

Dear Reviewer #2 (RC2)

Thank you for your comments and suggestions. Based on your suggestion, I have made certain revisions to the manuscript.

Q1: misuse of desertification, monsoon and other geographic terms throughout the manuscript. The study region, defined by the authors, is not “areas affected by desertification”, neither “monsoon climate region”. Please check in detail.

A1: Thank you very much for this suggestion. The study area is mostly "temperate continental climate", not“monsoon climate region”. It was deleted. (Line 104)

For “areas affected by desertification”. Our study area is provided by the National Forestry and Grassland Administration (China). We checked with the appropriate manager recently, and they thought that it is more scientific to replace “areas affected by desertification” with “desertification areas”. All of these words have been changed in the manuscript. Please see the modified version.

Q2: Line 18-19: are you sure “very sensitive to SM”?

A2: Thank you very much for this suggestion. The ultimate purpose of our study is vegetation restoration in the study area, where soil moisture is a key indicator. "very sensitive to SM " is a bit absolute, we have modified it to " sensitive to SM". (Line 20)

Q3: Line 32: provide references for GLDAS.

A3: Thank you very much for this suggestion. Related literature has been added. (Line 33)

Q4: Line 36-40: data assimilation products may be produced with satellite data as inputs. Thus, it is not independent on remote sensing. Modify your statements.

A4: Thank you very much for this suggestion. In order to avoid ambiguity, we have made some modifications. (Line 40)

Q5: Line 44: what does the “very stable” mean here? Passive microwave radiometer data are sensitive to more influences, such as atmospheric effects and surface vegetation.

A5: Thank you very much for this suggestion. We would have liked to express that passive remote sensing products are generally more stable compared to active remote sensing. In order to avoid ambiguity, we deleted it.

Q6: Line 56: what does ‘directly retrieve’ mean?

A6: It has been modified (Line 59). Hope to help you understand.

Q7: Each method produces a dataset. That does not mean the multiple machine learning methods produce the datasets following the normal distribution. In this sense, statistical mean may be biased, which is well-known to climate community.

A7: Thank you very much for this suggestion. Our study uses a combination of multiple machine learning to select the best regression model for each period and not by taking an average for the SM result.

Q8: Line 91-92: wrong description of the region with “monsoon climate”. So is the desertification.

A8: Thank you very much for your suggestions. We have made modifications. Refere to Question (1).

Q9: Line 93: “water-vapor-ecosystem”, what does it mean?

A9: Thank you very much for your suggestions. It has been modified.(Line 105)

Q10: Line 115: Give the full spelling for NDWI, LSW, ECMWF, EVI, geotiff and many others for their first appearance in text.

A10: Thank you very much for your suggestions. We have carefully checked all abbreviations, and some abbreviations that do not appear are added with full spelling. Please see the modified version.

Q11: The parameters used for ML are linearly correlated. Does it affect your results?

A11: Thank you very much for your suggestions. Collinearity between variables will affect the simulation results, which is not considered in the description process of this paper. We add some content in Section 4.2. In general, except for ensemble algorithms (including RF and XGB), collinearity is more or less affected. Due to this advantage of the ensemble algorithm, many studies generally do not consider multicollinearity problems when using random forests for regression or classification. We have added discussion in section 4.2. Please see the modified version.

Q12: Line 177: incomparable?

A12: Thank you very much for your suggestions. The expression is a bit absolute, replaced by “prominent”. (Line 206)

Q13: Equations for RMSE (6) and (8) are wrongly expressed.

A13: Thank you very much for your suggestions. There should be nothing wrong with these two equations. The following is the equations from the literature (Hu et al., 2020).

$$\text{RMSE} = \sqrt{E[(\theta_{\text{SMAP}} - \theta_{\text{insitu}})^2]} \quad (7)$$

$$\text{ubRMSE} = \sqrt{E\{[(\theta_{\text{SMAP}} - E[\theta_{\text{SMAP}}]) - (\theta_{\text{insitu}} - E[\theta_{\text{insitu}}])]^2\}} \quad (8)$$

where  $E[\cdot]$  represents the mean operator,  $\theta_{\text{insitu}}$  is the in situ SM,  $\theta_{\text{SMAP}}$  is the downscaled SM,  $\sigma_{\text{SMAP}}$  is the standard deviation of SMAP SM, and  $\sigma_{\text{insitu}}$  is the standard deviation of the in situ SM.

Q14: Figures 4 and 5: there are clearly seasonal variation in correlation coefficient and RMSE. It means significant systematic errors in the products. Give scientific explanation to the data reliability.

A14: Thank you very much for your suggestions. Regarding the difference in seasons, we think that it is mainly affected by the sample size (Line 344), and we also added discussion in the new version (Line 539).

Q15: Line 251-260: The errors are large between the Maqu and the Bbaso network, which need substantial investigation.

A15: Thank you very much for your suggestions. The evaluation accuracy of Babao network is generally lower than that of Maqu Network, the first reason is the measured soil depth. The soil depth measured by SMAP is 5 cm, the same as that measured by Maqu Network, while that measured by Babao Network is 4 cm. Another reason is that there is a bias between site measurements and remote sensing data itself. The relationship between the 1 km×1 km grid and the site itself will have a large error. In this study, even SM of some sites from Maqu network did not match well the Remote sensing SM.

To make the results more convincing, we added some comparison between the grid SM data and the in-situ SM data. (Figure 8: Comparison of gridded products and in situ observation SM of the Maqu Network; Figure S2: Comparison of gridded products and in situ observation SM of the Babao Network). Please see the modified version.

Q16: Line 283: “due to spatial resolution” is a superficial reasoning. Insightful clarification should be given.

A16: Thank you very much for your suggestions. For a detailed discussion, please refer to the revised version (Section 4.4: Uncertainty)

Q17: Line 291: here appears ‘process of vegetation growth’. SMAP SM data are subject to vegetation cover, which is known in the field, but the authors failed to address it.

A17: Thank you very much for your suggestions. It was deleted.

Q18: Line 295: “little variation”? change the words.

A18: Thanks for your advice! It has been modified. (Line 429)

Q19: There are too many “some” in text. Vague expression.

A19: Thank you very much for your suggestion. The unnecessary "some" of the manuscript have been deleted. Some expressions are also further modified.

Q20: Line 327: strange subtitle.

A20: Thank you very much for your suggestion. It is modified to “Variable importance assessment”.  
(Line 469)

Q21: Line 335: “influence of soil texture (sand, silt and clay) is relatively weak, but it cannot be completely ignored.”. why?

A21: Thank you very much for your suggestion. It caused some ambiguity and we deleted the latter part.  
(Line 424)

Q22: Line 347: IncNodePurity? What is it?

A22: Thank you very much for your suggestion. It reflects an important indicator, and we added its full spelling. (Line 436)

Q23: Line 350: various noises? How many?

A23: Thank you very much for your suggestion. My expression was not clear, we deleted this sentence.  
(Line 450)

Q24: Line 367: mainly significantly. Remain one.

A24: Thank you very much for your suggestion. “mainly” was modified to "more"

Q25: Line 382-383: delete it.

A25: Thank you very much for your suggestion. It has been removed.

Q26: Line 391: “a framework was proposed”? It does not make sense.

A26: Thank you very much for your suggestion. “a framework” was modified to "an approach".( Line 523)

Finally, thanks a lot for your careful review and invaluable advices. Looking forward to the opportunity to learn from you! I also made some other revisions, please refer to other review results.

Looking forward to your next suggestions. Thank you!

Sincerelly!

Pinzeng Rao

Address: Beijing, China.

Email: 578264905@qq.com