Reply to Referee #1 interactive comment

The author assesses the performance of the surface and root-zone soil moisture (SSM and RZSM) estimates by SMAP Level-4 DA system using an open loop (OL) simulations and two years in situ profile soil moisture observations at 2474 sites over mainland China. The anomaly Spearman's rank rather than Pearson correlation coefficient is calculated for comparisons and evaluations. In the following, to evaluate the efficiency of SMAP L4 DA system, the author chooses eight factors and uses methods of random forest regression and box plot comparisons to do the attribution analysis. Results show the improvement of SSM and RZSM estimates through the increased anomaly with in situ measurements, compared to OL based results. Three factors namely the standard deviation of the observation-minus-forecast Tb residuals, errors in precipitation forcing data and microwave soil roughness parameter H are found dominantly affecting the efficiency for SSM and RZSM estimates by SMAP Level-4 DA system. Furthermore, the SSM-RZSM coupling strength characterizing the surface to subsurface physics in CLSM is evaluated based on in situ measurements and OL and DA estimates.

Although it is enough to understand what 'went on', the scientific and English language is imprecise in various places as well as some cited information. I have given some examples below and labeled some in the attachment, but the authors should go throughout the entire manuscript carefully, and check that the descriptions and citations are as exact as possible. On the other hand, the author often uses different tenses in a paragraph even in one sentence, making presentations a bit messy. Additionally, too many brackets are used to present information. Please do the appropriate revisions, as a reader, I tend to get accurate information rather than having a hesitation on whether I shall ignore/keep the information, and thereby guess how does each step be carried on and may lose interest. I am sorry. I would say, maybe some main contents are ignored by the reviewer because of the weak presentation.

We sincerely thank the reviewer for the constructive and thoughtful comments.

Many of the reviewer's comments are, unfortunately, based on misunderstandings. We apologize if the original text was not sufficiently clear. To address the comments, we have undertaken major revisions of the text, including a careful revision of the imprecise expressions and the tenses throughout the manuscript.

Major comments:

1. In line 77, please specify key CLSM parameters and give the reason why you choose these parameters.

The accuracy of CLSM simulated soil moisture is affected by a range of vegetation and soil parameters. For example, Dong et al. (2019) demonstrated that soil moisture DA within the CLSM system is strongly affected by LAI. Therefore, LAI is used here to represent the error impacts from vegetation.

In addition, root-zone soil moisture dynamics are controlled by their connection with the surface soil moisture, soil hydraulic properties, and soil water transport. The bulk relationship between root-zone and surface soil moisture can be captured by vertical coupling strength of soil moisture (Kumar et al., 2009).

Therefore, this study uses LAI and surface-rootzone coupling strength to characterize the error impacts from vegetation and soil parameters on CLSM. We have further clarified this in the revised manuscript.

[1] Dong, J., Crow, W.T., Reichle, R., Liu, Q., Lei, F., and Cosh, M.: A global assessment of added value in the SMAP Level 4 soil moisture product relative to its baseline land surface model, Geophys. Res. Lett., 46, 6604-6613, doi:10.1029/2019GL083398, 2019.

[2] Kumar, S.V., Reichle, R.H., Koster, R.D., Crow, W.T., and Peters-Lidard, C.D.: Role of subsurface physics in the assimilation of surface soil moisture observations, J. Hydrometeorol., 10, 1534-1547, doi:10.1175/2009JHM1134.1, 2009.

2. In line 105, please clarify whether the OL run is conducted in this study? In line 29, I am sorry I cannot understand, what does "error in Tb observation space" mean? Please also explicitly clarify Tb error. In line 108-111, please clarify/specify "microwave soil roughness parameters, a vegetation structure parameter, and the microwave scattering albedo". "Soil roughness parameters" are used throughout the paper but without explaining what they are. Does it refer to both h and N, or s and L, or others? Additionally, please keep the cited information correct (equation A1 instead of B1). Please carefully check throughout the manuscript.

Yes, the OL run is conducted in this study.

Tb error refers to the difference between Tb observations from SMAP sensor and the Tb simulations obtained via the tau-omega radiative transfer model.

The microwave soil roughness parameters refers to parameter h that accounts for soil dielectric properties that vary at the subwavelength scale, the vegetation structure parameter refers to the nadir vegetation opacity τ, and the microwave scattering albedo ω refers to the single-scattering albedo defined as the fractional power scattered within the vegetation layer.

We have further explained these expressions in the revised manuscript. In addition, we have checked throughout the manuscript and make corresponding corrections.

3. In line 120, LH and SH are mentioned. LH error is seen, please if possible, give the reason why SH error is disregarded.

Note that note that since LH and SH are generally (strongly) anti-correlated, it is not appropriate to include both in a single random forest analysis – since including both would yield biased (high) regression weights for LH and SH. We have further clarified this in the corresponding paragraph of Section 2.3 in the revised manuscript.

4. In line 127, if possible, please give the figure plotting the distribution of CASMOS as new Fig. 1.

We have added a separate figure showing the distribution of CASMOS as new Fig. 1.

5. In line 133-134, I cannot be convinced by the described reason about the use of Spearman correlation rather than Pearson correlation. Could you explain more? Wikipedia says that the Spearman correlation concerns the rank and Pearson correlation the mean. Do you calculate Pearson correlation based results? Please give the definition of outliers excluded in this study. In line 144-147, please clarify why these five control factors are chosen, and why the difference in clay fraction across the vertical can be used to quantify vertical variability in soil properties.

Note that Pearson correlation assumes the linear consistency of underlying variables. However, this assumption may be adversely affected by outliers. To avoid ad-hoc thresholds, we do not exclude any soil moisture outliers and employ Spearman's rank correlation, which is less sensitive to such outliers. Nonetheless, we repeat the analysis based on Pearson correlation (see Figs. 1-2 below). The Pearson-based results are quantitatively consistent with the results using Spearman's correlation. We have further clarified this in the corresponding paragraph of Section 2.2 in the revised manuscript.

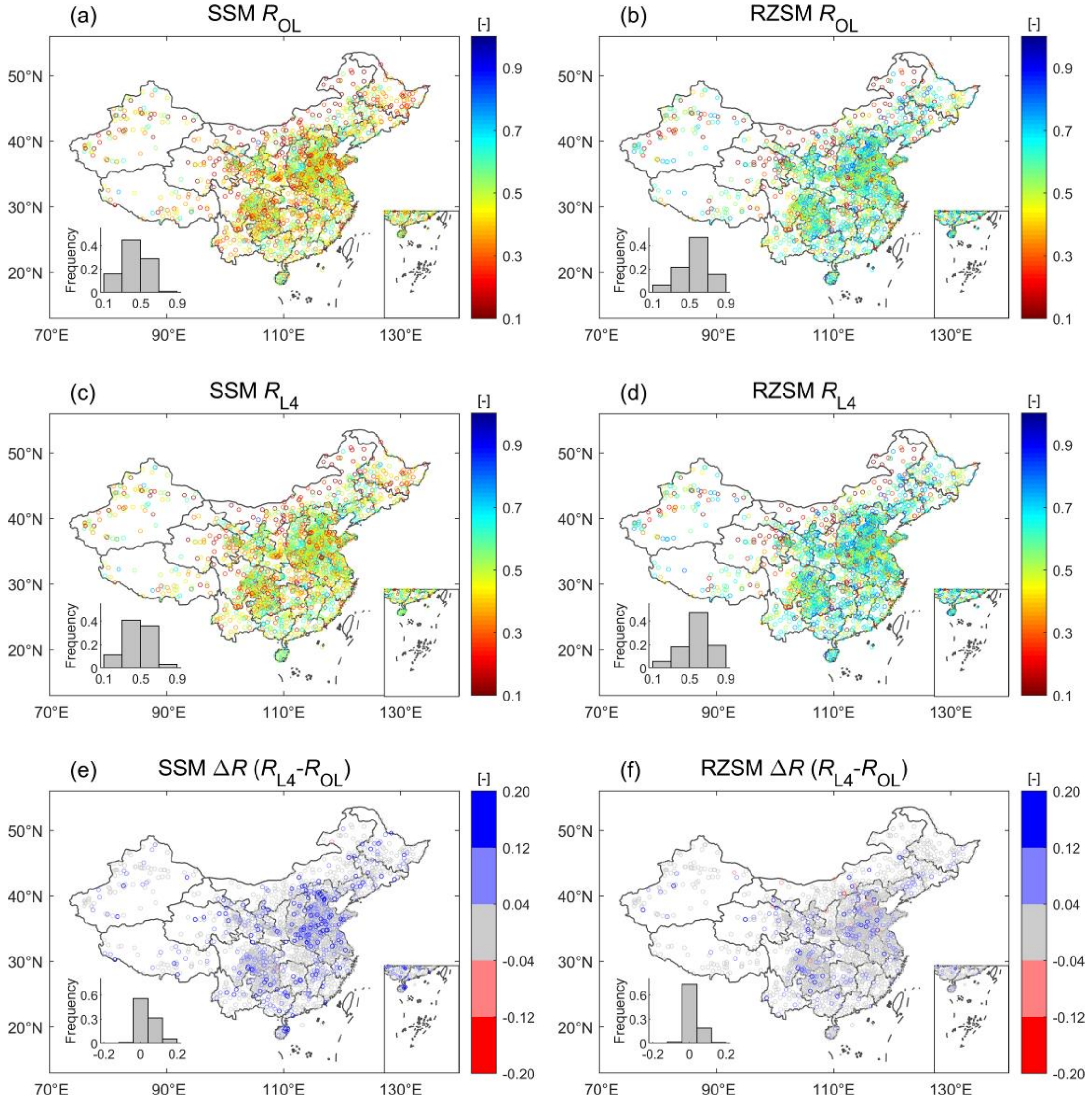


Fig. 1 Same content as in Fig. 2 of the manuscript, except that the correlation between in-situ soil moisture measurements and SMAP is measured using Pearson correlation.

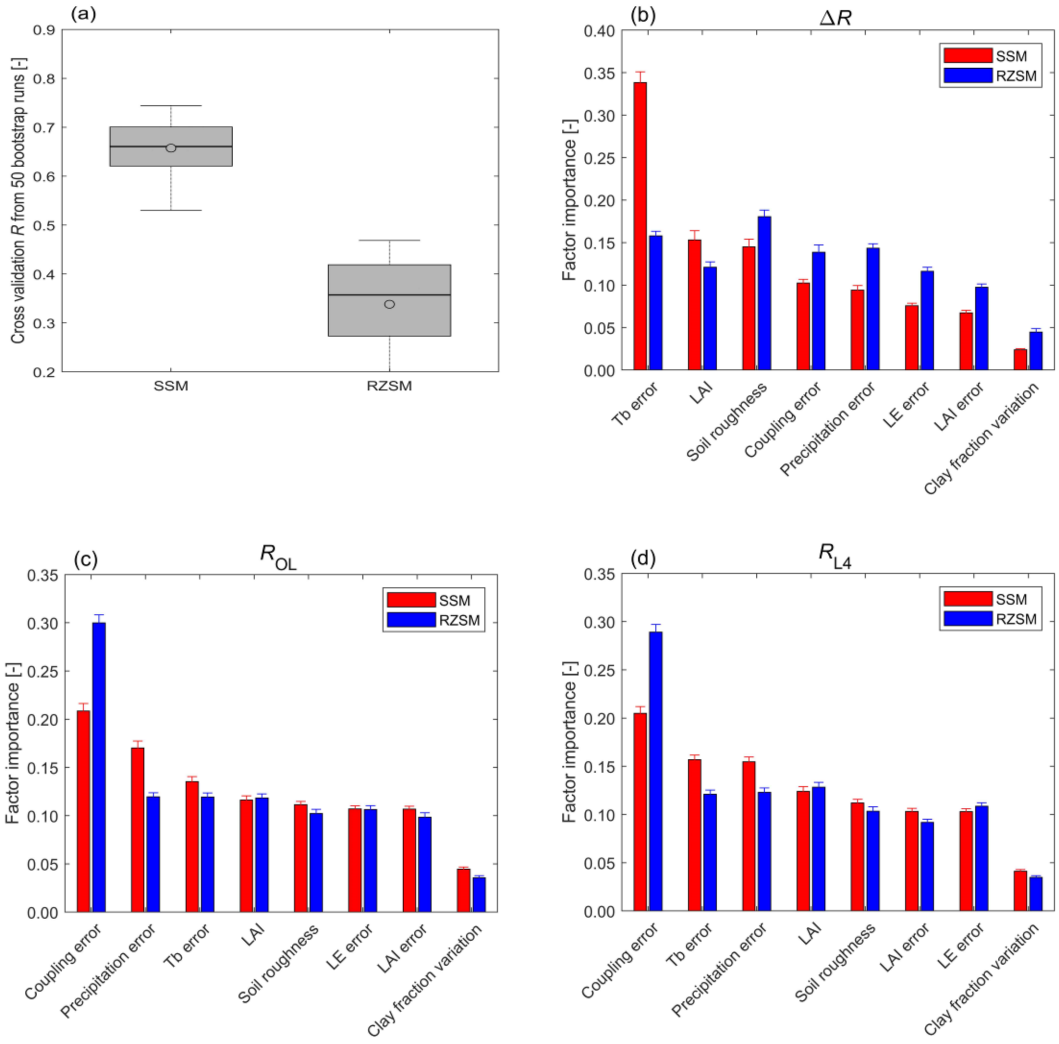


Fig. 2 Same content as in Fig. 4 of the manuscript, except that the correlation between in-situ soil moisture measurements and SMAP is measured using Pearson correlation.

The above response also applies to the Major comment #7 from Reviewer #2.

Re. the comment about the choice of the control factors: As mentioned in the abstract, the modeling portion of the SMAP L4 system consists of two components: land surface modelling (LSM) and radiative transfer modeling (RTM). Therefore, we select control factors from each of the two components.

For the LSM component, the errors can be attributed to potential factors including: 1) model input forcing errors of a) precipitation and b) LAI; 2) model structure errors in a) characterizing SSM-RZSM coupling strength and b) the presence of vertical variability in soil properties; 3) model output error of LE.

For the RTM component, errors are characterized by: 1) DA innovation, i.e., SMAP-observed minus RTM-simulated Tb; 2) the environmental factors that complicate the DA analysis when assimilating Tb observations, which include the magnitude of a) microwave soil roughness and b) LAI.

These eight control factors from the above-mentioned five aspects determine the crucial aspects of both the LSM and RTM components in the L4 system and are readily quantifiable using remote sensing products. Thus, they are selected to investigate the mechanism underlying the L4 improvement observed in this study. We have further clarified this in the corresponding paragraph of Section 1 in the revised manuscript.

The above response also applies to the Major comment #1 by Reviewer #2 and Major comment #1 by Reviewer #3.

Re. the <u>comment about the difference in clay fraction across the vertical</u>: With the exception of some surface-layer moisture transport parameters, CLSM assumes soil texture and associated soil parameters are vertically homogeneous within the soil column. However, the Harmonized World Soil Database (HWSD) often captures distinct vertical variations in soil properties. Therefore, since it is largely neglected by CLSM, the magnitude of vertical heterogeneity in soil texture may be an effective proxy for overall CLSM soil moisture accuracy. We have further clarified this in the corresponding paragraph of Section 2.3 in the revised manuscript.

6. In Table 1, please clarify why different LAI products are used? What is the relationship between these two LAI datasets? Why does SMAP L4 LAI be used for LSM rather than RTM, which simulates Tb that is used for comparisons to SMAP Tb.

The inherent LAI in SMAP L4 system is merged from a MODIS/Geoland-based data product (Mahanama et al., 2015; Reichle et al., 2017).

To correctly characterize error in LAI of SMAP L4, we use LAI product from an entirely independent source, i.e. from the SPOT satellite. The prominent difference between SMAP L4 LAI and SPOT LAI is that the former uses an LAI climatology from the period 1999-2011, while the latter is the actual LAI time series with inter-annual variation.

Note that besides the LAI from SMAP L4 system, we only use one external LAI product of SPOT VGT. We have correctly listed both LAI datasets in Table 1, and further clarified in Section 2.3: “Note that the LAI used in the L4 system is a merged climatology from Moderate Resolution Imaging Spectroradiometer (MODIS) and Geoland data based on satellite observations of the Normalized Difference Vegetation Index (Mahanama et al., 2015; Reichle et al., 2017a)”.

[1] Mahanama, S. P., and Coauthors: Land boundary conditions for the Goddard Earth Observing System model version 5 (GEOS-5) climate modeling system–Recent updates and data file descriptions. NASA/TM-2015-104606, Vol. 39, 55 pp. NASA Goddard Space Flight Center, Greenbelt, MD. Available at https://ntrs.nasa.gov/search.jsp?R=20160002967, 2015.

[2] Reichle, R. H., and Coauthors: Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. J. Hydrometeorol. 18(10), 2621–2645, 10.1175/JHM-D-17-0063.1, 2017.

The above response also applies to Major comment #5 by Reviewer #2.

7. In line 153, why is there a joint error in SMAP Tb observations and RTM Tb simulations? Sorry if I misunderstood something, what does "joint" mean? How do you quantify this joint error and what is the rationality behind?

The expression of "joint" is meant to refer to the combined error in the SMAP Tb observations and RTM Tb simulations. The DA innovation is estimated by subtracting the SMAP Tb observations from RTM Tb simulations. In the revised manuscript, we have modified this expression as "1) estimates of the DA innovation, namely difference between SMAP Tb observations and RTM Tb simulations".

8. In line 154, "the magnitude of LAI (as a proxy for the vegetation optical depth at microwave frequencies, which modulates the sensitivity of the observed Tb to SSM conditions)". The description is inaccurate. LAI should be as a proxy for the estimation of vegetation optical depth. Please clarify how vegetation optical depth modulates the sensitivity of the observed Tb to SSM conditions, it is hard to make the audience understand who does not be familiar with the zero-order RTM.

We thank the reviewer for bringing this to our attention. In the revised manuscript, we have clarified this expression.

9. In line 156-160, please make expressions precise. You give "e.g.," may I ask what else do you use, please list every item as accurate as possible, as such, readers and the author are on the same page. In line 160, I fully doubt "because increased LAI is associated with decreased soil moisture information content in microwave observations", is it true? How do you explain, for example, when vegetation is mature, the soil experiences drying and wetting processes? Please make expressions accurate.

In line 156-160, the first category of factors addresses errors fed into the L4 system include: 1) error in CLSM rainfall forcing data; 2) error in SSM-RZSM coupling strength; 3) vertical variability of clay fraction; 4) SMAP L4 LAI error; 5) output LE error; 6) Tb error. The second category of factors is based on the magnitude of the variable itself and include microwave soil roughness and annual mean LAI. We have made it clearer in the revised manuscript.

Regarding to the comment for Line 160, please see our reply to Major comment #8.

10. In line 204-205, please clarify the reason.

We have clarified why using the difference in clay fraction across the vertical soil profile, i.e., the clay fraction difference between topsoil and deep-layer soil to quantify vertical variability in soil properties. Please see our reply to Major comment #5.

11. In line 210, why the anomaly SSM and RZSM are not used for Eq. 1, because in previous it is mentioned that anomaly Spearman's rank correlation is calculated with in-situ observations.

Indeed, the anomaly SSM and RZSM are used in the Eq. 1. We have made it clearer in the revised manuscript.

12. In line 214, "Cases with negative CP do not exist." I have litter doubt whether the in situ measurements will show that α is greater than 2.0, then CP can be negative? Please confirm this.

Based on the in-situ measurements during our 2-year study period, we do not observe any negative CP.

13. In line 227, please explicitly clarify "error" in FLUXCOM LE. Does this error refer to the uncertainties mentioned in line 186?

The FLUXCOM LE product is generated via merging energy flux measurements from FLUXNET eddy covariance towers with remote sensing and meteorological data. The error of FLUXCOM LE could stem from each data source and also from the merging process. This error also refers to the uncertainties mentioned in Line 186.

14. In line 235, please clarify "three independent sources (x, y and z)", does it refer to geographic location or one of the variables mentioned in your study? Please also explicitly explain two instrumental variables I and J. I did not see the time information mentioned in Eq. 2. Please is "(I and J, i.e., It = αxPt-1 + Bx + εxt-1, Jt = αyPt-1 + By + εyt-1)" important in the calculation, if so, please list it as an independent equation. Please clarify ε_(xt-1) or do you mean ε_(x,t-1)? Additionally, too much information is listed in brackets, shall readers ignore/keep this information? Please do revisions.

In Line 235, the expression of "three independent sources (x, y and z)" refer to any of three geophysical variables that are not linearly correlated in each of their time series.

The instrumental variable I refer to the lag-1 time series of variable x, and instrumental variable J refer to the lag-1 time series of variable y.

To be clearer, we have listed the following equation originally listed in the bracket: $I_t = \alpha_x P_{t-1} + B_x + \varepsilon_{x,t-1}$, $J_t = \alpha_y P_{t-1} + B_y + \varepsilon_{y,t-1}$ as new Eqs. (3) and (4). In addition, ε_{xt-1} and ε_{yt-1} have been more precisely denoted as $\varepsilon_{x,t-1}$ and $\varepsilon_{y,t-1}$ respectively.

In correspondence with Eqs. (3) and (4), the Eq. (2) have been shown with time information, which is $x_t = \alpha_x P_t + B_x + \varepsilon_{x,t}$

In addition, in the revised manuscript content within the brackets are rearranged to be clearer to readers in Section 2.5.

15. In line 255, "based on the output of RF", as a reviewer, I do not know more about RF, what are inputs for RF? I think the introduction of RF is too general and not informative. Please do revisions. Taking this paragraph as a case, past and present tenses are mixed used. Please do revisions.

As a machine learning based regression approach, RF uses the selected eight control factors as regressors to estimate, or regress the DA improvement (i.e., the difference of OL and DA soil moisture accuracy) for both SSM and RZSM estimates. Therefore, the input for RF is our eight control factors (see Table 1) that cover two perspectives of L4, and the output of RF is the DA improvement in L4 SSM and RZSM sampled at 2474 sites. Note that, by training the eight control factors to capture observed DA improvement, RF can also summarize the relative importance of each control factor in controlling the L4 DA improvement.

In addition, we have revised tenses throughout the manuscript.

16. In Fig. 1a-d, what is the maximum value for R? Can it reach 0.9? If not, please adjust the scalar. Please rewrote the caption of Fig.1.

The maximum of R in original Fig. 1a-d can reach to approximately 0.9, so their common maximum have not been adjusted.

17. In line 261, "an increasing trend of SSM estimation skill moving from northwest to southeast China", if possible, please write a short sentence to explain the reason.

The reason for the observed "an increasing trend of SSM estimation skill moving from northwest to southeast China" is most likely due to the similar spatial pattern of gauge density. We have briefly explained this in the revised manuscript.

18. In line 280-281," Errors in the CLSM precipitation forcing are relatively higher in northern and northwestern areas of China (Fig. 2a), where the gauge density is generally more sparse than southern China." I agree with this point. But I am sorry if I misunderstood. The magnitude of precipitation on the northwestern part may be smaller than on the southern part, as such, there is a possibility that errors may present a reverse trend, is this a case? Please confirm.

By comparing the gauge density of northern and southern China, we can clearly observe that the former is sparser than the latter, which inevitably results in higher interpolating error of precipitation forcing in northern China.

*19. Figure 2g, please revise the title as "soil roughness parameter *". In Fig. 2h, the maximum value of LAI is 2.0 m2/m2, please confirm. Fig. 2f, please revise the title as "the standard deviation of O-F Tb residuals". I think the meaning of "O-F Tb residuals" is different from Tb error itself.*

We have revised the titles of the subplot as suggested.

The maximum of annual mean LAI is higher than 2.0 m^2/m^2. In the original manuscript, we set the colorbar maximum to be 2.0 m^2/m^2, so that the spatial difference in LAI magnitude can be easily observed in Fig. 2h. We have recovered the colorbar maximum of 4.0 m^2/m^2 and change the colormap to be nonlinear to reconcile the two issues. In addition, it should be noted that the SMAP LAI time series during growing season frequently exceed 4.0 m^2/m^2 as expected, whereas our Fig. 2h shows the annual mean LAI covering both growing and non-growing seasons and hence shows lower maximum value of LAI.

20. In line 297-298, "The 2017-2018 mean of soil roughness and the 2017-2018 mean LAI show higher values in southwest and southeast China (Fig. 2g-h)." The sentence is not informative. Please revise.

We have revised the original expression as: "The 2017-2018 mean of soil roughness shows a relatively scattered spatial pattern (Fig. 3g), while the 2017-2018 mean LAI shows higher values in southwest and southeast China (Fig. 3h)".

21. In line 335-336, OL run does not implement DA, why "Tb error (microwave soil roughness)" are involved. Please clarify. I am sorry if I misunderstood something.

Indeed, the OL run does not involve the implementation of DA, and its errors are therefore not related to Tb error or microwave soil roughness. We had tried to explain this point in the subsequent text of original manuscript, which stated that the observed high correlation between OL skill and these two factors does not imply causality. However, showing the feature importance of Tb error and microwave soil roughness to the OL skill (R_{OL}) in original Fig. 3c seems to be misleading anyway. Therefore, to avoid confusion, we set the bars that represent RTM-related feature importance to be semi-transparent in original Fig. 3c.

22. In line 389, "OL does get worse with increasing roughness, there is more room for improvement as the roughness increases", please clarify whether the increase of roughness is physically reasonable.

In this context, we are not implying to increase soil roughness, which would be physically implausible, as it is a function of soil moisture. The logic should be that in areas with higher soil roughness, the possibility of improving OL skill is higher. We have further clarified this point in the revised manuscript.

23. In line 441, "it is unclear whether or not the observed SSM-RZSM coupling strength biases are real in an absolute sense – or simply reflect inconsistencies in the depth of modelled versus observed SSM and RZSM time series". I am sorry, I am confused whether the coupling strength based on in situ measurements can represent the real?

Ideally, when comparing the SSM-RZSM coupling strength of in-situ measurement and that of CLSM, SSM and RZSM data for identical depths from both in-situ measurements and CLSM should be used. However, the depth of first-layer SSM measurement is 0~10cm, which is thicker than CLSM SSM of 0~5cm. This discrepancy could inherently result in higher fluctuation of CLSM SSM simulation than that of SSM measurement, and consequently lower SSM-RZSM coupling of CLSM simulation (CP_{OL}) than that of measurement (CP_{obs}). Therefore, we cannot conclude that the observed lower CP_{OL} compared to CP_{obs} is due to the negative bias of SSM-RZSM coupling strength, or the depth inconsistencies of CLSM modelled versus observed SSM and RZSM time series.

24. In Conclusions, the second and fourth paragraphs have duplicate content. Please do revisions.

We have removed the duplication in the revised manuscript.

25. In line 451-452, "the partitioning of the available energy into latent and sensible heat (LE error) and the microwave radiative transfer modeling (Tb error)." is not informative.

We have revised the original expression as follows: "…additional focus should thus be placed on improving the model's characterization of the microwave radiative transfer modeling (Tb error), together with the partitioning of the available energy into latent and sensible heat (LE error)."

Minor comments:

1. Please give the full name for abbreviations when they appear for the first time. The examples are SPOT VGT and EASE. Please carefully check throughout the manuscript.

We have carefully checked the first-time abbreviations and made corresponding revisions throughout the manuscript.

2. In line 216, SMAP L4 CP estimates (CP_{OL}), please confirm. You mentioned SMAP L4 is the assimilation experiment.

We have revised the original expression as "CP estimates of OL (CP_{OL})…"

3. Please confirm the use of RTM-related, R-values, and so on throughout the whole paper, as well as the use of "their" and "our".

We have carefully checked those occurrences and made corresponding revisions throughout the manuscript.

Re. the comments in the annotated manuscript pdf file:

We have made corresponding revisions in the manuscript.

Reply to Referee #2 interactive comment

The paper evaluates the data assimilation efficiency of SMAP brightness temperature data by updating the root-zone soil moisture with CLSM (the Catchment Land Surface Model) model and an RTM model (radiative transfer modeling). The result of soil moisture filed delta_R increments then identifies substantial factors that control this data assimilation efficiency, such as precipitation error and SSM-RZSM coupling strength error. I appreciate the motivation of this paper, and its conclusion and inference are probably attractive to the L-band TB data assimilation community. However, I cannot agree on the methodology part of this paper, and I don't think the findings would help with the further development of RZSM DA improvements.

We thank the reviewer for the constructive criticisms and helpful comments.

Note that the results using Pearson and Spearman's correlation are qualitatively consistent, as will be detailed in our response to Major comment #7.

We would also like to clarify that SMAP L4 is the only operational global DA system that assimilates L-band Tb and provides near-real-time root-zone soil moisture information. However, SMAP L4 has been mainly evaluated across sparse in-situ sites and SMAP cal/val sites (mainly in the US, Europe and Australia) and factors affecting its accuracy are still largely unknown, which is particularly true for RZSM.

Therefore, based on soil moisture observations from 2474 in-situ sites in China, this is the first study that comprehensively quantifies the SMAP L4 SSM and RZSM skill improvements and their major error sources, and further identifies the key priorities for future L4 development. For instance, SSM-RZSM coupling strength is identified as the most important factor in determining the L4 RZSM accuracy. Therefore, the appropriate representation of SSM-RZSM coupling strength should be considered as a priority for developing next generation of L4 system.

Additionally, we would like to point out that LSMs, RTMs and different variants of DA algorithms typically share similar structures. Therefore, our findings are not limited to the SMAP L4 system but are expected to be transferable for diagnosing and improving general soil moisture DA systems and enhancing their RZSM accuracies. We have added more emphasis on these aspects in the revised manuscript.

Major comments:

1. Line 23 & Line 75-79, Line 144 & Line 152: how do the authors select these eight control factors?

Re. the <u>comment about the choice of the control factors</u>: We provide the same response to Major comment #5 by Reviewer #1. For easy reference, please see below:

As mentioned in the abstract, the modeling portion of the SMAP L4 system consists of two components: land surface modelling (LSM) and radiative transfer modeling (RTM). Therefore, we select control factors from each of the two components.

For the LSM component, the errors can be attributed to potential factors including: 1) model input forcing errors of a) precipitation and b) LAI; 2) model structure errors in a) characterizing SSM-RZSM coupling strength and b) the presence of vertical variability in soil properties; 3) model output error of LE.

For the RTM component, errors are characterized by: 1) DA innovation, i.e., SMAP-observed minus RTM-simulated Tb; 2) the environmental factors that complicate the DA analysis when assimilating Tb observations, which include the magnitude of a) microwave soil roughness and b) LAI.

These eight control factors from the above-mentioned five aspects determine the crucial aspects of both the LSM and RTM components in the L4 system and are readily quantifiable using remote sensing products. Thus, they are selected to investigate the mechanism underlying the L4 improvement observed in this study. We have further clarified this in the corresponding paragraph of Section 1 in the revised manuscript.

2. Line 80, please show which part of the paper corresponds to each sentence. For instance, "Next, the in-situ measurements…" As I see, only figure 1 is about the in-situ measurements.

In the revised manuscript, we have specifically clarified which section of the paper corresponds to each sentence of the guidance paragraph in original Line 80-85.

To be clearer, we have added an illustrative figure showing the spatial distribution of in-situ sites as new Fig. 1 in the revised manuscript.

3. Still, Line 80-81, the soil moisture profile measurements from CMA networks can reach 100 cm. please refer to: Han Shuai, Shi Chunxiang, Jiang Lipeng, Zhang Tao, Liang Xiao, Jiang Zhiwei, Xu Bin, Li Xianfeng, Zhu Zhi, Lin Hongjin. The Simulation and Evaluation of Soil Moisture Based on CLDAS[J]. Journal of Applied Meteorological Science, 2017, 28(3): 369-378. I suggest the authors separate the sites that contain measurements with 100 cm and the rest in the analysis.

We have included the suggested reference in the revised manuscript.

Re. the <u>comment about separating the sites with 100 cm measurements</u>: According to the in-situ measurements data provider, constrained by soil profile condition, only a fraction of 2474 sites have complete soil moisture measurements up to 100 cm. In addition, in the SMAP L4 validation procedure using SMAP core validation sites, although a few of the sparse-network sites have deeper-layer measurements (typically in the 80 cm to 100 cm range), Reichle et al. (2019) used measurements from sensors placed within 50 cm from the surface, as they found the spotty in-situ measurements time series at the deeper depths of relatively little use. Therefore, we routinely used in-situ data only down to 50 cm to evaluate the RZSM estimates of L4 and OL.

[1] Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J., Kimball, J. S.,

and Kolassa, J.: Version 4 of the SMAP Level-4 soil moisture algorithm and data product, J. Adv. Model Earth Sy., 11(10), 3106-3130, doi:10.1029/2019MS001729, 2019.

4. Section 2.2, the repetitiveness of in-situ soil moisture measurements is questionable. Line 128, the atmospheric elements such as air temperature, humidity, etc. of these stations cover different land-use types. When it comes to soil moisture, due to high spatial difference, the in–situ soil moisture profile measurements may vary a lot to the station outside. In standard, all these CMA stations should only have grassland or bare soil land types. Other land covers are impossible, and this affects precipitation, evaporation, draining, etc., as authors said in Line 144-146. The scale mismatch between CLSM outputs and in-situ measurement would exceed the accuracy indicates in the evaluation.

We thank the reviewer for the insight into the CMA data.

For the random point-to-grid upscaling errors of sampling, we agree that using in-situ soil moisture data to validate the large-scale L4 product is challenged by spatial representativeness error. However, as demonstrated in Dong et al. (2020), the representativeness error is essentially random and can be averaged out when sampled across multiple sites. Therefore, it mainly affects the absolute soil moisture evaluation metrics, and has no impacts on the relative accuracy. Given that this study is mainly interested in the relative accuracy of OL and L4 data and approximately 2474 sites have been used, representativeness error is expected to have a minor impact on our conclusions.

For the systematic sampling errors, we analyze the site "representativeness" using the 500m MODIS Land Cover product (MCD12Q1 v6) in 2017, IGBP dataset. First, we take the land cover (LC) type of the MODIS grid cell where a given in-situ site is located as the ground-based LC type. Next, we search all the MODIS grid cells that fall within the SMAP 9km EASE grid cell where this in-situ site is located. The latter area consists of about 20 x 20 = 400 MODIS grid cells. We calculate the fraction of these 400 MODIS grid cells that have the same LC type as the ground-based LC and define this fraction as the site representativeness. Fig. 1 shows a histogram of the site representativeness of all 2474 sites used. Similar results were found for 2018 (not shown).

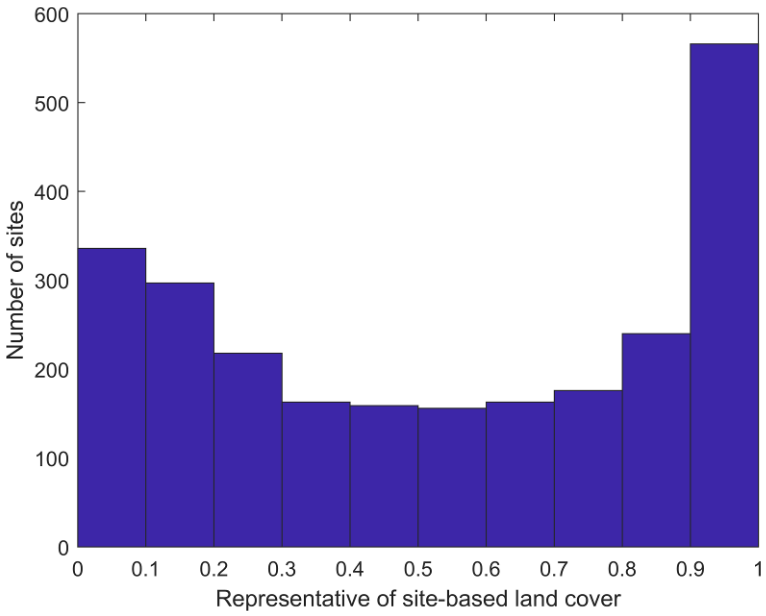


Fig. 1 Site representativeness of all 2474 sites. See text above for details.

It shows 59% of sites have site representativeness higher than 40%, and this fraction is 52% for site representativeness higher than 50%, and 46% for site representativeness higher than 60%. When we only use sites with site representativeness higher than 50% for the RF attribute analysis, the top 3 factors controlling skill improvement ($R_{\mathrm{L4}} - R_{\mathrm{OL}}$) L4 skill ($R_{\mathrm{L4}}$), and OL skill ($R_{\mathrm{OL}}$) are still the same, although the precipitation error becomes the top influencer for R_{L4} (Fig. 2).

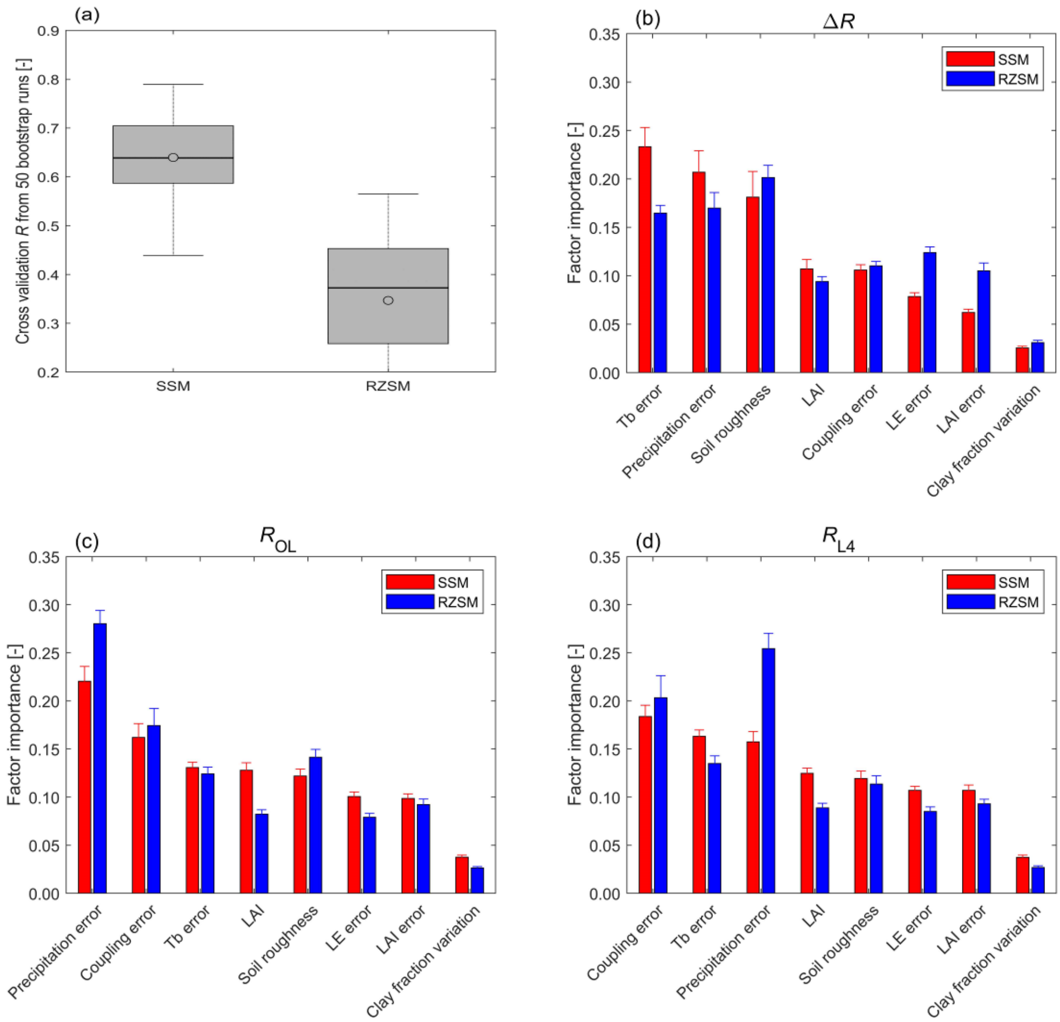


Fig. 2 Same content as in Fig. 4 of the manuscript, except that we use sites with site representativeness higher than 50%.

We have further clarified this in Section 3.3.1 of the revised manuscript.

[1] Dong, J., Crow, W.T., Tobin, J. K., Cosh, H. M., Bosch, D. D., Starks, J. P., Seyfried, M., and Collins, H. C. Comparison of microwave remote sensing and land surface modeling in surface soil moisture climatology estimation, Remote Sens. Environ., 242, 111756, doi :10.1016/j.rse.2020.111756, 2020.

5. Line 161, Table 1 & Line 192, could we use the same LAI data? As well as the rainfall data.

Re. the comment about LAI dataset: We provide the same response to Major comment #6 by Reviewer #1, for easy reference, please see below:

The inherent LAI in SMAP L4 system is merged from a MODIS/Geoland-based data product (Mahanama et al., 2015; Reichle et al., 2017).

To correctly characterize error in LAI of SMAP L4, we use LAI product from an entirely independent source, i.e. from the SPOT satellite. The prominent difference between SMAP L4 LAI and SPOT LAI is that the former uses an LAI climatology from the period 1999-2011, while the latter is the actual LAI time series with inter-annual variation.

Note that besides the LAI from SMAP L4 system, we only use one external LAI product of SPOT VGT. We have correctly listed both LAI datasets in Table 1, and further clarified in Section 2.3: "Note that the LAI used in the L4 system is a merged climatology from Moderate Resolution Imaging Spectroradiometer (MODIS) and Geoland data based on satellite observations of the Normalized Difference Vegetation Index (Mahanama et al., 2015; Reichle et al., 2017a)".

[1] Mahanama, S. P., and Coauthors: Land boundary conditions for the Goddard Earth Observing System model version 5 (GEOS-5) climate modeling system–Recent updates and data file descriptions. NASA/TM-2015-104606, Vol. 39, 55 pp. NASA Goddard Space Flight Center, Greenbelt, MD. Available at https://ntrs.nasa.gov/search.jsp?R=20160002967, 2015.

[2] Reichle, R. H., and Coauthors: Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements. J. Hydrometeorol. 18(10), 2621–2645, 10.1175/JHM-D-17-0063.1, 2017.

Re. the comment about rainfall dataset: The result of precipitation error presented in the manuscript only involves rain gauge (CGDPA) data, other rainfall data are used to prove the robustness of analysis results from CGDPA.

6. Line 248, what is the DA efficiency? Line 309, it says "the efficiency of SMAP L4 DA (i.e., R = RL4 - ROL)". The data efficiency is not that simple, indeed. Please refer to (Nearing et al. 2018) for the definition of data assimilation efficiency, or provide where this "the efficiency of SMAP L4 DA (i.e., R = RL4 - ROL)" comes from in citations. Nearing, G., Yatheendradas, S., Crow, W., Zhan, X., Liu, J., & Chen, F. (2018). The Efficiency of Data Assimilation. Water Resources Research, 54, 6374-6392.

We have changed "DA efficiency" to "DA skill improvement", which is a more precise terminology for quantifying R_{L4} and R_{OL} differences.

7. Figure 1 & Section 3.1, one of the conclusions in this paper that RZSM is improved by assimilating brightness temperature. Figure 1 & Section 3.1 are the only evidence to support this view, which is vital to the following paragraphs. By any two datasets, the increased correlation coefficient is hard to address improvement. unRMSE, bias, and other characters shall also be accounted as SMAP evaluate its soil moisture *products. As in Line 52-54, "observations-minus-forecast residuals" may not be sufficient, but it doesn't mean it is unnecessary. Besides, Spearman's rank correlation coefficient is very loose in statistics. Pearson correlation can assess linear relationships hypothesized in ordinary DA filters. Line 133-134 is not solid for support the advantage of Spearman's and it should clarify what the outliers are.*

We thank the reviewer for the constructive comments. We would like to clarify that the goal of this manuscript is not to prove increased RZSM accuracy, but to evaluate L4 RZSM and understand the mechanism that controls L4 RZSM accuracy.

Nonetheless, it is indeed our finding that L4 DA system improves OL RZSM accuracy, which is supported by evaluation sampled from 2474 sites at $p = 0.05$ significance level (based on a 1000-member bootstrapping analysis). However, as stated in the abstract, that we do agree the improvements in RZSM is slight ($\Delta R = 0.034$, or 7% in relative term) over 74% of soil moisture in-situ sites. Moreover, this statistically significant improvement in RZSM – albeit small – is also true for metrics of ubRMSE (2.3% decrease in relative term over 65% of in-situ sites). We have further clarified this in the revised manuscript.

Re. the <u>comment on using Spearman's rank correlation vs. Pearson correlation</u>: We provide the same response to Major comment #5 by Reviewer #1, which is reproduced here for easy reference:

Note that Pearson correlation assumes the linear consistency of underlying variables. However, this assumption may be adversely affected by outliers. To avoid ad-hoc thresholds, we do not exclude any soil moisture outliers and employ Spearman's rank correlation, which is less sensitive to such outliers. Nonetheless, we repeat the analysis based on Pearson correlation (see Figs. 1-2 below). The Pearson-based results are quantitatively consistent with the results using Spearman's correlation. We have further clarified this in the corresponding paragraph of Section 2.2 in the revised manuscript.

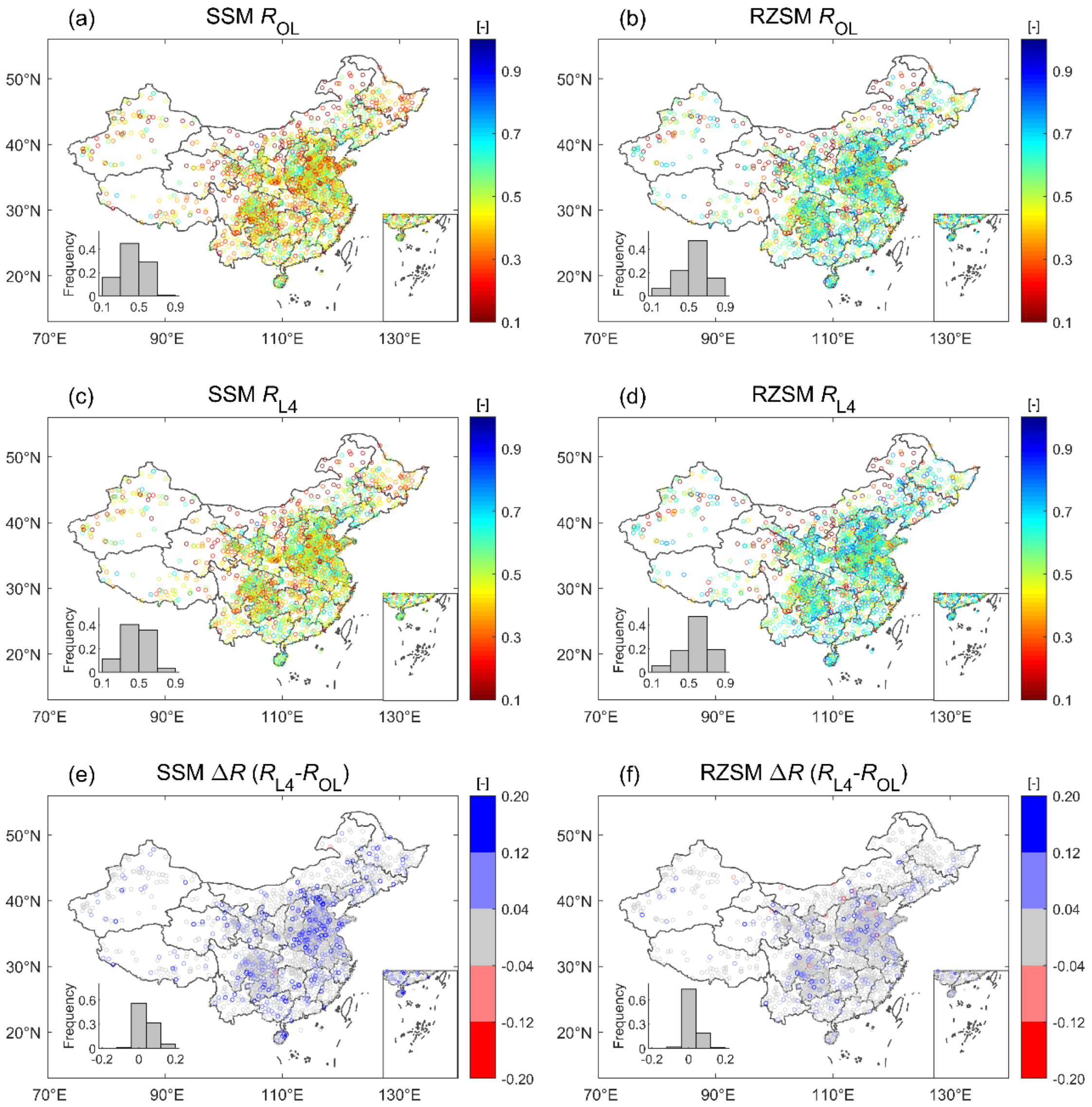


Fig. 3 Same content as in Fig. 2 of the manuscript, except that the correlation between in-situ soil moisture measurements and SMAP is measured using Pearson correlation.

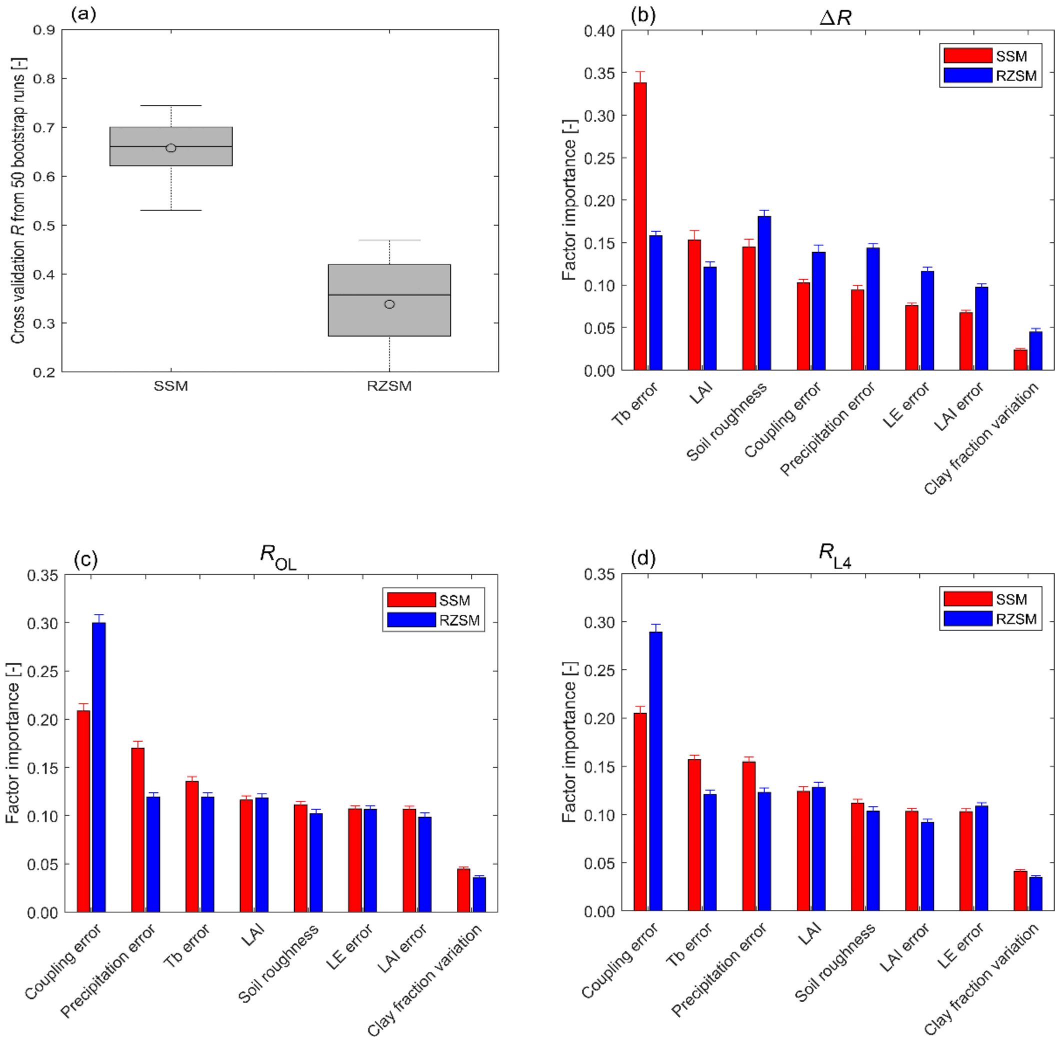


Fig. 4 Same content as in Fig. 4 of the manuscript, except that the correlation between in-situ soil moisture measurements and SMAP is measured using Pearson correlation.

Minor comments:

1. Line 1, please clarify what is the added value, is it a correlation coefficient or DA efficiency? The term "added value of ... soil moisture..." is misleading because the study is based on delta_R, not soil moisture increment.

We have changed the term “added value” to “benefit” in the revised manuscript.

2. Line 25, it should be "Spearman’s rank correlation coefficient" instead of "Spearman rank correlation skill". Skill is more sophisticated.

We have revised this expression into “Spearman’s rank correlation” in the revised manuscript.

3. Line 27, "the same percentage" is not clear.

We have revised this expression into specific number of “74%” in the revised manuscript.

4. Line 172-174, a citation is needed.

We have added the following reference in original Line 174 of the manuscript.

[1] Shen, Y. and Xiong, A.: Validation and comparison of a new gauge-based precipitation analysis over mainland China, Int. J. Climatol., 36(1), 252-265, doi:10.1002/JOC.4341, 2015.

5. Line 258, clarify what kind of anomalies it is.

We defined anomaly in Line 117-118: “…we performed our analysis using anomaly time series, derived by subtracting a seasonally-varying (daily) climatology from each raw time series.” To avoid redundancy, we did not repeat this again in original Line 258 or the following Section.

Reply to Referee #3 interactive comment

In this study, Qiu et al assess the performance of SMAP L4 DA system using 2 years of in-situ soil moisture profile observations at 2474 sites across mainland China. They then apply a random forest (RF) regression to identify the dominant factors (preselected by the authors) that control the spatial distribution of the data assimilation efficiency. This is an interesting study that could potentially lead to improvement in the SMAP L4 data assimilation system. I have annotated a pdf document with some suggestions as an attempt to help. In particular it would be interesting to try to justify more the choice of the studied dominant factors and then to discuss perspectives, what can be built upon this study? Without such proper discussion I have the feeling that the conclusion of the study is a bit weak (but you may want to prove me wrong!) with outcomes we could have guessed beforehand (e.g. precipitation is the dominant factor for explaining the skill of the OL results).

Please also note the supplement to this comment: https://hess.copernicus.org/preprints/hess-2020-407/hess-2020-407-RC3-supplement.pdf

We sincerely thank Dr. Albergel for his constructive comments.

Re. the <u>comment about the choice of the control factors</u>: We provide the same response to Major comment #5 by Reviewer #1 and Major comment #1 by Reviewer #2. For easy reference, please see below:

As mentioned in the abstract, the modeling portion of the SMAP L4 system consists of two components: land surface modelling (LSM) and radiative transfer modeling (RTM). Therefore, we select control factors from each of the two components.

For the LSM component, the errors can be attributed to potential factors including: 1) model input forcing errors of a) precipitation and b) LAI; 2) model structure errors in a) characterizing SSM-RZSM coupling strength and b) the presence of vertical variability in soil properties; 3) model output error of LE.

For the RTM component, errors are characterized by: 1) DA innovation, i.e., SMAP-observed minus RTM-simulated Tb; 2) the environmental factors that complicate the DA analysis when assimilating Tb observations, which include the magnitude of a) microwave soil roughness and b) LAI.

These eight control factors from the above-mentioned five aspects determine the crucial aspects of both the LSM and RTM components in the L4 system and are readily quantifiable using remote sensing products. Thus, they are selected to investigate the mechanism underlying the L4 improvement observed in this study. We have further clarified this in the corresponding paragraph of Section 1 in the revised manuscript.

Re. the <u>comment about discussing the possible application of this study's conclusion</u>: It should be noted that, among the RF analysis results of this study, there are conclusions that fall squarely within expectation; for instance, the OL skill is predominately determined by precipitation error, and L4 skill improvement (i.e., R_{L4} - R_{OL}) is mostly determined by Tb error. On the other hand, there are also some more surprising results. For instance, we found that SSM-RZSM coupling error and precipitation error have a comparable impact on OL. For L4 skill, however, the impact of SSM-RZSM coupling error exceeds that of precipitation error. More specifically, L4 DA contributes the most benefit for cases where CLSM underestimates SSM-RZSM vertical coupling strength. These findings could be used for L4 product development. In addition, this study pinpoints that the L4 skill improvement is not heavily impacted by LAI magnitude, which gives confidence for using the L4 product over densely vegetated areas. We have further clarified this in Section 4 of the revised manuscript.

Re. the comments in the annotated manuscript pdf file: We have made corresponding revisions in the manuscript.

Specifically, regarding the comment of adding one time series in Section 3.1, please see the time series in figure below.

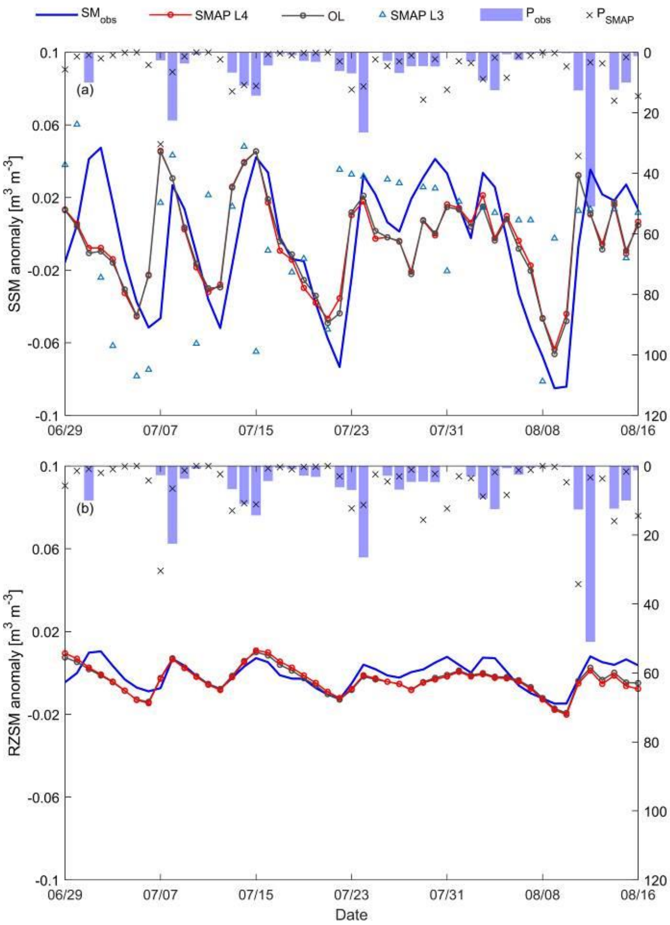


Fig. 1 The time series of gauge-based precipitation CGDPA (P_{obs}), SMAP L4 forcing precipitation (P_{SMAP}), and anomaly time series of in-situ soil moisture measurements (SM_{obs}), SMAP L4, OL, and SMAP L3 at site-collocated grid cell for: (a) SSM, (b) RZSM.

The ~~added value~~benefit of brightness temperature assimilation for the SMAP Level-4 surface and root-zone soil moisture analysis over mainland China

Jianxiu Qiu[1,2], Jianzhi Dong[3], Wade T. Crow[3], Xiaohu Zhang[4,5], Rolf H. Reichle[6], Gabrielle J. M. De Lannoy[7]

[1]Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China

[2]Southern Laboratory of Ocean Science and Engineering (Guangdong, Zhuhai), Zhuhai, 519000, China

[3]USDA ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD 20705, USA

[4]National Engineering and Technology Center for Information Agriculture, Nanjing Agricultural University, Nanjing, China

[5]Jiangsu Key Laboratory for Information Agriculture, Nanjing Agricultural University, Nanjing, China

[6]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA

[7]Department of Earth and Environmental Sciences, KU Leuven, Heverlee, Belgium

Correspondence to: Jianxiu Qiu (qiujianxiu@mail.sysu.edu.cn)

Abstract. The Soil Moisture Active Passive (SMAP) Level-4 ~~Surface Soil Moisture and Root-Zone Soil Moisture~~ (L4) product provides global estimates of surface soil moisture (SSM) and root-zone soil moisture (RZSM) via the assimilation of SMAP brightness temperature (Tb) observations into the Catchment Land Surface Model (CLSM). Here, using in-situ measurements from 2474 sites in mainland China, we evaluate the performance of soil moisture estimates from the L4 data assimilation (DA) system and from a baseline "open-loop" (OL) simulation of CLSM without Tb assimilation. Using random forest regression, the efficiency of the L4 ~~data assimilation (~~DA~~)~~ system (i.e., the performance improvement in ~~L4~~ DA relative to OL) is attributed to eight~~8~~ control factors related to the ~~land surface modelling (~~CLSM~~)~~ and as well as tau-omega radiative transfer model ~~modeling~~ (RTM) components of the L4 system. Results show that ~~77% of the 2287 9-km EASE grid cells in mainland China that contain at least one ground station exhibit an increase in~~ the Spearman rank correlation ~~skill~~ (*R*) for L4 SSM with in-situ measurements increases for 77% of the in-situ measurement locations ~~for L4 SSM compared~~(relative to that of OL), with an average *R* increase of approximately 14% (Δ*R* = 0.056). RZSM skill is improved for about ~~the same percentage~~74% of the in-situ measurement locations~~9-km EASE grid cells~~, but the average *R* increase for RZSM is only 7% (Δ*R* = 0.034). Results further show that the SSM DA ~~efficiency~~ skill improvement is most strongly related to the difference between~~error in~~ the RTM-simulated Tb ~~observation space~~and the SMAP Tb observation, followed by the error in precipitation forcing and microwave soil roughness *h*. For the RZSM DA ~~efficiency~~skill improvement, these three dominant control factors remain the same, although the importance of soil roughness exceeds that of the Tb simulation error, as the soil roughness strongly affects the ingestion of DA increments and further propagation to the subsurface. For the skill of the L4 and OL estimates themselves, the top two control factors are the precipitation error and the SSM-RZSM coupling strength error ~~(in descending order of factor importance for R_{OL})~~, both of which are related to the CLSM component of the L4 system. Finally, we find that the L4 system can effectively filter out errors in precipitation. Therefore, future development of the L4 system should focus on improving the characterization of the SSM-RZSM coupling strength.

Keywords. SMAP Level 4, soil moisture, data assimilation, attribute analysis, random forest regression

1 Introduction

Soil moisture modulates water and energy feedback~~s~~ between the land surface and the lower atmosphere by determining the partitioning of incoming net radiation into latent and sensible heat (Seneviratne et al., 2010, 2013). High-quality, global-scale soil moisture products have become increasingly available in recent years ~~(Gruber et al., 2020)~~. In particular, the L-band NASA Soil Moisture Active Passive (SMAP) satellite mission (Entekhabi et al., 2010; Piepmeier et al., 2017) has significantly improved the skill of available, global-scale soil moisture products. However, the SMAP observations contain temporal data gaps and are only representative of conditions within only the ~~top~~ first 5 cm of the vertical soil moisture column (Entekhabi et al., 2010). To address these limitations, the SMAP Level-4 Surface and Root-Zone Soil Moisture (L4) algorithm assimilates SMAP brightness temperature (Tb) observations into the NASA Catchment Land Surface Model (CLSM) to derive an analysis of surface (0–5 cm) and root-zone (0–100 cm) soil moisture estimates with global, 3-hourly coverage (Reichle et al., 2017a; Reichle et al., 2017b; Reichle et al., 2019).

However, the performance of a land data assimilation (DA) system is sensitive to ~~its~~ the DA parameterization and requires careful assessment. For instance, Reichle et al. (2008) demonstrate that DA based on incorrect assumptions of modeling errors and observation errors can degrade soil moisture estimates, compared with the case of not performing ~~any~~ DA, which is commonly referred to as the "open-loop" (OL) baseline. Theoretically, the optimality of DA can be evaluated using so-called "innovations", or observation~~s~~-minus-forecast residuals; however, an investigation of the innovations alone is often insufficient to determine if the soil moisture analysis is optimal, as the innovations are affected by multiple factors (Crow and Van Loon, 2006).

Recently, Dong et al. (2019a) proposed a novel statistical framework for evaluating the performance of a soil moisture DA system. Specifically, they demonstrated that the relative skill of surface soil moisture (SSM) estimates acquired with and without DA can be estimated using the ratio of their correlations with just one noisy but independent ancillary remote sensing product. This approach was applied to the SMAP L4 system using Advanced Scatterometer~~ASCAT~~ soil moisture retrievals. Their results show that the ~~added value~~benefit of SMAP DA is closely related to densities of both rain gauge and vegetation~~ density~~. Generally, higher rain gauge density indicates lower error in precipitation forcing, and lower vegetation density indicates higher background model performance - both conditions lead to reduced SMAP DA benefit. However, due to the limited availability of independent root-zone soil moisture (RZSM) products for performing statistical error estimation, this method is only applicable for SSM estimates.

Relative to SSM, the efficiency of assimilating land surface observations to improve RZSM is complicated by model structural error that affects the ability of the DA to update unobserved model states. For instance, Kumar et al. (2009) identified the surface–root zone coupling strength, which is the result of a model-dependent representation of processes related to the partitioning of rainfall into infiltration, runoff, and evaporation components, as an important factor for determining RZSM improvement associated with the assimilation of SSM retrievals. Their synthetic experiments suggest that~~,~~— faced with unknown true subsurface physics~~,~~— overestimating the surface–root zone coupling in the land model is a more robust strategy for obtaining skill improvements in the root zone than under-estimating the coupling. Likewise, Chen et al. (2011) suggested that ~~their~~ the Soil and Water Assessment Tool significantly under-predicts the magnitude of vertical soil water coupling in the Cobb Creek Watershed in southwestern Oklahoma, USA, and this lack of coupling impedes the ability of DA to effectively update ~~deep-layer~~ soil moisture in deep layers, groundwater flow and surface runoff. In the context of the present paper, the evaluation of L4 RZSM estimates has been limited to ~~relatively few~~ SMAP core validation and sparse network sites (Reichle et al., 2017a; Reichle et al., 2017b; Reichle et al., 2019). With such limited ~~sample sizes~~validation sites, the RZSM skill of the L4 product at the global scale remains uncertain.

The primary objective of this study is to ~~determine~~ assess the DA ~~efficiency~~skill improvement, i.e., the performance improvement in DA results relative to the ~~open-loop (~~OL~~)~~ baseline of the L4 product, and to further determine how DA skill improvement varies as a function of the major ~~a variety of system~~ aspects in the system.~~,~~ As mentioned above, the modeling portion of the L4 system consists of two components: land surface modelling (LSM) and radiative transfer modelling (RTM). Therefore, we select control factors from each of the two components. For the LSM component, the errors can be attributed to potential factors including: 1) model input forcing errors of a) precipitation and b) LAI;

2) model structure errors in a) characterizing SSM-RZSM coupling strength and b) the presence of vertical variability in soil properties; 3) model output error of LE. For the RTM component, errors are characterized by: 1) DA innovation, i.e., SMAP-observed minus RTM-simulated Tb; 2) the environmental factors that complicate the DA analysis when assimilating Tb observations, which include the magnitude of a) microwave soil roughness and b) LAI. These eight control factors from the above-mentioned five aspects determine the crucial aspects of both the LSM and RTM components in the L4 system and are readily quantifiable using remote sensing products. Thus, they are selected to investigate the mechanism underlying the L4 improvement observed in this study.~~including errors in CLSM forcing (e.g., precipitation), errors in key CLSM parameters (e.g., relating to vegetation), mean errors in CLSM structure (e.g., surface and root-zone coupling), and errors in the radiative transfer modeling (RTM) that links the modeled soil moisture and temperature estimates to the observed Tb.~~

~~To this end~~Therefore, to achieve the two major objectives, we first evaluate the performance of L4 SSM and RZSM estimates using ~~a very large number (n =~~ 2474~~) of~~ sites in mainland China with soil moisture profile measurements ~~sites~~ (generally acquired at sub-surface depths between 10 and 50 cm)~~ within mainland China,~~ during the two-year period of 2017 to 2018. Next, the in-situ measurements are used to assess the DA ~~efficiency~~skill improvement of the L4 system, which is defined as the skill difference between the L4 estimates and the OL baseline~~model-only estimates derived without SMAP Tb assimilation~~. Additionally, we apply a machine-learning technique to quantify by how much ~~various~~ the eight potential control factors drive the spatial variations in the efficiency of the L4 system. In this way, we seek to prioritize future enhancements to the L4 system.

2 Data and Methods

In t~~T~~his section, we briefly describe~~s~~ the SMAP L4 soil moisture product (Section 2.1), the ~~extensive~~ network of in-situ soil moisture observations ~~over~~ in mainland China (Section 2.2), ~~and~~ the above-mentioned control factors and ancillary data sources (Section 2.3), and the vertical coupling metric~~s~~ used in the skill assessment (Section~~s 2.3 and~~ 2.4). Next, we introduce the double instrumental variable (IVd) method employed to determine the errors in control factors that cannot be determined using ground observations (Section 2.5). Finally, we describe the random forest (RF) regression method used to identify the main factor(s) (out of the eight~~8~~ control factors from both CLSM and RTM aspects) that affect the spatial variations in SMAP L4 DA ~~efficiency~~ skill improvement and L4 performance (Section 2.6).

2.1 SMAP L4 soil moisture product

The SMAP L4 soil moisture product (version 4; Reichle et al., 2019) is generated by assimilating the SMAP L1C Radiometer half-orbit 36 km Equal-Area Scalable Earth (EASE) Grid ~~brightness temperature (~~Tb~~)~~ observations (Version 4 SPL1CTB; Chan et al., 2016) into the CLSM. The SMAP Tb observations are assimilated at 3-h intervals using a spatially distributed, 24-member ensemble Kalman filter (Reichle et al. 2017b). The surface meteorological forcing data are from the global Goddard Earth Observing System (GEOS) Forward Processing atmospheric analysis (Lucchesi, 2013), with precipitation corrected using the daily, 0.5-degree, gauge-based Climate Prediction Center Unified (CPCU) product (Xie et al. 2007). The L4 product provides global, 9-km, 3-hourly surface (0–5 cm) and root-zone (0–100 cm) soil moisture estimates along with related land surface fields and analysis diagnostics. For the present study, we aggregate~~d~~ all soil moisture estimates to daily ~~-~~averaged (00:00 to 23:59 UTC) data. The OL~~A~~ baseline is a ~~,~~model-only, ensemble CLSM simulation without the assimilation of SMAP Tb observations ~~(~~but otherwise using the same configuration, including perturbations, as in the L4 system (Reichle et al., 2020)~~ is referred to as the "open-loop" (OL) run~~.

The SMAP L4 assimilation system includes a zero-order "tau-omega" forward RTM (De Lannoy et al., 2013) that converts SSM and surface soil temperature into L-band brightness temperature estimates. Select~~ed~~ parameters of the L4 RTM, including the: microwave soil roughness parameter~~s~~ *h*, ~~a~~ vegetation structure parameter τ, and the microwave scattering albedo ω, ~~were~~ are calibrated using multi-angular L-band brightness temperature observations from the Soil Moisture Ocean Salinity (SMOS) mission (De Lannoy et al., 2014). The L4 RTM parameterizes microwave soil roughness as a function of SSM (De Lannoy et al., 2013, their equation B1). Here, we use~~d~~ this parameterization to compute the 2017-2018 daily averaged~~time-averaged~~ microwave soil roughness estimates as one potential indicator of DA ~~efficiency~~ skill improvement (Section 2.3). The necessary parameters ~~were~~ are obtained from L4 "Land-Model-Constants" output Collection (last access: 8 July 2020; DOI: https://doi.org/10.5067/KGLC3UH4TMAQ; Reichle et al., 2018a). The L4 "Analysis-Update-Data" output Collection includes RTM predictions of Tb and the assimilated SMAP Tb observations (last access: 8 July 2020; DOI: https://doi.org/10.5067/60HB8VIP2T8W; Reichle et al., 2018b). To avoid the impact of seasonality, we perform~~ed~~ our analysis using anomaly time series, derived by subtracting a seasonally ~~-~~varying (daily) climatology from each raw time series. The climatology of a given time series ~~was~~ is obtained by sampling the mean value of all soil moisture estimates that fall within a 31-day moving window centered on a particular day-of-year. Moreover, L4 estimates of land latent heat flux (LE), land sensible heat flux (SH) and the climatological LAI inputs to CLSM and the RTM, ~~were~~ are obtained from the L4 "Geophysical-Data" output Collection (last access: 6 April 2020; DOI: https://doi.org/10.5067/KPJNN2GI1DQR; Reichle et al., 2018c). These datasets ~~were~~ are also used to compute control factors to explain spatial variations in the DA ~~efficiency~~skill improvement of the L4 system (Section 2.3).

2.2 Soil moisture validation data

In-situ soil moisture measurements during 2017 and 2018 ~~were~~ are collected from a national network of Chinese Automatic Soil Moisture Observation Stations (CASMOS) maintained by the Chinese Meteorological Administration (CMA; Han et al., 2017). In total, soil moisture measurements from 2474 separate stations array~~ed~~ across mainland China, and covering different land use types, ~~were~~ are collected. At each CASMOS site, frequency domain reflectometry-based instruments ~~were~~ (DNZ1, DNZ2, and DNZ3) are used to record hourly volumetric soil moisture content within the following vertical depth ranges: 0–10, 10–20, 20–30, 30–40, and 40–50 cm below the surface. These hourly estimates (at multiple depths) ~~were~~ are then aggregated into daily values and linearly averaged (vertically) to produce 0-10 cm (SSM) and 0-50 cm (RZSM) in situ soil moisture measurements – which ~~were~~ are subsequently used to validate the L4 and OL SSM (0-5 cm) and RZSM (0-100 cm) estimates. Note that Spearman correlation rather than Pearson correlation is used for L4 and OL validation because~~, in order to avoid impact of outliers in the time series and prior assumptions about soil moisture distributions.~~ Pearson correlation assumes linear consistency of the underlying variables and is more sensitive to outliers. By employing Spearman's rank correlation, we avoid introducing ad-hoc thresholds and do not exclude soil moisture outliers. Nonetheless, we repeat the analysis based on Pearson correlation (not shown) and find that the results are qualitatively consistent with the results using Spearman's correlation.

Ground observations ~~falling~~ within the same 9-km EASE grid were averaged for comparisons against the collocated 9-km L4 and OL soil moisture estimates. A total of 2287 individual 9-km EASE grid cells within mainland China are included in the analysis. Among them, 92.35% of grid cells contain one in-situ site, 7.26% contain two sites, 7 grid cells contain three sites, and the remaining two grid cells contain four and five sites respectively. Figure 1 shows the number of in-situ CASMOS sites within each 9-km EASE grid.

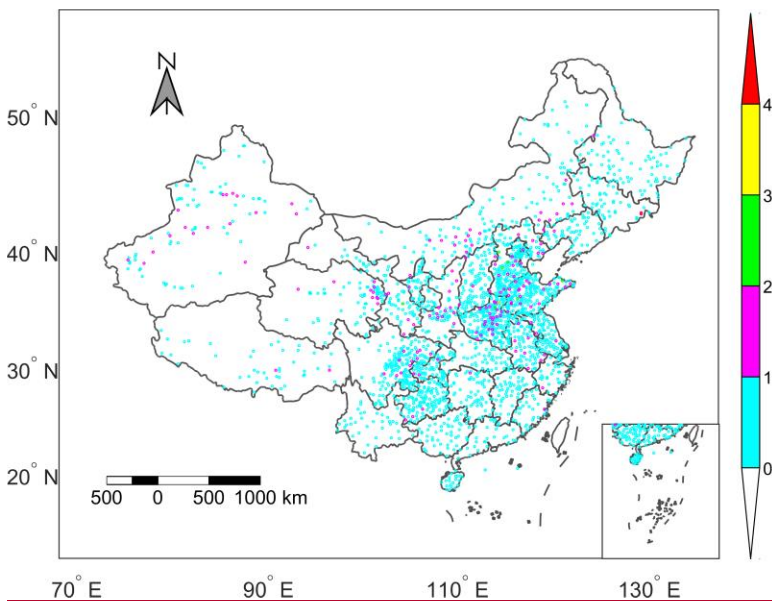


Figure 1: The number of in-situ CASMOS sites within each 9-km EASE grid across mainland China.

2.3 Explanatory data products

As discussed above, our hypothesis is that the efficiency of the SMAP L4 system will be sensitive to the ability of the ensemble-based L4 analysis in filtering errors that exist in ~~the OL (that is,~~ CLSM~~)~~, ~~in~~ the ~~model~~ RTM forecast Tb ~~(that is, the RTM)~~, and ~~in~~ the assimilated SMAP Tb observations. We therefore consider~~ed~~ two separate categories of factors that potentially control spatial variations in DA ~~efficiency~~skill improvement. The factors are summarized in Table 1.

The first category represents a range of factors known to affect the skill of soil moisture estimates derived from the LSM (in this case, CLSM). The five control factors in this category are: ~~i~~1) the error in precipitation forcing, ~~ii~~2) the error in (input) LAI, ~~iii~~3) the error in (output) LE, ~~iv~~4) the magnitude of mean error in CLSM SSM-RZSM coupling strength, and ~~v~~5) the presence of vertical variability in soil properties (defined as the difference in clay fraction across the vertical soil profile). Note that such variability represents a potential source of error because, with the exception of some surface-layer moisture transport parameters, CLSM assumes ~~that~~ soil texture and ~~the~~ associated soil parameters are vertically homogeneous within the soil column~~, with the exception of some surface-layer moisture transport parameters~~. However, the Harmonized World Soil Database (HWSD) often captures distinct vertical variations in soil properties. Therefore, since it is largely neglected by CLSM, the magnitude of vertical heterogeneity in soil texture may be an effective proxy for overall CLSM soil moisture accuracy~~The soil texture information is from Harmonized World Soil Database (HWSD) v1.2~~. In addition, note that since LH and SH are generally (strongly) anti-correlated, it is not appropriate to include both in a single random forest analysis – since including both would yield biased (high) regression weights for LH and SH.

The second category contains three factors that affect radiative transfer modeling (RTM) and therefore DA updates. These include: ~~i~~1) estimates of the DA innovation, namely difference between~~ joint error in~~ SMAP Tb observations and RTM Tb simulations, ~~ii~~2) the magnitude of microwave soil roughness, and ~~iii~~3) the magnitude of LAI (as a proxy for the vegetation optical depth at microwave frequencies, which modulates the contribution of surface soil to ~~sensitivity of~~ the observed Tb~~ to SSM conditions~~).

The control factors take a variety of forms. Some factors are based on estimates of the errors fed into the L4 system~~ as (e.g.~~, namely: 1) the error in CLSM rainfall forcing data; 2) error in SSM-RZSM coupling strength; 3) vertical variability of clay fraction; 4) SMAP L4 LAI error; 5) output LE error; 6) Tb error~~)~~. Other factors consist of the magnitude of the variable itself, ~~(~~namely the magnitude of microwave soil roughness and annual mean LAI~~e.g., the vertical variability of clay fraction)~~. Note that LAI is used in both ways: LAI error is used to predict OL performance (because LAI is an important input into CLSM), while mean LAI is used to explain DA performance (because increased LAI is associated with decreased soil moisture information ~~content~~ in microwave observations).

Note that the LAI used in the L4 system is a merged climatology from Moderate Resolution Imaging Spectroradiometer (MODIS) and Geoland data based on ~~derived from~~ satellite observations of the Normalized Difference Vegetation Index (Mahanama et al., 2015; Reichle et al., 2017a). Therefore, to indicate the magnitude by which ~~each grid cell's~~the LAI of each grid cell typically deviates from its long-term climatology, we use the temporal standard deviation for the~~of~~ anomaly time series of ~~the~~ a benchmark LAI ~~(from SPOT VGT product)~~time series as a measure of the error in the LAI value used in the L4 system. This benchmark LAI is from the SPOT-Vegetation (SPOT VGT) product and includes inter-annual variations (Section 2.3.3). Owing to the lack of reference Tb observations at similar satellite overpass times and locations, Tb errors are gauged using the time series standard deviation of the observation-minus-forecast (O-F) Tb residuals, which indicate the typical misfit between the model forecast Tb and the ~~(~~rescaled~~)~~ SMAP Tb observations. This rescaling process ensures zero-mean differences between Tb observations and forecasts and involves a seasonal multiyear-mean bias correction, which makes sure that the DA only corrects for errors in short-term and inter-annual variations and not for errors in the climatological seasonal cycles of the modeled soil moisture or other land surface fields. Th~~is metric~~e standard deviation of the O-F Tb residuals measures the total error in Tb observation space.

The exact data sets and the metrics utilized for evaluating all~~these~~ eight~~8~~ control factors are summarized in Table 1.

Table 1 Benchmark data sets and metrics used for evaluating control factors of SMAP L4

Factor category	Control factor	Dataset/Benchmark	Temporal resolution	Spatial resolution	Data range	Metrics
LSM	Precipitation error	Rain gauge (CGDPA)	daily	0.25 °	2017-2018	Spearman's rank correlation *R*
	SSM-RZSM coupling strength error	CASMOS	daily	NA	2017-2018	ΔCP (see Section 2.4)
	Vertical variability of clay fraction	HWSD	NA	9 km	NA	Difference in clay fraction between topsoil (0-30 cm) and root-zone (0-100 cm) layers
	SMAP L4 LAI error	SPOT-VGT LAI	10 d	1 km	2017-2018	Temporal standard deviation of SPOT VGT LAI anomaly
	LE error	FLUXCOM	daily	(1/120) °	2017-2018	IVd-based *R*
RTM	Tb error	SMAP L4	daily	9 km	2017-2018	Temporal standard deviation of O-F Tb residuals
	Microwave soil roughness	SMAP L4	daily	9 km	2017-2018	Temporal average based on De Lannoy et al. (2013)
	Annual mean LAI	MODIS/Geoland-based product	daily	9 km	2017-2018	Climatological mean

2.3.1 Gauge-based precipitation gridded product

Errors in the ~~GEOS~~ precipitation data used to force the CLSM within the SMAP L4 system ~~were~~ are estimated via Spearman's rank correlation with available rain-gauge observations. These network observations are based on an analysis of ~2400 rain gauge stations distributed ~~unevenly over~~across mainland China (Shen et al., 2015). Recently, the China Gauge-based Daily Precipitation Analysis (CGDPA) with a spatial resolution of 0.25°×0.25° based on this network was constructed and has been made operational over mainland China (last access: 28 April 2020; http://data.cma.cn/data/cdcdetail/dataCode/SEVP_CLI_CHN_PRE_DAY_GRID_0.25.html). CGDPA uses a modified ~~interpolation method of~~version of climatology-based optimal interpolation (OI) with topographic correction proposed by Xie et al. (2007). In this process, the daily precipitation climatology over mainland China is optimized and ~~is~~ rebuilt using the 30-year average precipitation observations from ~2400 gauges of the period 1971–2000 (Shen et al., 2010). CGDPA is shown to have smaller bias and root mean square error (for instance, 13.51 mm day^{-1} vs. 17.02 mm day^{-1} for precipitation of 25.0–50.0 mm day^{-1}) than the CPCU product used in the SMAP L4 system, which is based on fewer than 400 gauge sites over mainland China (Shen et al., 2015).

2.3.2 FLUXCOM LE estimates

The FLUXCOM ensemble of global land-atmosphere energy fluxes ~~was~~ is used to evaluate ~~the error of the~~error in L4 LE estimates. This ensemble merges energy flux measurements from FLUXNET eddy covariance towers with remote sensing and meteorological data based on ~~a~~ four broad categories of machine learning method (namely tree-based methods, regression splines, neural networks, and kernel methods) to estimate global gridded net radiation, latent and sensible heat and their related uncertainties (Jung et al., 2019). The resulting FLUXCOM database has a 0.0833° spatial resolution when applied using MODIS remote sensing data. The monthly energy flux data of all ensemble members, as well as the ensemble estimates from the FLUXCOM initiative, are freely available (CC4.0 BY license) from the Data Portal (http://fluxcom.org/), while the daily- and 8-day FLUXCOM products are available upon request from dataset provider Martin Jung (last access: 14 April 2020). To calculate the LE error, we~~’ve~~ collected the daily, high spatial resolution FLUXCOM product and extracted the LE estimates where in-situ soil moisture sites are located.

2.3.3 SPOT VGT LAI

The data set used as a benchmark for assessing leaf area index (LAI) errors present in the SMAP L4 analysis ~~was~~ is derived from the SPOT/VEGETATION and PROBA-V LAI products (version 2) that ~~are~~ generated every 10 days (at best) with a~~at~~ spatial resolution of 1 km. The SPOT LAI version 2 product GEOV2 is provided by the Copernicus Global Land Service (last access: 15 April 2020; https://land.copernicus.eu/global/products/LAI; Baret et al., 2013). It capitalizes on the development ~~and validation~~ of already existing products: CYCLOPES version 3.1 and MODIS collection 5 ~~and the use of~~based on neural networks (Baret et al., 2013; Verger et al., 2008). ~~The version 2 products are derived from top of canopy daily (S1-TOC) reflectances instead of normalized top of canopy 30-day composited reflectances as in the version 1.~~ Compared to version 1, the version 2 products are derived from top of canopy daily reflectances, ~~the compositing step is performed at the biophysical variable level instead of reflectance level. This~~which ensures reduced sensitivity to missing observations and avoids the need for a bidirectional reflectance distribution function~~BRDF~~ model.

2.3.4 HWSD soil texture

The soil texture information is from the ~~The~~ HWSD attribute database (v1.2; FAO/IIASA/ISRIC/ISSCAS/JRC, 2012), which is a 30 arc-second raster database with 15773 different soil-mapping units worldwide. It provides information on the standardized soil parameters for topsoil (0–30cm) and subsoil (30-100 cm) separately. In this study, we use the difference of clay fractions between topsoil (0-30cm) and the aggregated 0-100cm layer to measure the vertical clay fraction variation at each 9-km grid cell.

2.4 Vertical coupling metric

The RZSM time series generally show decreased temporal dynamics relative to SSM. As a result, overestimated SSM-RZSM coupling tends to spuriously increase the (correlation-based) similarity of SSM and RZSM time series, and thereby, overestimate RZSM temporal variability. Therefore, analogous to Kling-Gupta efficiency (Gupta et al., 2009), we define~~d~~ the SSM-RZSM coupling strength (CP) as:

$$\mathrm{CP} = 1-\sqrt{(R-1)^2 + (\alpha-1)^2} \tag{1}$$

where R is the Spearman's rank correlation between SSM and RZSM, and α is the ratio of temporal standard deviation of SSM to that of RZSM. The CP estimation is based on anomaly time series of both SSM and RZSM. A CP value of one represents the extreme case where RZSM is identical to SSM, i.e., a strongly coupled case. Likewise, a CP of zero represents the opposing case of completely uncoupled time series. Cases with negative CP do not exist in this study. Observed CP ($\mathrm{CP_{obs}}$) was based on comparisons between 0-10 cm "surface" ~~estimates~~ and 0-50 cm "root-zone" ~~in~~ in-situ observations and used as a benchmark. In contrast, ~~SMAP L4~~ CP estimates of OL ($\mathrm{CP_{OL}}$) was based on the comparison of 0-5 cm "surface" ~~estimates~~ and 0-100 cm "root-zone" estimates. Therefore, the surface versus root-zone storage contrast in the observation time series is less than that of the L4 estimates. This will likely cause the observed correlation between surface and root-zone time series to be systematically higher than the analogous vertical correlation calculation for L4 estimates. However, this bias is partially corrected for by the second term in Eq. (1) – since the observed α ratio will, by the same token, tend to be smaller (i.e. closer to one) than α sampled from the L4 analysis. Such ability to compensate for vertical depth differences is a key reason we apply CP, rather than *simple correlation*, as a vertical coupling strength metric. Nevertheless, it should be noted that our main interest here lies in describing spatial variations in ($\mathrm{CP_{OL}}$ - $\mathrm{CP_{obs}}$) and care should be taken when interpreting raw ($\mathrm{CP_{OL}}$ - $\mathrm{CP_{obs}}$) differences as an *absolute* measure of L4 vertical coupling bias.

2.5 Double instrumental variable (IVd) method

The benchmark data set of FLUXCOM LE described above contains error that is ~~(likely)~~assumed to be of a similar order of magnitude as the L4 LE dataset it is applied to evaluate. Therefore, in an attempt to correct for the impact of this error, the LE error used here as a control factor is obtained via a double instrumental variable (IVd; Dong et al., 2019b) analysis approach that minimizes the spurious impact of random errors in benchmark data sets. As shown in Dong et al. (2019b), for the evaluation of two time series containing~~with~~ auto~~-~~correlated errors~~ion in both of them~~, IVd is more robust than a single instrumental variable based algorithm, therefore we apply IVd to evaluate the LE error. IVd is a modified version of triple collocation (TC) analysis. In TC analysis (McColl et al., 2014), geophysical variables obtained from three independent sources (x_t, y_t and z_t) at time t are assumed to be linearly related to the true signal P_t as:

$$x_t = \alpha_x P_t + B_x + \varepsilon_{x,t} \tag{2}$$

where the α_x is a scaling factor; B_x is a temporal constant bias and $\varepsilon_{x,t}$ is zero-mean random error. As opposed to the TC method, IVd uses only two independent products (x, y) to characterize geophysical data product errors. This method introduces two instrumental variables ~~(~~I, which is the lag-1 time series of x, and J~~, i.e., $I_t = \alpha_x P_{t-1} + B_x + \varepsilon_{xt-1}$, $J_t = \alpha_y P_{t-1} + B_y + \varepsilon_{yt-1}$)~~, which is ~~are based on~~ the lag-1 ~~(day)~~ time series ~~(at day t)~~ of ~~x and~~ y, respectively.

$$I_t = \alpha_x P_{t-1} + B_x + \varepsilon_{x,t-1} \tag{3}$$

$$J_t = \alpha_y P_{t-1} + B_y + \varepsilon_{y,t-1} \tag{4}$$

Therefore, assuming that the errors of two independent products are serially white, the covariance between instrumental variables and products can be written as follows:

$$C_{Ix} = \alpha_x^2 L_{PP}$$ (~~3~~5)

$$C_{Jy} = \alpha_y^2 L_{PP}$$ (~~4~~6)

where C represents the covariance of the subscript products. For instance, C_{Ix} represents the covariance of x and its instrumental variable I. Variable L_{PP} is the lag-1 auto-covariance of the true signal. Combining Eqs. (~~3~~5) and (~~4~~6), the scaling ratio s_{ivd} of the two products x and y can be written as:

$$s_{ivd} = \sqrt{\frac{C_{Ix}}{C_{Jy}}}$$ (~~5~~7)

Based on Eq. (~~5~~7), their correlation with truth can be estimated as:

$$R_{Px}^2 = \frac{C_{xy} s_{ivd}}{C_{xx}}$$ (~~6~~8)

$$R_{Py}^2 = \frac{C_{xy}}{C_{yy} s_{ivd}} \qquad (\cancel{7}9)$$

In this way, the error in the L4 LE (measured by IVd-based correlation with truth) can be estimated robustly using the FLUXCOM LE product described in Section 2.3.2.

2.6 Random forest regression

A random forest (RF) regression approach ~~was~~ is used to rank and quantify the importance of the eight~~8~~ control factors introduced above (Table 1) for describing spatial patterns in DA ~~efficiency~~skill improvement for both SSM and RZSM estimates. The RF method is a supervised learning algorithm based on an averaged ensemble of decision trees (Breiman, 2001). Unlike linear regression approaches, RF can capture non-linear interactions between the features and the target. In addition, the normalization (or scaling) of data is not necessary in RF application. Another advantage of the RF algorithm is that it can readily measure the relative importance of each feature on the estimates, which makes it highly suitable for an attribution analysis. Therefore, based on the output of RF, key control factors determining the ~~efficiency~~skill improvement of SMAP DA ~~were~~ are evaluated and ranked. The RF estimates are based on a 10-fold cross-validation approach.

3 Results

3.1 Validation of SMAP L4 and OL estimates of SSM and RZSM anomalies

Figure ~~1~~ 2 maps validation results (i.e., anomaly Spearman's rank correlation with in-situ observations, R) for SMAP L4 and associated OL soil moisture estimates. The skill patterns for OL and L4 are, in general, quite spatially consistent. Both are characterized by an increasing trend of SSM estimation skill moving from northwest to southeast China (Fig. ~~1a~~ 2a and ~~1b~~2b) that matches the increasing density of the rain gauge network. In relative terms, the L4 product surpasses the baseline OL's SSM skill ~~within~~ for 77% of the 2287 9-km EASE grid cells containing ground observations – with a mean R increase of $\Delta R = 0.056$ [-] and mean relative improvement versus R_{OL} of 14%.

Similar spatial patterns are observed for RZSM skill. As with SSM, generally higher consistency with in-situ RZSM measurements is found in southeast China relative to northern and northwestern China. However, relative to SSM, the ~~added value~~benefit of SMAP data assimilation (i.e., L4) is reduced for RZSM and the mean relative R improvement ~~falls to~~is only 7% ($\Delta R = 0.034$ [-]) (compare Fig. ~~1e~~ 2e and ~~1f~~2f). This reduction is ~~not surprising~~expected since assimilated SMAP Tbs are primarily sensitive to soil moisture conditions in the surface (0-5 cm) layer.

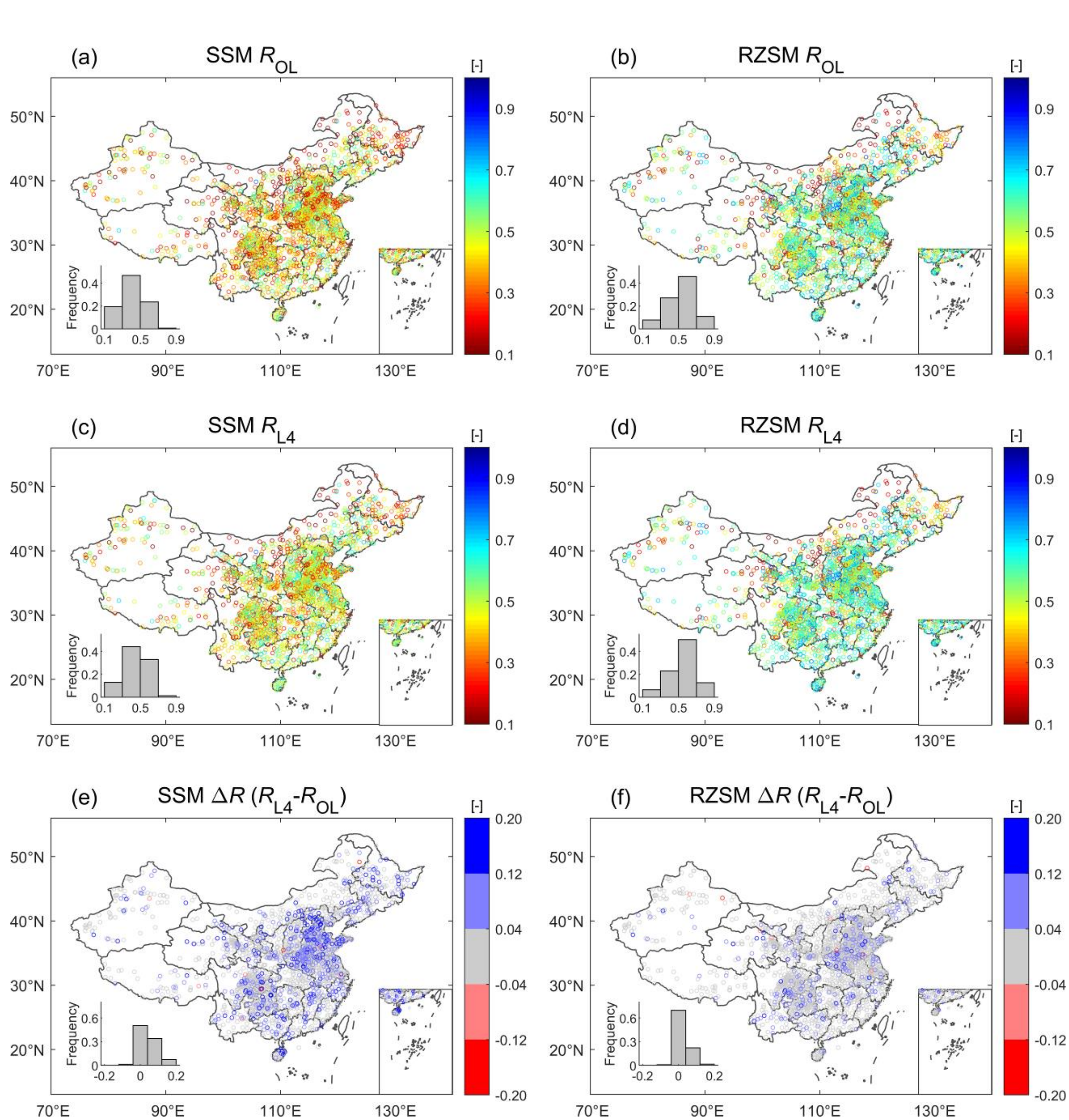
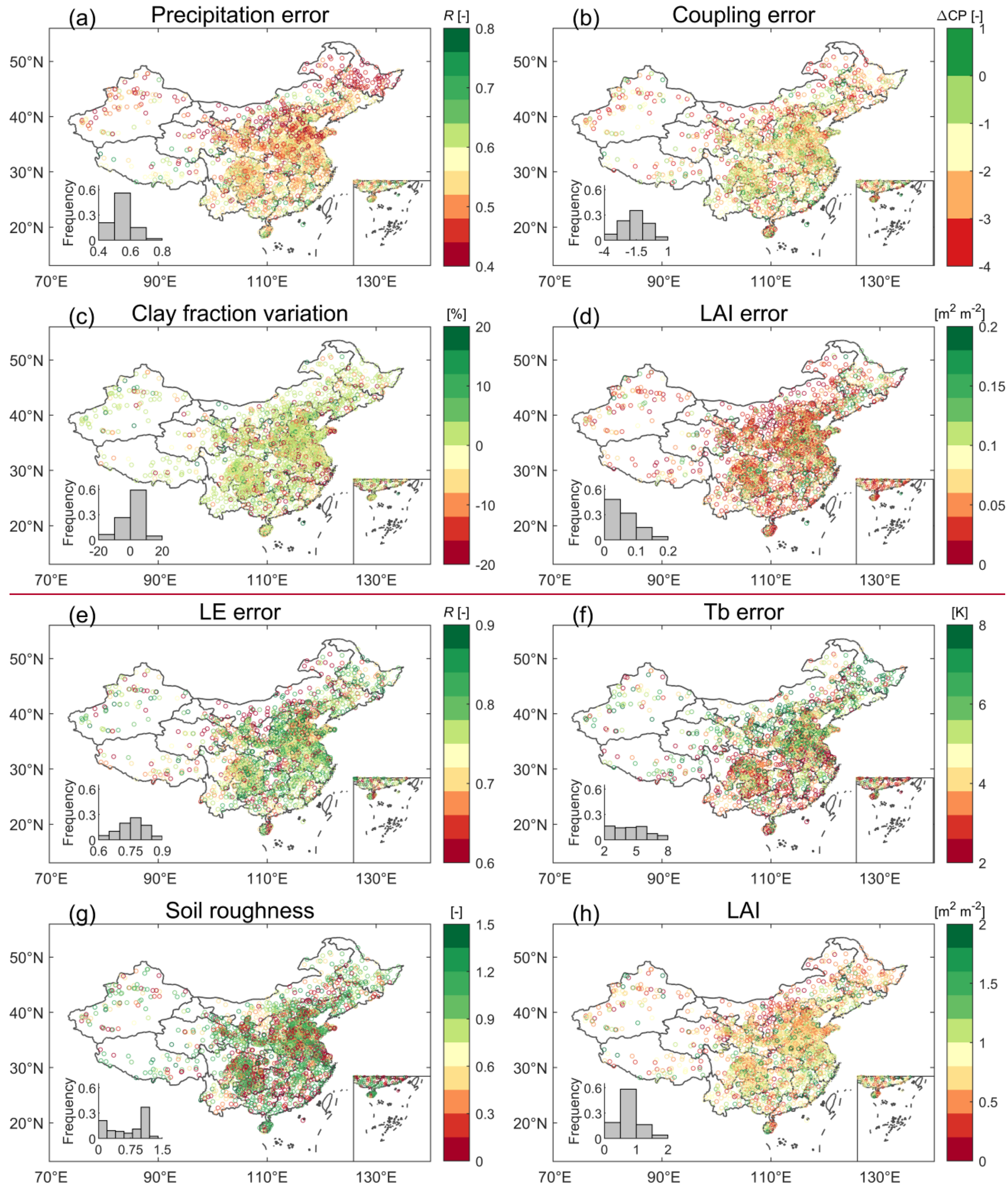


Figure ~~1~~2: OL (a, b) and L4 (c, d) skills (*R* values) for SSM (left column) and RZSM (right column). DA ~~efficiency~~skill improvement ($\Delta R = R_{L4} - R_{OL}$) for (e) SSM and (f) RZSM. Blue (red) colors in (e) and (f) indicate grid cells where L4 estimates are better (worse) than OL. Non-significant differences (based on a 1000-member bootstrapping analysis) are shaded~~colored~~ grey. The lower left inset in each subplot indicates the frequency of binned *R*-values across all 9-km EASE grid cells containing ground observations.

3.2 Spatial distribution of potential factors controlling SMAP L4 DA performance

As described in Section 2.3, we select~~ed~~ eight~~8~~ control factors that potentially influence the skill of SMAP L4 soil moisture estimates. Using the attribution analysis described in Section 2.6, these factors are~~will be~~ used to explain the spatial variations in skill and DA ~~efficiency~~ skill improvement seen in Fig. ~~1~~2. As a first step, this section examines the spatial patterns inherent in the eight~~8~~ control factors. Errors in the CLSM precipitation forcing are relatively higher in northern and northwestern areas of China (Fig. ~~2a~~3a), where the gauge density is generally ~~more sparse~~sparser than in southern China. Among the factors representing CLSM structural errors, a pre-dominantly negative bias is observed in SSM-RZSM coupling strength generally across China (i.e., lower CP_{OL} compared to CP_{obs}), while a very small number of grid cells show a positive coupling strength bias in eastern China (dark green dots in Fig. ~~2b~~3b). This is expected since the coupling strength generally decreases with ~~at the~~ coarser resolution, i.e., the ~~model's~~ vertical coupling strength of model is assumed ~~should be~~ much ~~less than at~~lower than that of any single ~~point~~site. In addition, this may be partially~~partly~~ attributed to ~~the~~ layer depth~~s~~ differences, since CLSM represents surface and root-zone depths of 0-5 cm and 0-100 cm, respectively, whereas the corresponding in-situ observations represent the 0-10 cm and 0-50 cm layers. Therefore, ~~and it can be expected that~~ CP_{OL} ~~should~~ is likely to be systematically~~thus be~~ smaller than CP_{obs}. In addition, the vertical variability of the clay fraction seems to show little spatial variation across mainland China (Fig. ~~2c~~3c). With respect to CLSM LAI error, regions in southern China that have generally higher LAI show larger standard deviations in SPOT LAI time series (Fig. ~~2d~~ 3d and ~~2h~~3h). The IVd-based estimates of SMAP L4 LE error, which represent a potential control factor for water-balance errors in CLSM, generally show a low ~~-~~level of error across mainland China (Fig. ~~2e~~3e).

For O-F Tb residuals describing RTM-related error, a higher standard deviation of O-F Tb residuals is observed in the North China Plain (Fig. ~~2f~~3f), which is very consistent in spatial distribution with areas displaying the highest and most significant SSM prediction improvement (Fig. ~~1c~~2c). This is expected, as mentioned above, because O-F Tb residuals are the basis for the soil moisture corrections (or increments) that are applied in the DA system as part of the L4 analysis. The 2017-2018 mean of soil roughness shows a relatively scattered spatial pattern (Fig. 3g), while ~~and~~ the 2017-2018 mean LAI shows higher values in southwest and southeast China (Fig. ~~2~~3~~g~~h).


(a) Precipitation error
R [-]
(b) Coupling error
ΔCP [-]
(c) Clay fraction variation
[%]
(d) LAI error
[m2 m-2]
(e) LE error
R [-]
(f) Tb error
[K]
(g) Soil roughness
[-]
(h) LAI
[m2 m-2]
Frequency
50°N
40°N
30°N
20°N
70°E
90°E
110°E
130°E

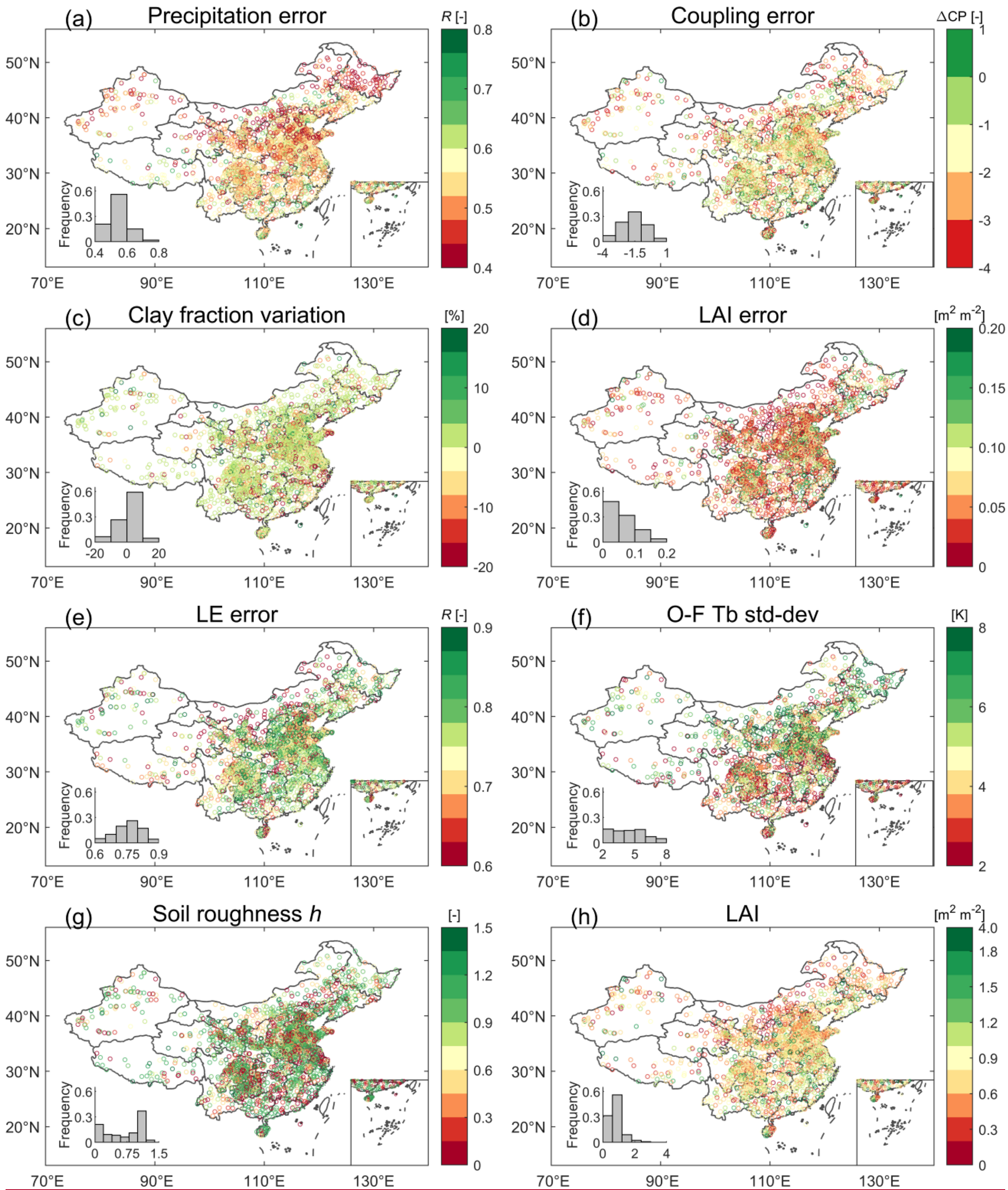


Figure ~~2~~3: Factors potentially influencing SMAP L4 performance over mainland China: (a) CLSM precipitation error measured by the Spearman's rank correlation between CLSM precipitation and ground observations; (b) SSM-RZSM coupling strength error (CP_{OL} minus CP_{obs}); (c) clay fraction variation (difference) across the soil profile; (d) error in LAI input to L4; (e) IVd-based error of LE from L4; (f) ~~Tb error~~ O-F Tb standard deviation; (g) L4 microwave soil roughness; (h) climatology mean of LAI input to L4. The last row shows factors that consist of the magnitude of the variable itself, while the other rows show factors based on estimates of the errors that are fed into the L4 system.

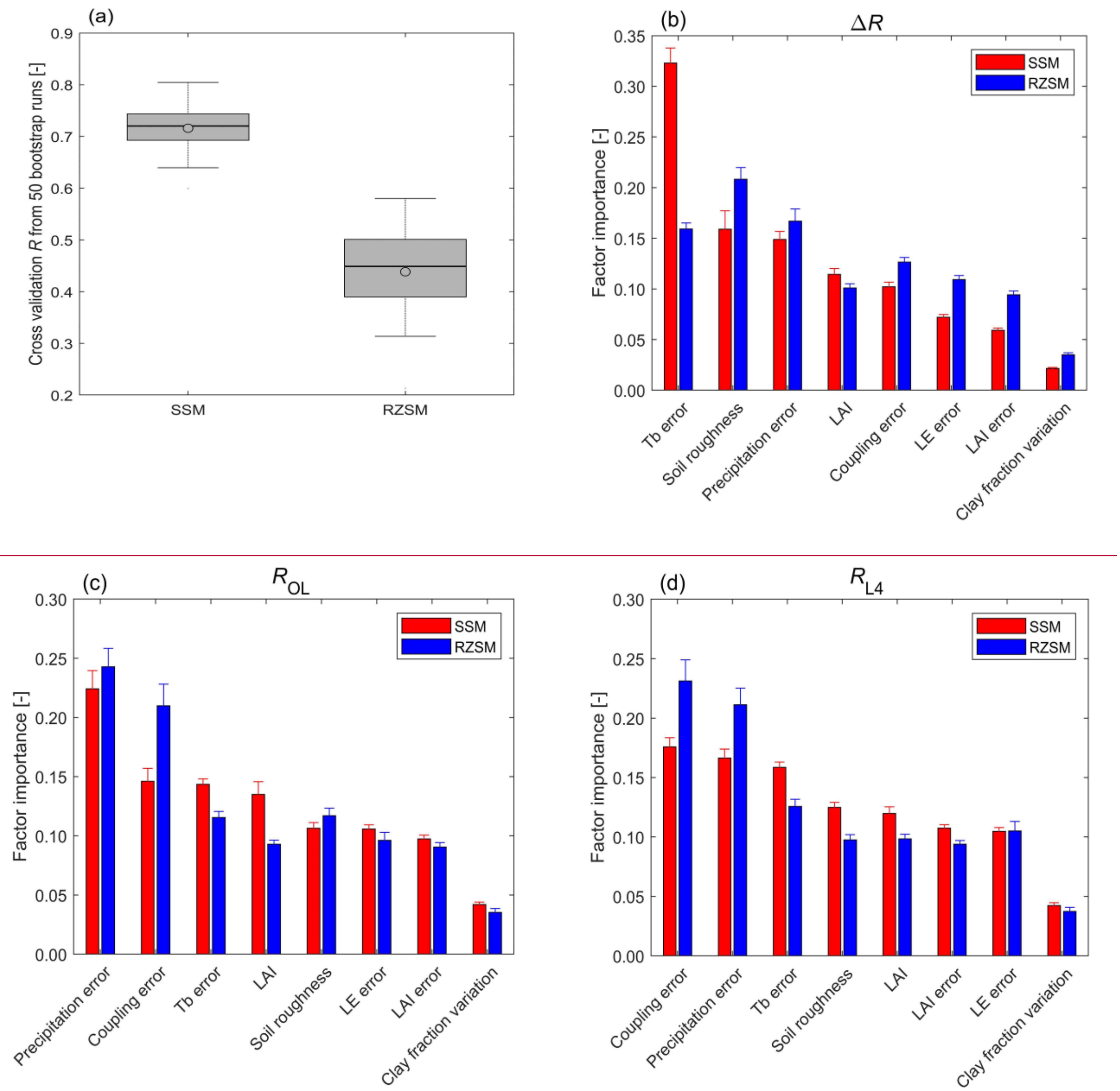
3.3 Attribution of SMAP L4 versus OL performance to control factors

3.3.1 Attribution using random forest regression

As mentioned above, RF regression ~~was~~is used to identify the relative importance of our eight~~8~~ control factors for determining the ~~efficiency~~ improvement of SMAP L4 DA (i.e., ΔR = R_{L4} - R_{OL})~~,~~ and also ~~L4 (~~R_{L4}~~)~~ and ~~OL performances (~~R_{OL}~~)~~. ~~To start, we~~We first investigate the robustness of RF for predicting ΔR. To estimate the magnitude of randomness in the RF algorithm, we use 50 bootstrap runs. As shown in Fig. ~~3a~~4a, the 10-fold cross-validation test (228 validation samples) shows that the predicted and in-situ-based ΔR have a mean correlation of 0.72 and 0.46 for SSM and RZSM, respectively. In Fig. 4a, the mean and median of the cross-validation correlation are shown in black circle and black line respectively within the boxes, while the second and third quartiles of the cross-validation correlation are shown as the edges of boxes.

Given the sampling errors of ΔR, which is based on a two-year validation period, and the relatively low spatial variability in RZSM skill (Figs. ~~1f~~2f), the performance of RF is acceptable. In addition, ground-measurement upscaling error is likely a significant contributor to unexplainable spatial variability for ΔR in Fig. ~~1~~2. In fact, Chen et al. (2016) found large spatial variability in the ability of point-scale SSM ground observations to describe grid cell-scale SSM dynamics. In-situ observations sites associated with larger random point-to-grid upscaling errors will introduce a spurious low bias into sampled estimates of ΔR values (see Appendix B in Dong et al., 2020). Therefore, ~~some~~ part of the ΔR spatial variability observed in Fig. ~~1~~ 2 is unrelated to any aspect of the L4 system and, therefore, ~~is therefore~~ unexplainable via ~~the~~ our eight selected~~8~~ control factors~~ we have selected~~.

Independent representativeness errors have an equal impact on both the L4 and OL skill assessments and should therefore not bias the relative skill assessments of L4 versus OL, particularly when these assessments are based on averaging across multiple grid cells. This holds if the location of ground-based measurements sites (within a footprint) is purely random. For the systematic sampling errors, we analyze the site "representativeness" using the 500m MODIS Land Cover product (MCD12Q1 v6) in 2017, IGBP dataset. First, we take the land cover (LC) type of the MODIS grid cell where a given in-situ site is located as the ground-based LC type. Next, we search all the MODIS grid cells that fall within the SMAP 9km EASE grid cell where this in-situ site is located. The latter area consists of about 20 x 20 = 400 MODIS grid cells. We calculate the fraction of these 400 MODIS grid cells that have the same LC type as the ground-based LC and define this fraction as the site representativeness. We find that 52% of the 2474 sites have site representativeness higher than 50%. When we use only these sites for the RF attribute analysis, the top three factors controlling skill improvement (R_{L4} – R_{OL}), L4 skill (R_{L4}), and OL skill (R_{OL}) are still the same, although the precipitation error becomes the top influencer for R_{L4} (not shown).


(a)
Cross validation R from 50 bootstrap runs [-]
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
SSM
RZSM
(b)
ΔR
Factor importance [-]
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
SSM
RZSM
Tb error
Soil roughness
Precipitation error
LAI
Coupling error
LE error
LAI error
Clay fraction variation
(c)
R OL
Factor importance [-]
0.30
0.25
0.20
0.15
0.10
0.05
0.00
SSM
RZSM
Precipitation error
Coupling error
Tb error
LAI
Soil roughness
LE error
LAI error
Clay fraction variation
(d)
R L4
Factor importance [-]
0.30
0.25
0.20
0.15
0.10
0.05
0.00
SSM
RZSM
Coupling error
Precipitation error
Tb error
Soil roughness
LAI
LAI error
LE error
Clay fraction variation

398

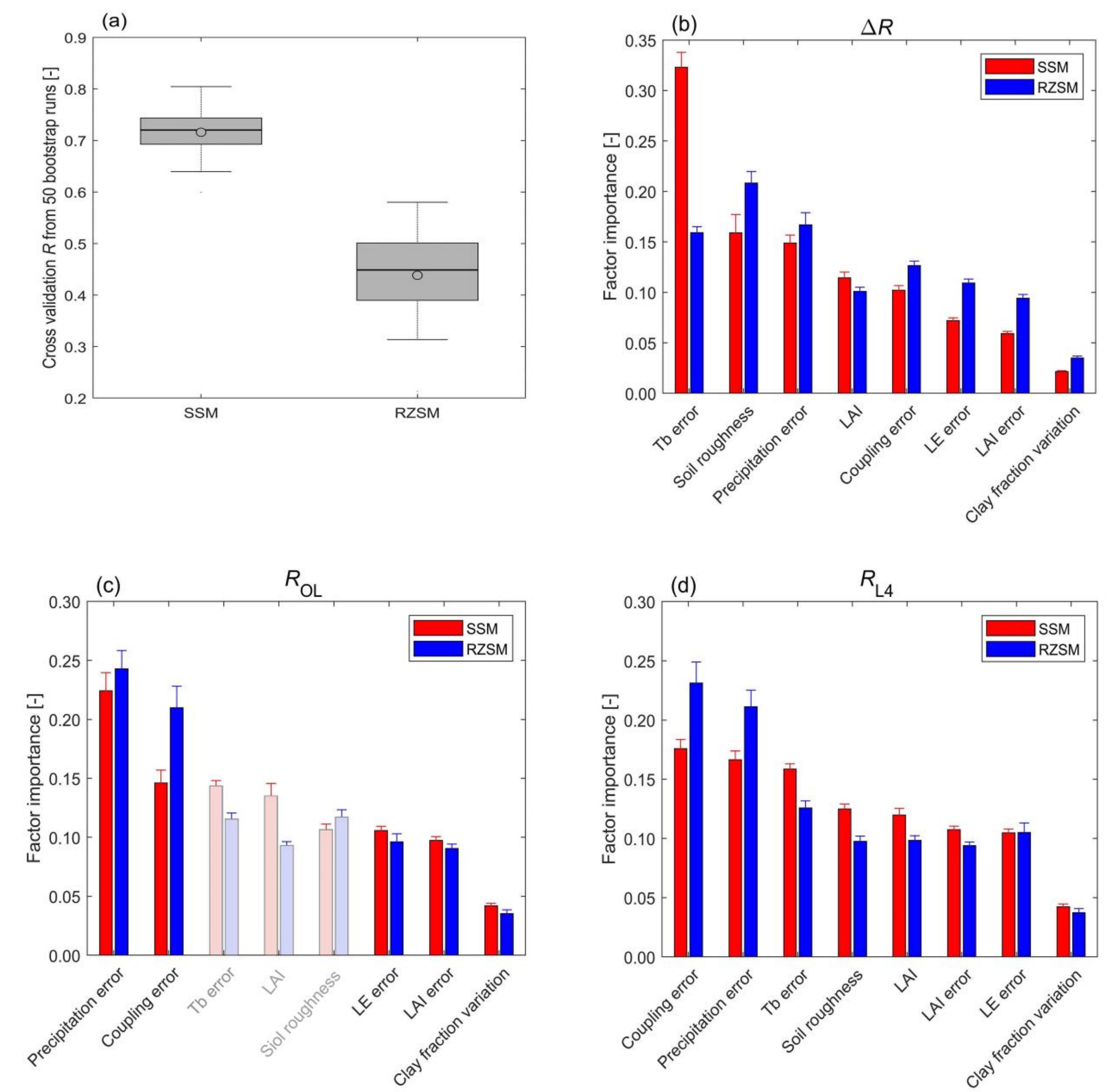


Figure ~~3~~4: Attribution analysis of SMAP L4 DA ~~efficiency~~skill improvement: (a) Cross-validation of RF regression method in predicting DA ~~efficiency~~skill improvement $\Delta R = R_{L4} - R_{OL}$ based on our eight~~8~~ control factors (Table 1). Relative importance of eight~~8~~ control factors determining spatial patterns in (b) DA ~~efficiency~~skill improvement (ΔR), (c) OL performance (R_{OL}), and (d) L4 performance (R_{L4}). Red (blue) bars represent predictor importance for SSM (RZSM). Error bars reflect the standard deviation from 50-member bootstrapping of the RF importance estimates.

Based on the RF results, the Tb error is quantified as the most prominent factor in determining DA ~~efficiency~~ skill improvement (i.e., $\Delta R = R_{L4} - R_{OL}$) – followed by precipitation error and microwave soil roughness (Fig. ~~3b~~4b). The RF-derived ranking of control-factor importance for RZSM is similar to that of SSM in that the same three factors are still the most explanatory. However, relative~~in contrast~~ to SSM, the importance of Tb error for RZSM decreased dramatically from >30% to ~15%. Other modeling error sources (e.g., the vertical variability of soil properties) have only very limited impacts on SMAP DA improvement.

As seen in Fig. ~~3c~~4c, for the OL performance (R_{OL}), the most important factors identified by RF include precipitation error, SSM-RZSM coupling error, and Tb error (microwave soil roughness) for SSM (RZSM). Note that although the Tb error is identified as third-most important factor for R_{OL} in SSM skill, this is an instance where ~~there is~~ correlation (i.e., poorer skill happens to coincide with higher Tb error)~~, but this~~ does not imply a causal relationship. Specifically, it is expected~~normal~~ that Tb (O-F) errors are higher in areas where the OL performs worse, but a high Tb error is not the cause of a low OL performance. The same argument applies to the relationship between microwave soil roughness and OL skill for RZSM estimation. To retain the consistency with analysis of R_{L4} and avoid the misconnection between RTM-related factors and R_{OL}, the bars representing the importance of RTM-related factors to R_{OL} are set semi-transparent in Fig. 3c. The SMAP L4 system is able to reduce ~~the predominant~~ impact of precipitation errors on both SSM and RZSM estimation skill, rendering SSM-RZSM coupling error the most important factor for R_{L4} (Fig. ~~3d~~4d). In addition, in the L4 system, the high vegetation density effect on SSM and RZSM estimation is clearly reduced, as the fourth-most important factor of LAI magnitude is replaced by Tb error.

The qualitative rankings provided by the RF analysis in Fig. ~~3~~ 4 are relatively robust to our particular choice of the benchmark data set to define the 'error' of various control variables. For instance, we replace~~d~~ the CGDPA precipitation benchmark with the Climate Prediction Center Morphing (CMORPH) ~~-~~merged product (Version 1, last access: 6 April 2020; DOI: https://doi.org/10.25921/w9va-q159; Xie et al., 2019), which is the 0.1 degree merging product of CMORPH and observations from more than 30,000 automatic weather stations in mainland China. In~~For~~ this case, the predictive power of the regression model established by the RF is not affected (similar to Fig. ~~3a~~4a), and the qualitative rankings of the precipitation error in R_{OL} and R_{L4} are not impacted (similar to Fig. ~~3c~~4c-d).

3.3.2 Attribution using box plot comparisons

As stated in Section 2.5, the RF method is adept at summarizing the impact of multiple (co-varying) control factors simultaneously in the established regression model, and thus provides more comprehensive insights than the examination of how the target variable (DA improvement) fluctuates with each individual control factor. However, it does not allow the investigation of the sign of the relationship between DA improvement and each control factor – which is important for understanding ~~exactly~~ how each factor influences the DA system. In addition, since the net impact of various factors can enhance DA ~~efficiency~~ skill improvement by either degrading the OL or enhancing the ability of DA to add more value, it is important to decompose the source of variations in ΔR. Therefore, in addition to examining how SMAP DA ~~efficiency~~skill improvement, i.e., $\Delta R = R_{L4} - R_{OL}$, varies as a function of the most prominent control factors identified ~~in the~~ above in Section 3.3.1 (i.e., Tb error, precipitation forcing error, and microwave soil roughness). W~~, w~~e also examine how precipitation error as a control factor affects the OL performance, i.e., R_{OL}.

To minimize the uncertainty caused by large errors in each of the control factors, we exclude samples with errors (separately for each control factor) ranking above the 80th percentile in the following analysis. The relationship between Tb errors and L4 DA ~~efficiency~~skill improvement is straightforward: higher Tb errors are associated with higher ΔR, with ΔR generally larger for SSM than for RZSM (Fig. ~~4a~~5a-b).

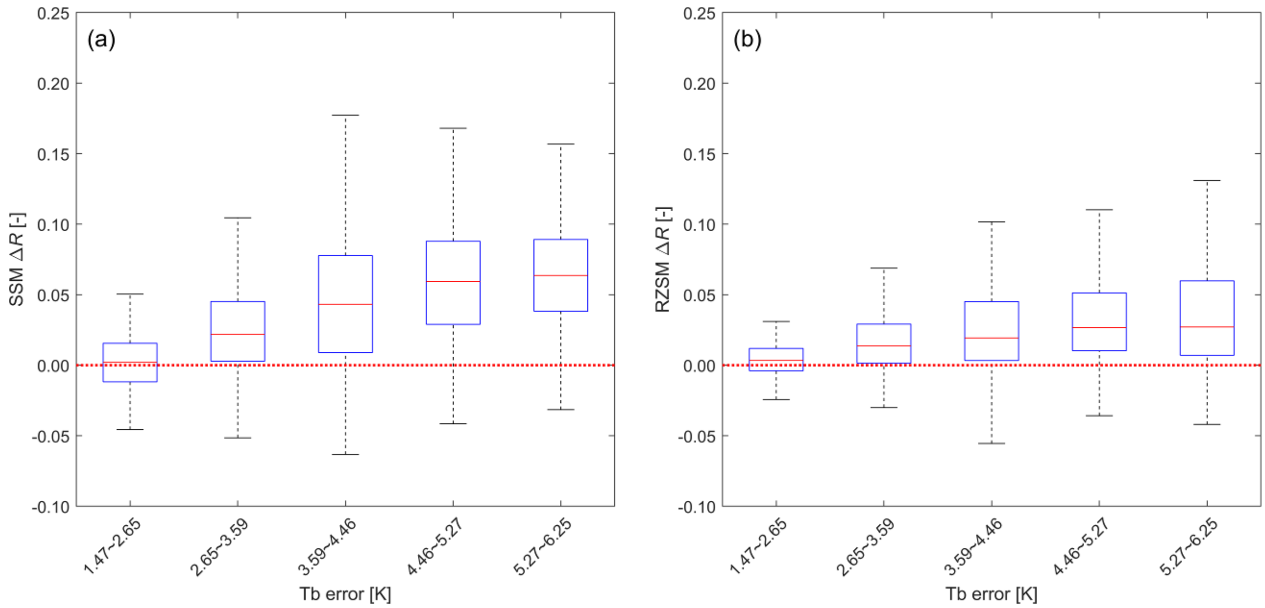


Figure ~~4~~5: SMAP L4 DA ~~efficiency~~skill improvement ($\Delta R = R_{L4} - R_{OL}$) as a function of Tb error for (a) SSM and (b) RZSM. Samples with Tb error ranking above the 80th percentile are excluded from the analysis.

For precipitation, this decomposition is illustrated in Fig. ~~5~~6. Note that, as expected, low-quality precipitation tends to degrade the skill (i.e., correlation versus ground observations) of OL SSM and RZSM estimates (see Fig. ~~5a~~6a-b). This degradation provides an enhanced opportunity for SMAP L4 DA to provide ~~added value~~benefit. As a result, ΔR tends to be a proportional function of precipitation skill (i.e., higher precipitation skill leads to lower ΔR, see Fig. ~~5c~~6c-d). This inverse relationship is a well-known tendency for land data assimilation systems (Liu et al., 2011; Bolten and Crow, 2012; Dong et al., 2019a). Precipitation quality has a diminished impact on RZSM estimation skill compared to SSM estimation skill. This is expected since RZSM is (essentially) the result of applying a low-pass time series filter to precipitation. As such, it is less sensitive to high-frequency errors in precipitation products than SSM is.

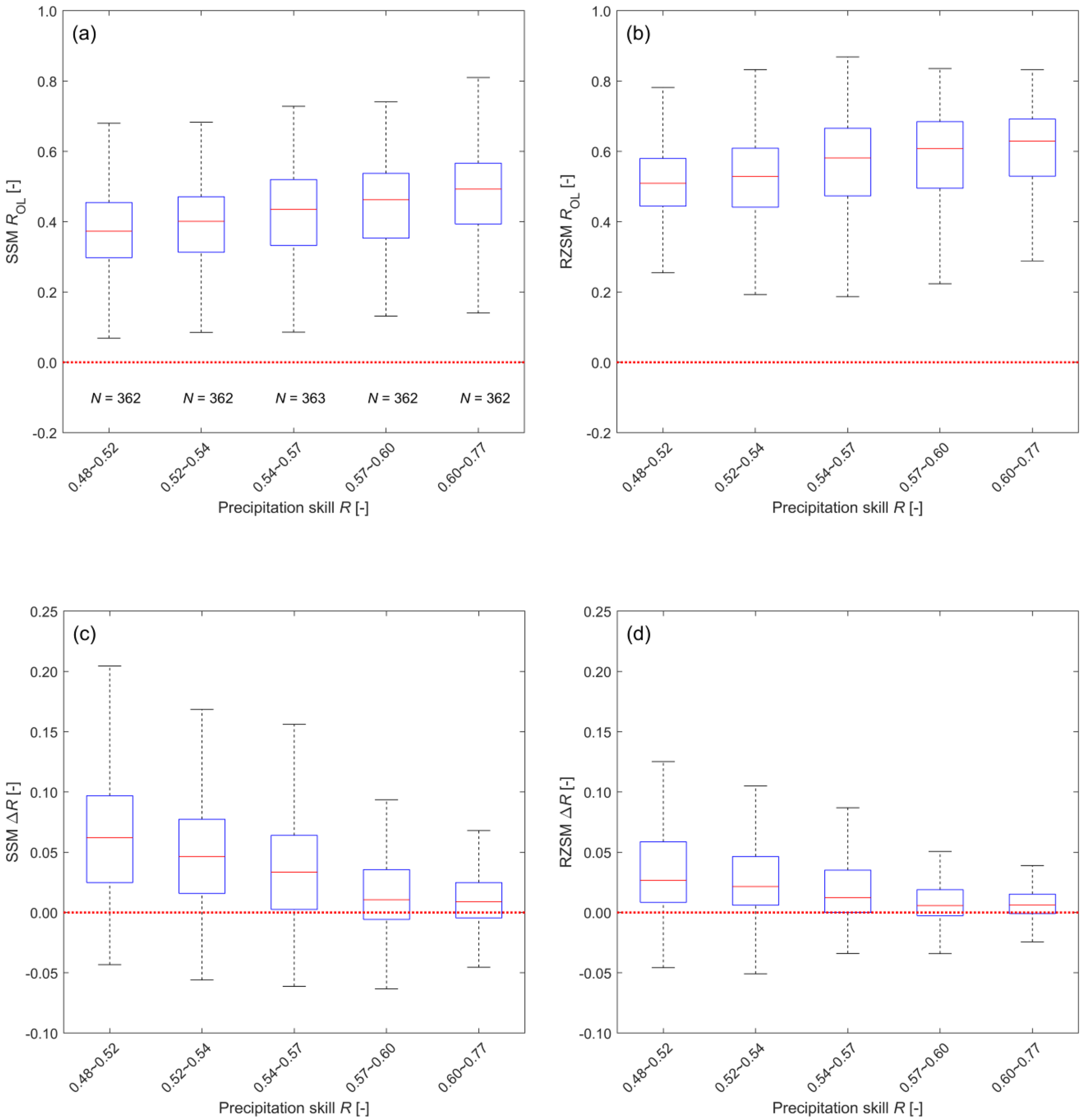


Figure ~~5~~6: OL performance (R_{OL}) as a function of precipitation forcing skill R for (a) SSM and (b) RZSM. SMAP L4 DA ~~efficiency~~skill improvement ($\Delta R = R_{L4} - R_{OL}$) as a function of precipitation skill for (c) SSM and (d) RZSM. Samples with precipitation skill ranking below the 20th percentile are excluded from the analysis.

Figure ~~6~~ 7 is analogous to Fig. ~~4~~ 5 but shows skill differences ΔR as a function of microwave soil roughness. Similar to Tb errors, it is as expected that this control factor of microwave soil roughness has little impact on the OL performance, except that R_{OL} shows slight decreasing tendency with increasing soil roughness (not shown). Given the fact that the OL does get worse with increasing roughness, there is more room for improvement in areas with higher soil roughness~~as the roughness increases~~, which makes it plausible that ΔR increases with increasing soil roughness (see Fig. ~~6a~~7a-b).

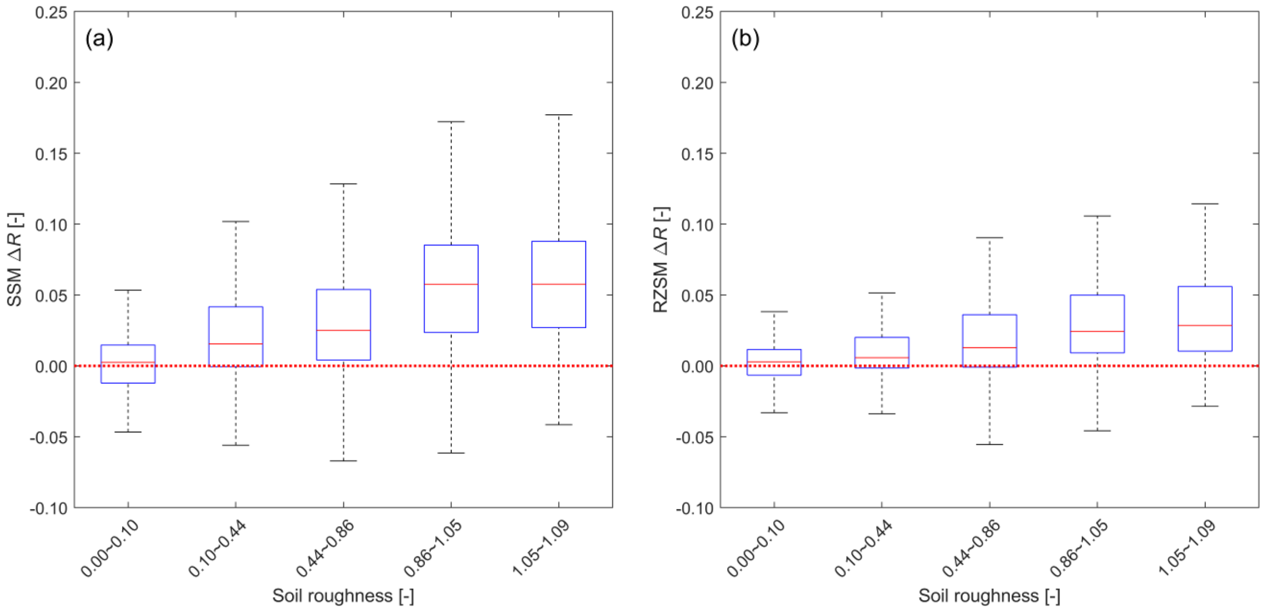


Figure ~~6~~7: As in Fig. ~~4~~ 5 but for ΔR as a function of microwave soil roughness.

Besides the above three control factors that dominate the DA ~~efficiency~~skill improvement, we also examine the top factor that affects SMAP L4 performance, i.e., vertical-coupling errors (Fig. ~~7~~8). As expected, larger (absolute) bias in SSM-RZSM coupling in CLSM tends to be associated with degraded OL estimates of both SSM and RZSM (see Figs. ~~7a~~8a-b), although the analysis does not prove such a causal relationship. Similar to precipitation errors above, decreased OL skill (seen on the left-hand-side of the figures) provides an opportunity for increased DA ~~efficiency~~skill improvement – which is clearly seen in Fig. ~~7~~8. However, such increases are much larger for SSM than for RZSM. For RZSM, SSM-RZSM coupling bias exerts both positive and negative effects on estimation accuracy~~represents a double-edged sword~~. While such bias leads to an enhanced opportunity to improve upon a degraded OL, it should also hamper the ability of DA to transfer SSM increments into the root-zone – particularly when, like here, the bias reflects the lack of vertical coupling in the model (Kumar et al., 2009). This means that some of the opportunity presented by the larger ~~OL~~ RZSM errors in OL is squandered by sub-optimal DA. As a result, the increase in RZSM DA ~~efficiency~~skill improvement associated with biased SSM-RZSM coupling (Fig. ~~7d~~8d) is smaller than the analogous increase in SSM DA ~~efficiency~~skill improvement (Fig. ~~7c~~8c).

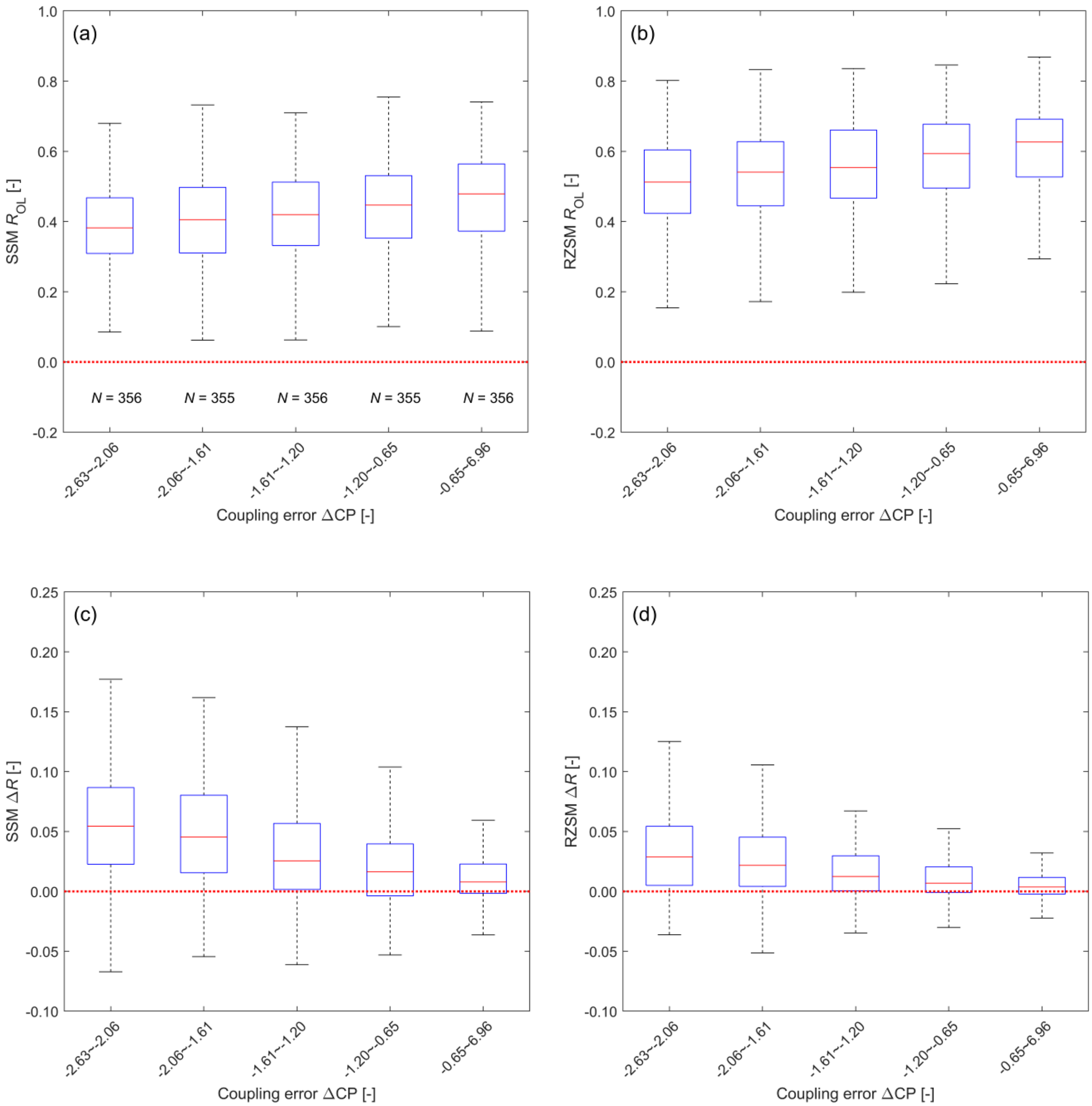


Figure ~~7~~8: As in Fig. ~~5~~ 6 but for R_{OL} and ΔR as a function of SSM-RZSM coupling error indicated by the CP difference ($\Delta CP = CP_{OL} - CP_{obs}$).

For the three strongest control factors that determine DA ~~efficiency~~skill improvement ΔR, i.e., Tb error, precipitation error and microwave soil roughness, we further conducted paired one-way analysis of variance ~~(not shown)~~. Results indicates that for each of the five binned groups separated by each of the above-mentioned three control factors, the inter-group difference in ΔR caused by each control factor is significant (p<0.01) for both SSM and RZSM. In addition, except for the groups with lowest mean ΔR in Fig. ~~4a~~ 5a and Fig. ~~6a~~7a, the averages of ΔR from all groups are significantly higher than 0 (p<0.01).

~~As expected, precipitation error is the dominant factor for explaining the skill of the OL estimates. In contrast, the SSM-RZSM coupling error is the dominant factor for explaining the skill of the L4 results, which shows DA is able to correct for precipitation errors.~~

4 Conclusions

The SMAP L4 algorithm assimilates L-band Tb observations into the Catchment Land Surface Model~~,~~ to provide surface and root-zone soil moisture estimates (i.e., SSM, RZSM) with global, 3-hourly coverage at 9-km resolution. The performance of the L4 soil moisture estimates compared to a baseline model-only simulation (OL) is influenced by multiple control factors associated with ~~the~~ C~~land surface modelling (~~LSM~~)~~ and the tau-omega ~~radiative transfer modeling~~ ~~(~~RTM~~)~~ components of the L4 system. In this study, we assess the performance of SMAP L4 DA system using two~~the 2~~ years of in-situ soil moisture profile observations at 2474 sites across mainland China. We apply a random forest (RF) regression to identify the dominant factors (from a pre-defined list) that control the spatial distribution of the DA ~~efficiency~~skill improvement (defined as the skill difference between the L4 and OL estimates of SSM and RZSM as measured by their Spearman rank correlation with in-situ measurements). Results show that L4 improves SSM prediction skill by 14% on average, with over 77% of the 2287 9-km EASE grid cells showing an increase in Spearman's rank correlation with in-situ observations. Similarly, widespread, though~~but~~ smaller, improvements are ~~also~~ observed in RZSM, with averaged R improvement of 7%.

Based on the RF regression analysis, the ~~added value~~benefit of SMAP L4 DA for SSM is primarily determined by Tb error (measured by standard deviation of O-F Tb residuals), followed by microwave soil roughness and daily precipitation error. These three factors are also the most prominent factors controlling SMAP DA improvement for RZSM, albeit with the Tb error being the least important of these three factors for RZSM DA ~~efficiency~~skill improvement.

Generally, the OL performance clearly decreases with increasing precipitation error, whereas for L4 performance precipitation error is not identified as the most dominant control factor. This indicates that the L4 system is able to correct for errors in precipitation forcing. In addition, our results demonstrate that SMAP DA contributes the most ~~added value~~benefit for cases where CLSM underestimates SSM-RZSM vertical coupling strength. However, due to the difference in top-layer soil depth between the in-situ observations (10 cm) and the L4 analysis (5 cm), it is unclear whether or not the observed SSM-RZSM coupling strength biases are real in an absolute sense – or simply reflect inconsistencies in the depth of modelled versus observed SSM and RZSM time series. Nevertheless, it is worth stressing that, despite the ambiguity ~~with regards to~~about their absolute magnitude/sign, relative variations in apparent SSM-RZSM coupling biases explain a significant amount of the observed spatial variation in L4 performance. Therefore, this finding clearly underpins the importance of properly specifying SSM-RZSM coupling strength in CLSM as a way to improve the SMAP L4 product.

For SMAP L4 SSM skill, the next-most important factors (after SSM-RZSM coupling) are the precipitation error, the Tb error and microwave soil roughness (Fig. ~~3d~~4d). For L4 RZSM skill, the next-most important factors (after SSM-

RZSM coupling) are the precipitation error, the Tb error and the LE error, with the latter two factors of comparable importance (Fig. ~~3d~~4d). To enhance the L4 performance, additional focus should thus be placed on improving the model's characterization of the microwave radiative transfer modeling (Tb error), together with the partitioning of the available energy into latent and sensible heat (LE error)~~ and the microwave radiative transfer modeling (Tb error)~~. Some of our RF analysis results fall squarely within expectation; for instance, the OL skill is predominately determined by precipitation error, and L4 skill improvement (i.e., R_{L4} - R_{OL}) is mostly determined by Tb error. On the other hand, there are also some more surprising results. For instance, we found that SSM-RZSM coupling error and precipitation error have a comparable impact on OL. For L4 skill, however, the impact of SSM-RZSM coupling error exceeds that of precipitation error. More specifically, L4 DA contributes the most benefit for cases where CLSM underestimates SSM-RZSM vertical coupling strength. These findings could be used for L4 product development. In addition, this study pinpoints that the L4 skill improvement is not heavily impacted by LAI magnitude, which gives confidence for using the L4 product over densely vegetated areas.

Data availability

The SMAP L4 datasets are available from https://nsidc.org/data/SPL4SMAU/versions/4. Gauge-based precipitation dataset CGDPA is from http://data.cma.cn/data/cdcdetail/dataCode/SEVP_CLI_CHN_PRE_DAY_GRID_0.25.html. The availabilities of other datasets are stated in their corresponding subsections.

Author contributions

Jianxiu Qiu and Jianzhi Dong conceptualized the study. Jianxiu Qiu carried out the analysis and wrote the first draft manuscript, Wade Crow refined the work, Jianzhi Dong, Rolf Reichle, and Gabrielle De Lannoy helped with the analysis. All authors contributed to the analysis, interpretation of the results and writing.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant Nos. 41971031, 41501450). Rolf Reichle was supported by the NASA SMAP mission. Gabrielle De Lannoy was supported by KU Leuven C1 (C14/16/045). The findings, conclusions and representations of fact in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy.

References

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., and Smets, B.: GEOV1: LAI, FAPAR Essential Climate Variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production, Remote Sens. Environ., 137, 299-309, doi:10.1016/j.rse.2013.02.030, 2013.

Bolten, J.D. and Crow, W.T.,: Improved prediction of quasi-global vegetation conditions using remotely-sensed surface soil moisture, Geophys. Res. Lett., 39(19), doi:10.1029/2012GL053470, 2012.

Breiman, L.: Random forests, Mach. Learn., 45(1), 5–32, doi:10.1023/A:1010933404324, 2001.

Chan, S., Njoku, E. G. and Colliander A.: SMAP L1C radiometer half-orbit 36 km EASE-Grid brightness temperatures, version 3. NASA National Snow and Ice Data Center Distributed Active Archive Center, 10.5067/E51BSP6V3KP7, 2016.

Chen, F., Crow, W.T., Starks, P.J. and Moriasi, D.N.: Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, Adv. Water Resources., 34(4), 526-536, doi:10.1016/j.advwatres.2011.01.011, 2011.

Chen, F., Crow, W.T., Colliander, A., Cosh, M.H., Jackson, T.J., Bindlish, R., Reichle, R.H., Chan, S.K., Bosch, D.D., Starks, P.J., and Goodrich, D.C.: Application of triple collocation in ground-based validation of Soil Moisture Active/Passive (SMAP) level 2 data products, IEEE JSTARS., 99, 1-14, doi:10.1109/JSTARS.2016.2569998, 2016.

Crow, W.T. and Van Loon, E.: The impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture, J. Hydrometeorol., 8(3), 421-431, doi:10.1175/jhm499.1, 2006.

De Lannoy, G. J. M., Reichle, R. H., and Pauwels, V. R. N.: Global calibration of the GEOS-5 L-band microwave radiative transfer model over nonfrozen land using SMOS observations, J. Hydrometeorol., 14(3), 765–785, doi:10.1175/JHM-D-12-092.1, 2013.

De Lannoy, G. J. M., Reichle, R. H., and Vrugt, J. A.: Uncertainty quantification of GEOS-5 L-band radiative transfer model parameters using Bayesian inference and SMOS observations, Remote Sens. Environ., 148, 146–157, doi :10.1016/j.rse.2014.03.030, 2014.

Dong, J., Crow, W.T., Reichle, R., Liu, Q., Lei, F., and Cosh, M.: A global assessment of added value in the SMAP Level 4 soil moisture product relative to its baseline land surface model, Geophys. Res. Lett., 46, 6604-6613, doi:10.1029/2019GL083398, 2019a.

Dong, J., Crow, W.T., Duan, Z., Wei, L., and Lu, Y.: A double instrumental variable method for geophysical product error estimation, Remote Sens. Environ., 225, 217-228, doi:10.1016/j.rse.2019.03.003, 2019b.

Dong, J., Crow, W.T., Tobin, J. K., Cosh, H. M., Bosch, D. D., Starks, J. P., Seyfried, M., and Collins, H. C.: Comparison of microwave remote sensing and land surface modeling in surface soil moisture climatology estimation, Remote Sens. Environ., 242, 111756, doi :10.1016/j.rse.2020.111756, 2020.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., and Edelstein, W. N.: The soil moisture active passive (SMAP) mission, P. IEEE., 98(5), 704–716, doi:10.1109/jproc.2010.2043918, 2010.

FAO/IIASA/ISRIC/ISSCAS/JRC (2012), Harmonized World Soil Database (version 1.2), Food and Agric. Organ., Rome. Available at: http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML.
~~Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J. C., and Draper, C.: Validation practices for satellite soil moisture retrievals: What are (the) errors?, Remote Sens. Environ., 244, 111806, doi:10.1016/j.rse.2020.111806, 2020.~~

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrometeorol., 377(1-2), 80-91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Han, S., Shi, C. X., Jiang, L. P., Zhang, T., Liang, X., Jiang, Z. W., Xu, B., Li, X. F., Zhu, Z., Lin, H. J.: The simulation and evaluation of soil moisture based on CLDAS, J. Applied Meteorol. Sci., 28(3), 369-378, doi:10.11898/1001-7313.20170310, 2017.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, Sci. Data., 6(1), 1-14, doi:10.1038/s41597-019-0076-8, 2019.

Kumar, S.V., Reichle, R.H., Koster, R.D., Crow, W.T., and Peters-Lidard, C.D.: Role of subsurface physics in the assimilation of surface soil moisture observations, J. Hydrometeorol., 10, 1534-1547, doi:10.1175/2009JHM1134.1, 2009.

Lucchesi, R.: File specification for GEOS-5 FP, NASA GMAO Office Note 4 (version 1.0), 63 pp. Available at https://ntrs.nasa.gov, 2013.

Mahanama, S. P., Koster R. D., Walker G. K., Takacs L. L., Reichle R. H., De Lannoy G., Liu Q., Zhao B., and Suarez M. J.: Land boundary conditions for the Goddard Earth Observing System model version 5 (GEOS-5) climate modeling system–Recent updates and data file descriptions. NASA/TM-2015-104606, Vol. 39, 55 pp. NASA Goddard Space Flight Center, Greenbelt, MD. Available at https://ntrs.nasa.gov/search.jsp?R=20160002967, 2015.

McColl, K., Vogelzang, J., Konings, A.G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, Geophys. Res. Lett., 41(17), 6229-6236, doi:10.1002/2014gl061322, 2014.

Piepmeier, J. R., Focardi, P., Horgan, K. A., Knuble, J., Ehsan, N., Lucey, J., Brambora, C., Brown, P. R., Hoffman, P. J., French, R. T., Mikhaylov, R. L., Kwack, E. Y., Slimko, E. M., Dawson, D. E., Hudson, D., Peng, J., Mohammed, P. N., de Amici, G., Freedman, A. P., Medeiros, J., Sacks, F., Estep, R., Spencer, M. W., Chen, C. W., Wheeler, K. B., Edelstein, W. N., O'Neill, P. E., and Njoku, E. G.: SMAP L-band microwave radiometer: Instrument design and first year on orbit, IEEE T. Geosci. Remote., 55(4), 1954–1966, doi:10.1109/TGRS.2016.2631978, 2017.

Liu, Q., Reichle, R., Bindlish, R., Cosh, M.H., Crow, W.T., de Jeu, R., de Lannoy, G., Huffman, G.J. and Jackson, T.J.: The contributions of precipitation and soil moisture observations to the skill of soil moisture estimates in a land data assimilation system, J. Hydrometeorol., 12(5), 750-765, doi:10.1175/JHM-D-10-05000.1, 2011.

Reichle, R.H., Crow, W.T., Koster, R. D., Sharif, H. and Mahanama, S.: Contribution of soil moisture retrievals to land data assimilation products, Geophys. Res. Lett., 35(1), doi:10.1029/2007GL031986, 2008.

Reichle, R. H., de Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., Crow, W., Jackson, T. J., Jones, L. A., Kimball, J. S., Koster, R. D., Mahanama, S. P., Smith, E. B., Berg, A., Bircher, S., Bosch, D., Caldwell, T. G., Cosh, M., González-Zamora, Á., Holifield Collins, C. D., Jensen, K. H., Livingston, S., Lopez-Baeza, E., Martínez-Fernández, J., McNairn, H., Moghaddam, M., Pacheco, A., Pellarin, T., Prueger, J., Rowlandson, T., Seyfried, M., Starks, P., Su, Z., Thibeault, M., van der Velde, R., Walker, J., Wu, X., and Zeng, Y.: Assessment of the SMAP Level-4 surface and root-zone soil moisture product using in situ measurements, J. Hydrometeorol., 18(10), 2621–2645, doi:10.1175/JHM-D-17-0063.1, 2017a.

Reichle, R. H., de Lannoy, G. J. M., Liu, Q., Koster, R. D., Kimball, J. S., Crow, W. T., Ardizzone, J. V., Chakraborty, P., Collins, D. W., Conaty, A. L., Girotto, M., Jones, L. A., Kolassa, J., Lievens, H., Lucchesi, R. A., and Smith, E. B.: Global assessment of the SMAP Level-4 surface and root-zone soil moisture product using assimilation diagnostics, J. Hydrometeorol., 18(12), 3217–3237, doi:10.1175/jhm-d-17-0130.1, 2017b.

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., and Liu, Q.: SMAP L4 Global 9 km EASE-grid surface and root zone soil moisture land model constants, Version 4, NASA National Snow and Ice Data Center DAAC, https://doi.org/10.5067/KGLC3UH4TMAQ, 2018a.

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q.: SMAP L4 global 3-hourly 9 km EASE-grid surface and root zone soil moisture analysis update data, version 4, NASA National Snow and Ice Data Center DAAC, https://doi.org/10.5067/60HB8VIP2T8W, 2018b.

Reichle, R. H., de Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., & Liu, Q.: SMAP L4 global 3-hourly 9 km EASE-grid surface and root zone soil moisture geophysical data, version 4, NASA National Snow and Ice Data Center DAAC, https://doi.org/10.5067/KPJNN2GI1DQR, 2018c.

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J., Kimball, J. S., and Kolassa, J.: Version 4 of the SMAP Level-4 soil moisture algorithm and data product, J. Adv. Model Earth Sy., 11(10), 3106-3130, doi:10.1029/2019MS001729, 2019.

Reichle, R. H., and Coauthors (2020), The contributions of gauge-based precipitation and SMAP brightness temperature observations to the skill of the SMAP Level-4 soil moisture product, J. Hydrometeorol., accepted, doi:10.1175/JHM-D-20-0217.1.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., and Lehner, I.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Sci. Rev., 99, 125–161, doi:10.1016/j.earscirev.2010.02.004, 2010.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., Hurk, B., Hagemann, S., and Berg, A.: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACECMIP5 experiment, Geophys. Res. Lett., 40(19), 5212-5217, doi:10.1002/grl.50956, 2013.

Shen, Y., Xiong, A., Wang, Y., and Xie, P.: Performance of high-resolution satellite precipitation products over China, J. Geophys. Res-Atmos., 115(D2), doi:10.1029/2009JD012097, 2010.

Shen, Y. and Xiong, A.: Validation and comparison of a new gauge-based precipitation analysis over mainland China, Int. J. Climatol., 36(1), 252-265, doi:10.1002/JOC.4341, 2015.

Verger, A., Baret, F., and Weiss, M.: Performances of neural networks for deriving LAI estimates from existing CYCLOPES and MODIS products, Remote Sens. Environ., 112, 2789-2803, doi:10.1016/j.rse.2008.01.006, 2008.

Xie, P., Yatagai, A., Chen, M., Hayasaka, T., Fukushima, Y., Liu, C., and Yang, S.: A gauge-based analysis of daily precipitation over East Asia, J. Hydrometeorol., 8, 607-626, doi:10.1175/JHM583.1, 2007.

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., Lin, R..: NOAA CDR Program: NOAA Climate Data Record (CDR) of CPC Morphing Technique (CMORPH) High Resolution Global Precipitation Estimates, Version 1. NOAA National Centers for Environmental Information, 2019.