Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-504-RC1, 2019

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "The Influence of Assimilating Leaf Area Index in a Land Surface Model on Global Water Fluxes and Storages" by Xinxuan Zhang et al.

Anonymous Referee #1

Received and published: 11 November 2019

The authors would like to thank the reviewer for their time, effort, and detailed comments. All suggestions were incorporated into the manuscript and explained in this response to reviewer document. We also thoroughly proofread and revised the whole manuscript.

General comments

The authors aim to assess to what extent the Noah-MP model can be optimized through the assimilation of leaf area index (LAI) observations at global scale. By utilizing two observing system simulation experiments (OSSEs) and the EnKF algorithm, the efficiency of assimilating LAI and model performance for water related variables are discussed. At first in my opinion this manuscript needs to be proofread/revised carefully for academic writing.

We would like to thank the reviewer. We have carefully proofread the revised manuscript.

Something that I do not understand is that the authors use the simulated LAI from the nature run as the 'truth' instead of observations. If nature run can achieve the "truth", why did the authors conduct assimilation based on different conditions (wet or dry)?

We chose to use an Observing System Simulation Experiment (OSSE) to quantify the potential impact of LAI assimilation on water variables simulated by the Noah-MP model while the forcing precipitation is affected by severe biases.

The forcing precipitation is usually provided by either reanalysis or satellite products. Such products are often affected by large biases (and random errors), which consequently affect the accuracy of the modeled variables. The question we want to answer here is: *When the forcing precipitation is biased, is LAI data assimilation able to improve the model estimates?* A real case study would certainly be of interest but in-situ observations (taken as reference) would also be affected by uncertainty, making it difficult to draw meaningful conclusions regarding the methodology itself. The proposed OSSE should serve as a feasibility test to quantify the potential of the proposed framework.

In an OSSE, i) the nature run (NR) intends to mimic the true input (including *unbiased* precipitation), LAI, and all water variables, ii) the open-loop run (OL) adds biases to the forcing precipitation (i.e., double or half the original value) to mimic the error in the precipitation product which will also produce biased model outputs of LAI and water variables; iii) the data assimilation (DA) run applies LAI DA to the OL run. We named the model run with double precipitation as wet condition, and named the run with half precipitation as dry condition to describe the wet/dry bias that these two runs represent.

We modified the manuscript to clarify the OSSE design in section 2.2:

"*First, the Noah-MP model is spun-up for a 10-year period (2001-2010) to ensure a physically realistic state of equilibrium. Second, the model is run for a 29-month period (January 2011 – May 2013) to conduct the Nature Run (NR) with the same configuration as the spin-up one. By definition, an OSSE is a controlled experiment that does not assimilate any real observation. Instead, it treats all the model outputs from the NR as the "true" condition (denoted as the "synthetic truth"). The "true" LAI (i.e., the LAI output from NR) is then perturbed via a simple additive error model to produce the synthetic observations to be assimilated into the DA runs. The spin-up run and NR are forced by the original MERRA-2 precipitation data. Third, two Open Loop (OL) runs (no DA) are conducted for the same 29-month period under two conditions: i) "extremely dry" condition (the model is forced by halving the MERRA-2 precipitation data; OL-dry), and ii) "extremely wet" condition (the model is forced by doubling the MERRA-2 precipitation; OL-wet). The biased forcing precipitation data in OL mimic typical precipitation biases in current precipitation reanalysis and satellite products (e.g., Ghatak et al. 2018; Yoon et al. 2019).The two DA runs are then conducted under the two same conditions (DA-dry and DA-wet) using a one-dimensional EnKF assimilation algorithm, which is a built-in DA method in LIS ……*"

Other important comment is that why did the authors use the precipitation which are extremely biased instead of using a more precise precipitation forcing. Furthermore, did the authors run the assimilation experiment using the MERRA-2 precipitation instead of halving or doubling the value?

The 10-year spin-up run and the nature run are forced by the original MERRA-2 precipitation data. The OL and DA runs are forced by a perturbed (i.e., biased) version of the MERRA-2 precipitation. As described above, the OSSE uses this input in the OL and DA runs to mimic common biases in currently available precipitation products.

In conclusion, the manuscript in its current form suffers from several issues that prevent it to be published as is. In my opinion the paper still worth to be published after addressing all these issues, and a major revision is asked.

Specific comments

1. P3L56-57: As far as I know, LSMs not only couple with dynamic vegetation models, but also involve some dynamic vegetation modules. So the statement is not appropriate.

We changed "*dynamic vegetation model*" to "*dynamic vegetation module*" in the manuscript.

2. Section 2.2: Why do you use the precipitation forcing data which are strongly biased.

In the OSSE study, we use biased precipitation data in OL and DA runs to mimic precipitation biases that are very common in current precipitation reanalysis and satellite products (e.g., Ghatak et al. 2018, Yoon et al. 2019. These two references have been added to text in section 2.2:

Ghatak, D., Zaitchik, B., Kumar, S., Matin, M. A., Bajracharya, B., Hain, C., & Anderson, M. (2018). Influence of Precipitation Forcing Uncertainty on Hydrological Simulations with the NASA South Asia Land Data Assimilation System. Hydrology, 5(4), 57. https://doi.org/10.3390/hydrology5040057

Yoon, Y., Kumar, S. V., Forman, B. A., Zaitchik, B. F., Kwon, Y., Qian, Y., Rupper, S., Maggioni, V., Houser, P., Kirschbaum, D., Richey, A., Arendt, A., Mocko, D., Jacob, J., Bhanja, S., & Mukherjee, A. (2019). Evaluating the Uncertainty of Terrestrial Water Budget Components Over High Mountain Asia. Frontiers in Earth Science, 7. https://doi.org/10.3389/feart.2019.00120

3. Why did you choose the LAI simulations from the nature run as the “truth” instead of using the LAI observations? As you have described the reasons from P9L171 to L172, there are many other LAI products without missing data which can be used for assimilation.

By definition, an OSSE is a controlled experiment that does not assimilate any real observation. Instead, it treats all the model output from the nature run as the “true” condition. The LAI from the nature run is also considered as the true. We then perturbed it with a simple additive error model to produce synthetic observations to be assimilated into the model (DA run). Some explanation was added in section 2.2:

“*Second, the model is run for a 29-month period (January 2011 – May 2013) to conduct the Nature Run (NR) with the same configuration as the spin-up one. By definition, an OSSE is a controlled experiment that does not assimilate any real observation. Instead, it treats all the model outputs from the NR as the “true” condition (denoted as the “synthetic truth”). The “true” LAI (i.e., the LAI output from NR) is then perturbed via a simple additive error model to produce the synthetic observations to be assimilated into the DA runs. The spin-up run and NR are forced by the original MERRA-2 precipitation data.”*

4. Did you evaluate the LAI or other variables from the natural run by using remote sensing LAI datasets or other kinds of observations?

As mentioned above, The LAI from the nature run is considered as the truth in the OSSE framework. The same LAI is perturbed with a simple additive error model to produce synthetic observations of LAI that are assimilated in the DA experiment. The LAI from OL or DA run is evaluated against the synthetic LAI observation from the nature run.

5. P9L178-P9L184: How did you determine the values of multiplicative perturbations (such as, the shortwave radiation and precipitation with a mean of 1 and standard deviations of 0.3 and 0.5, the standard deviation for longwave radiation of 50 W/m2, the standard deviation for LAI of 0.1)?

The forcing data perturbation applied here used the same perturbations as found in the literature below. We mentioned these past studies in section 2.2 “*Similar to previous work (Kumar et al. 2014, 2019a, 2019b), the MERRA-2 forcing inputs such as shortwave/longwave radiations and precipitation are perturbed hourly……*”.

Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., Xia, Y., Ek, M., Riggs, G., Livneh, B. and Cosh, M.: Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation, J. Hydrometeorol., 15, 2446-2469, https://doi.org/10.1175/JHM-D-13-0132.1, 2014.

Kumar, S. V., Jasinski, M., Mocko, D. M., Rodell, M., Borak, J., Li, B., Beaudoing, H. K. and Peters-Lidard, C. D.: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system

for the National Climate Assessment, J. Hydrometeorol., 20, 1571-1593, https://doi.org/10.1175/JHM-D-17-0125.1, 2019.

Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D. and Borak, J.: 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 US, J. Hydrometeorol., 20, 1359-1377, https://doi.org/10.1175/JHM-D-18-0237.1, 2019.

6. Have the evaluation and error metrics been used in former studies? If so, please list at least one references.

The equation for the Normalized Information Contribution (NIC) index is similar to the NIC used by Kumar et al. 2016. We added this reference to the text:

Kumar, S.V., Zaitchik, B.F., Peters-Lidard, C.D., Rodell, M., Reichle, R., Li, B., Jasinski, M., Mocko, D., Getirana, A., De Lannoy, G. and Cosh, M.H.: Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System, J. Hydrometeorol., 17, 1951-1972, https://doi.org/10.1175/JHM-D-15-0157.1, 2016

7. How did you determine the initial conditions?

Initial conditions are obtained by a 10-year spin-up run. The spin-up run is described in the second paragraph of section 2.2.

8. The discussion section should include the discussion of the results in the context of other papers dealing with the same of similar subjects.

We added some discussion of past work on similar subjects:

"*Overall the improvement of water variables through LAI assimilation is not remarkable enough to compensate the model degradation caused by the biased precipitation forcing data. Previous studies (Pauwels et al. 2007; Sabater et al. 2007; Barbu et al. 2011; Fairbairn et al. 2017; Albergel et al. 2017) have tested the performance of the joint assimilation of LAI and soil moisture over regional domains and showed promising results. However, no experiment was performed at the global scale. Future work could investigate a multi-variate data assimilation system that concurrently merges both LAI and soil moisture (or TWS) observations globally.*"

9. A more in-depth analysis of the results is necessary. In this paper the authors only talk about the statistical characteristic variables (such as the NCRMSE, NIC, etc) of LAI and water related variables. Why not focus on the LAI and water related variables themselves?

Sections 3.1 and 3.2 have been modified to add more analyses. Time series of global averaged LAI and water variables (Figure 3) were also added to the manuscript to provide more information on the actual variables (rather than anomalies). The discussion section (3.3) was also modified to provide more in-depth interpretation of the results.

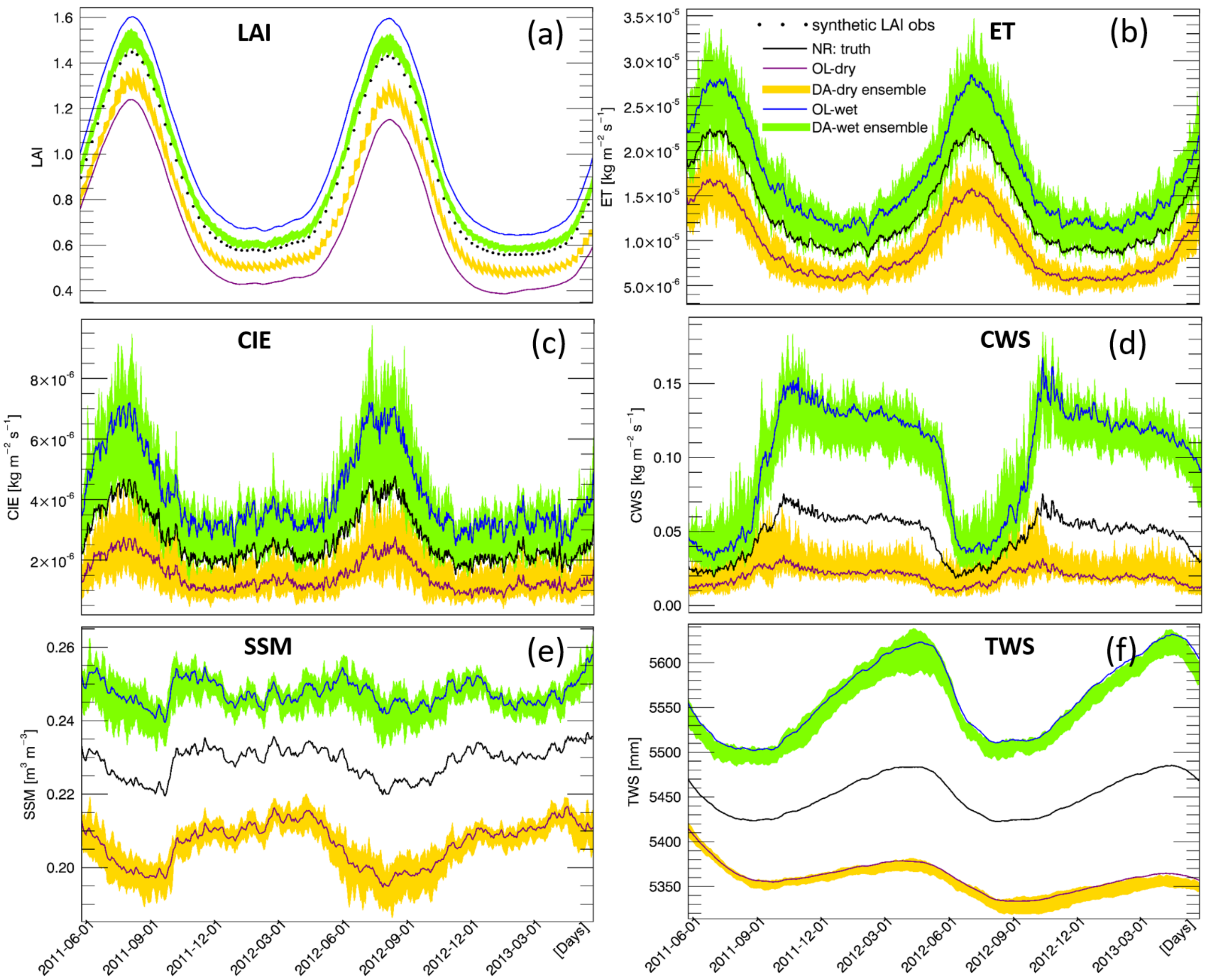


Figure 3. Global averaged daily values of LAI and five water variables (2011-06-01 to 2013-05-30).

10. Why only perturb the meteorological forcing and not the initial conditions and/or model parameters?

We perturbed precipitation and radiation forcings because deemed dominant in water variable simulated by land surface models. Perturbing initial condition and model parameters is certainly an option that could be investigated in future studies. This recommendation has been added to the conclusion section.

"*Future research should focus on alternative DA methods, such as updating other related model states while assimilating LAI observations, perturbing the model initial condition and model* *parameters, and/or assimilating actual satellite-based LAI observations (e.g., MODIS, GLASS) at the global scale to verify the efficiency of the proposed vegetation DA framework.*"

11. How sensitive is LAI with respect to the meteorological forcing?

LAI is very sensitive to the forcing precipitation data. The wet and dry conditions have large impacts on the magnitude of LAI. The revised manuscript shows the time series of LAI values and anomalies. Below are the figures and description we added:

Section 3.1 LAI

"*Figure 3a and Figure 4a show time series of global averaged LAI values and corresponding anomalies, respectively. As expected, LAI values are largely impacted by the extreme precipitation conditions. The wet condition introduces more vegetation, while the dry condition limits the vegetation growth throughout the two-year period. The DA procedure effectively corrects the LAI errors caused by the biased precipitation input.*"

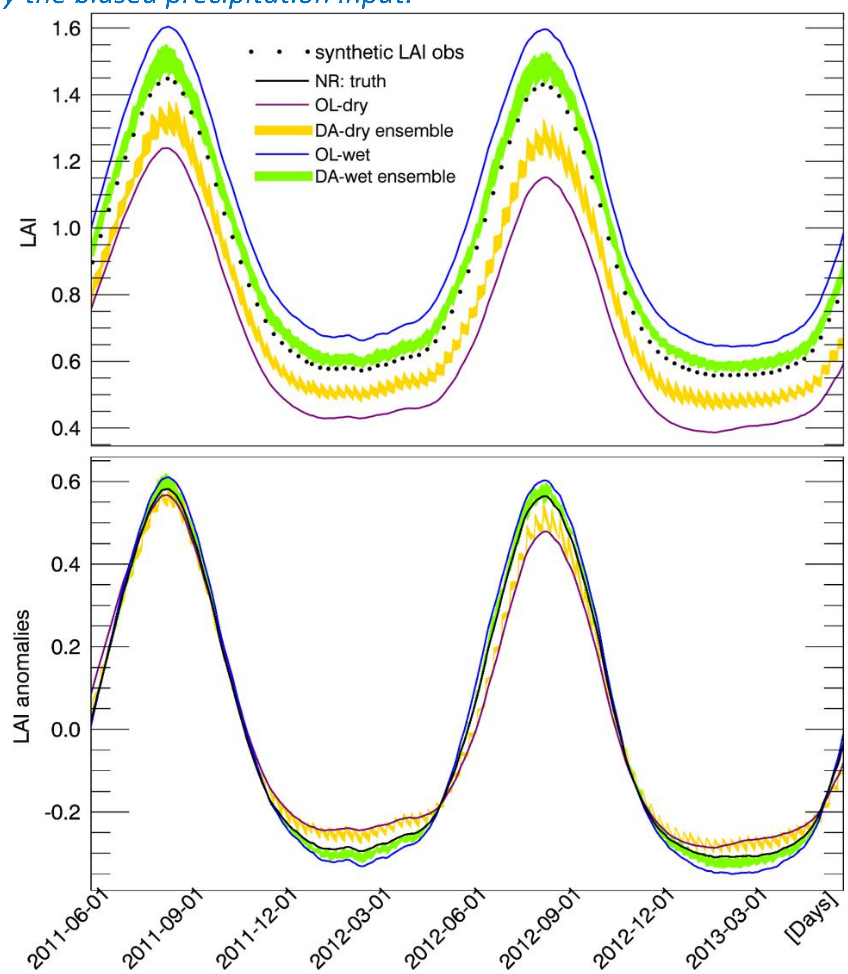


Figure 3a (top) and Figure 4a (bottom): Global averaged daily values of LAI and LAI anomalies

Technical corrections

1. P2L27-L28: Can you illustrate which land surface model you use here? And the same to P2L38, P5L104, and so on.

We added "Noah-MP" to these three sentences.

2. P2L28-L29: Remove "the" from the phrase of "at the global scale", and the same to P5L100, P5L100, P22L361, and so on.

"the" was removed from "at the global scale" phrases.

3. P3L44: Do not need to leave two blank spaces here.

This has been fixed.

4. P3L46: It's not appropriate to use "between" among vegetation, precipitation, and soil moisture.

"between" was changed to "among".

5. P3L51: The related references cited here are not enough to illustrate the phenomenon that "these land surface processes and feedbacks have been examined through numerical modeling experiments". List more ......

More references have been listed: "*Foley et al. 1996; Kim and Wang 2007; Druel et al. 2019*"

6. P3L54: You needn't capitalizes the first letter for leaf area index.

This was fixed.

7. P4L67: "the Moderate Resolution Imaging Spectroradiometer" has been abbreviated to "MODIS" before.

This was removed.

8. P4L88-P5L90: Please refine this sentence.

The sentence was rewritten as "*Some water budget variables were improved through the assimilation procedure. The improvement is remarkable in agricultural areas because the assimilation added harvesting information to the model.*"

9. P5L95: Change "model simulated LAI" to "simulated LAI".

Removed.

10. P5L97: Please refine the statement of "focused on small regions".

Changed to "*...... and most of them are small region studies*".

11. P5L106-L107: Please define the abbreviation of all the water related variables when they first appear in this manuscript. Furthermore, "evapotranspiration" has been abbreviated to "ET" in P5L93.

The revised manuscript defined the abbreviation of all the water variables when they first appeared.

12. P5L110: Please specify which land surface model.
The title of section 2.1 was changed to "2.1. Land surface model (Noah-MP)".

13. P6L116-120: Please refine this sentence as it is too long.
The long sentence was divided into two sentences. "*Specifically, the prognostic vegetation growth combines a Ball-Berry photosynthesis-based stomatal resistance (Ball et al. 1987) with a dynamic vegetation model (Dickinson et al. 1998). The dynamic vegetation model calculates the carbon storages in various parts of the vegetation (leaf, stem, wood, and root) and the soil carbon pools.*"

14. P6L121: Please define "NASA".
Defined NASA as National Aeronautics and Space Administration.

15. P6L126: Keep the tense consistent.
Changed to "*The …….. (MERRA-2 ……) dataset serves as the meteorological forcings for Noah-MP.*"

16. P6L133-P7L138: Please define the abbreviation of all the water related variables when they first appear in this manuscript.
The revised manuscript defined the abbreviation of all the water variables when they first appeared.

17. P7L150: I am not sure whether the state of "a LAI EnKF" is appropriate.
Changed to "*the EnKF LAI assimilation*".

18. P7L153: The phase of "on a global scale" is not appropriate.
Changed to "*The proposed framework is evaluated through a global experiment (Antarctica excluded) at the 0.625° × 0.5° spatial resolution of the MERRA-2 forcing dataset (Figure 1).*"

19. P10L188-L189: Keep the tense consistent.
The sentence was changed to "*Thus, all the DA simulations are run for 20 members.*"

20. P10L194-L195: The water related variables have been defined before, and you can use their acronyms.
The variable names were changed to their acronyms.

21. P10L203: What does i and N in Equation 1 mean?
This explanation was added to the manuscript: "*N is the total number of X values, and i represents the index of each X value.*"

22. P10L208: the word “O” in the denominator looks like “zero” in Equation 2.
The letter “*O*” in “*OL*” does look similar to “zero”, though “zero” is thinner. Hope the readers won’t get confused.

$$C = \frac{E_{DA} - E_{OL}}{0 - E_{OL}}$$

23. P11L209: There are two periods.
Removed.

24. P12L220-L222: As Figure 3 shows the GLOBAL averaged LAI anomalies, it is better to use the statement of month (or JJA and SON seasons) instead of winter/summer season.
Changed to “*Moreover, the seasonality of LAI anomalies is evident, showing larger variations in DJF and JJA than during the transition periods (MAM and SON).*”

25. P12L229: Please refine this sentence.
The sentence was rewritten as “*Moreover, the DA runs show lower NCRMSEs than the corresponding OL runs across the globe (Figure 4) especially over shrublands and grasslands (refer to Figure 1 for land covers).*”

26. P12L241: Remove “the”. Furthermore, this sentence is a little too long in my opinion.
Changed “*In the summer*” to “*In JJA*”.
The long sentence was divided into two short ones. “*In JJA, the vegetation leaves in the north hemisphere are fully developed and the plants can use stomatal closure to preserve water under water limited condition (dry condition). Thus, the NCRMSE of dry condition becomes smaller and does not show much difference from the wet condition.*”

27. P14L263: Please change the “has higher chance” into “is more likely to”.
Modified.

28. P14L268-269: I think this is the first appearance that positively biased is wet condition (or negatively biased is dry condition), or maybe earlier, and this statement does not need to be repeated each time it appears in this paper (see P14L277, P16L296-L297, P21L337, P21L339).
Most of the repetitions were removed. We kept the “wet/dry” in the conclusion section in case some readers check the conclusion before going through the whole manuscript.

29. P15L282-L287: It is better to use the statement of month instead of season.
All the season names in the manuscript were substituted by month.

30. Please add the description for the Y-coordinate for Figure 7, 8 and 9.
The Y-axis titles are added to Figure 7, 8 and 9 in the revised manuscript.

31. P21L357: Please specify which land surface model.

This was added to manuscript: "*This study evaluates the efficiency of assimilating vegetation information (i.e., LAI synthetic observations) within a land surface model (Noah-MP 3.6) when the precipitation forcing data are strongly biased (either positively or negatively).*"

Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-504-RC2, 2019

Hydrology and Earth System Sciences Open Access
Discussions

Interactive comment on "The Influence of Assimilating Leaf Area Index in a Land Surface Model on Global Water Fluxes and Storages" by Xinxuan Zhang et al.

Anonymous Referee #2
Received and published: 5 December 2019

The authors would like to thank the reviewer for their time, effort, and detailed comments. All suggestions have been incorporated into the manuscript and explained in this document. We also thoroughly proofread and revised the whole manuscript.

General comments

Synthetic observations are used to assess the impact of assimilating satellite-derived LAI estimates into the Noah land surface model. A major shortcoming of the assimilation system used in this study is that LAI assimilation has no direct impact on soil moisture. As a result, dry precipitation biases cannot be compensated for. This issue was at least partly solved in other assimilation systems. Unfortunately, the relevant literature is not completely cited. This paper is not well written, not complete for understanding, and cannot be published in the present form. Methods description is incomplete. Interpretation of results is made in the Result section instead of the Discussion section.

In the dry condition simulation, the amount of vegetation is less than in the reference simulation (what we call synthetic truth), due to a decrease (or even lack) in precipitation. When assimilating observations of LAI, we introduce more vegetation into the model, bringing it closer to the synthetic truth and consequentially improving CIE (canopy interception evaporation) and CWS (canopy water storage), which are directly related to LAI. However, variables that are not directly impacted by LAI, such as SSM (surface soil moisture), can hardly be improved by LAI assimilation solely. The poor performance of SSM in the dry condition experiment is mainly attributed to the fact that the amount of water in the model is conservative. Specifically, LAI assimilation introduces more vegetation, which requires more water than what available in the system (i.e., soil). Past work attempted to solve this problem by jointly assimilating LAI and soil moisture (Pauwels et al. 2007; Sabater et al. 2007; Barbu et al. 2011; Fairbairn et al. 2017; Albergel et al. 2017). We added some discussion on this topic and cited all these articles in discussion section of the revised manuscript (more detail is provided in our response to the Reviewer's specific comments below).

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., ... & Le Moigne, P. (2017). Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8. 0: LDAS-Monde assessment over the Euro-Mediterranean area. Geoscientific Model Development, 10(10), 3889-3912.

Barbu, A. L., Calvet, J. C., Mahfouf, J. F., Albergel, C., & Lafont, S. (2011). Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: grassland case study. Biogeosciences, 8(7), 1971-1986.

Fairbairn, D., Barbu, A., Napoly, A., Albergel, C., Mahfouf, J. F., & Calvet, J. C. (2017). The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. Hydrology and Earth System Sciences, 21(4), 2015-2033.

Pauwels, V. R., Verhoest, N. E., De Lannoy, G. J., Guissard, V., Lucau, C., & Defourny, P. (2007). Optimization of a coupled hydrology–crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter. Water Resources Research, 43(4).

Sabater, J. M., Rüdiger, C., Calvet, J. C., Fritz, N., Jarlan, L., & Kerr, Y. (2008). Joint assimilation of surface soil moisture and LAI observations into a land surface model. Agricultural and forest meteorology, 148(8-9), 1362-1373.

Recommendation: major revision.

Particular comments:

- L. 39-40: Examples of joint assimilation of LAI and soil moisture in a land surface model can be found in the literature.

We have reviewed several studies that used LAI-soil moisture joint assimilation (Pauwels et al. 2007; Sabater et al. 2007; Barbu et al. 2011; Fairbairn et al. 2017; Albergel et al. 2017) and cited them in the discussion section:

"*Overall the improvement of water variables through LAI assimilation is not remarkable enough to compensate the model degradation caused by the biased precipitation forcing data. Previous studies (Pauwels et al. 2007; Sabater et al. 2007; Barbu et al. 2011; Fairbairn et al. 2017; Albergel et al. 2017) have tested the performance of the joint assimilation of LAI and soil moisture over regional domains and showed promising results.*"

All the cited LAI-SM joint DA studies were conducted over regional domains. We emphasized our study is "at global scale" in the end of the discussion section to make the statement more accurate: "*However, no experiment was performed at the global scale. Future work could investigate a multi-variate data assimilation system that concurrently merges both LAI and soil moisture (or TWS) observations globally.*"

- L. 95-97: In the same context and at the continental scale, Albergel et al. showed that sequential LAI assimilation can be used to analyse soil moisture at various depth, in addition to vegetation biomass (https://doi.org/10.5194/gmd-10-3889-2017). This property is particularly useful in dry conditions, when surface soil moisture tends to be decoupled from deeper soil layers.

Thank you for pointing us to this reference. We added it to the manuscript:

"*Only a few studies discussed the influences of LAI assimilation on the estimation of water variables such as soil moisture or streamflow (Pauwels et al. 2007; Sabater et al. 2008) and most of them focused on limited regions. Most recently, Albergel et al. (2017) conducted a study on a much larger domain – Europe and the Mediterranean basin –and showed improvement in soil moisture at various depths thanks to LAI assimilation.*"

We also added this study to the discussion section as shown in the answer above.

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbu, A. L., ... & Le Moigne, P. (2017). Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8. 0: LDAS-Monde assessment over the Euro-Mediterranean area. Geoscientific Model Development, 10(10), 3889-3912.

- L. 109 (Section 2): A section describing the DA method is needed. What are the analysed variables? Does LAI DA impacts soil moisture?

The DA method in this study is implemented within the NASA Land Information System (LIS). The method has been applied in many sequential data assimilation studies. We added brief description and references for the DA method in section 2.2:

"*The two DA runs are then conducted under the two same conditions (DA-dry and DA-wet) using a one-dimensional EnKF assimilation algorithm, which is a built-in DA method in LIS. The EnKF DA algorithm is suitable for non-linear and intermittent land surface processes (Reichle et al. 2002a, 2002b). Details of the algorithm can be found in numerous previous studies (Reichle et al. 2010; De Lannoy et al. 2012; Liu et al. 2015; Kumar et al. 2019a).*"

- L. 138: How are subsurface waters represented? Do you represent inundations plains? Lakes?

TWS is the sum of snow water equivalent, surface water, soil moisture, and groundwater. So, subsurface water (i.e., groundwater) is included. Lakes and inundation plains are considered as surface water, which is also included in TWS. This information was added to Section 2.1 as follows:

"…and TWS (defined as the sum of all water storage on the land surface and in the subsurface, including snow water equivalent, surface water, soil moisture, and groundwater [mm])."

- L. 188 (ensemble members): How is this ensemble generated?

The model ensemble is generated by perturbing the meteorological forcing inputs (precipitation and shortwave/longwave radiations). Section 2.2 discusses all the details:

"*Similar to previous work (Kumar et al. 2014, 2019a, 2019b), the MERRA-2 forcing inputs such as shortwave/longwave radiations and precipitation are perturbed hourly. Multiplicative perturbations are applied to the shortwave radiation and precipitation with a mean of 1 and standard deviations of 0.3 and 0.5, respectively. The longwave radiation is perturbed via an additive perturbation with a standard deviation of 50 W/m2. The perturbations of the three meteorological forcing variables also include cross correlations: cross correlation between shortwave radiation and precipitation is -0.8, cross correlation between longwave radiation and precipitation is 0.5; and cross correlation between shortwave and longwave radiations is -0.5."*

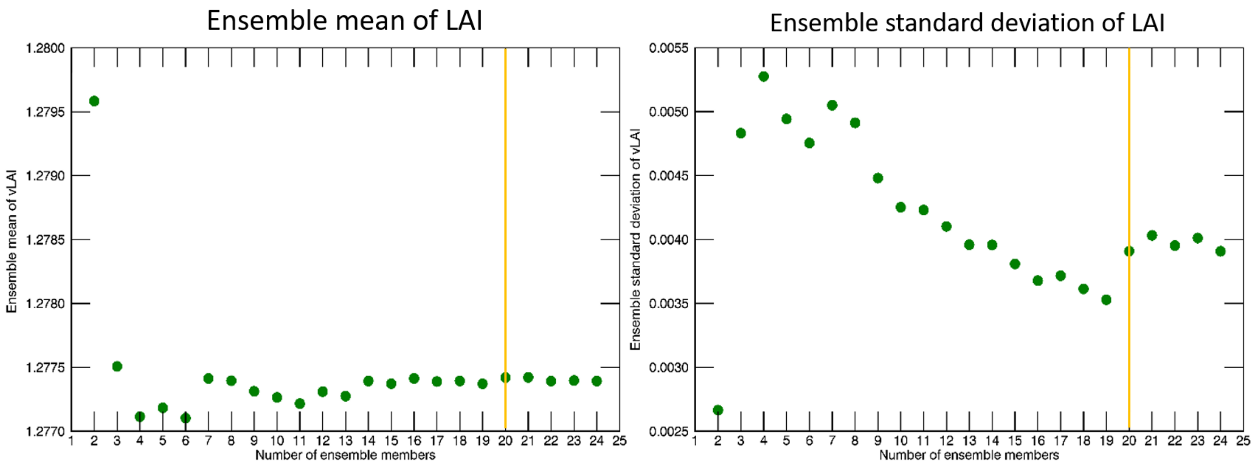
Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., Xia, Y., Ek, M., Riggs, G., Livneh, B. and Cosh, M.: Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation, J. Hydrometeorol., 15, 2446-2469, https://doi.org/10.1175/JHM-D-13-0132.1, 2014.

Kumar, S. V., Jasinski, M., Mocko, D. M., Rodell, M., Borak, J., Li, B., Beaudoing, H. K. and Peters-Lidard, C. D.: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment, J. Hydrometeorol., 20, 1571-1593, https://doi.org/10.1175/JHM-D-17-0125.1, 2019.

Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D. and Borak, J.: 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 US, J. Hydrometeorol., 20, 1359-1377, https://doi.org/10.1175/JHM-D-18-0237.1, 2019.

Moreover, an ensemble size sensitivity test was conducted to choose the number of ensemble members needed in this study (please refer to the Figure below). We omitted this figure in the manuscript, but mentioned the sensitivity test in section 2.2:

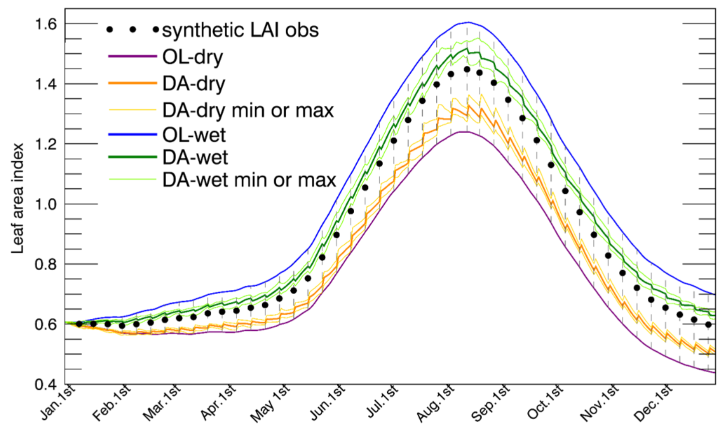
"To select the optimal ensemble size, a sensitivity test is performed for ensemble sizes spanning from 2 to 24 members. The number of ensemble members has a strong impact on the model results at small sizes, while the model performance tends to become steady when more than 20 ensemble members are considered. Thus, all the DA simulations are run for 20 members."



- L. 196-197: Why are these instabilities generated by DA?

The initial condition for the OL and DA runs is generated by a 10-year spin-up run that uses the original MERRA-2 precipitation. The OL and DA runs are forced by either doubled or halved precipitation that is not consistent with the spin-up run. So, the model needs to run for a certain time before stabilizing. The figure below shows the global averaged LAI time series from the beginning of the simulation (Jan. 1st, 2011) to Dec. 31st, 2011. The LAI simulated by OL and DA runs does not get stable until around May. Therefore, we decided to eliminate the first 5-month model outputs in the analyses. We added this explanation in the manuscript in section 2.3.

"*The initial condition for the OL and DA runs is generated by a spin-up run that uses the original MERRA-2 precipitation as input. However, the OL and DA runs are forced by either doubled or halved precipitation, which is not consistent with the spin-up run and the model needs some time to stabilize. The first 5-month model outputs are therefore eliminated from the evaluation to avoid the model systematic instability at the beginning of the OL and DA simulations and the evaluation, thus, focused only on model outputs from 2011-06-01 to 2013-05-31.*"



- L. 203 (Eq. 1): Why do you use NCRMSE and not standard score metrics such as RMSE or ubRMSE (i.e. standard deviation of differences)?

In the manuscript, we compared the result of LAI and five water related variables (ET, CIE, CWS, SSM, and TWS). Units of these variables are very different, which is why we decided to adopt unitless statistical metrics. UbRMSE is certainly another valid option.

- L. 217 (Figure 3): Please change evaporation units. Since these time series are daily, should be per day instead of per second. It seems that CWS anomalies are 3 order of magnitude larger than ET anomalies. Why? Define here what you mean by "anomaly" (not defined in the text). NR anomalies: with respect to what? Is NR the benchmark or not? Real values have to be showed at some stage. Not only anomalies.

The ET value shown in the figure refers to the model output, which is an average over the day with the unit “kg m-2 s-1”. CWS is the canopy water storage which include the water stored in the leaves and the intercepted water. So, it is much larger than the ET.

We analyzed anomalies (rather than actual values) because they are unitless and this is good practice when comparing the impact DA has on different variables. The anomalies are defined in section 2.3. “Each of the anomaly time series is computed relative to the mean of its respective model run.” The NR anomalies are calculated with the respect to the mean of NR run. OL anomaly is calculated with the respect to the mean of OL. DA anomalies follow the same rule.

Nevertheless, we understand the value of showing actual values and, in the revised manuscript, we added these time series as Figure 3. Please check below for the figures of variable actual values and anomalies and the related descriptions we added in the revised manuscript.

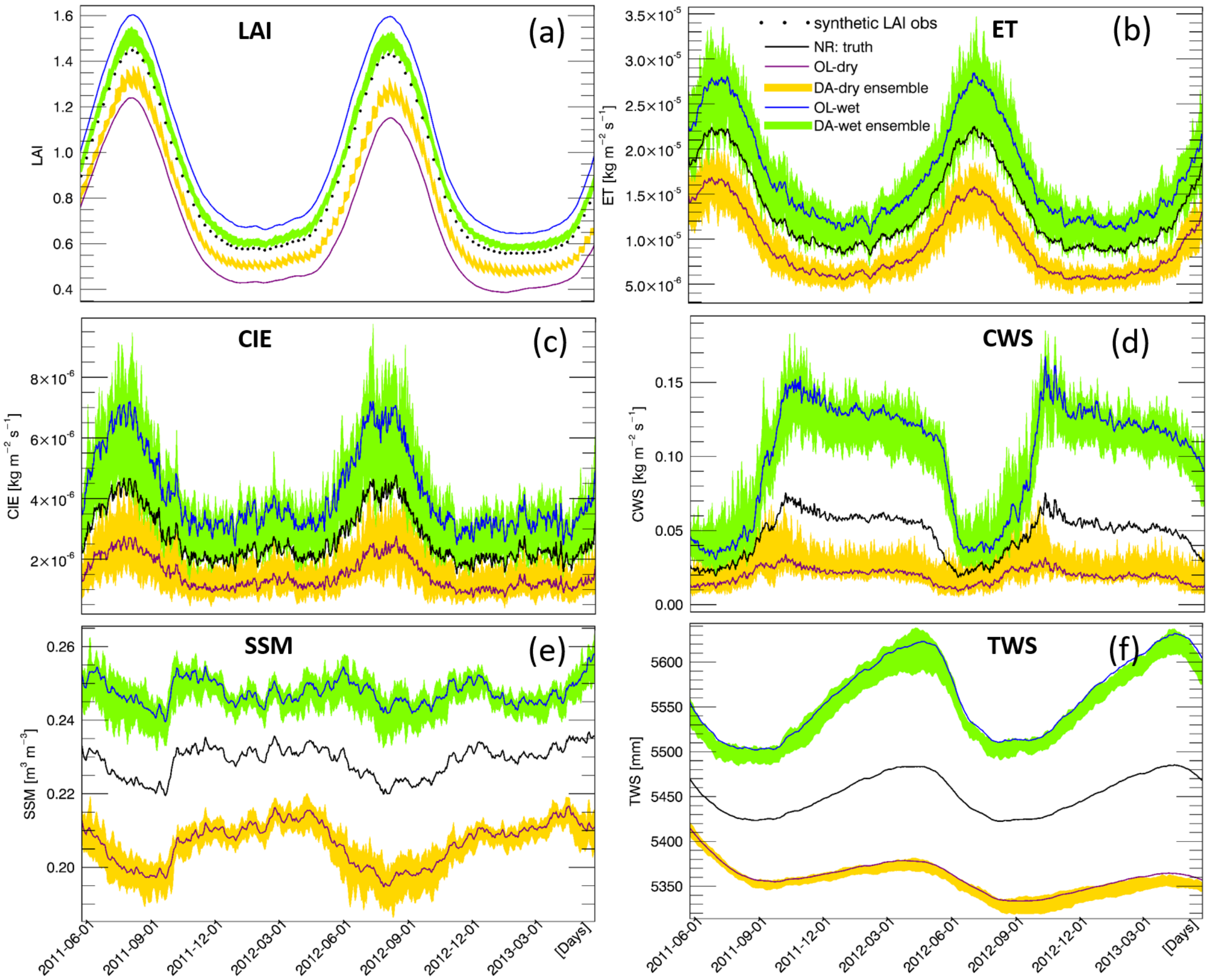


Figure 3. Global averaged daily values of LAI and five water variables (2011-06-01 to 2013-05-30).

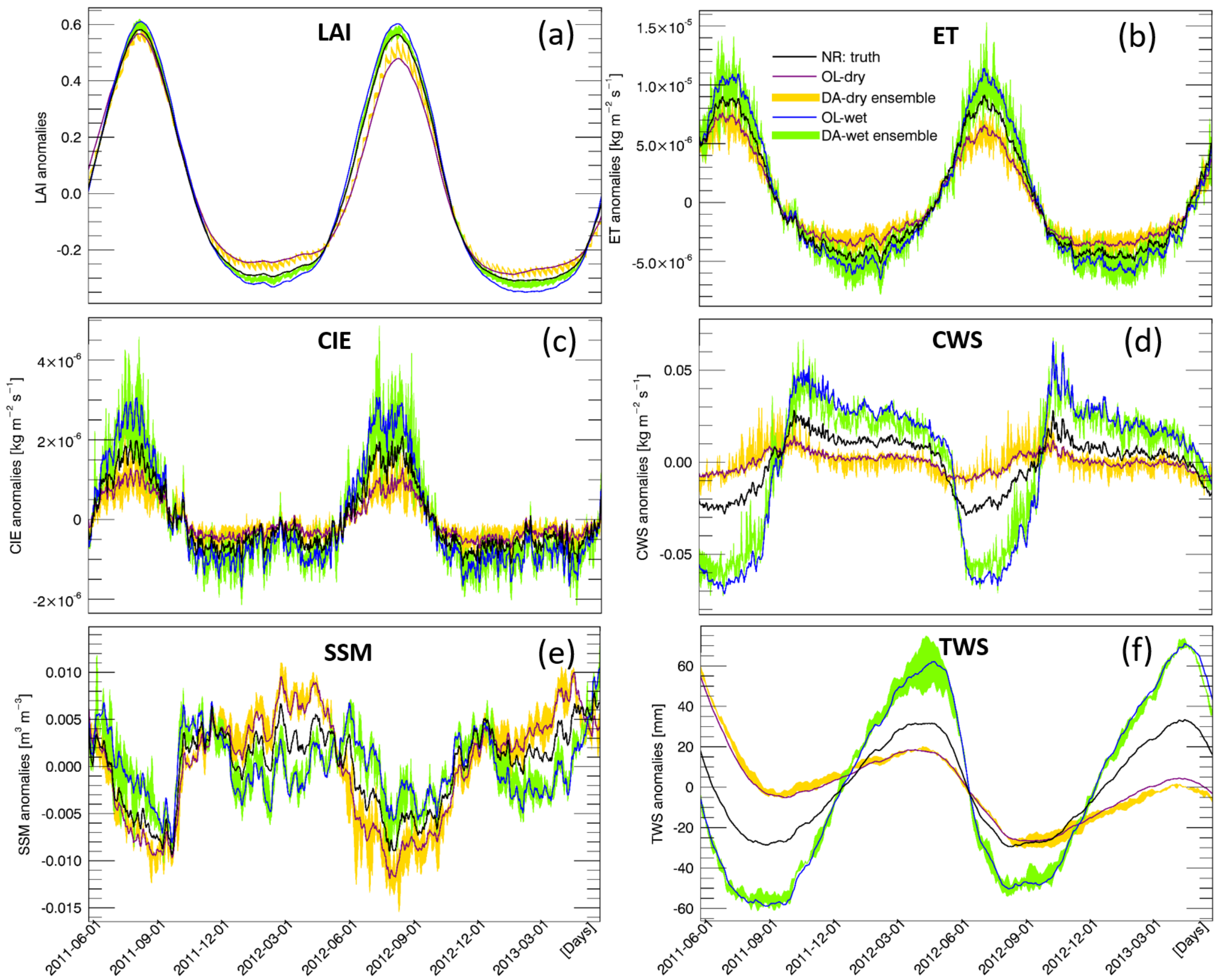


Figure 4. Global averaged daily anomalies of LAI and five water variables (2011-06-01 to 2013-05-30).

Section 3.1 LAI

"*Figure 3a and Figure 4a show time series of global averaged LAI values and corresponding anomalies, respectively. As expected, LAI values are largely impacted by the extreme precipitation conditions. The wet condition introduces more vegetation, while the dry condition limits the vegetation growth throughout the two-year period. The DA procedure effectively corrects the LAI errors caused by the biased precipitation input.*"

Section 3.2 Water fluxes and storages

"*Daily time series of global averaged values and corresponding anomalies of the five water variables are shown in Figure 3(b-f) and Figure 4(b-f), respectively. The model well simulates the seasonality of all water fluxes/storages considered here. The OL runs reveal that global average values of all five variables are impacted by the highly biased precipitation conditions.*"

- L. 243 ("thus the NCRMSE . . . becomes smaller"): Why?

In JJA, the stomatal closure can help to preserve water. So, the system does not lose too much water under the dry condition which result in smaller difference between DA-dry and the NR truth, and consequently shows smaller NCRMSE.

- L. 276-277 (LAI assimilation unable to correct for dry precipitation bias): Why?

A dry precipitation bias means that the system has (erroneously) has less water than in reality (NR in the synthetic experiment). Since no water is otherwise added to the system, LAI DA cannot fully correct water-related model states (such as soil moisture). The manuscript has been modified as below:

"*However, LAI assimilation is not able to correct the model when the input precipitation is negatively biased (dry condition). A dry precipitation bias means that the system has (erroneously) less water than in reality (NR in the synthetic experiment). Since no water is otherwise added to the system, LAI DA cannot fully correct water-related model states (such as soil moisture). The NCRMSEs of DA runs are either the same as in the OL runs (ET/CIE/CWS) or worse (SSM/TWS).*"

- L. 320-322: I don't see the logics. I would expect that large water-holding capacity would enhance the impact of LAI DA.

Our thought is that the LAI can affect soil moisture by changing the model's surface water condition. Over forest and woodland, the surface water condition is not changing much due to the large soil reservoir.

- L. 323 (forests and woodlands): Is this because of large rooting depth?

Large rooting depth is an important fact. Some discussion has been added to the manuscript:

"*In other words, forest and woodland tend to have lower sensitivity in response to the change of precipitation conditions because of their large rooting depth.*"

- L. 331: Water-holding? Do you mean interception reservoir or soil reservoir?

It is soil reservoir. We changed it in the manuscript:

"*This is due to large soil reservoir of forests and woodlands ……*"

- L. 374-375: This could be because the used DA system is not able to analysed RZSM from LAI observations. Please explain.

In the Noah-MP model, the relationship between LAI and soil moisture is very complex and indirect. So, the current LAI DA system is not able to have much of an effect on surface moisture at all depth.

Editorial comments:

- L. 251 (Figure 4): Color scale is difficult to interpret. Please use several colors (e.g. blue in addition to red).

We changed the color scale of Figure 4.

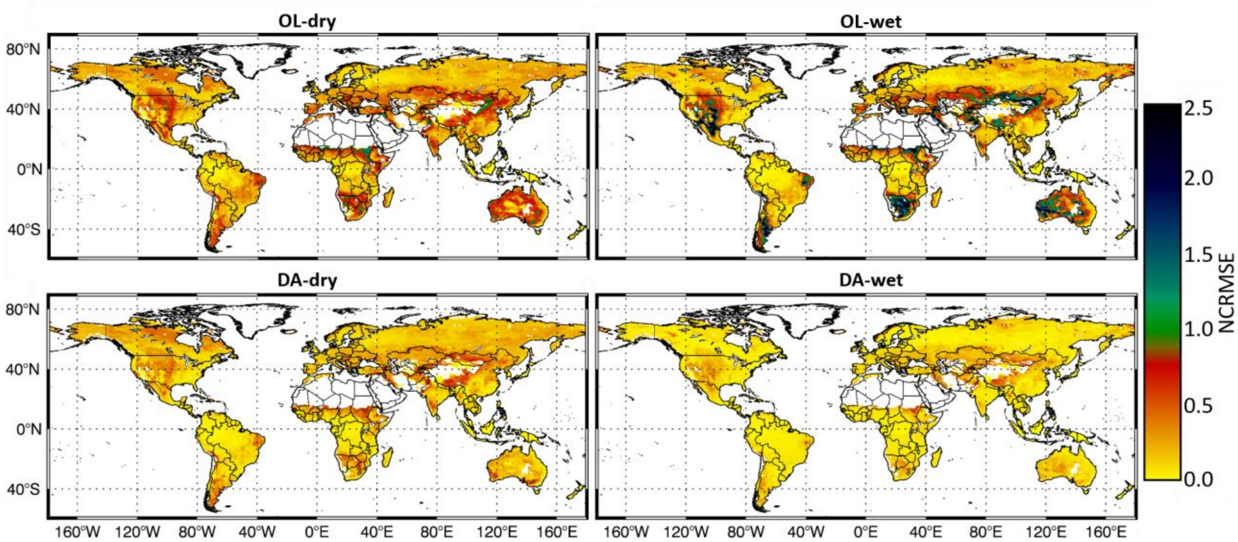


Figure 5. Maps of LAI NCRMSE for the OL and DA runs.

- L. 289 (Figure 5): Time axis labels are not readable. Please improve!

We enlarged the font size of the axis label and showed the time less frequently (every 3-month). Please check below. We also enlarged the axis font size of all figures in the manuscript.

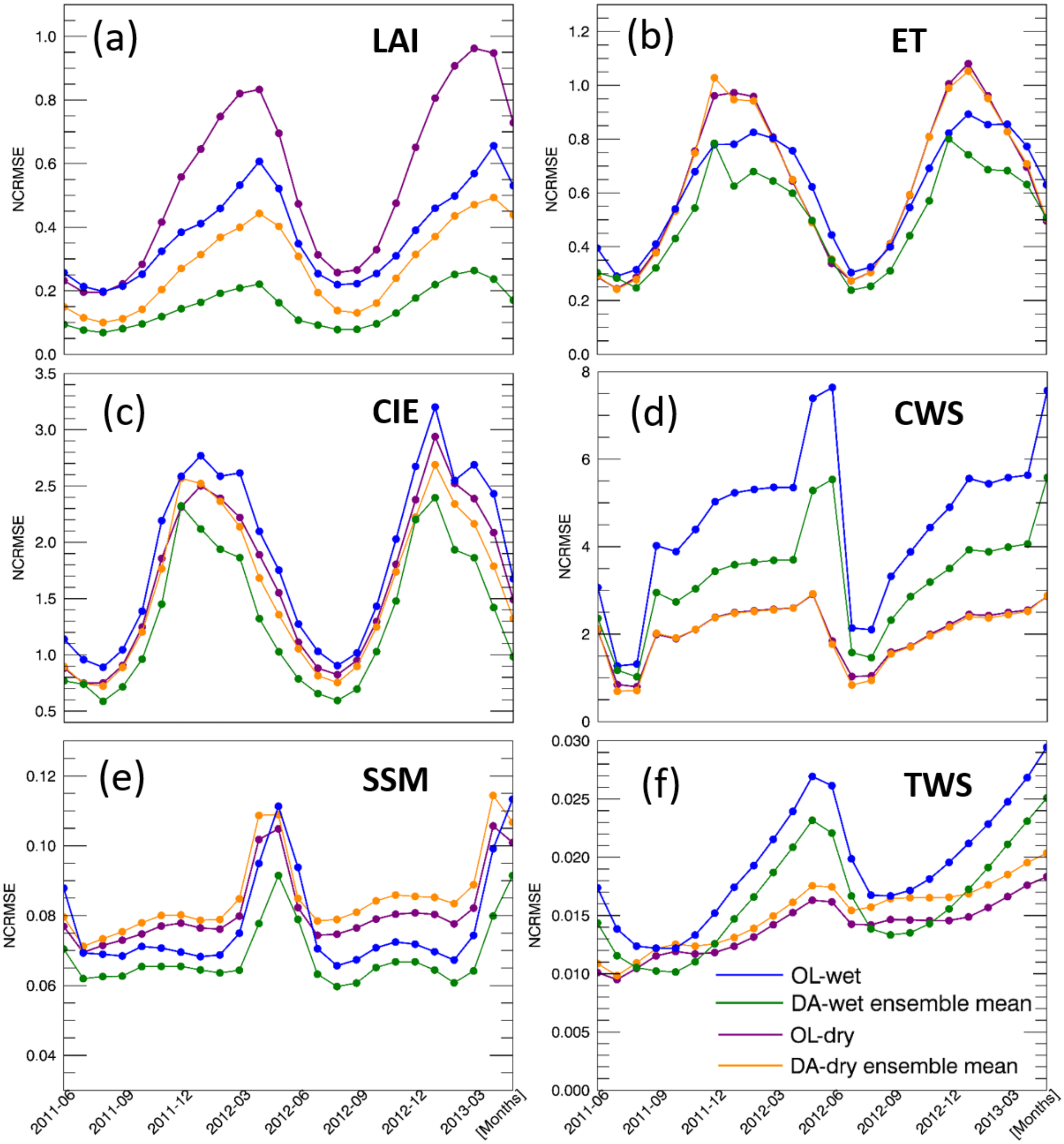


Figure 6. Monthly averaged NCRMSE for LAI and five water variables over the Northern hemisphere.

- L. 291 (Figure 6): Time axis labels are not readable. Please improve!

We enlarged the font size of the axis label and showed the time less frequently (every 3-month). Please check below.

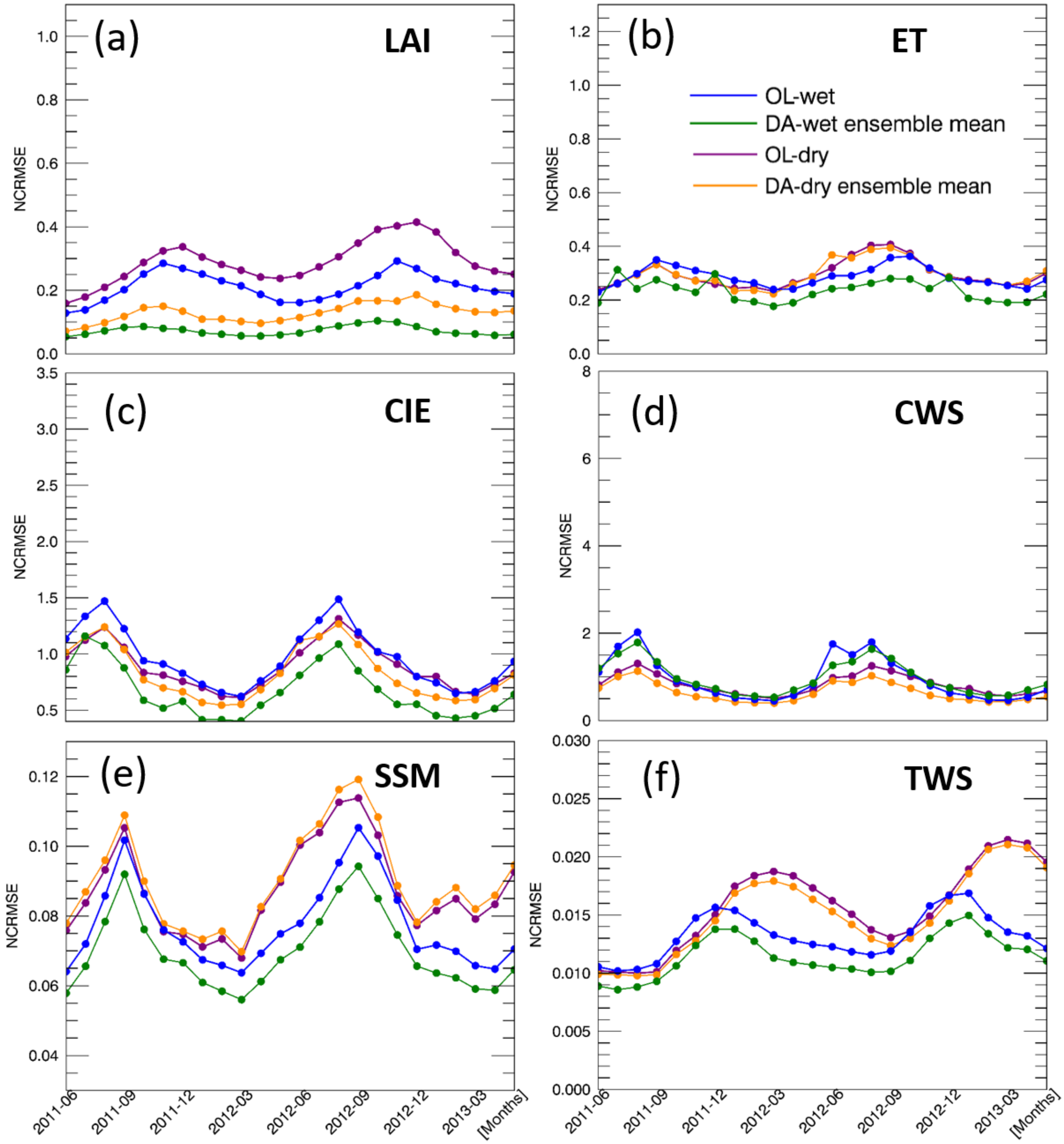


Figure 7. Same as in Figure 6, but for the Southern hemisphere.

The Influence of Assimilating Leaf Area Index in a Land Surface Model on Global Water Fluxes and Storages

Xinxuan Zhang[1], Viviana Maggioni[1], Azbina Rahman[1], Paul Houser[1], Yuan Xue[1], Timothy Sauer[1], Sujay Kumar[2] and David Mocko[2]

1. George Mason University, Fairfax, VA, USA

2. Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Submit to:

Hydrology and Earth System Sciences

March, 2020

Abstract

Vegetation plays a fundamental role not only in the energy and carbon cycles, but also in the global water balance by controlling surface evapotranspiration (ET). Thus, accurately estimating vegetation-related variables has the potential to improve our understanding and estimation of the dynamic interactions between the water, energy, and carbon cycle~~s~~. This study aims to assess to what extent a land surface model (LSM) can be optimized through the assimilation of leaf area index (LAI) observations at the ~~the~~ global scale. Two O~~o~~bserving ~~s~~System S~~s~~imulation E~~e~~xperiments (OSSEs) are performed to evaluate the efficiency of assimilating LAI into ~~Noah~~an LSM ~~with multi-parameterization options (Noah-MP)~~ through an Ensemble Kalman Filter (EnKF) to estimate LAI, ~~evapotranspiration (~~ET~~)~~, canopy interception evaporation (CIE), canopy water storage (CWS), surface soil moisture (SSM), and terrestrial water storage (TWS). Results show that the LAI data assimilation framework not only effectively reduces errors in LAI model simulations, but ~~. LAI assimilation~~ also improves ~~the model estimates of~~ all the modeled water flux and storage variables considered in this study (ET, CIE, CWS, SSM, and TWS), even when the forcing precipitation is strongly positively biased (extremely wet condition). However, it tends to worsen some of the modeled ~~estimated~~ water-related variables (SSM and TWS) when the forcing precipitation is affected by a dry bias. This is attributed to the fact that the amount of water in ~~Noah-MP~~the LSM is conservative and the LAI assimilation introduces more vegetation, which requires more water than what available within the soil. ~~Future work should could investigate a multi-variate data assimilation system that concurrently merges both LAI and soil moisture (or TWS) observations at global scale.~~

1. Introduction

Terrestrial vegetation plays a vital role in the global water cycle, as it controls the surface evapotranspiration (ET) and the state of the carbon cycle. As shown in past literature, there exists a strong relationship ~~between~~ among vegetation, precipitation, and soil moisture (Di et al., 1994; Farrar et al., 1994; Richard and Poccard, 1998; Adegoke and Carleton, 2002). Nevertheless, the role that vegetation and its dynamics play in the water cycle (for instance on the variability of precipitation) is extremely complex (Wang and Eltahir 2000; Wang et al. 2011). In the past half-century, these land surface processes and feedbacks have been examined through numerical modeling experiments (~~e.g.,~~ Foley et al. 1996; Kim and Wang 2007; Druel et al. 2019). In early generation land surface models (LSMs), the development stage of vegetation was prescribed by regularly updating vegetation variables, based on fixed lookup tables to simplify the model computation (Foley et al. 1996). This approach uses constant vegetation indices, e.g., the ~~L~~leaf ~~A~~area ~~I~~index (LAI), ~~throughout a certain period,~~ while in reality the growth of vegetation continuously changes in response to weather and climate conditions. To overcome this deficiency, new generation LSMs are coupled with dynamic vegetation ~~models~~ modules that comprehensively simulate several biogeochemical processes (Woodward and Lomas 2004; Gibelin et al. 2006; Fisher et al. 2018) and that~~.~~ ~~LSMs with a dynamic vegetation module~~ are able to capture more detailed variations in plant productivity than traditional ~~LAI~~ methods (Kucharik et al. 2000; Arora 2002; Krinner et al. 2005).

LAI can also be estimated through observations from satellite sensors, such as the Moderate Resolution Imaging Spectroradiometer (MODIS, Pagano and Durham 1993; Justice et al. 2002), the Système Probatoire d’Observation de la Terre VEGETATION (SPOT-VGT, Baret et al. 2007), and the National Oceanic and Atmospheric Administration (NOAA) Advanced Very

High Resolution Radiometer (AVHRR, Cracknell 1997). LAI products retrieved from different satellite missions and sensors provide spatially and temporally varying LAI fields on a routine basis at regional and global scales, including the ~~Moderate Resolution Imaging Spectroradiometer (~~MODIS~~)~~ LAI (Myneni et al. 2002), the Global Land Surface Satellite (GLASS) LAI (Xiao et al. 2013), and the GLOBMAP LAI dataset (Liu et al. 2012), among others. Satellite-derived LAI products were found to be affected by uncertainties due to the limitation of retrieval algorithms and vegetation type sampling issues (Cohen and Justice 1999; Privette et al. 2002; Tian et al. 2002; Morisette et al. 2002).

A method to combine the inherently incorrect estimates from satellite observations and model simulations is data assimilation (DA). One of the most common DA systems — the Ensemble Kalman Filter (EnKF; Evensen 2003) — dynamically updates the model error covariance information by producing an ensemble of model predictions, which are individual model realizations perturbed by the assumed model error (Reichle et al. 2007). The ensemble approach is widely used in hydrology~~ic DA~~ because of its flexibility with respect to the type of model error (Crow and Wood 2003) and well suited to the nonlinear nature of land surface processes (Reichle et al. 2002a, 2002b; Andreadis and Lettenmaier 2006; Durand and Margulis 2008; Kumar et al. 2008; Pan and Wood 2006; Pauwels and De Lannoy 2006; Zhou et al. 2006). However, the use of an EnKF for the assimilation of LAI in LSMs has not been thoroughly investigated in the past. Pauwels et al. (2007) proposed an ~~observing~~ Observing ~~system~~ System S~~s~~imulation E~~e~~xperiment (OSSE) to evaluate the performance of assimilating LAI in a hydrology-crop growth model ~~by~~ with an EnKF algorithm. Other studies have also tested simplified 1D-VAR and extended Kalman filter methods for LAI assimilation (e.g., Sabater et al. 2008; Barbu et al. 2011; Fairbairn et al. 2017). Recently, Kumar et al. (2019) assimilated GLASS LAI ~~assimilation~~

in a land surface model with an EnKF across the ~~Continental~~ continental U.S. Some ~~model simulated~~ water budget ~~terms~~variables were improved through the assimilation procedure, particularly~~.~~ ~~The improvement is~~ ~~, especially~~~~remarkable~~ in agricultural areas ~~because~~ where the assimilation added harvesting information to the model. Ling et al. (2019) assimilated global LAI information ~~at the global scale~~ with an Ensemble Adjust Kalman Filter (EAKF) algorithm and found that the assimilation is more effective during the growing season. LAI assimilation also had a positive impact on gross primary production (GPP) and ~~evapotranspiration (~~ET~~)~~ in low latitude regions.

Nevertheless, most of the aforementioned studies mainly focused on the impact of LAI assimilation on the ~~model~~ simulated LAI or vegetation biomass. Only a few studies discussed the influences of LAI assimilation on the estimation of water variables such as soil moisture or streamflow (Pauwels et al. 2007; Sabater et al. 2008) and most of them focused on ~~small~~ limited regions. Most recently, Albergel et al. (2017) conducted a study on a much larger domain − Europe and the Mediterranean basin − and showed ~~that LAI assimilation can be used to improve~~ improvement in soil moisture at various depths thanks to LAI assimilation.

This work leverages upon these studies but aims to assess to what extent a land surface model, especially the ~~model estimations~~simulation of water-related variables, can be optimized through the assimilation of LAI observations at the ~~the~~ global scale. As this study serves as a feasibility test to quantify the impact of LAI assimilation on water cycle variables, an OSSE is chosen to investigate the model's behavior. This guarantees that reference variables (often referred to as the "truth"), which are synthetically produced, are available for quantifying the performance of the proposed framework. Specifically, two OSSEs that apply an EnKF algorithm to ~~an~~the Noah LSM with multi-parameterization options ~~model~~ (Noah-MP, Niu et al. 2011; Yang et al. 2011) are performed to evaluate the efficiency of assimilating LAI observations for estimating ~~evapotranspiration~~ET, canopy interception evaporation (CIE), canopy water storage (CWS), surface soil moisture (SSM), and terrestrial water storage (TWS).

2. Methods and materials

2.1. Land surface model (Noah-MP)

The ~~Noah LSM with multi-parameterization options (~~Noah-MP 3.6~~,~~ (Niu et al. 2011; Yang et al. 2011) is adopted in this study. Noah-MP contains a separate vegetation canopy defined by a canopy top and bottom, crown radius, and leaves with defined dimensions, orientation, density, and radiometric properties (Niu et al. 2011). Multiple options are available for surface water infiltration, runoff, groundwater transfer and storage including water table depth to an unconfined aquifer (Niu et al. 2007), dynamic vegetation, canopy resistance, and frozen soil physics. Specifically, the prognostic vegetation growth combines a Ball-Berry photosynthesis-based stomatal resistance (Ball et al. 1987) with a dynamic vegetation model (Dickinson et al. 1998). The dynamic vegetation model ~~which~~ calculates the carbon storages in various parts of the vegetation (leaf, stem, wood, and root) and the soil carbon pools.

The Noah-MP 3.6 LSM has been implemented into the National Aeronautics and Space Administration (NASA) Land Information System (LIS; Peters-Lidard et al. 2007; Kumar et al. 2006). LIS is a software that provides an interagency test bed for land surface modeling and data assimilation that allows customized systems to be built, assembled and reconfigured easily, using shared plugins and standard interfaces. All the experiments ~~of Noah-MP~~ in this study are setup through LIS. The Modern-Era Retrospective analysis for Research and Applications Version 2

(MERRA-2; Gelaro et al. 2017) dataset serves~~d~~ as the meteorological forcings ~~for~~ of Noah-MP. MERRA-2 is the latest atmospheric reanalysis produced by the NASA Global Modeling and Assimilation Office (GMAO) and includes updates from the Goddard Earth Observing System (GEOS). The meteorological ~~forcing~~ variables selected from MERRA-2 include surface pressure, surface air temperature, surface specific humidity, incident radiations, wind speed, and precipitation rate.

Five model output variables that describe terrestrial water fluxes and storages are investigated in this work: ~~Evapotranspiration~~ ET (~~ET,~~ defined as the sum of evaporation and the plant transpiration [kg/m^{2}s]), ~~Canopy Interception Evaporation~~CIE (~~CIE,~~ defined as the evaporation of the canopy intercepted water [kg/m^{2}s]), ~~Canopy Water Storage~~CWS (~~CWS,~~ defined as the amount of canopy intercepted water in both liquid and ice phases [kg/m^2]), ~~Surface Soil Moisture~~SSM (~~SSM,~~ defined as the water content in the top 10 cm of the soil column [m^3/m^3]), and ~~Terrestrial Water Storage~~TWS (~~TWS,~~ defined as the sum of all water storage on the land surface and in the subsurface [mm]).

2.2. Experimental design

An OSSE is designed to understand the efficiency of assimilating LAI within Noah-MP version 3.6 using a one-dimensional EnKF algorithm (Reichle et al. 2010), when the precipitation forcing data are strongly biased. Being the major driving force of the hydrological cycle, the quality of input precipitation is critical for the accuracy of ~~a~~ land surface model outputs. However, global precipitation datasets are far from being perfect and often affected by large regional biases. For example, the MERRA-2 precipitation dataset shows a widespread relative bias greater than 100% in South Asia (Ghatak et al. 2018). Although an ~~an~~ ~~the~~ EnKF ~~data assimilation~~ is optimal only under the assumption of unbiasedness (which is not met in the proposed experimental setup), we want to investigate here to what extent the ~~a LAI~~ EnKF LAI assimilation (even if sub-optimal) can improve water storages and fluxes under two extreme conditions, i.e., a very dry and a very wet precipitation bias, knowing that such biases are very plausible in the real world and often unknown (and therefore difficult to remove). The proposed framework is evaluated ~~byon a~~through a global ~~scale~~ experiment~~s~~ (Antarctica excluded) at the 0.625° × 0.5° spatial resolution of the MERRA-2 forcing dataset (Figure 1).

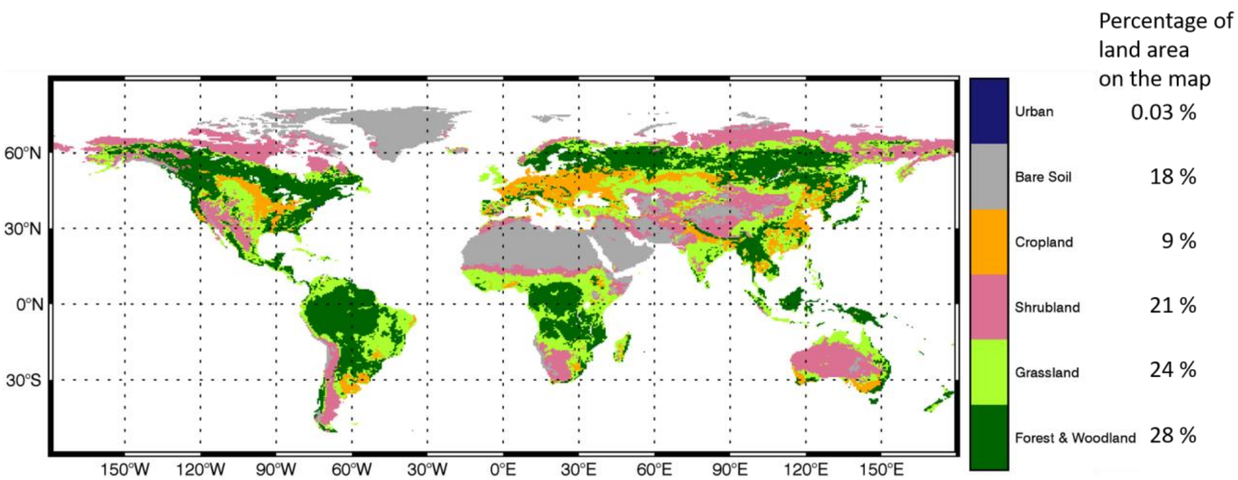


Figure 1. Study domain and land cover types (Hansen et al. 2000).

Figure 2 shows a schematic diagram of the experiments. First, the Noah-MP model is spun-up for a 10-year period (2001-2010) to ensure a physically realistic state of equilibrium. Second, the model is run for a 29-month period (January 2011 – May 2013) to conduct the Nature Run (NR) with the same configuration as the spin-up one. By definition, an OSSE is a controlled experiment that does not assimilate any real observation. Instead, it treats all the model outputs from the NR as the “true” condition (denoted as the “synthetic truth”).~~.~~ ~~Thus, the output LAI from NR is considered as t~~The “tru~~the~~”~~,~~ LAI (i.e., the LAI output from NR) ~~which~~ is then perturbed via a simple additive error model to produce the synthetic observations to be assimilated into the DA runs. The spin-up run and NR are forced by the original MERRA-2 precipitation data. Third, two Open Loop (OL) runs (no DA) are conducted for the same 29-month period under two conditions: i) “extremely dry” condition (the model is forced by halving the MERRA-2 precipitation data; OL-dry), and ii) “extremely wet” condition (the model is forced by doubling the MERRA-2 precipitation; OL-wet). The biased forcing precipitation data in OL mimic typical precipitation biases in current precipitation reanalysis and satellite products (e.g., Ghatak et al. 2018; Yoon et al. 2019).

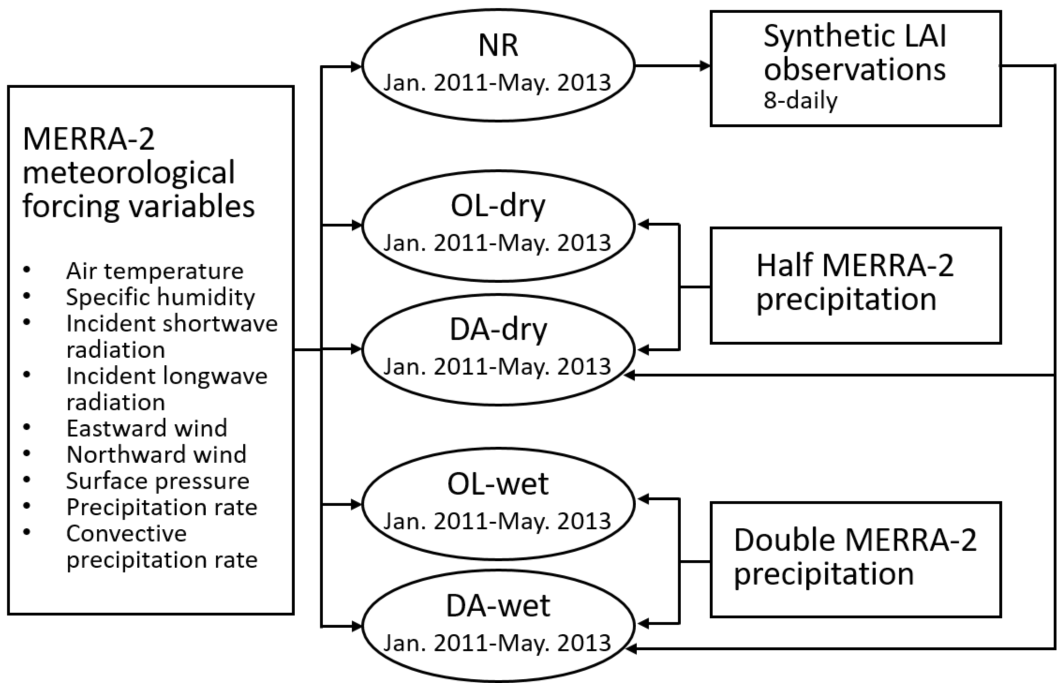


Figure 2. Schematic diagram of the OSSE design.

The two DA runs are then ~~produced~~ conducted under the two same conditions (DA-dry and DA-wet) using a~~n~~ one-dimensional EnKF assimilation algorithm, which is a built-in DA method in LIS. The EnKF DA algorithm is suitable for non-linear and intermittent land surface processes (Reichle et al. 2002a, 2002b). Details of the algorithm ~~have been illustrated~~ can be found in numerous previous studies (Reichle et al. 2010; De Lannoy et al. 2012; Liu et al. 2015; Kumar et al. 2019a).

The model ensemble is generated by perturbing a set of meteorological forcing. To select the optimal ensemble size, a sensitivity test is performed for ensemble sizes spanning from 2 to 24 members (not shown here). The number of ensemble members has a strong impact on the model results at small sizes, while the model performance tends to become steady when more than 20 ensemble members are considered. Thus, all the DA simulations are run for 20 members.

The synthetic LAI observations are obtained from the NR and assimilated to the DA system at 8-daily frequency. The synthetic LAI observation has the same temporal resolution as the MODIS LAI product but with full coverage over the ~~entire~~ study domain. In real case studies, satellite LAI products contain a substantial amount of missing data, mainly due to the cloud obscuration gaps. Based on the vegetation type in the model, the leaf mass fields are also updated. Random perturbations of MERRA-2 meteorological forcings and synthetic LAI observations are applied to create an ensemble of land surface conditions that represent the uncertainties of in the LSM.

Similar to previous work (Kumar et al. 2014, 2019a, 2019b), the MERRA-2 forcing inputs such as shortwave/~~ and~~ longwave radiations and~~as well as~~ precipitation are perturbed hourly. Multiplicative perturbations are applied to the shortwave radiation and precipitation with a mean of 1 and standard deviations of 0.3 and 0.5, respectively. The longwave radiation is perturbed via an additive perturbation with a standard deviation of 50 W/m^2. The perturbations of the three meteorological forcing variables also include cross correlations: cross correlation between shortwave radiation and precipitation is -0.8, cross correlation between longwave radiation and precipitation is 0.5; and cross correlation between shortwave and longwave radiations is -0.5. The synthetic LAI observations are perturbed via an additive model with a standard deviation of 0.1.

2.3. Evaluation and error metrics

Output variables from the OL and DA runs are evaluated against the "truth" from the NR at daily, monthly, and seasonal temporal scales. Besides LAI, five more water fluxes and storages are evaluated in the results section: ~~evapotranspiration~~ET, CIE~~interception evaporation~~, CWS~~canopy water storage~~, SSM~~surface soil moisture~~, and TWS~~terrestrial water storage~~.

The initial condition for the OL and DA runs is generated by a spin-up run that uses the original MERRA-2 precipitation as input. However, the OL and DA runs are forced by either doubled or halved precipitation, which is not consistent with the spin-up run and the model needs some time to stabilize. The first 5-month model outputs are therefore eliminated from the evaluation to avoid the model systematic instability at the beginning of the OL and DA simulations and the evaluation, thus, focused only on model outputs from 2011-06-01 to 2013-05-31. Results are discussed using ~~both~~ maps and ~~anomaly~~ time series of global averaged values and anomalies. Each of the anomaly time series is computed relative to the mean of its respective model run. Moreover, two error metrics are employed to quantify the difference between OL (and DA) with respect to the reference variables (from the NR). The first one is the Normalized and Centered Root Mean Square Error (NCRMSE), defined as follows:

$$E = \frac{\left\{\frac{1}{N}\sum_{i=1}^{N}\left[\left(X_i - mean(X)\right) - \left(O_i - mean(O)\right)\right]^2\right\}^{\frac{1}{2}}}{mean(O)} \qquad \text{Eq. 1}$$

where E is the NCRMSE, O is the NR output variable, and X is the output variable from the OL runs or DA runs. N is the total number of X values, and i represents the index of each X value. Second, to investigate the improvement (or degradation) due to the DA of LAI observations, we adopt the Normalized Information Contribution (NIC, similar to the NIC in Kumar et al. 2016) index based on NCRMSE and defined as:

$$C = \frac{E_{DA} - E_{OL}}{0 - E_{OL}} \qquad \text{Eq. 2}$$

where *C* represents the NIC index and *E* is the NCRMSE for OL or DA runs.~~,~~ NIC equals to 1 means that DA realizes the maximum possible improvement over the OL; NIC equals to zero means that DA and OL show the same performance skills; and negative NIC indicates a model degradation through DA.

3. Results and discussion

3.1. LAI

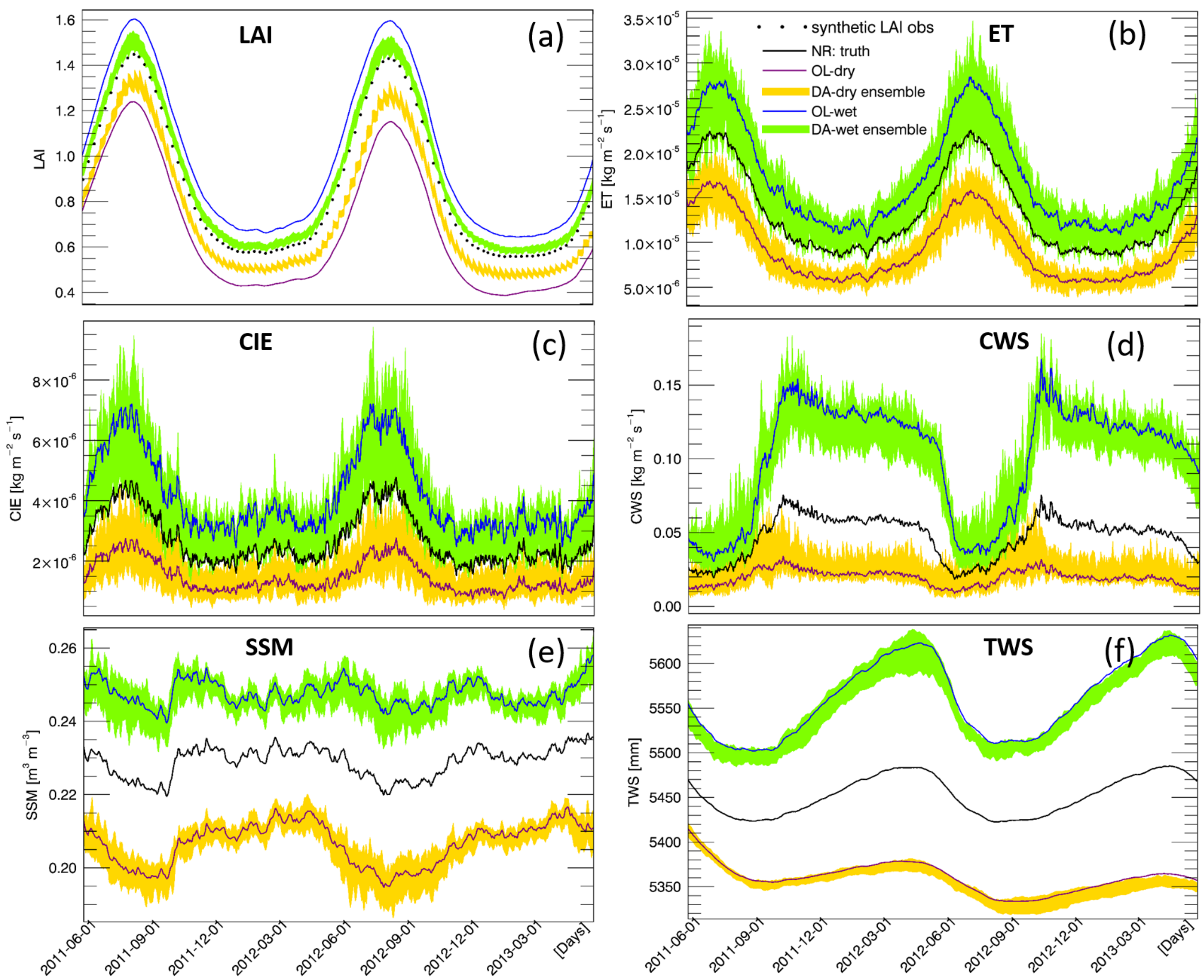


Figure 3. Global averaged daily values of LAI and five water variables (2011-06-01 to 2013-05-30).

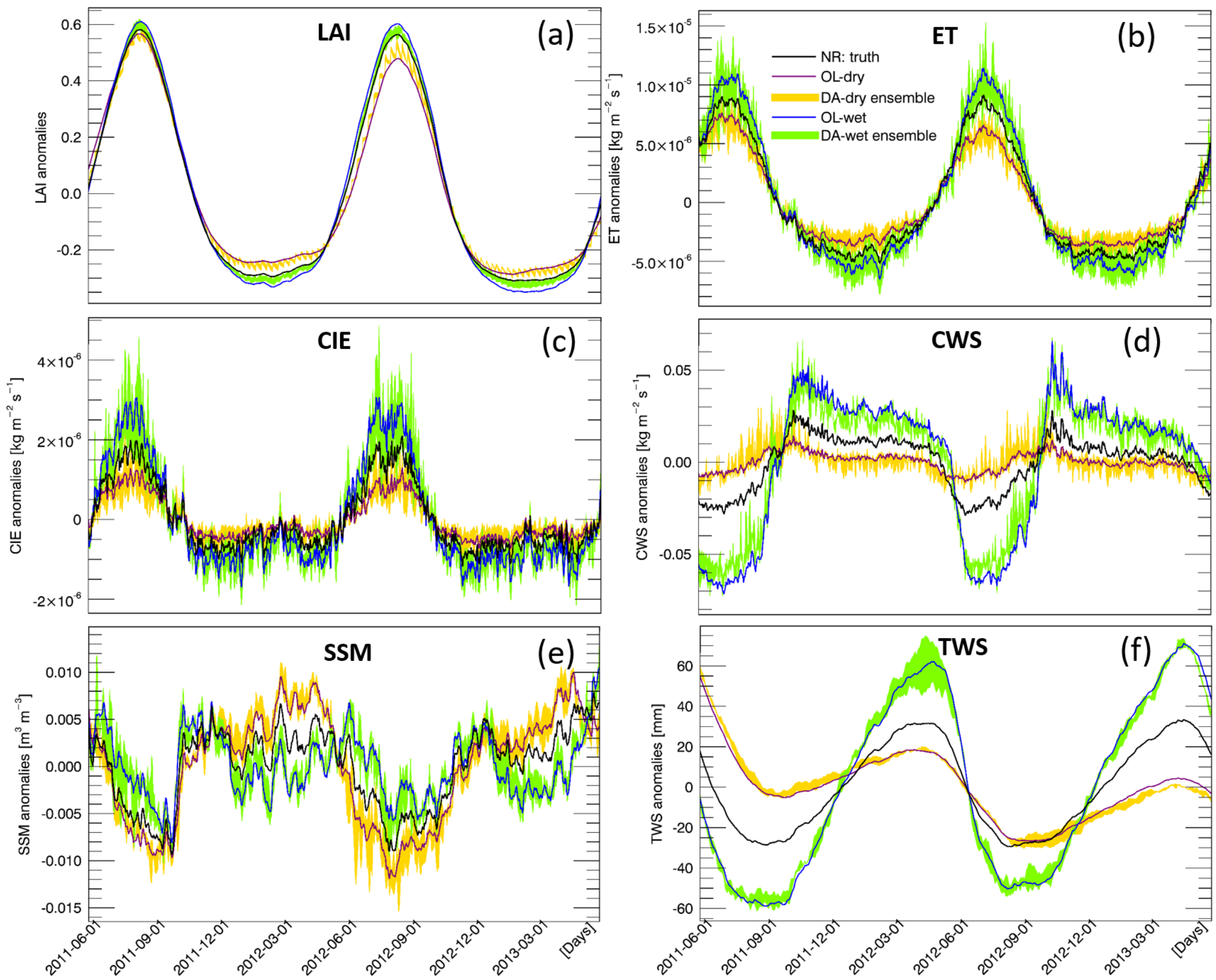


Figure 4~~3~~. Global averaged daily anomalies of LAI and five water variables (2011-06-01 to 2013-05-30).

Figure 3a and Figure 4a show time series of global averaged LAI values and corresponding anomalies, respectively. As expected, LAI values are largely impacted by the extreme precipitation conditions. The wet condition introduces more vegetation, while the dry condition limits the vegetation growth throughout the two-year period. The DA procedure effectively corrects the LAI errors caused by the biased precipitation input. ~~Figure 3a shows time series of global averaged LAI anomalies.~~ The seasonality of LAI anomalies is evident, showing larger variations in ~~winter~~DJF and ~~summer~~JJA than during the transition periods (~~spring~~MAM and ~~fall~~SON). The OL-wet condition simulation ~~(blue line)~~ shows larger LAI anomalies than the NR reference ~~(black line)~~, while the OL-dry condition ~~(purple line)~~ has smaller LAI anomalies than NR. ~~The green and yellow shaded areas represent the 20 ensemble members of the DA runs.~~ The LAI anomalies obtained from DA runs ~~The LAI DA procedure~~ under both wet and dry conditions are closer ~~effectively corrects the LAI anomalies comparing~~ to the reference anomalies than the corresponding OL runs. In general, DA performs better in the wet condition experiment than in the ~~DA~~ dry case.

Moreover, the DA runs show lower NCRMSEs than the corresponding OL runs ~~in several regions~~ across the globe (Figure 5~~4~~a),~~,~~ especially~~with larger~~ over shrublands and grasslands ~~areas~~ (refer to Figure 1 for land covers).

In order to illustrate how LAI assimilation performs for different seasons, Figure 6~~5~~a and Figure 7~~6~~a show monthly averages of NCRMSE for LAI across the northern and southern hemispheres, respectively. In the northern hemisphere (Figure 6~~5~~a), the NCRMSE time series follow clear seasonal patterns. First, the NCRMSE is higher in DJF/MAM~~winter/spring~~ and is lower in ~~summer~~JJA/SON~~/fall~~ for both extreme precipitation conditions. The highest NCRMSE values are in March and April ~~(spring)~~, and the lowest values are in July, August, and September. The differences of NCRMSE between OL and the corresponding DA runs tend to be much larger in ~~spring~~MAM than in any other seasons, which means that LAI assimilation is more effective in the vegetation growth period. Moreover, the NCRMSE is constantly higher in the dry condition runs than the wet ones, which is due to the fact that the growth of vegetation is sensitive to the lack of water. Differences between wet and dry conditions are much smaller in ~~summer~~JJA than in other seasons. In ~~the summer~~JJA, the vegetation leaves in the north hemisphere are fully developed and the plants can use stomatal closure to preserve water under water limited condition (dry condition).~~;~~ ~~t~~Thus, the NCRMSE of dry condition becomes smaller and does not show much difference from the wet condition. The southern hemisphere (Figure 7~~6~~a), which does not have a strong climate seasonality, shows more modest seasonal NCRMSE patterns than the northern regions. In general, the NCRMSEs in the southern hemisphere are smaller than the ones in the northern hemisphere all year around. Specifically, NCRMSEs in the southern hemisphere are slightly higher in October, November, and December, when the differences between OL and DA runs are also larger.

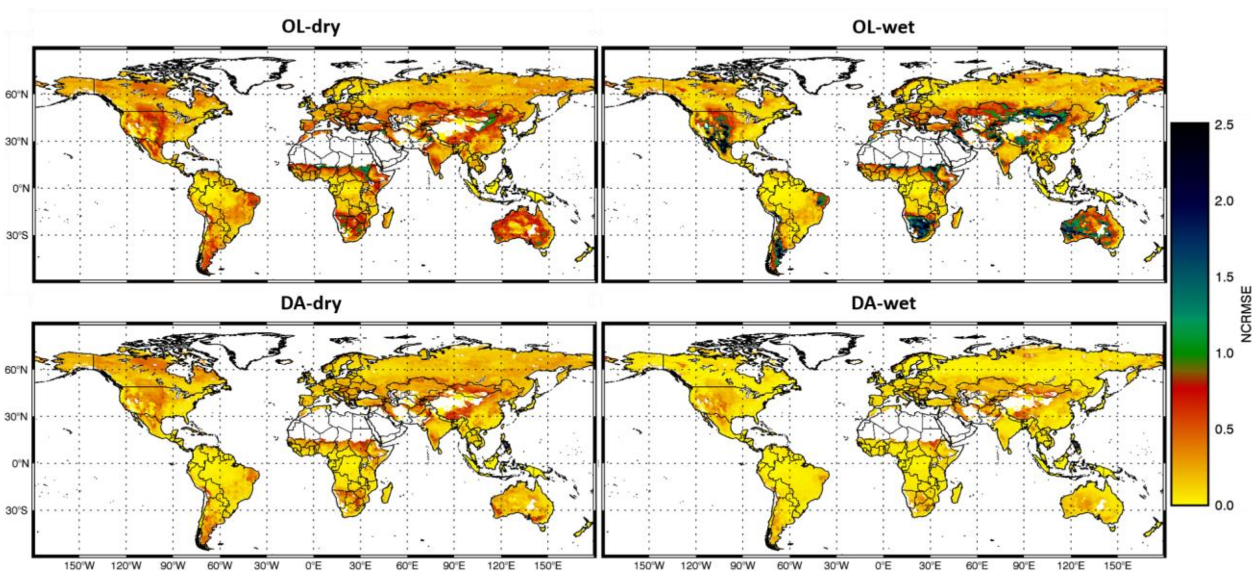


Figure 5~~4~~. Maps of LAI NCRMSE for the OL and DA runs.

3.2. Water fluxes and storages

As mentioned in section 2.3, we focus on five water-related variables from the Noah-MP output to evaluate the impact of LAI assimilation on simulating the water cycle (ET, CIE, CWS, SSM, and TWS). Daily time series of global averaged values and corresponding anomalies of the five water variables are shown in Figure 3(b-f) and Figure 4(b-f), respectively. The model well simulates the seasonality of ~~anomalies for~~all water fluxes/storages considered here. The OL runs reveal that global average values of all ~~the~~ five variables are impacted by the highly biased precipitation conditions~~ (dry and wet)~~. ~~Specifically, t~~The variations of anomalies for~~of~~ ET, CIE, CWS, and TWS tend to be amplified by the wet condition and tend to be dampened by the dry condition. On the contrary, the anomalies of SSM become larger in dry conditions and become smaller in wet conditions, which is probably due to the limited soil water capacity. The surface soil ~~has higher chance~~is more likely to get saturated in wet conditions when the precipitation doubles the original amount, but SSM cannot get larger once the soil is saturated, even if there is more precipitation added to the system. Thus, the range of SSM anomalies in the wet experiment is limited and narrower than in the dry condition. The green and yellow shaded areas represent the ensemble of the DA runs. The anomaly ensembles of the five water variables show slight improvements through DA when precipitation is severely positively biased (wet condition). However, none of these variables shows improvement when the precipitation is severely negatively biased (dry condition) – the anomalies either have no change through the LAI DA (ET, CIE, and CWS) or worsen the OL-dry run (SSM and TWS).

To further investigate the efficiency of assimilating LAI in Noah-MP, time series of monthly NCRMSE averages are shown in Figure 6~~5~~(b-f) and Figure 7~~6~~(b-f) for all five water variables. The five variables can be divided into two main groups based on their performances: ET/CIE/CWS and SSM/TWS. For the wet bias experiment, DA improves the NCRMSE for all variables. However, LAI assimilation is not able to correct the model when the input precipitation is negatively biased (dry condition). A dry precipitation bias means that the system has (erroneously) less water than in reality (NR in the synthetic experiment). Since no water is otherwise added to the system, LAI DA cannot fully correct water-related model states (such as soil moisture). ~~and t~~The NCRMSEs of DA runs are either the same as in the OL runs (ET/CIE/CWS) or worse (SSM/TWS). Specifically, ET/CIE/CWS have larger NCRMSE in the northern hemisphere and much smaller NCRMSEs in the southern hemisphere, but SSM/TWS do not show large differences between north and south. Moreover, ET/CIE/CWS in the northern hemisphere follow a seasonal pattern: NCRMSEs are lower in ~~summer~~warm season (JJA) and higher in the colder season~~s~~ (DJF and March~~December, January, February, and March~~). In the southern hemisphere the three variables also have relative higher NCRMSE in the colder season (JJA~~June, July, and August~~). On the contrary, SSM/TWS show a different seasonal pattern that NCRMSEs are larger in the warmer season (April, May, and June) over northern hemisphere. In southern hemisphere, TWS also has larger NCRMSEs in warmer season (October to April), but SSM shows higher NCRMSEs in colder season (similar to the ET/CIE/CWS group).

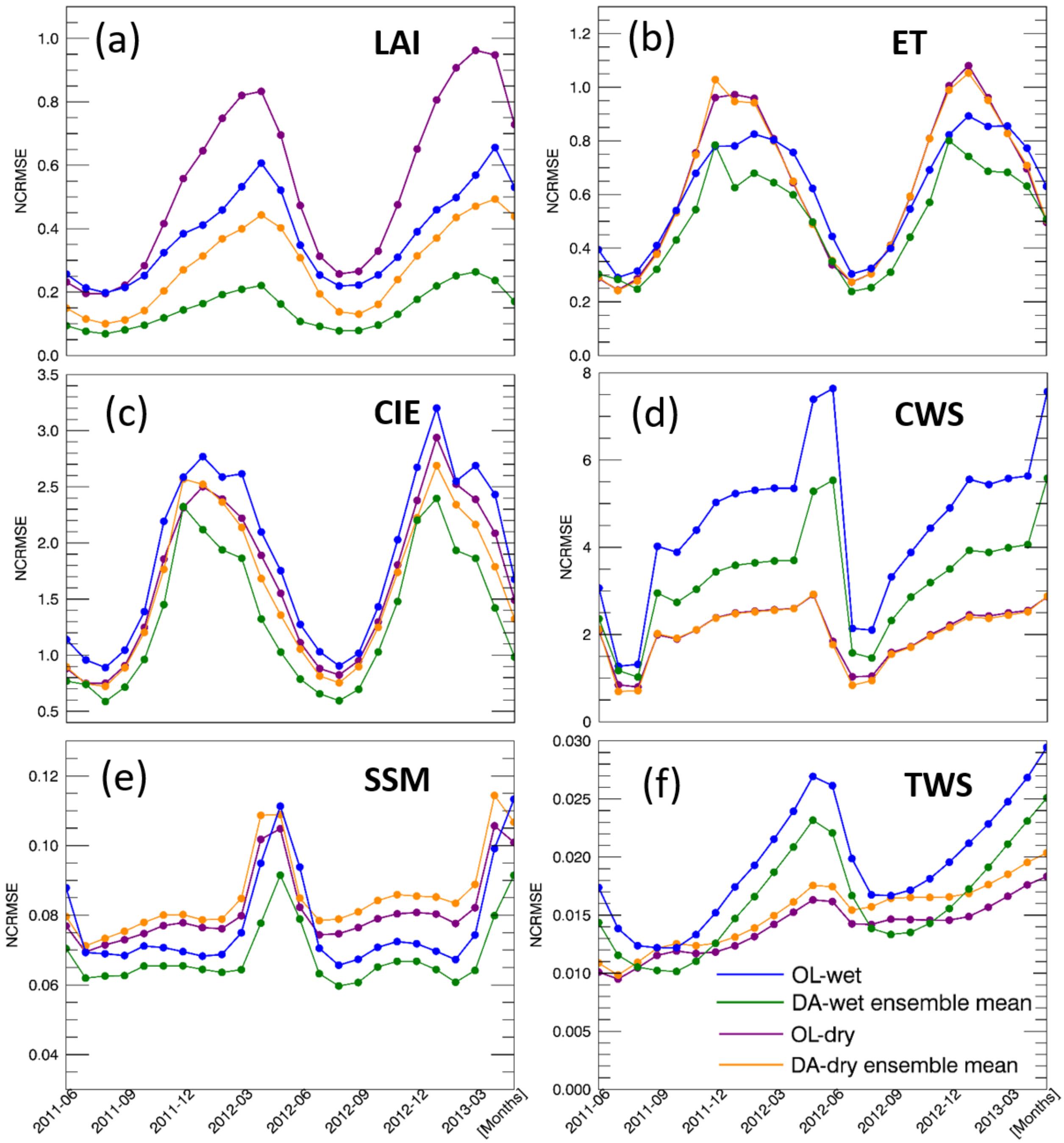


Figure ~~6~~5. Monthly averaged NCRMSE for LAI and five water variables over the Northern hemisphere.

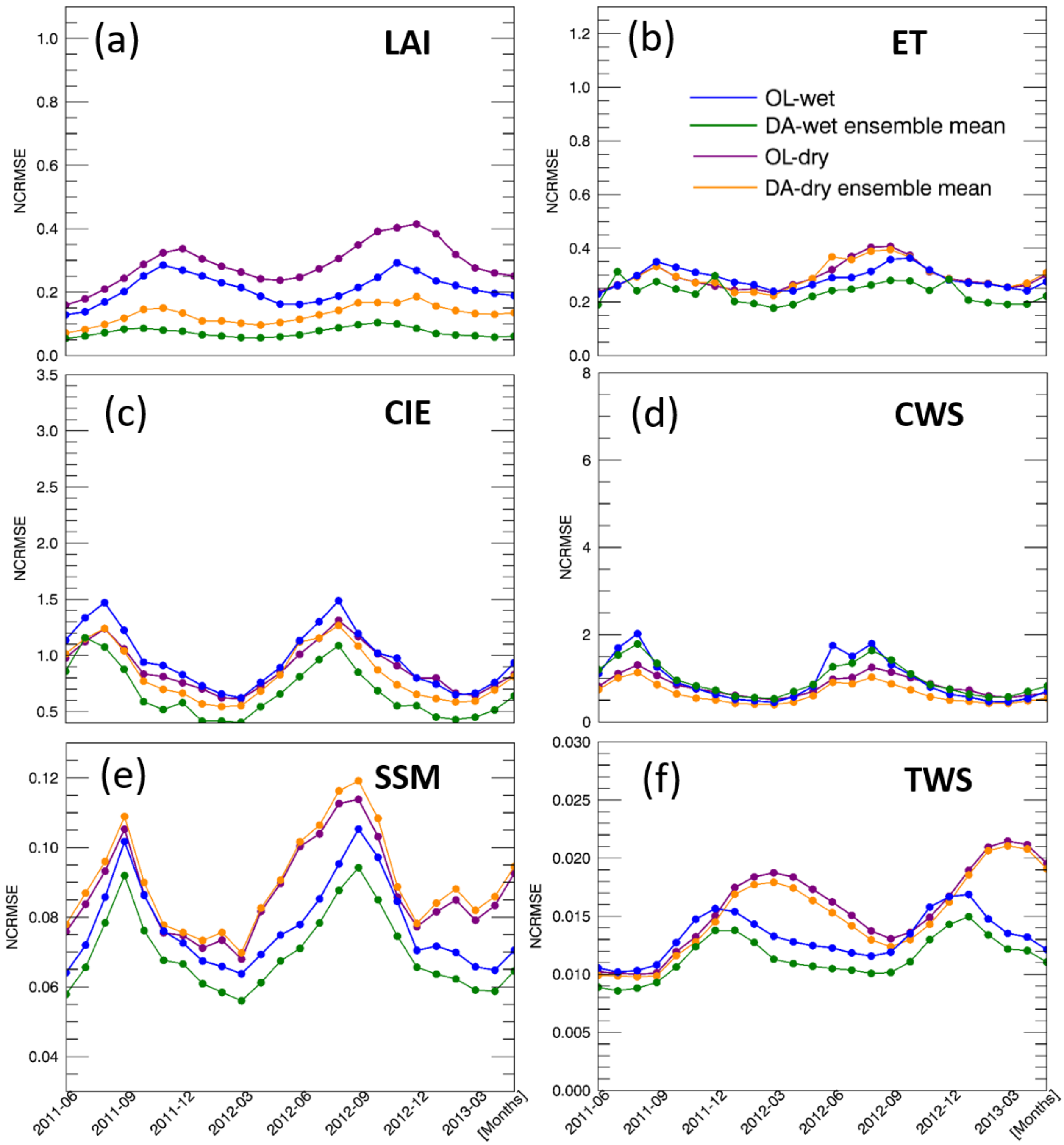


Figure 7~~6~~. Same as in Figure 6~~5~~, but for the Southern hemisphere.

The improvements in the model water fluxes and storages through LAI DA are also quantified by the NIC index (defined in Eq. 2). Figure 8~~7~~ presents comparisons among NIC indices for each water variable analyzed in this study across areas with four different land cover types:

forest & woodland, grassland, shrubland, and cropland. In general, LAI DA improves the NIC indices with positively biased input precipitation (DA-wet ~~condition~~) but worsens the NIC when negatively biased input precipitation (DA-dry ~~condition~~) is considered. Specifically, in wet condition, ET, CIE, and CWS have higher variability over areas with different land cover types, while SSM and TWS have similar NIC values across different land covers. Shrubland and cropland tend to perform better in wet condition except for TWS. In dry condition, the NICs of ET, CIE, and TWS have higher variability than the ones of CWS and SSM. SSM and TWS show very low NIC values in dry condition for almost all land covers. Overall the NIC values of ET, CIE, and CWS are better than the ones of SSM and TWS for all land cover types, though the NICs of ET and CIE over forest & woodland perform very poorly.

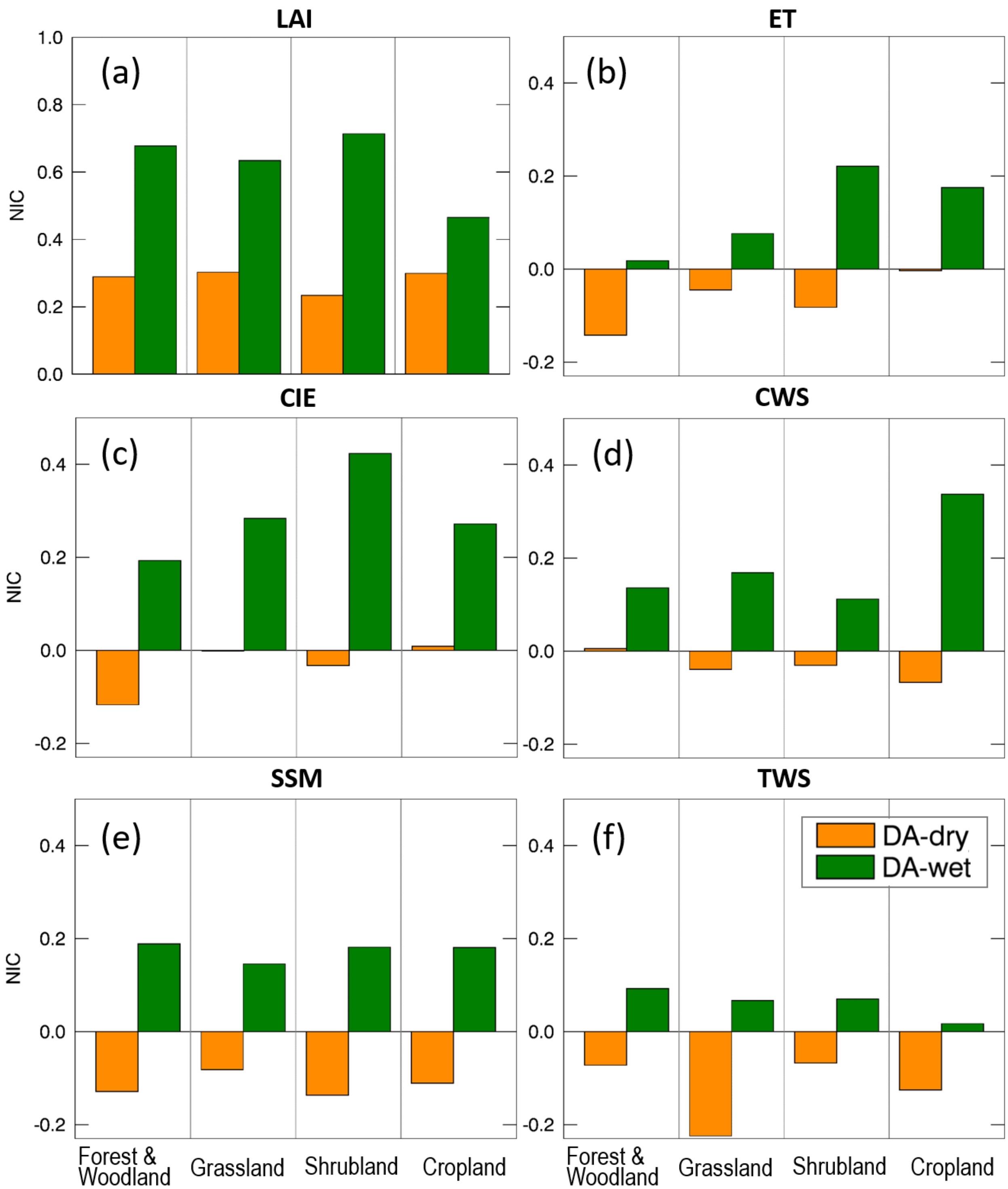


Figure ~~7~~8. NIC for different variables and different land cover types for the two DA runs.

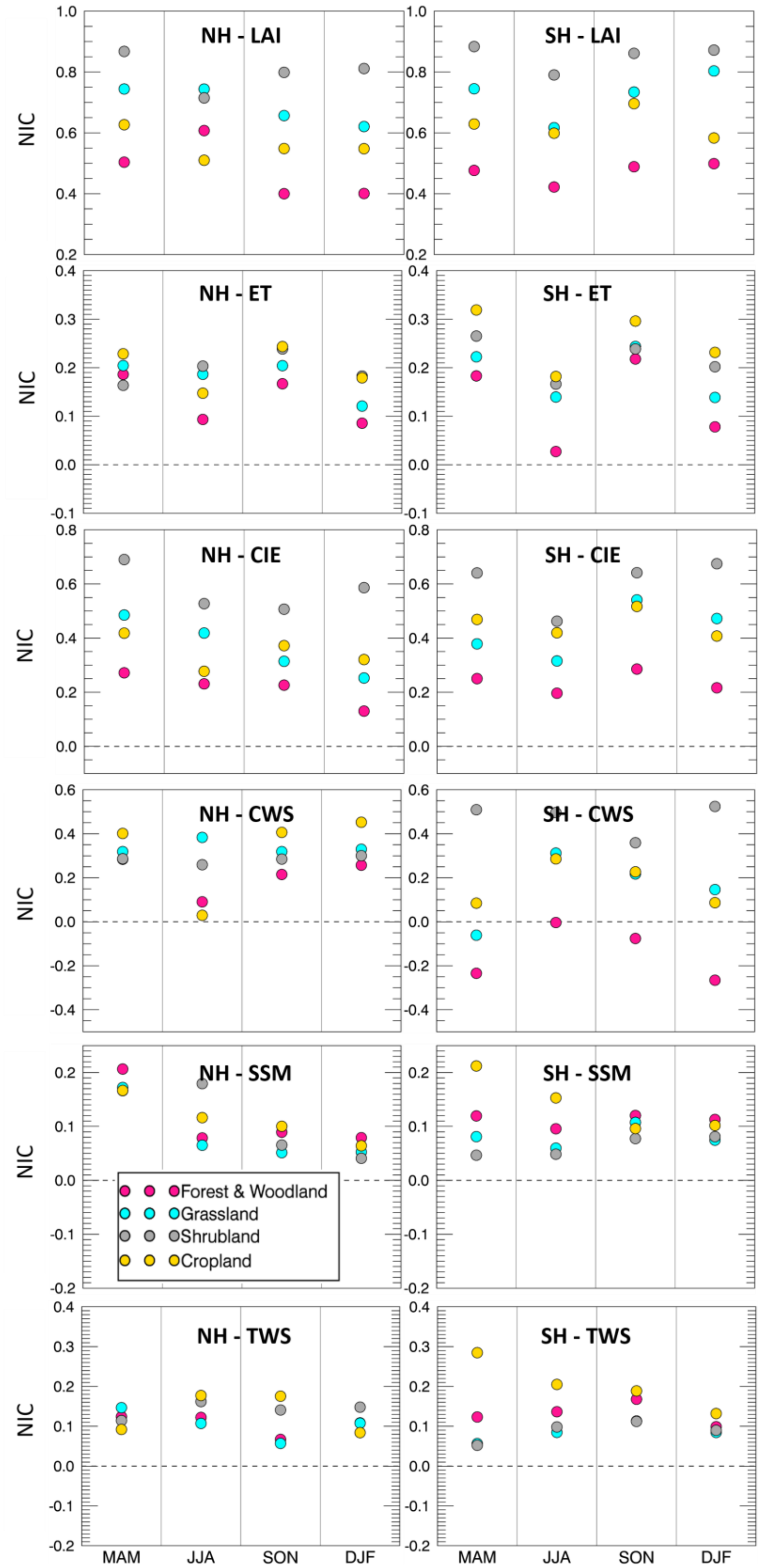


Figure 9~~8~~. NIC of five water variables under wet precipitation conditions over northern and southern hemispheres (NH and SH) during different seasons (MAM, JJA, SON, and DJF)

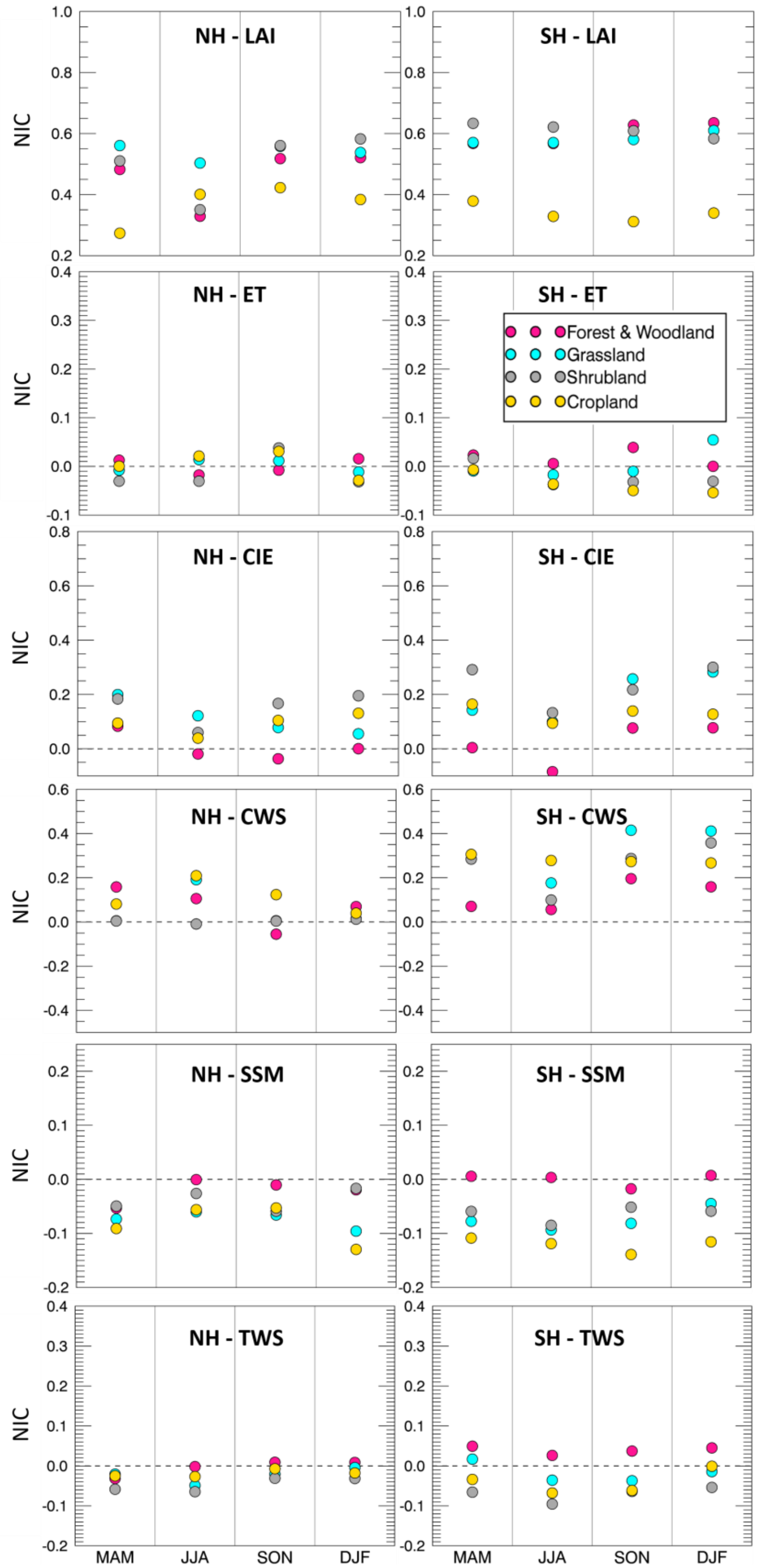


Figure ~~9~~10. Same as in Figure ~~8~~9, but for the dry precipitation experiment.

The effectiveness of LAI DA therefore varies across the northern and southern hemispheres, different land cover types, as well as different input precipitation biases. To further investigate the influence of LAI assimilation, Figures 8 and 9 present NIC values for each hemisphere, each season, and each of the input precipitation conditions – wet and dry, respectively. For the wet case (Figure 9~~8~~), NIC is positive in most cases, which means that the five water variables benefit from the LAI assimilation in all seasons and in both hemispheres. The only exception is CWS which has negative NIC values in the southern hemisphere over grassland (in MAM season) and over forest & woodland (in all seasons). In fact, the forest & woodland region tends to have the least improvement through the LAI assimilation among all land cover types. This is probably because forests and woodlands have large water-holding capacity; thus, the change of water amount caused by LAI DA is not enough to improve the water-related variables. In other words, forest and woodland ~~areas~~ tend to have lower sensitivity in response to the change of precipitation conditions because of their large rooting depth. On the contrary, cropland is very sensitive to precipitation and it benefits the most from the assimilation of LAI for most of the variables. Moreover, NICs of ET/CIE/CWS tend to be smaller than the NICs of SSM and TWS. There is no clear seasonality in the NIC values, though it has a weak tendency to be lower in warm seasons.

For the dry condition case (Figure 10~~9~~), NIC values are much lower than in the wet bias case. Nearly half of the NIC values for the five water-related variables are negative, meaning that DA degrades the OL estimates. Nevertheless, the forest & woodland regions tend to perform better than other land covers in dry condition for SSM and TWS. This is due to large soil reservoir~~water-holder capacity~~ of forests and woodlands, which keeps the model water storage more stable when the input precipitation is affected by large negative biases.

3.3. Discussion

Results presented in sections 3.1 and 3.2 indicate that assimilating LAI in Noah-MP improves the model estimates of water fluxes and storages under positively biased precipitation input~~ (wet case)~~, but does not benefit most of the selected water variables when the precipitation input is characterized by a negative bias~~ (dry case)~~.

In the dry condition runs, Noah-MP is fed by only half of the original MERRA-2 precipitation used in the NR. Considering that the amount of water in Noah-MP is conservative (since based on a water balance equation), the model has no additional water source in the system, even though the LAI assimilation pushes the model towards more vegetation (that should result in more water). As a matter of fact, introducing more vegetation in the system results in more ~~evapotranspiration~~ET and more root water uptake from the soil, which is most likely the cause for the poor performance of most water fluxes and storages in the DA-dry experiment.

On the other hand, the LAI assimilation is found to improve the original OL runs when the input precipitation is positively biased~~ (DA-wet vs. OL-wet)~~. This is because LAI assimilation is able to help constrain the partitioning of model water storage when there is abundant water in the system, thus, improving the performance of water-related variables. In summary, although the EnKF is run here in a sub-optimal mode (not satisfying the unbiasedness assumption), the assimilation of LAI is shown to have a positive impact on multiple variables and in several regions of the world.

Overall the improvement of water variables through LAI assimilation is not remarkable enough to compensate the model degradation caused by the biased precipitation forcing data. Previous studies (Pauwels et al. 2007; Sabater et al. 2007; Barbu et al. 2011; Fairbairn et al. 2017;

Albergel et al. 2017) have tested the performance of the joint assimilation of LAI and soil moisture over regional domains and showed promising results. However, no experiment was performed at the global scale. Future work could investigate a multi-variate data assimilation system that concurrently merges both LAI and soil moisture (or TWS) observations globally.

4. Conclusions

This study evaluates the efficiency of assimilating vegetation information (i.e., LAI synthetic observations) within a land surface model (Noah-MP 3.6) when the precipitation forcing data are strongly biased (either positively or negatively). Two OSSEs that use an EnKF algorithm for LAI assimilation are performed at ~~the~~ global scale during June 2011 – May 2013. The experiments use ~~Noah-MP as a land surface model and~~ MERRA-2 as meteorological forcing data. The OL and DA runs are evaluated against a synthetic "truth" from a nature run, in which the MERRA-2 precipitation is neither perturbed nor biased. The performance of the proposed framework is evaluated for several model output, including LAI estimates and five water-related variables (ET, CIE, CWS, SSM, and TWS).

Overall the EnKF LAI assimilation procedure effectively reduces the LAI error under positively (wet case) and the negatively (dry case) biased precipitation conditions. For the five selected water flux or storage variables, LAI DA improves the model estimates when the model input precipitation is positively biased ~~(wet)~~, but tends to worsen the OL estimates for some of those variables when the input precipitation is negatively biased ~~(dry)~~. Specifically, SSM and TWS estimates are degraded in the DA-dry run with respect to the OL-dry run, while ET, CIE, and CWS do not present large changes when LAI is assimilated in the dry bias run. The poor performance of LAI DA under dry condition is mainly attributed to the fact that the amount of water in Noah-

MP is conservative. The LAI assimilation in dry condition introduces more vegetation, which requires more water in the system to replenish the soil water supply. However, the model has no additional source of water, since the input precipitation is negatively biased.

Although a blind bias case (e.g., unknown biases in the precipitation forcing dataset) is presented here in which the EnKF is run in a sub-optimal mode, the assimilation of LAI observations is proven useful to improve several model output variables. Future research should focus on alternative DA methods~~ to run the DA system in a more optimal way~~, such as updating other related model states while assimilating LAI observations, perturbing the model initial condition and model parameters, and/or assimilating actual satellite-based LAI observations (e.g., MODIS, GLASS) at the global scale to verify the efficiency of the proposed vegetation DA framework. This may be particularly useful in agricultural areas, where the vegetation conditions are largely impacted by cropping schedules (Kumar et al. 2019b). Moreover, future work ~~should~~ could investigate multi-variate DA techniques that combine the assimilation of several variables (such as LAI, soil moisture, and TWS) at the global scale~~once~~.

Acknowledgements: This research is sponsored by the NASA Modeling, Analysis, and Prediction (MAP) Program (80NSSC17K0109). We would also like to acknowledge the computational resources and support from the ARGO HPC Cluster team at George Mason University.

References

Adegoke, J. O. and Carleton, A. M.: Relations between soil moisture and satellite vegetation indices in the US Corn Belt, J. Hydrometeorol., 3, 395-405, https://doi.org/10.1175/1525-7541(2002)003<0395:RBSMAS>2.0.CO;2, 2002.

Albergel, C., Munier, S., Leroux, D.J., Dewaele, H., Fairbairn, D., Barbu, A.L., Gelati, E., Dorigo, W., Faroux, S., Meurey, C. and Le Moigne, P.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8. 0: LDAS-Monde assessment over the Euro-Mediterranean area, Geosci. Model Dev., 10, 3889-3912, https://doi.org/10.5194/gmd-10-3889-2017, 2017.

Andreadis, K. M. and Lettenmaier, D. P.: Assimilating remotely sensed snow observations into a macroscale hydrology model, Adv. Water Resour., 29, 872-886, https://doi.org/10.1016/j.advwatres.2005.08.004, 2006.

Arora, V.: Modeling vegetation as a dynamic component in soil - vegetation - atmosphere transfer schemes and hydrological models, Rev. Geophy., 40, 3-1, https://doi.org/10.1029/2001RG000103, 2002.

Ball, J. T., Woodrow, I. E. and Berry, J. A.: A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, in: Progress in photosynthesis research, Springer, Dordrecht, 221-224, https://doi.org/10.1007/978-94-017-0519-6_48, 1987.

Barbu, A. L., Calvet, J. C., Mahfouf, J. F., Albergel, C., and Lafont, S.: Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: grassland case study, Biogeosciences, 8, 1971-1986, https://doi.org/10.5194/bg-8-1971-2011, 2011.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O. and Roujean, J. L.: LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm, Remote Sens. Environ., 110, 275-286, https://doi.org/10.1016/j.rse.2007.02.018, 2007.

Cohen, W. B. and Justice, C. O.: Validating MODIS terrestrial ecology products: linking in situ and satellite measurements, Remote Sens. Environ., 70, 1-3, 1999.

Cracknell, A. P.: Advanced very high resolution radiometer AVHRR, CRC Press, 543, 1997.

Crow, W. T. and Wood, E. F.: The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97, Adv. Water Resour., 26, 137-149, https://doi.org/10.1016/S0309-1708(02)00088-X, 2003.

De Lannoy, G.J., Reichle, R.H., Arsenault, K.R., Houser, P.R., Kumar, S., Verhoest, N.E. and Pauwels, V.R.: Multiscale assimilation of Advanced Microwave Scanning Radiometer–EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado, Water Resour. Res., 48, https://doi.org/10.1029/2011WR010588, 2012

Di, L., Rundquist, D. C. and Han, L.: Modelling relationships between NDVI and precipitation during vegetative growth cycles, Int. J. Remote Sens., 15, 2121-2136, https://doi.org/10.1080/01431169408954231, 1994.

Dickinson, R. E., Shaikh, M., Bryant, R. and Graumlich, L.: Interactive canopies for a climate model, J. Climate, 11(, 2823-2836, https://doi.org/10.1175/1520-0442(1998)011<2823:ICFACM>2.0.CO;2, 1998.

Druel, A., Ciais, P., Krinner, G., and Peylin, P.: Modeling the vegetation dynamics of northern shrubs and mosses in the ORCHIDEE land surface model, J. Adv. Model Earth Sy., 11, 2020-2035, https://doi.org/10.1029/2018MS001531, 2019.

Durand, M. and Margulis, S. A.: Effects of uncertainty magnitude and accuracy on assimilation of multiscale measurements for snowpack characterization, J. Geophys. Res.: Atmos., 113, D02105, https://doi.org/10.1029/2007JD008662, 2008.

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation, Ocean dynam., 53, 343-367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Fairbairn, D., Barbu, A., Napoly, A., Albergel, C., Mahfouf, J. F., and Calvet, J. C.: The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. HESS, 21, 2015-2033, https://doi.org/10.5194/hess-21-2015-2017, 2017

Farrar, T. J., Nicholson, S. E. and Lare, A. R.: The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. II. NDVI response to soil oisture, Rremote Sens. Environ., 50, 121-133, https://doi.org/10.1016/0034-4257(94)90039-6, 1994.

Fisher, R. A., Koven, C. D., Anderegg, W. R., Christoffersen, B. O., Dietze, M. C., Farrior, C. E., Holm, J. A., Hurtt, G. C., Knox, R. G., Lawrence, P. J. and Lichstein, J. W.: Vegetation demographics in Earth System Models: A review of progress and priorities, Global Change Biol., 24, 35-54, https://doi.org/10.1111/gcb.13910, 2018.

Foley, J. A., Prentice, I. C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S. and Haxeltine, A.: An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics, Global Biogeochem. Cy., 10, 603-628, https://doi.org/10.1029/96GB02692, 1996.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R. and Wargan, K.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), J. Climate, 30, 5419-5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Ghatak, D., Zaitchik, B., Kumar, S., Matin, M., Bajracharya, B., Hain, C. and Anderson, M.: Influence of Precipitation Forcing Uncertainty on Hydrological Simulations with the NASA South Asia Land Data Assimilation System, Hydrology, 5, 57, https://doi.org/10.3390/hydrology5040057, 2018.

Gibelin, A. L., Calvet, J. C., Roujean, J. L., Jarlan, L. and Los, S. O.: Ability of the land surface model ISBA - A - gs to simulate leaf area index at the global scale: Comparison with satellites products, J. Geophys. Res.: Atmos., 111, D18102, https://doi.org/10.1029/2005JD006691, 2006.

Hansen, M. C., DeFries, R. S., Townshend, J. R. and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, Int. J. Remote Sens., 21, 1331-1364, https://doi.org/10.1080/014311600210209, 2000.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., Roy, D. P. and Morisette, J. T.: An overview of MODIS Land data processing and product status, Remote Sens. Environ., 83, 3-15, https://doi.org/10.1016/S0034-4257(02)00084-6, 2002.

Kim, Y. and Wang, G.： Impact of vegetation feedback on the response of precipitation to antecedent soil moisture anomalies over North America J. Hydrometeorol., 8, 534-550, https://doi.org/10.1175/JHM612.1, 2007.

Krinner, G., Viovy, N., de Noblet‐Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere‐biosphere system, Global Biogeochem. Cy., 19, GB1015, https://doi.org/10.1029/2003GB002199, 2005.

Kucharik, C. J., Foley, J. A., Delire, C., Fisher, V. A., Coe, M. T., Lenters, J. D., Young‐Molling, C., Ramankutty, N., Norman, J. M. and Gower, S. T.: Testing the performance of a dynamic global ecosystem model: water balance, carbon balance, and vegetation structure, Global Biogeochem. Cy., 14, 795-825, https://doi.org/10.1029/1999GB001138, 2000.

Kumar, S. V., Peters-Lidard, C. D., Tian, Y., Houser, P. R., Geiger, J., Olden, S., Lighty, L., Eastman, J. L., Doty, B., Dirmeyer, P. and Adams, J.: Land information system: An interoperable framework for high resolution land surface modeling, Environ. Modell. Softw., 21, 1402-1415, https://doi.org/10.1016/j.envsoft.2005.07.004, 2006.

Kumar, S. V., Peters-Lidard, C., Tian, Y., Reichle, R., Geiger, J., Alonge, C., Eylander, J. and Houser, P.: An integrated hydrologic modeling and data assimilation framework, Computer, 41, 52-59, https://doi.org/10.1109/MC.2008.475, 2008.

Kumar, S. V., Peters-Lidard, C. D., Mocko, D., Reichle, R., Liu, Y., Arsenault, K. R., Xia, Y., Ek, M., Riggs, G., Livneh, B. and Cosh, M.: Assimilation of remotely sensed soil moisture and snow depth retrievals for drought estimation, J. Hydrometeorol., 15, 2446-2469, https://doi.org/10.1175/JHM-D-13-0132.1, 2014.

Kumar, S.V., Zaitchik, B.F., Peters-Lidard, C.D., Rodell, M., Reichle, R., Li, B., Jasinski, M., Mocko, D., Getirana, A., De Lannoy, G. and Cosh, M.H.: Assimilation of gridded GRACE terrestrial water storage estimates in the North American Land Data Assimilation System, J. Hydrometeorol., 17, 1951-1972, https://doi.org/10.1175/JHM-D-15-0157.1, 2016

Kumar, S. V., Jasinski, M., Mocko, D. M., Rodell, M., Borak, J., Li, B., Beaudoing, H. K. and Peters-Lidard, C. D.: NCA-LDAS land analysis: Development and performance of a multisensor, multivariate land data assimilation system for the National Climate Assessment, J. Hydrometeorol., 20, 1571-1593, https://doi.org/10.1175/JHM-D-17-0125.1, 2019.

Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D. and Borak, J.: 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 US, J. Hydrometeorol., 20, 1359-1377, https://doi.org/10.1175/JHM-D-18-0237.1, 2019.

Ling, X. L., Fu, C. B., Guo, W. D. and Yang, Z. L.: Assimilation of remotely sensed LAI into CLM4CN using DART, J. Adv. Model. Earth Sy., https://doi.org/10.1029/2019MS001634, 2019.

Liu, Y., Liu, R. and Chen, J. M.: Retrospective retrieval of long‐term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data, J. Geophys. Res.: Biogeo., 117, G04003, https://doi.org/10.1029/2012JG002084, 2012.

Liu, Y., Peters‐Lidard, C.D., Kumar, S.V., Arsenault, K.R. and Mocko, D.M.: Blending satellite‐based snow depth products with in situ observations for streamflow predictions in the Upper Colorado River Basin, Water Resour. Res., 51, 1182-1202, https://doi.org/10.1002/2014WR016606, 2015

Morisette, J. T., Privette, J. L. and Justice, C. O.: A framework for the validation of MODIS land products, Remote Sens. Environ., 83, 77-96, https://doi.org/10.1016/S0034-4257(02)00088-3, 2002.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. R. and Lotsch, A.: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83, 214-231, https://doi.org/10.1016/S0034-4257(02)00074-3, 2002.

Niu, G. Y. and Yang, Z. L.: An observation‐based formulation of snow cover fraction and its evaluation over large North American river basins, J. Geophys. Res.: Atmos., 112, D21101, https://doi.org/10.1029/2007JD008674, 2007.

Niu, G. Y., Yang, Z. L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E. and Tewari, M.: The community Noah land surface model with multiparameterization options (Noah‐MP): 1. Model description and evaluation with local‐scale measurements, J. Geophys. Res.: Atmos., 116, D12109, https://doi.org/10.1029/2010JD015139, 2011.

Pagano, T. S. and Durham, R. M.: Moderate resolution imaging spectroradiometer (MODIS), in: Proceedings of SPIE 1939 Sensor Systems for the Early Earth Observing System Platforms, International Society for Optics and Photonics, Orlando, FL, United States, 25 August 1993, 2-17, https://doi.org/10.1117/12.152835, 1993

Pan, M. and Wood, E. F.: Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter, J. Hydrometeorol., 7, 534-547, https://doi.org/10.1175/JHM495.1, 2006.

Pauwels, V. R. and De Lannoy, G. J.: Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, J. Hydrometeorol., 7, 458-477, https://doi.org/10.1175/JHM490.1, 2006.

Pauwels, V. R., Verhoest, N. E., De Lannoy, G. J., Guissard, V., Lucau, C. and Defourny, P.: Optimization of a coupled hydrology–crop growth model through the assimilation of observed soil moisture and leaf area index values using an ensemble Kalman filter, Water Resour. Res., 43, W04421, https://doi.org/10.1029/2006WR004942, 2007.

Peters-Lidard, C. D., Houser, P. R., Tian, Y., Kumar, S. V., Geiger, J., Olden, S., Lighty, L., Doty, B., Dirmeyer, P., Adams, J. and Mitchell, K. High-performance Earth system modeling with

NASA/GSFC's Land Information System, Innovations Syst. Softw. Eng., 3, 157-165, https://doi.org/10.1007/s11334-007-0028-x, 2007.

Privette, J. L., Myneni, R. B., Knyazikhin, Y., Mukelabai, M., Roberts, G., Tian, Y., Wang, Y. and Leblanc, S. G.: Early spatial and temporal validation of MODIS LAI product in the Southern Africa Kalahari, Remote Sens. Environ., 83, 232-243, https://doi.org/10.1016/S0034-4257(02)00075-5, 2002.

Reichle, R. H., McLaughlin, D. B. and Entekhabi, D.: Hydrologic data assimilation with the ensemble Kalman filter, Mon. Weather Rev., 130, 103-114, https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2, 2002.

Reichle, R. H., Walker, J. P., Koster, R. D. and Houser, P. R.: Extended versus ensemble Kalman filtering for land data assimilation, J. Hydrometeorol., 3, 728-740, https://doi.org/10.1175/1525-7541(2002)003<0728:EVEKFF>2.0.CO;2, 2002.

Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P., Njoku, E. G. and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR‐E) and the Scanning Multichannel Microwave Radiometer (SMMR), J. Geophys. Res.: Atmos., 112, D09108, https://doi.org/10.1029/2006JD008033, 2007.

Reichle, R. H., Kumar, S. V., Mahanama, S. P., Koster, R. D. and Liu, Q.: Assimilation of satellite-derived skin temperature observations into land surface models, J. Hydrometeorol., 11, 1103-1122, https://doi.org/10.1175/2010JHM1262.1, 2010.

Richard, Y. and Poccard, I.: A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa, Int. J. Remote Sens., 19, 2907-2920, https://doi.org/10.1080/014311698214343, 1998.

Sabater, J. M., Rüdiger, C., Calvet, J. C., Fritz, N., Jarlan, L. and Kerr, Y.: Joint assimilation of surface soil moisture and LAI observations into a land surface model, Agr. Forest Meteorol., 148, 1362-1373, https://doi.org/10.1016/j.agrformet.2008.04.003, 2008.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanov, N. V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B. and Häme, T.: Multiscale analysis and validation of the MODIS LAI product: I. Uncertainty assessment, Remote Sens. Environ., 83, 414-430, https://doi.org/10.1016/S0034-4257(02)00047-0, 2002.

Wang, G. and Eltahir, E. A.: Role of vegetation dynamics in enhancing the low‐frequency variability of the Sahel rainfall, Water Resour Res., 36, 1013-1021, https://doi.org/10.1029/1999WR900361, 2000.

Wang, G., Sun, S. and Mei, R.: Vegetation dynamics contributes to the multi‐decadal variability of precipitation in the Amazon region, Geophys. Res. Lett., 38, L19703, https://doi.org/10.1029/2011GL049017, 2011.

Woodward, F. I. and Lomas, M. R.: Vegetation dynamics–simulating responses to climatic change, Biol. Rev., 79, 643-670, https://doi.org/10.1017/S1464793103006419, 2004.

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L. and Song, J. Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance, IEEE T. Geosci. Remote, 52, 209-223, https://doi.org/10.1109/TGRS.2013.2237780, 2013.

Yang, Z. L., Niu, G. Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah‐MP): 2. Evaluation over global river basins, J. Geophys. Res.: Atmos., 116, D12110, https://doi.org/10.1029/2010JD015140, 2011.

Yoon, Y., Kumar, S.V., Forman, B.A., Zaitchik, B.F., Kwon, Y., Qian, Y., Rupper, S., Maggioni, V., Houser, P., Kirschbaum, D. and Richey, A.: Evaluating the uncertainty of terrestrial water budget components over High Mountain Asia, Front. Earth Sci., 7, https://doi.org/10.3389/feart.2019.00120, 2019

Zhou, Y., McLaughlin, D. and Entekhabi, D.: Assessing the performance of the ensemble Kalman filter for land surface data assimilation, Mon. Weather Rev., 134, 2128-2142, https://doi.org/10.1175/MWR3153.1, 2006.