The file corresponding to the position of all modifies in the manuscript is the "HESS-2021-131-highlight.pdf"

Referee 1:

The paper by Li et al. proposes using a four-dimensional variational assimilation method (PODEn4DVar) with soil moisture (SM) and leaf area index (LAI) observations for calculating the evapotranspiration (ET) and gross primary production (GPP). The paper is rich in content but there are a lot of confusions. Therefore, major revisions are required before proceeding with your paper.

As you said, the spatial scale mismatch between the ground observed footprint size and satellite-derived footprint size were the vital factors affecting assimilation performance. The spatial resolution of assimilated GLASS LAI data is 5 km, and thus the resolution of ET and GPP estimates is 5 km. The spatial footprint of EC measurement of ET and GPP is at most about 100-500 m. Hence, there is a mismatch between the footprint of model estimates (5 km) and field EC measurements (100 m). Larger spatial representation mismatches can lead to completely wrong conclusions, especially on heterogeneous surfaces. Perhaps you discussed it in the manuscript, but such uncertainty will strongly change the result.
Response: Yes, the scale mismatch between the field EC measurements and satellite-derived footprint or model estimates exists and cannot be avoided. However, it is common to use EC measurements to validate model outputs at different scales. In addition, the comparison of results across multiple models is also used to validate our products. Many studies have shown that GLASS LAI products (Xiao et al.,2014; Xu et al.,2018; Liu et al.,2018), SMOS and SMAP products (Burgin et al.,2017; Colliander et al.,2017) are consistent with the in situ measurements in terms of temporal variations, that is, the consistency of tendency, and that overall bias is controlled within reasonable limits. Overall, GLASS LAI products, SMOS and SMAP products can accurately capture the daily variation of the observations, and the absolute errors can be partially eliminated when they are considered as observation errors during assimilation.

Lines 47: Maybe it should be "GLASS LAI"?
Response: The "GLASS" has been modified to" GLASS LAI".

Lines 73: "ENKF" or "EnKF", please unify.
Response: ENKF in this manuscript has been unified as ENKF.

Lines 74: There have been many applications of four-dimensional variational method (4DVar).
Response: Yes, there have been indeed many applications. More references about the application of four-dimensional variational method (4DVar) have been added in Line 82.

Lines 165: What does this equation mean? What are the constraints of SM on ET?

Response: This equation is an implicit soil moisture constraint. Although soil moisture data did not directly constraint ET, both RH and VPD in this equation are related to soil moisture. Thus, it can reflect the condition of the soil moisture where the vegetation is located to a certain extent. We have clarified it in the manuscript. (Lines 219)

Lines 170, 179, 180: Confusing expression! Do you assimilate ET in the manuscript?
Response: We did not assimilate ET. the SSM was assimilated to LPJ-PM. The sentence was revised to clarify this issue (Lines 192-193).

Lines 187: The flow chart of the LPJ-VSJA assimilation program is complicated and needs to be simplified. What are "section 3.1.1" and " section 3.2" in the Figure 1?
Response: The flow chart of the LPJ-VSJA assimilation program has been simplified in section 2.2. The section 3.1.1 and section 3.2 was deleted.

Lines 196: RMSD and ubRMSD have similar meanings, maybe you only need to keep one.
Response: The ubRMSD was retained and the associated charts and analysis have been adjusted accordingly.

Lines 216: Observation errors should be defined based on the errors of instruments measuring LAI. Why did you choose arbitrary values for LAI observation errors?
Response: In this study, we set the model and observation errors at a given time as 20% and 10% (optimal scale factors for assimilation performance) of the LAI value and the observed LAI value, respectively (Lines 285-286). They are not the arbitrary values.

According to some previous studies, the observation error of LAI is usually a fixed value determined by the errors of instruments measuring LAI. However, in fact, errors have a strong correlation with the vegetation and environmental conditions (e.g. vegetation types, season, geospatial distribution), (Fang et al.,2013). For example, GLASS LAI products have been confirmed that it has high accuracy in the growing season and for tropical forest vegetation types, as well as in tropical and northern high latitude regions (Xiao et al.,2013(a); Xiao et al.,2013(b)). The analysis increments depend not only on perturbation sizes and defined errors, but also on the plant physiological activity. As perturbation sizes have been defined as proportional to the amount of active biomass, the background error definition may require a similar adaptation. Sabater et al. (2008) defined the observation and background errors as a constant 1 m^2/m^2, while Jarlan et al. [2008] introduced a variable model error of 20% of the forecast LAI state. For LAI, perturbations are set to a fraction (0.001) of the modelled LAI following Rüdiger et al. (2010). LAI background error SD is set to 20% of the LAI value for modelled values above 2.0m^2 m^{-2} and to a constant 0.4m^2 $m^{-2}$for modelled values below 2.0m^2 m^{-2}. This SEKF configuration is the same as the one detailed in Albergel et al. (2017). Thus, the scale factor method was adopted to confirm the optimal LAI error.

Lines 217: "(R, RMSD, BIAS,)" remove the comma.
Response: The comma has been removed.

Lines 229: What is the time step and spatial resolution of the model running?

Response: The time step is daily and the spatial resolution is 0.25°（Line 457）

Lines 301: I can't find ubRMSD in the Taylor diagram of Figure 10.
Response: We modified the description of ubRMSE in the Taylor's diagram and Figure 10.

Lines 318: Please add the reference about Liu et al. (2018). You should add the download URL about Heihe data.
Liu, S., et al., 2018. The Heihe Integrated Observatory Network: A Basin-Scale Land Surface Processes Observatory in China. Vadose Zone Journal 17, 180072.
https://doi.org/10.2136/vzj2018.04.0072
Response: The reference and the URL about Heihe data were added in Line 407.

Lines 337: You should add download URLs for LAI, SM, and other remote sensing data.
Response: The URLs were added in Table 2.

Lines 367: The spatial resolution of MODIS ET is 1km, and the spatial resolution of GLASS ET is 5km. To be consistent, have you resampled?
Response: Yes, in order to maintain consistency with the SMAP Enhanced 3 Level product (Entekhabi et al. 2010), model-forcing data were resampled to a 9 km spatial resolution based on EASE-2 projection grid for site simulation. In the global spatial simulation, the model-forcing data were resampled to 0.25° using bilinear interpolation to ensure consistent temporal and spatial resolution of the driving data required for this study. According to Yang et al. (2020), the spatial resolution is determined by considering the model's output efficiency of carbon and water flux simulation. The corresponding modification have been added to Lines 445-447.

Lines 387: You should mention Figure 3 in the description.
Response: The Figure 3 has been mentioned in Line 482.

Lines 395: The unit format is not uniform, you need to double check. "g C/m^2/mon" or " g C m^{-2} mon^{-1} ", please unify. The x-axis in Figure 3 and 6 is confusing, or you can express it in years.
Response: The unit format in the manuscript was unified and the x-axis in Figures 3 and 6 was modified to the “year”.

Lines 408: Over cropland, the data assimilation scheme tends to generally underestimate GPP. What is the reason for the underestimation?
Response: In cropland, human interference, especially seasonal irrigation and fertilization, is an essential factor for cropland production. In the LPJ-DGVM model, the artificial management module and the nitrogen cycle were missing, which led to the underestimation of GPP simulations. Although the leaf area index and soil moisture assimilation could improve GPP simulation, the leaf area index and soil moisture products were also underestimated

compared with the field footprint observations (Xiao et al.,2013; Zhang et al.,2019; Liao et al.,2015). Therefore, the GPP for croplands is generally underestimated.

Lines 411: “Fig.4”, Full name or abbreviation? please check carefully!
Response: Modified.

Lines 413: There are a lot of up-scaled ET and GPP products (Jung et al., 2009, 2020) that can be used for regional validation.
Jung, M., et al., 2009. Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model. Biogeosciences 6 (10), 2001–2013.
Jung, M., et al., 2020. Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. Biogeosciences 17, 1343–1365. https://doi.org/10.5194/bg-17-1343-2020
Response: FLUXCOM GPP products (2010-2015) were used to compare with all other products, including LPJ-VSJA products studied in this manuscript (Table S6). FLUXCOM GPP products were included in above referred literature.

Lines 504: The X-axis label is the same as the y-axis label in the Taylor chart.
Response: The X-axis label in Figure 10 has been modified to avoid confusion with the code in the Taylor diagram.

Lines 516: The full name should be used the first time you use it (NSD).
Response: The full name of NSD has been added in Line 376.

Lines 548: In column A, GPP is low in the tropics and high in arid regions?
Response: The legend of Column A has been reversed, and so now GPP is high in the tropics and low in arid regions.

Figure 13: The TC method has been used to quantify uncertainties of gridded datasets. The TC method can only calculate uncertainty based on three (ET/GPP) products. How did you calculate the five (ET/GPP) products?
Response: In this study, the five products were divided into three product categories, including satellite product (MODIS, GOSIF GPP), reanalysis product (GLASS, GLDAS) and data assimilation product (GLEAM ET, LPJ-VSJA). One product in each category was selected to form a group to calculate their error. The LPJ-VSJA product was set as the reference.

For GPP products, GOSIF, GLASS, and LPJ-VSJA were treated as a group, and MODIS, GLASS and LPJ-VSJA were treated as another group to calculate the errors; the final errors were determined by the average of these two.

Similarly, to calculate the errors for ET, GLEAM, GLASS, and MODIS were chosen as a group; LPJ-VSJA, GLDAS, and MODIS were treated as a group; LPJ-VSJA, GLASS and MODIS were considered as a group. In order to reduce the influence of orthogonality

hypothesis of error, the first and third groups are for indirect and effective comparison between LPJ-VSJA product and GLEAM product.

The above explanation was added in section 2.4 (Lines 396-406)

Lines 579: “Except for MODIS, GLASS, and LPJ-DGVM (0–60 mm month-1), the σ of other products was generally between 0-20 mm month-1.”. This is hard to see in Figure 14.

Response: In the Figure 14, it is clear to see the range and distribution of σ for each product. The starting and ending values on the x-coordinate reflect the approximate range in the histograms of error standard deviation (σ) of Figure 14. The X-axis was unified to 0-60 units makes the values more comparable.

Referee 2:

Abstract:

1.Confusing mix of LPJ-something used here. First LPJ-Vegetation, then LPJ-DGVM, then LPJ-PM. Also LPJ-VSJA, but I appreciate that is the DA system (although the VSJA acronym is not explained). Then in line the introduction the authors talk about LSMs, not DGVMs and at line 66 terrestrial biosphere models are mentioned. Please be clear and consistent throughout the manuscript.

Response: The abbreviation of model and assimilation framework mentioned in this article is explained in Table 1.

For the second question, LSMs are the land surface models, which includes the terrestrial biosphere models, and the terrestrial biosphere models includes DGVMs. The DGVMs (dynamic global vegetable models) are process-based dynamic terrestrial biosphere models, which can simulate material exchange between vegetation and different conditions from the perspective of vegetation physiological processes, and is widely used to estimate carbon and water fluxes of terrestrial vegetation at various scales. This sentence was added in Lines 56-59 and the description for these model keeps consistent throughout the manuscript to avoid confusion.

2.Please explain all acronyms. Once you’ve explained an acronym then use that throughout.

Response: All acronyms in this manuscript have been interpreted at first use and remain consistent throughout the manuscript.

3. A clear explanation of which is the model version that has been optimized with the DA framework and which not would be helpful in the abstract.

Response: We explicated the LPJ-DGVM (version 3.01) that has been optimized with the DA framework in Line 26. It is currently the only one version of LPJ-VSJA.

4.Line 34: “The assimilated GPP and ET” suggests that GPP and ET data have been assimilated. I suggest “posterior GPP and ET” would be better.

Response: “The assimilated GPP and ET” has been modified to “posterior GPP and ET”.

Introduction:

5. Line 65: Probably more appropriate references here. See Scholze et al. (2017) or Exbrayat et al. (2019) for further references.

Response: The relevant references have been added in Line 69-70.

6. Line 68: You also need the underlying model, not just these three components.

Response: Yes, the underlying model was the component of the assimilation system and it "the was described in Line 73.

7. Line 71-73: I would re-write this sentence as "which significantly improve simulations by periodically updating state variables (e.g., LAI and soil moisture) using remote sensing data without changing the model structure".

Response: We agree with you and have revised the relevant description.

8.Line 74: "obtain the dynamic balance of the estimation window" ï I would explain fully what is meant by this for non DA specialists. It might also be useful to add an additional sentence explaining the difference between EnKF and 4DVar either before or after this set of sentences.

Response: EnKF relies on the instantaneous observations to update the state variable at the current time, and gives the predicted value at the next time based on the forward integration of the updated state variable. This explanation was added in Lines 78-80 for explaining the difference between EnKF and 4DVar.

9. Line 79: Please can the authors be more specific when they say "satisfactory performance in land DA" beyond what is specified for a different paper later in the sentence (that the method does better at estimating GPP and ET with ENKF)?

Response: we specific this sentence in Lines 86-87.

10. Line 85: I am not sure you want to reference Liu et al here because they talk about how different LAI products have inconsistent estimates; therefore, that is a disadvantage for using LAI data to evaluate or optimize models, as how do we know which LAI product is more accurate? This actually is in contrast to lines 94-96.

Response: We referred Liu et al. to demonstrate the high sensitivity of the estimation of carbon and water fluxes to the LAI, so that a more accurate assimilation of LAI products into the model can obtain a more accurate simulation of carbon and water fluxes. An explanation was added in Line 92.

GLASS LAI products are designed by combining ground observation with MODIS and CYCLOPES LAI information, and MODIS reflectance was used to train and generate fused LAI using General Regression Neural Networks (GRNNs). It has been verified that this method can improve the LAI inversion accuracy of long time series (Liang et al. 2013; Xiao et al. 2016). GLASS LAI product has been verified to be more accurate than MODIS and CYCLOPES, with stronger temporal continuity and spatial integrity (Xiao et al. 2013). Considering the temporal and spatial continuity and accuracy, GLASS LAI products were selected as observation data, and the error of LAI products was analyzed in the Discussion section (Lines 776-783).

Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., Zhang, X., Liu, Q., Cheng, J., & Tang, H. (2013). A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. International Journal of Digital Earth, 6, 5-33

Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., & Song, J. (2016). Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance. IEEE Transactions on Geoscience and Remote Sensing, 54, 5301-5318

Liu, L., Gudmundsson, L., Hauser, M., Qin, D., Li, S., & Seneviratne, S.I. (2020). Soil moisture dominates dryness stress on ecosystem production globally. Nature communications, 11, 1-9

11. Line 88: Do the authors mean more accurate SM data assimilated into models can improve accuracy? And if the authors are not talking about assimilating SM data here, then how was SM data used to improve accuracy of models and is that relevant to a DA study? Same comment for the references used on Line 85. From the sentence they're referencing I assume these references demonstrate how LAI has been used to improve models, but I am not sure that is the case. If instead these references are to demonstrate uncertainty in these variables in models then that should be better specified.

Response: Yes, the concept is that more accurate SM data assimilated into models can improve accuracy. Similarly, more accurate LAI can improve model accuracy. We have deleted these references above to avoid ambiguity.

12. Line 104: Maybe the authors could explain why microwave RS instruments are used to detect soil moisture, and how that differs to the type of RS instruments that are used to derive LAI data, for the purposes of consistency.

Response: Microwave satellite data have a strong correlation with soil dielectric constant, and therefore microwave remote sensing is considered as an effective tool to measure soil water content (Petropoulos et al. 2015). Because atmospheric effects can be minimized and less energy is absorbed or reflected by vegetation at L-band, the L-band (12 GHz) is considered the best band for soil moisture retrieval. SMAP and SMOS (Jacquette et al. 2010) are the only two soil moisture specific satellites that are currently in orbit and are equipped with L-band microwave instruments. A verification analysis based on soil moisture measurements from 231 sites across the globe (Cui et al. 2018; Kim et al. 2018) showed that SMAP and SMOS products are superior to other soil moisture products (e.g., Advanced Scatterometer (ASCAT) and Advanced Microwave Scanning Radiometer 2 (AMSR2)). We chose the SMOS-L2 product and the SMAP-L3-Enhanced product, which both provide global coverage every three days for soil depth of 5 cm. By contrast, the GLASS LAI product used in this paper is generated from the optical reflectance data derived from the Advanced Very High Resolution Radiometer (AVHRR) and moderate resolution imaging spectrometer (MODIS) (Xiao et al. (2013)).

Cui, C., Xu, J., Zeng, J., Chen, K.-S., Bai, X., Lu, H., Chen, Q., & Zhao, T. (2018). Soil moisture mapping from satellites: An intercomparison of SMAP, SMOS, FY3B, AMSR2, and ESA CCI over two dense network regions at different spatial scales. Remote Sensing, 10, 33

Kim, H., Parinussa, R., Konings, A.G., Wagner, W., Cosh, M.H., Lakshmi, V., Zohaib, M., & Choi, M. (2018). Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products. Remote Sensing of Environment, 204, 260-275

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., & Song, J. (2013). Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance. IEEE Transactions on Geoscience and Remote Sensing, 52, 209-223

Petropoulos, G.P., Ireland, G., & Barrett, B. (2015). Surface soil moisture retrievals from remote sensing: Current status, products & future trends. Physics and Chemistry of the Earth, Parts A/B/C, 83, 36-56

13. Lines 122-124: Do the authors imply that they are assimilating global data, i.e. every grid cell of the products? This needs to be made clearer in Section 3.2. There have been other studies assimilating LAI and SM, even if they have not See Wu et al. (2018) as well as other papers from the same authors/group as the Bonan et al. (2020) paper. The introduction needs to be expanded beyond to reflect this history and how this study builds on that beyond just the assimilation of global data. Or at least, their hypothesis for how the assimilation of global data will be a step beyond those previous studies, but that that hypothesis needs to be evaluated in their analysis/results. In short, the authors need to do better at explaining, or demonstrating via analysis, why their study goes beyond the previous land DA studies assimilating LAI and SM. The authors need to answer the question "what do we learn from this study beyond what past studies have told us?". Points could be added to discussion too. This will help the modeling and DA community more widely discern the best practices and possible pitfalls for assimilation of these two datasets. If it is purely a technical advance (e.g. sheer scale of obs etc), then those advances and lessons learned should be highlighted more in this manuscript. The authors could add specific questions that they are trying to answer to the final paragraph of the introduction.

Response: No, they did not assimilate global data. Bonan et al. (2020) assimilated LAI and SM regionally, not at global scale. A description of this history was added to the introduction (Lines 125-142) and we summarized the difference of our study from previous studies and the progress made in this study (e.g., the use of datasets, assimilation methods, and regional analyses).

Methods:

14. Table 1: Is LPJ-VSJA used for assimilating data into LPJ-DGVM or LPJ-PM? I would have thought the latter?

Response: Yes, it is LPJ-PM. The relevant part of the manuscript has been revised.

15. Lines: 147-149: Not sure I understand here. There is or is not soil stratification in LPJ? And please could the authors explain how that connects to simulating water limited regions? I also think this sentence might be better after the authors have explained LPJ more generally.

Response: Due to the characteristics of water limitation in semi-arid regions, many studies have shown that surface soil moisture is the main factor controlling vegetation productivity

(Liu et al. 2020), and introducing surface soil moisture (SSM) into the model can significantly improve the simulation accuracy of GPP and ET in arid and semi-arid regions (He et al. 2017; Li et al. 2020).

There is soil stratification in LPJ. In the LPJ model, the soil is assumed to be barrel-shaped. The soil is vertically divided into two layers with a thickness of 0.5 m (upper layer) and 1m (lower layer).

$$E_S = Ep \times Wr_{20} \times (1 - fv)$$

In this module, it is assumed that the soil layer above 20 cm produces water through evaporation, and Wr_{20} is the relative water content of the soil above 20 cm, which is used as the only soil water limit for calculating vegetation transpiration and soil evaporation. In the evapotranspiration estimation, the over-simplification of soil structure and soil water limitation can lead to a large error (Sitch et al. 2003), while LPJ-DGVM cannot directly assimilate surface soil water due to the limitation of soil layer stratification. In addition, the model is driven by monthly data. The simulated daily soil moisture could not accurately reflect its diurnal variation, thus causing propagation errors for the simulated daily GPP and ET.

The above description of soil stratification in the LPJ model and the limitation of soil moisture are explained in Lines 200-209.

He, L., Chen, J.M., Liu, J., Bélair, S., & Luo, X. (2017). Assessment of SMAP soil moisture for global simulation of gross primary production. Journal of Geophysical Research: Biogeosciences, 122, 1549-1563

Li, S., Zhang, L., Ma, R., Yan, M., & Tian, X. (2020). Improved ET assimilation through incorporating SMAP soil moisture observations using a coupled process model: A study of US arid and semiarid regions. Journal of hydrology, 590, 125402

16. Line 152: Need much more information than this: "the GPP is calculated by implementing coupled photosynthesis and water balance" with references.

Response: The canopy GPP is updated daily:

$$GPP = \frac{(J_E + J_C - \sqrt{(J_E + J_C)^2 - 4\theta J_E J_C})}{2\theta} \quad (2.1)$$

where J_C is the Rubisco limiting rate of photosynthesis, J_E is the light limiting rate of photosynthesis, and the empirical parameter θ represents the common limiting effect between the two terms. J_E is related to APAR (absorbed photosynthetic radiation, product of FPAR and PAR), while J_C is related to Vcmax (canopy maximum carboxylation capacity, μ mol CO^2/m ^{2}/s) :

$$J_E = C_1 APAR \quad (2.2)$$

$$J_C = C_2 V_{C\,max} \quad (2.3)$$

where C_1 and C_2 are determined by a variety of photosynthetic parameters and the intercellular

partial pressure of CO_2, which is related to atmospheric CO_2 content and further altered by leaf stomatal conductance (Sitch et al. 2003). APAR and FPAR are directly related to LAI.

More detailed explanations and related formulae have been added in the manuscript. (Lines 183-197).

17. Lines 147-161: I feel like the reader needs a lot more basic information on LPJ and the PTJPL models. Perhaps they could have their own sections before describing how, and why, the models are combined?

Response: Similar to the ET calculation in the LPJ-DGVM model, the Priestley Taylor-Jet Propulsion Laboratory (PT-JPL) method provides three components of LE, i.e., soil evaporation (LE_S), vegetation transpiration (LE_T), and vegetation leaf evaporation (LE_I); the sum of which can be used to determine the total evapotranspiration (LE) based on the relationship between energy and water fluxes. In the updated PT-JPL model (hereafter referred to as PT-JPL$_{SM}$), Purdy et al. (2018) added a constraint (0-1) on the SM in transpiration and soil evaporation (Eq. 6) and included the calculation of soil heat (G), which helps to avoid the implicit definition of soil water control in the previous PT-JPL model.

In the LPJ-PM model, the LAI, canopy height, Maximum annual photosynthetically active radiation and soil texture were inherited by LPJ model. ET_{PM} was calculated by PT-JPL$_{SM}$. A relationship between the assimilated ET and soil moisture in process models is required to construct the connection between the assimilated system and model. The soil water content was calculated from the nonlinear soil water availability function using the assimilated ET and soil parameters. The soil moisture modeled by LPJ-DGVM in the next time step is replaced by the soil water content.

A basic description of the LPJ and PT-JPL models including the important formulas involved in the coupled models was added to section 2.1.

18.Line 167: What do the authors mean when they say "The SMAP SM was applied to model global ET using PT-JPLSM"? Do they mean the data was assimilated?

Response: No, the SMAP SM was used as an input to the PT-JPL$_{SM}$ model (Purdy et al. 2018).

19.Line 170: The authors talk about "scheme 2" here before talking about scheme 1? This is confusing. Please resolve.

Response: This paragraph mainly describes how SM is assimilated into the LPJ-PM model, that is scheme 2. This paragraph that describe the SM assimilation was moved to the section 2.2.2 in Lines 302-307 .

20.Line 169-176: I am a bit confused by what is going on in this paragraph. Please make it more clear for the reader.

Response: This paragraph mainly describes how the SMAP SM was assimilated into the model in the SM assimilation scheme. The assimilated ET_{SM} is superior to the ET_{PM} and ET_{LPJ} through site-level validation. This paragraph was revised and moved to Section 2.2.2 (Lines 303-307).

21.Line 185: Earlier you say "PODEN4DVAR".

Response: The acronym has been standardized as "PODEn4DVar".

22.Lines 190-205: This whole paragraph is difficult to parse as there are no sentences and instead there are a confusing number of semi-colons. I know the authors are describing steps, but I strongly encourage them to split this up into sentences. You can always start a sentence with Step 1 or Step 2 etc.

Response: The paragraph has been re-organized according to the suggestion.

23. Lines 201-202: which dataset did the authors use to define humid, semi-arid etc?

Response: The basis for distinguishing arid and humid regions is the classification system of global arid and humid regions in Middleton and Thomas (1997) that uses the "drought index" to classify different arid and wet regions. The drought index is defined as the ratio of precipitation to potential evapotranspiration. Regions with aridity index between 0.2 and 0.65 are defined as semi-arid regions, regions between 0.05 and 0.2 are arid regions, and regions below 0.05 are severely arid regions, that is, desert areas. The drought index of humid area and sub-dry humid area is about 0.65-3. In this paper, arid zone, semi-arid zone, humid zone and sub-dry humid zones are selected to evaluate the assimilation results in different regions.
The reference was added in Material Supplement S2.

Middleton, N., & Thomas, D. (1997). World atlas of desertification.. ed. 2. Arnold, Hodder Headline, PLC

24. Line 210: “propagated by energy transmission and ecosystem processes in the dynamic model “ ï□ Not clear what this means. should provide an explanation and references. Same for lines 234-235.

Response: Assimilated LAI was involved in the calculation of ecological and physiological processes of vegetation in the LPJ-DGVM model, such as photosynthesis. The explanation was added in Line 280.

26. Lines 218-221: What are the scale factors? What are the integration members? These have not been explained. I am confused again at lines 246-249.

Response: A brief explanation of scale factors has been added to Line 287. The integration number (ensemble member) was one of the parameters to be set in the assimilation method, which provides estimates of the background error covariance similar to the ensemble method in EnKF. The details of the assimilation algorithm are referenced as Evensen et al.,2004 (Lines 361).

27.Section 2.2.3: this is really a step-wise assimilation, rather than a true “simultaneous” joint assimilation. There are advantages and disadvantages to that should be discussed, and assumptions explained. See MacBean et al. (2016) for discussion.

Response: Yes, this is really a step-wise joint assimilation. We adopted step-wise assimilation because of technical constraints. As a nonlinear dynamic processed-model, LPJ-DGVM was simulated according to numerous physiological processes of vegetation; when it is running, SM and LAI could only be assimilated step by step inevitably. In the discussion section, the

influence of assimilation sequences on assimilation results, the importance of error correlation between parameters, the density of spatio-temporal information of observations, and the deviation between model and observations to the step-wise joint assimilation performance has been added in Lines 817-843.

28.Line 244-245: "Finally, GPPCO and ETCO were output by joint assimilation based on the POD-En4DVar method." I am confused here. This sentence reads like a separate joint assimilation is done when from earlier in the section/paragraph it seems like the LAI and SM/ET have already been assimilated?

Response: Yes, this sentence summarized the results of joint assimilation, which has been explained step by step in the previous paragraph. This sentence was deleted to avoid confusion.

29.Line 251: Earlier you said the "PODEn4DVAR" reference was Tian and Feng 2015.

Response: This reference was revised as Tian and Feng 2015.

30.Line 252: Explain what "POD base" is. And at line 269 please explain "POD decomposition".

Response: The POD decomposition technique is adopted to transform the original ensemble coordinate system into an optimal one in the L^2 norm (Ly and Tran, 2001), which contributes greatly to its enhanced assimilation performance. The POD base is the Transformed OP (Observing Perturbation) and MP (Model Perturbation. This was explained in Lines 333-334

Ly, H. V. and Tran, H. T. 2001. Modeling and control of physical processes using proper orthogonal decomposition. Math. Comput. Model. 33, 223-236

31.Line 254: "flow-dependent error estimates" please explain what this is for the non DA specialist.

Response: By forecasting statistical characteristics, the EnKF through ensemble method can provide flow-dependent estimates of the background error covariance. The flow-dependent is the ensembles of forecasting statistical characteristics in the t time, which is also explained in Line 336.

32.In general the number of subtext acronyms is difficult to parse. I suggest the authors find a slightly different way to refer to all the variables. For example, GPP_prior, GPP_scheme1, GPP_scheme2 etc.

Response: We feel that the subscript of assimilated observation data can make the result comparison more intuitive for readers in the result analysis, while the scheme serial number may cause confusion for readers. So the original subscripts are still retained.

Results:

32. Figure 2: Hard to tell what a,b,c,d is for each set of metrics just by looking at the figures. Figure text should be larger.

Response: Text "a, b, c, d" in the figure has been enlarged.

33.Figures 3 and 6 are not referenced in the text.

Response: The Figures 3 and 6 have been referred in the relevant text.

34.Figure 8: would be useful to put the labels "semi-arid" etc inside the actual subplots.

Response: We have added a label at the bottom right corner of each figure.

35.Lines 468-476: this is nice but it would be great to see the prior model-data comparison to see how the "CO" optimization has improved things. Otherwise, the authors' claim at line 476 that SM data are needed for water-limited areas is an overreach. Actually, without comparing to schemes 1 and 2 it is hard to say whether it is SM or LAI data that have achieved a good result in water-limited areas. The authors do seem to discuss the prior in the paragraph lines 485-490 but I am having trouble seeing where this fits into the bigger picture.

Response: The assimilation results of the LPJ-DGVM and the three schemes in wet and dry regions are analyzed in Tables S2 and S3 of the Supplementary Material.

For ET, the R^2 and ubRMSE implied that the SM assimilation alone had a better performance than the LAI assimilation alone, especially for sites in arid areas. And the bias showed that the ET_{LAI} improved better than ET_{SM} for sites in humid and sub-dry humid areas.

For GPP, the R^2 and bias implied that the LAI assimilation alone had a better performance than the SM assimilation alone. However, for sites in arid and semi-arid areas, the RMSE and ubRMSE showed that the accuracy of the GPP_{SM} improved much better than GPP_{LAI}, which both demonstrated SM data are essential for data assimilation in water-limited regions. These analyses were added to section 4.2(Line 577-579;589-592).

36.Line 496: do you mean Figure 7 here?

Response: The "Figure 7" has been revised to "Figure 9".

37.Figure 9: GPP_SM and ET_PM? I am confused here? Labels on the subplots would also help here.

Response: We have added a label at the upper right corner of each figure.

38.Figure 10: the color coded grid is helpful here.

Response: We thereby retain the color coded grid in the figure.

Discussion:

Generally a well-rounded discussion of the advantages and caveats of the approach. I would appreciate more discussion on the inconsistency between LAI products in Section 5.3, and implications of the fact the assimilated products (LAI and SM) may be biased. What impact do the authors think that would have on the results? Also issues related to temporal sampling interval could be discussed somewhere in the discussion, as well as assumptions/caveats of the DA method that may affect the results.

Response: In the discussion section, the inconsistency of LAI products, the analysis of the assimilation results (reasons of bias), and the influence of ensemble size, error setting and temporal sampling interval on assimilation performance have been discussed in Lines 817-843.