Responses to RC #4:

The manuscript entitled "Spatio-temporal Monitoring of Agricultural Drought in China Based on Downscaled Soil Moisture Data" by Luo et al. investigates the downscaling of remotely sensed soil moisture products from a spatial resolution of approximately 0.25° to 0.05° for the purpose of agricultural drought monitoring. The authors present a comprehensive effort to design and assess the proposed downscaling framework. The subject is both timely and highly relevant, given the increasing importance of accurate drought monitoring for water resources management and policy decisions. Overall, the study fits well within the scope of Hydrology and Earth System Sciences and demonstrates clear potential to make a meaningful contribution to the field. Nevertheless, several methodological aspects and minor technical details would benefit from additional clarification and refinement to improve the transparency and robustness of the findings. The comments below are intended to support the authors in further strengthening the manuscript's clarity.

Reply: Thank you for your thoughtful review and valuable feedback. We appreciate your positive acknowledgment of the potential significance of our study and your constructive comments, and we are grateful for the time and effort you've dedicated to this review. Your comments and good suggestions are very important for us to improve the quality of the manuscript. We have carefully addressed all the issues raised by you and the response is presented below.

Specific comments:

1. Gap-filling strategy across AMSR-E, SMOS, and AMSR-2 datasets

①The general approach used to fill the gaps between AMSR-E, SMOS, and AMSR-2 soil moisture products requires further clarification.

Reply: We thank the reviewer for this constructive comment. In the revised manuscript, we have clearly specified that the same three-step gap-filling approach was uniformly applied to all three satellite soil moisture products (AMSR-E, SMOS, and AMSR-2). For each missing pixel in a given product, we first reconstructed its value using an enhanced Savitzky-Golay filter based on temporal dynamics, and then reconstructed the same pixel using an enhanced geographically weighted regression method based on spatial information. Subsequently, we applied least squares regression to integrate these two reconstructions (temporal-based and spatial-based) to produce the final reconstructed value. This method effectively fuses spatio-temporal dynamic information to reconstruct high-quality missing values. This description has been fully incorporated into the revised manuscript. Thank you again for your valuable guidance.

②In addition to the variations in ascending and descending times, diurnal variability may also impact the monthly mean soil moisture. These sensors use different methods to calculate soil moisture from microwave signals, including various algorithms, models, frequency bands, calibration steps, and processing systems. Such methodological differences may introduce structural inconsistencies and biases in the merged time series. The manuscript would benefit from a more detailed discussion of how these inter-sensor differences were addressed, harmonized, or evaluated prior to merging the datasets. In addition, clearer documentation and references describing the specific soil moisture products used would improve transparency and reproducibility.

Reply: We thank the reviewer for this constructive and detailed guidance. We apologize that the description of how we addressed inter-sensor differences was insufficiently clear in the original manuscript. Following your suggestion, we have now provided a more detailed clarification in the revised manuscript.

Specifically, we adopted the method used by Meng et al. (2021, Earth System Science Data) to harmonize the multi-sensor soil moisture products. Considering that differences in microwave frequency and satellite overpass time affect soil moisture retrievals, we used the longer-term AMSR-E product as the reference. For SMOS-IC, we established a pixel-wise linear regression between SMOS-IC and AMSR-E over their overlapping period (2003–2011), applied the derived calibration parameters to the SMOS-IC data from 2011–2012, and obtained the calibrated SMOSreg. Similarly, using the overlap between AMSR2 and SMOS-IC, we calibrated AMSR2 data for 2012–2023 to produce AMSR2reg. The method computes the mean, standard deviation, and correlation coefficient between the reference and the target dataset to determine a linear scaling factor, thereby effectively removing structural inconsistencies and biases caused by differences in sensor frequency, overpass time, etc. This procedure has been clearly described in the revised manuscript, with relevant references added. Thank you again for your detailed guidance.

2. Comparison with previous machine learning studies

The manuscript compares the performance of the proposed method with results reported in previous studies that applied various machine learning approaches. Although these comparisons offer valuable context, they do not ensure the accuracy or reliability of the proposed method. Differences in study area, input datasets, preprocessing procedures, spatial and temporal resolution, training strategies, and validation frameworks can substantially influence reported performance metrics. Therefore, a more cautious interpretation of these comparisons is recommended. The manuscript would benefit from a more detailed explanation of how similar the experimental settings are in different studies and a more direct discussion of the uncertainties that come with the modeling framework and input data. This would strengthen the validity of the conclusions drawn from the comparative analysis.

Reply: Thank you for this constructive guidance. Following your suggestion, we have separated the Discussion and Conclusion sections in the revised manuscript and thoroughly rewrote the Discussion. In the new Discussion, we provide a more detailed comparison of our downscaled SM product with benchmark datasets (e.g., the SM product from Meng et al., 2021), specifically focusing on their effectiveness in agricultural drought extraction. We also discuss how differences in input data (e.g., environmental predictors and their sources) and modeling choices influence the accuracy gains achieved by our downscaling algorithm. These revisions allow for a more balanced and cautious interpretation of our comparative results and better highlight the uncertainties associated with the modeling framework and input data. All these improvements have been incorporated into the updated manuscript. Thank you again for your valuable suggestion.

3. Interpretation of correlation strength between SSI and drought-affected areas

The manuscript reports statistically significant negative correlations between area-averaged SSI and governmental crop drought-affected areas (Pearson's R = 0.015–0.277, $p < 0.05$). While statistical significance is achieved, the reported correlation coefficients indicate very weak to weak relationships. In particular, an R value as low as 0.015 suggests negligible practical association, despite being statistically significant. Therefore, the conclusion that the downscaled SSI "can effectively capture agricultural drought patterns across China" appears somewhat overstated.

Reply: Thank you for this valuable comment. We fully agree that the originally reported correlation coefficients (Pearson's R = 0.015–0.277) indicated only weak relationships, despite statistical significance, and that the conclusion of "effectively capturing agricultural drought patterns" was overstated.

To substantially improve drought identification accuracy, we have made two major improvements following your guidance.

First, we recalculated the SSI using month-by-month means and standard deviations. Specifically, we now compute the long-term mean and standard deviation independently for each calendar month over the 2003–2023 study period. This approach properly removes the inherent seasonal cycle of soil moisture and ensures that the SSI reflects only anomalous deviations from the expected conditions for that specific month. We have fully recalculated the entire SSI dataset and updated all related results and figures. With this improved SSI, the consistency between the extracted drought-affected areas (EA) and the official drought statistics (BA) became very strong, achieving an coefficient of determination (R^2) of 0.7456. Moreover, we evaluated the relative error (RE) between EA and BA using the formula:

$$RE = |EA - BA| / BA \times 100\%.$$

The RE was 23%, confirming the superior ability of the revised SSI to identify true agricultural drought events.

Second, we optimized the drought thresholding method. In the original manuscript, we used a fixed threshold (SSI < 0) to identify drought. To make the threshold adaptive to region, month, and interannual variability, we developed a dynamic threshold method based on kernel density estimation (KDE) and historical drought records. Specifically, we divided China into four climatic zones (arid, semi-arid, sub-humid, humid). For each zone and each calendar month, we extracted all valid SSI values from 2003–2023, removed outliers (beyond ±3 standard deviations), and applied KDE to obtain a smoothed probability density curve. Using three complementary reference datasets – (i) government-published annual crop drought-affected area statistics (from the China Statistical Yearbook and China Flood and Drought Disaster Bulletin); (ii) the China Agricultural Meteorological Disaster Dataset compiled by the China Meteorological Administration (CMA), which contains disaster name, occurrence date, intensity, affected area, and damage percentage; and (iii) the daily agricultural drought comprehensive monitoring spatial distribution data from the Open Laboratory of the National Meteorological Center Forecast System − we identified the SSI intervals corresponding to agricultural drought. This yielded a set of dynamic drought thresholds that vary spatially (by climate zone) and temporally (by month). Subsequently, we extracted drought areas using both the original fixed threshold and our newly constructed dynamic threshold across three soil moisture products: the coarse-resolution SM product, the SM dataset proposed in this study, and the SM product developed by Meng et al. (2021). Validation results demonstrate that the proposed dynamic threshold method effectively improved drought identification accuracy across all three datasets (Fig. 1). Quantitatively, compared with the fixed threshold, the dynamic threshold achieved relative improvements in R^2 between estimated and observed drought areas of 23% for the coarse-resolution product, 19% for our proposed product, and 22% for Meng's product.

We have re-identified drought areas in the revised manuscript using this combined approach. All related figures and conclusions have been updated. These revisions will be visible in the manuscript when we upload it in due course. Thank you again for your rigorous and constructive guidance, which has substantially strengthened the validity of our drought assessment.

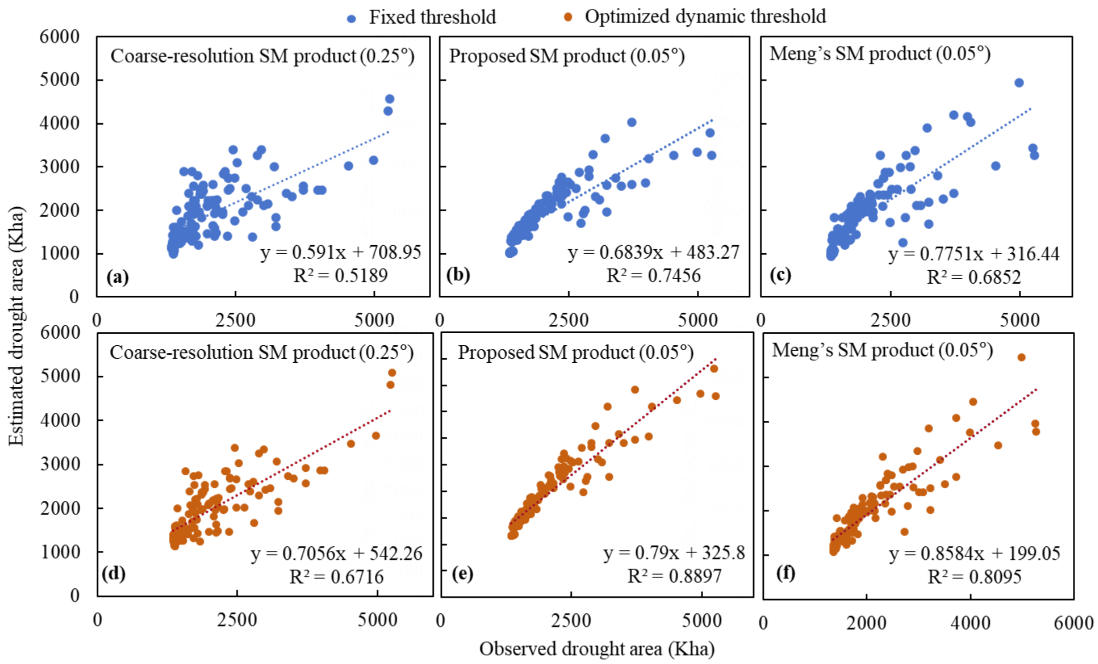

Fig.1 Validation of drought area identification using different SSI thresholds across three soil moisture products with varying resolutions. The top row (a-c) shows results from the fixed threshold SSI method (blue points), and the bottom row (d-f) shows results from the KDE-based optimized dynamic threshold SSI method (orange points). Linear fitting equations and coefficients of determination (R^2) are provided in each subplot.

Technical corrections:

In general, the figures in the manuscript would benefit from consistent fonts, sizes, and styles across all panels and labels in accordance with the journal's figure requirements.

Reply: Thank you for this suggestion. We have revised all figures in the manuscript to ensure consistent fonts, sizes, and styles across all panels and labels, in accordance with the journal's figure requirements. These revisions have been incorporated into the updated manuscript. Thank you again for your guidance.

1. Figure 1: The caption could be rephrased to clarify that the figure shows the locations of the in-situ soil moisture (SM) stations rather than the in-situ SM distribution. Additionally, it may be worth verifying the elevation color scale, as it appears to be reversed (e.g., coastal areas appear higher than the Tibetan Plateau and the Himalayas in the west). "Province boundary" is written twice in the legend

Reply: Thank you. Done as suggested.

2. Figure 2: It provides a comprehensive overview of the methodology but is somewhat dense due to the large number of formulas included. The text already explains many equations, so the figures could benefit from a simplified layout that omits some formulas to improve readability.

Reply: Done as suggested. Thank you again for your guidance.

3. Figure 4: It would be helpful to use the same maximum and minimum limits for the color bar in both the original and downscaled images for easier comparison. The axis of horizontal profile should include units (e.g., consider changing Data Value to Soil Moisture (m--3 m-3), Transect(unit)). While I assume the inset corresponds to the horizontal profile along the transect, it is currently difficult to relate the x-axis values to the transect location on the map. Is this profile also showing the average value along the latitudinal axis?

Reply: Thank you for your guidance. Recognizing that the original figure was not sufficiently intuitive in showing the spatial details of downscaled soil moisture, we have replotted Fig.4. Since our team has long-term in-situ soil moisture measurements in three regions—Harbin (Heilongjiang Province), the junction of Lüliang, Taiyuan and Xinzhou (Shanxi Province), and the Aba Tibetan and Qiang Autonomous Prefecture (marked as I, II, and III in the figure, respectively)—we selected these three regions to demonstrate the downscaling performance in detail. The land cover types are: Region I – agricultural land, forest, grassland, and residential land; Region II – residential land, agricultural land, and forest; Region III – forest and agricultural land. By comparing the spatial patterns before and after downscaling, the improved capability of our product in capturing local soil moisture heterogeneity is clearly visible. The revised Figure 4 has been included in the updated manuscript. Thank you again for your valuable suggestion.

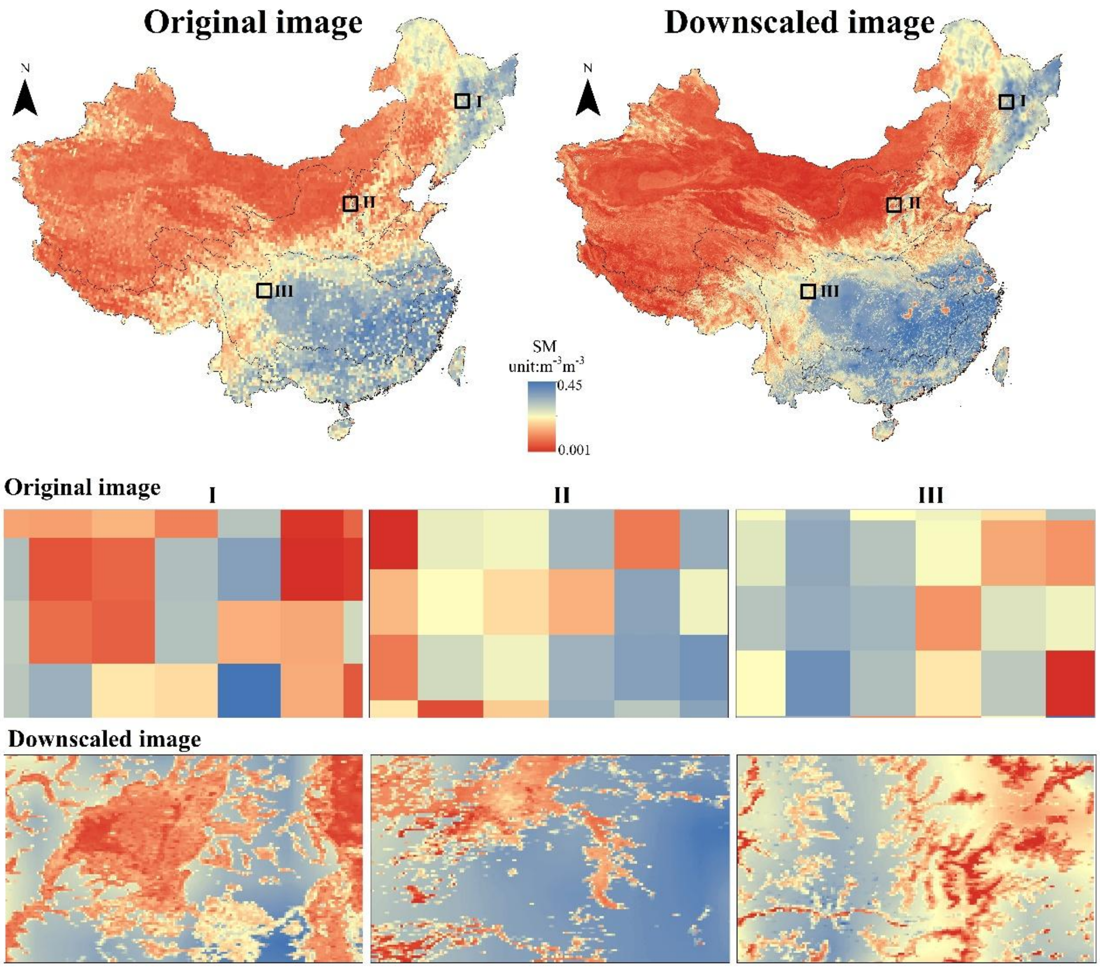

Fig.4 in the revised manuscript — Comparison of original images and downscaled SM images

4. Figure 5: The unit on the figure is currently written as m^3/m^3. For clarity and proper SI notation, it should be formatted as m-$^{-3}$ m^{-3}.

Reply: Thank you. Done as suggested.

5. Figure 7: It is recommended to name each panel with the river basin name and its abbreviation instead of using letters (e.g. (h) Shandong (YRB-HRB-HRYB)) as the current panel naming can be confusing.

Reply: We thank the reviewer for these insightful and constructive suggestions. we have modified it.

6. Figure 8: Please use the same maximum and minimum range to observe the drought evolution throughout the years.

Reply: Done as suggested. Thank you.

Additional remark:

Lines 758-759 in the revised manuscript: "The authors are grateful to the editor, five reviewers, and one reader for their constructive comments and suggestions on this paper." has been added to Acknowledgments.

Special thanks are extended to you for your valuable comments.

We are doing our best to improve the manuscript and are making substantial changes to address the concerns raised.

We greatly appreciate your help and hope that the revisions will meet with approval once we submit the updated manuscript.

Once again, we would like to extend our sincere gratitude and appreciation for your valuable comments and suggestions.