimilation, and joint vegetation and soil moisture assimilation has an impact on evapotranspiration, gross primary productivity and soil moisture over CONUS. The paper is of general interest to the scientific audience, but needs some clarifications and could also benefit from a careful language check.

We thank the referee #1 for her/his positive comments about our work and for her/his detailed review that has helped us to improve the quality of our manuscript. Responses to comments and subsequent changes are detailed below.

Major points:

Comment 1.1:

The intense massaging of the VOD data begs the question which signal is ultimately helping the assimilation system in this paper. It would be good to make this clearer in the manuscript. The rescaling of VOD to LAI needs to be spelled out more precisely. Is a relationship found pixel per pixel and per season, or was the relationship based on the CONUS clouds per vegetation class for all months or only for the growing season or per season presented in Fig 5, or still something else? Unless I missed it, there is also no mentioning of how the discrepancy in spatial resolution (and the spatial-temporal collocation in general) in VOD and LAI data is handled to obtain this linear relationship. After rescaling and applying a 90-day rolling average, the short-term variability is probably gone. Yet, the latter might be very important to catch the start of the growing season. In the end, it sounds like the only 'original' signal that can improve the assimilation system is of an interannual (and perhaps seasonal) nature. If that is correct, it needs to be explicitly mentioned in the paper. Finally, the 90-day rolling average means that the benefit of a filter is defeated: why not simply directly use a smoother and limit the observation preprocessing?

Response 1.1:

The authors agree that the original description of the VOD re-scaling is insufficient. To remedy this, the following text has replaced a phrase in section 2.3.3:

"Before linear re-scaling, the LAI and VOD observations are first scaled and matched to the same 0.25° x 0.25° grid. A linear monthly re-scaling was then performed using a 3-month moving window period to best match the two datasets over seasonal timescales. Over an entire year, this re-scaling is represented by 12 monthly equations each taking into account the climatologies of the months preceding and succeeding it, and it is applied on a per pixel basis. Each monthly equation is the same from one year to another. Each equation results from a first-order linear regression. In addition to this CDF-matching, a 30-day rolling average is applied after the re-scaling to smooth the resulting LAI proxy, and allow for better performance of the assimilated data. VOD is sensitive to short term changes in vegetation water content such as rainwater interception (Saleh et al. 2006). This day-to-day variability does not reflect changes in LAI."

Regarding the application of the re-scaling, the manuscript now describes that the LAI and VOD observations are first matched to the same spatial grid at 0.25 degree resolution. As this re-scaling is performed per monthly, with a 3-month period, the a and b parameters in the linear re-scaling are calculated with all the available observations of LAI and VOD. Additionally, several paragraphs in this section 2.3.3 have been reordered in order to first describe VODCA, and then discuss the processing of the VOD data.

It must be noticed that the original manuscript mistakenly described a 90-day rolling average applied after the re-scaling in order to smooth the results, allowing for better assimilation performance. However, in fact it was a 30-day rolling average that was applied, which is now accurately reflected in the manuscript. This 30-day rolling average does still remove some short term variability as you have noted, but it is also significantly less than the mistakenly written 90-day rolling average.

See also Response 2.1 (to Reviewer 2) regarding the difficulties in physically simulating VOD and justifying the use of a statistical re-scaling approach.

Comment 1.2:

Related to the above, the choice to evaluate the results only in terms of Pearson *correlation needs to be explained. Is it not more common to evaluate at least soil moisture in terms of anomaly correlations? And how about including an evaluation in terms of unbiased RMSD, or at least mention if the story remains the same for other metrics?*

Response 1.2:

A sentence is now added in section 2.5 explaining the choice of the correlation as a statistical metric, as well as saying that the Root Mean Squared Deviation (RMSD) was also taken, with very similar results, but is not shown:

"The correlation is chosen as it is a simple, yet effective, measure of proximity to reference datasets. The average correlation as well as distribution of correlations can allow the quick assessment of improvement or degradation, and are consistent with previous studies of LSMs and LSVs. Root mean squared deviation (RMSD) was also calculated for the comparisons to reference observations, showing the same results as with correlation, and are thus not shown."

In this work we preferred using R than anomaly correlation because the previous study from Albergel et al. (2018) has shown that NIC for anomaly correlation is too optimistic (i.e. less stations presented negative NIC values for anomaly correlation NIC).

Comment 1.3:

The monthly correlations (fig 7, 12) are not clear: (i) are all time steps included, i.e. both forecasts and analysis time steps, and at which temporal resolution, (ii) are these values spatial correlations between simulations and reference data, temporally averaged per month, or (iij) *are these values multi-year temporal correlations computed at each location and then spatially averaged? It would be nice to also (1) compute confidence intervals for these monthly correlations; (2) show the number of pixels involved per month (the high correlations for SSM in the winter month might be applicable to far less pixels, if QC screening was applied).*

Response 1.3:

These correlations are produced by combining all points in the domain into a single, long, time series, where the correlation is then computed against the observations processed in an identical way. This provides only one correlation score over the domain for each period. However, the significance of the score is strengthened due to the very large sample length (15 years are considered over a large domain containing more than 24000 land grid cells at a spatial resolution of 0.25 x 0.25 degrees). This description has also been added to the manuscript text in section 2.5.

"When calculating correlation to satellite-based observations of LAI, ET, GPP, and SSM, the correlations are produced by combining all points in the domain into a single, long, time series, where the correlation is then computed against the observations processed in an identical way. This provides only one correlation score over the domain for each period. However, the significance of the score is strengthened due to the far larger sample length (15 years are considered over a large domain containing more than 24000 land grid cells at a spatial resolution of 0.25 x 0.25 degrees)."

Comment 1.4:

Please read the manuscript thoroughly another time. There is some imprecise language and there are plenty of grammar issues. Random examples are listed here (the paper is full of issues; far too many to start noting):

- *L. 5-L.8: This capability -> this positive impact (implicit flow of thoughts)*
- *L. 5 difference between model simulations and forecasts (drop forecasts?)*
- *L. 8: due to the low temporal…, [which is] at best [,] every ten days, and can suffer*
- *L.13: far more .. than.. product*s* (or *an* optical product…)*
- *L.110: for nature tiles. What is "nature"? There is a hint on line 135, that the model converts urban to bare rock – and rocks are nature?*
- *L. 119: NIT option not explained.*
- *L. 176: LAI that has been of direct estimations… (rephrase?).*
- *L. 182: large -> long wavelengths*
- *L. 192: VODX and VODC are not known to everyone, introduce*
- *L. 290: mention spatial resolution of simulations?*
- *L. 387: , and all the model, and even all the observations… (rephrase)*

- *L. 553: first sentence is poorly constructed, rephrase.*
- *Throughout: use the same number of significant numbers in the text and figures (we have everything from 0.8, to 0.66 to 0.795 for R-values)*

Response 1.4:

Changes have been made to correct the noted problems. See below for each individual correction.

- *L. 5-L.8: This capability -> this positive impact (implicit flow of thoughts)*

“capability” has been changed to “positive impact”

- *L. 5 difference between model simulations and forecasts (drop forecasts?)*

“and forecasts” has been removed from this sentence

- *L. 8: due to the low temporal…, [which is] at best [,] every ten days, and can suffer*

L.8 now reads: “However, this positive impact does not reach its full potential due to the low temporal availability of optical-based LAI observations, which is at best, every ten days, and can suffer from months of no data over regions and seasons with heavy cloud cover such as winter or monsoon conditions.”

- *L.13: far more .. than.. product*s* (or *an* optical product…)*

L.13 now reads: “The Vegetation Optical Depth Climate Archive (VODCA) dataset provides near-daily observations of vegetation conditions, far more frequently than optical based products such as LAI.”

- *L.110: for nature tiles. What is “nature”? There is a hint on line 135, that the model converts urban to bare rock – and rocks are nature?*

Nature tiles are all non-urban surfaces. L. 110 now begins: “For nature (i.e. non-urban) tiles as determined by land use databases…”

- *L. 119: NIT option not explained.*

The NIT option allows for the simulation of non-woody above ground biomass, both leaf and structural, as well as transition the LAI variable from being prescribed to diagnostic based on the leaf biomass. This is described in L. 119-120. NIT itself is not an acronym.

- *L. 176: LAI that has been of direct estimations… (rephrase?).*

L. 176 now reads: “Previous implementations of LDAS-Monde have directly assimilated LAI products from optical observations.”

- *L. 182: large -> long wavelengths*

“large” has been replaced with “long”

- *L. 192: VODX and VODC are not known to everyone, introduce*

Section 2.3.3 now includes the following line: "VOD is separated into wavelength bands based on the radiation wavelengths from which they are derived. This study examines C-band (3.75 to 7.50 cm) and X-band (2.50 to 3.75 cm) VOD, while also discussing L-band (15 - 30 cm) VOD."

- *L. 290: mention spatial resolution of simulations?*

Section 2.5 now begins with the following line: "The experiments performed and reported in this study occur over the Contiguous United States (CONUS) from 2003 to 2018 at 0.25° x 0.25° spatial resolution."

- *L. 387: , and all the model, and even all the observations… (rephrase)*

L. 387 now reads: "On average, the month of May sees some of the fastest vegetation change of the year for CONUS."

- *L. 553: first sentence is poorly constructed, rephrase.*

L. 553, section 5, now reads: "This study finds a generally positive relationship between observations of LAI and VODX."

- *Throughout: use the same number of significant numbers in the text and figures (we have everything from 0.8, to 0.66 to 0.795 for R-values)*

All reported scores in the text and tables are now given with two significant figures.

Comment 1.5:

L. 58: Assimilation here assumes a dynamic vegetation model, which is not present in all LSMs. In the broad sense, LAI assimilation could also refer to an updating of input LAI parameters.

Response 1.5:

L. 58 now specifies that this is applicable with LSMs capable of dynamically simulating vegetation.

"LAI, for example, can be constrained indirectly in LSMs capable of dynamically simulating vegetation, through the assimilation of LSVs such as brightness temperature (Vreugdenhil et al., 2016; Sawada et al., 2020) and radar backscatter (Lievens et al., 2017; Shamambo et al., 2019)"

Comment 1.6:

L. 107: is the same 20% error applied to actual LAI observations and VOD observations that are rescaled to LAI? Or did you 'rescale' the observation error somehow? Figure 4 implicitly shows that the observation error (relative to the model LAI) will be different for both. It would be nice to check the error between the model LAI and the observed LAI and the LAI-rescaled VOD and at least correct the observation error accordingly to interpret the results.

Response 1.6:

In reality, in Barbu et al. (2011) and Fairbairn et al. (2017) both background and observation error standard deviations were represented in the same way for LAI. A 20% observation error was applied to LAI values larger than 2 m2m-2. A constant error value of 0.4 m2m-2 was used for LAI values below 2 m2m-2. In this study, we follow the approach proposed by Albergel et al. (2017): background LAI model errors are prescribed as in Barbu et al. (2011) and Fairbairn et al. (2017) but LAI observation errors are fixed as 20% of observed LAI values. The LAI proxy derived from the re-scaled VOD is also assimilated with these prescribed observation error standard deviation of 20%. This has been stated at the end of section 2.1 in the updated manuscript:
"In Barbu et al. (2011) and Fairbairn et al. (2017) both background and observation LAI error standard deviations were represented in the same way. A 20% error was applied to LAI values larger than 2 m2m-2. A constant error value of 0.4 m2m-2 was used for LAI values below 2 m2m-2. In this study, we follow the approach proposed by Albergel et al. (2017): background LAI model errors are prescribed as in Barbu et al. (2011) and Fairbairn et al. (2017) but LAI observation errors are fixed as 20% of observed LAI values. The LAI proxy derived from the re-scaled VOD observations is also assimilated with these prescribed observation error standard deviation of 20%. Further work would be required to assess to what extent this value of 20% is applicable to the re-scaled VOD."

Comment 1.7:

*L. 156: hard to believe that the CCI product provides *daily* data from 1978 onwards. If so, then some interpolation must have happened, and it would not be recommended to assimilate interpolated data.*

Response 1.7:

We agree. Not all time periods from 1978 are covered by daily data. "daily" was deleted. While the temporal sampling of the merged CCI SM product is 1 day (section 7.2.1 in https://esa-soilmoisture-cci.org/sites/default/files/documents/public/CCI%20SM%20v06.1%20documentation/ESA_CCI_SM_RD_D2.1_v2_ATBD_v06.1_issue_1.1.pdf), not all days have associated observations. Figure 6 in the CCI Product Validation and Intercomparison Report (https://esa-soilmoisture-cci.org/sites/default/files/documents/public/CCI%20SM%20v06.1%20documentation/ESA_CCI_SM_D4.1_v2_PVIR_v6.1_issue_1.0.pdf) shows the fractional number of valid soil moisture observations per month over the globe.

Comment 1.8:

ALEXI and FLUXCOM both use MODIS LAI-related data at some point. Would you expect even more consistency with these 'reference products' when assimilating MODIS LAI? What is the issue about data access for FLUXCOM? (this is really for the FLUXCOM developers - I want to raise awareness for open data access).

Response 1.8:

It is likely that using MODIS LAI would allow for more consistency with the ALEXI and FLUXCOM products. Assimilating CGLS instead of MODIS LAI allows a more independent evaluation from ALEXI and FLUXCOM. However, a comparison assimilating MODIS LAI as well as using it for the linear re-scaling could be an interesting future pursuit. Regarding the data access for FLUXCOM, when conducting the analysis, the data was not accessible, and only the data that had been already downloaded was used. However, it looks like the FTP is now running without any access issues.

Comment 1.9:

The text jumps from Fig. 9 to Fig 12; Fig 10-11 are only discussed later. Re-order the figures; perhaps the latter figures can even be removed and be presented in a table (~ Table 4).

Response 1.9:

Figures 10 and 11 were moved to the Supplement as Tables 3 and 4 have the relevant information regarding the average correlations and the number of improvements and degradations.

Comment 1.10:

The impact on SSM is negligible in this paper and not all in line with other studies. Is the system designed to minimally update SM, i.e. to avoid harm? How general is the conclusion that vegetation DA has a greater impact? Is it just for the ISBA model or would you expect it to be general for all LSMs?

Response 1.10:

To the best of our knowledge, many similar studies assimilating SSM using state of the art land surface models (e.g. Martens et al. 2007, de Rosnay et al. 2013) obtained the same kind of results. The following paragraph has been added to the Discussion (now in section 4.3):

"The small impact of assimilating SSM can be explained by the fact that we use a state-of-the-art land surface model able to represent diffusion processes into the soil. In dry conditions, the simulated SSM is decoupled from soil moisture of deeper soil layers. As a result, assimilating SSM has a limited impact on the model state variables (Parrens et al. 2014). On the other hand, directly assimilating LAI impacts deep soil layers and a more efficient analysis of root-zone soil moisture can be done than assimilating SSM alone (see also Fig. 4 in Albergel et al. 2017)."

References:

De Rosnay, P.; Drusch, M.; Vasiljevic, D.; Balsamo, G.; Albergel, C.; Isaksen, L. A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. Q. J. R. Meteorol. Soc. 2013, 139, 1199–1213.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Parrens, M., J.-F. Mahfouf, A. Barbu, and J.-C. Calvet: Assimilation of surface soil moisture into a multilayer soil model: design and evaluation at local scale, Hydrol. Earth Syst. Sci., 18, 673-689, https://doi.org/10.5194/hess-18-673-2014, 2014.

Comment 1.11:

L. 505: why is there a discussion about L-band VOD if no L-band VOD is used in this paper? Similarly, why is section 4.3 in this paper?

Response 1.11:

In this paper, we mention L-band VOD as there may be links between the research conducted with L-band and other microwave frequencies in the context of vegetation data assimilation. See also Response 2.1 to Reviewer 2.

Section 4.3 was originally in the article to identify concrete next steps taken for this research. However, as the specifics of future drought monitoring studies are not necessary for this article, we have removed this section, and instead added the following addition to the Discussion section 4.3, formerly 4.2:

"By improving initial conditions of the LDAS, next steps also include testing drought forecasting by combining these known improvements through more frequent and joint assimilation of observations with LDAS-Monde's forecast capacity. The analysis of drought forecast accuracy and potential warning time could prove useful for agricultural managers and stakeholders.