

Root zone soil moisture estimation over China

Jing Tian and **Yongqiang Zhang** 

Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China
Email: tianj.04b@igsrr.ac.cn

Abstract: Root zone soil moisture (RZSM) refers to the water available to plants. Compared to surface soil moisture, RZSM plays a more crucial role in various applications such as agricultural monitoring, drought and flood forecasting, water supply management, land-atmosphere feedbacks, and the hydrologic response of watersheds.

There are three commonly used methods to estimate RZSM. The first method involves in-situ soil moisture sensors. However, field measurements are only representative of a small spatial scale, which limits their applications. The second method is based on surface soil moisture (SSM). Microwave remote sensors are effective for measuring the global distribution of SSM. However, due to the limitation of microwave depth signal penetration, RZSM cannot be obtained based on it. Since SSM is closely related to RZSM through the diffusion process, some methods have been developed to estimate RZSM from SSM. For instance, the exponential filter (EF) method uses a simplified two-layer water balance approach to establish the relationship between SSM and RZSM (Albergel et al., 2008). Data-driven methods use SSM, meteorological data, land surface conditions, and soil properties as drivers to build machine learning models to estimate RZSM. The third method for obtaining RZSM involves modeling approaches that combine climate models or land surface models with data assimilation techniques. RZSM reanalysis data products such as ERA5, MERRA2, and CFSR rely on this method. Similarly, the RZSM products of SMAP level 4 and GLDAS are also model-based. However, large errors usually exist in the RZSM simulations due to uncertainties in model parameters and forcings (Xu).

This study first evaluates seven RZSM products, including ERA5, MERRA2, CFSR, SMAP level 4, GLDAS_NOAH2.1, GLDAS_Catchment2.2, and SMOS CATDS Level 4, using 1153 ground sites across China. Then, a data fusion method is developed based on machine learning method to integrate the seven RZSM products to produce a more accurate RZSM dataset. The evaluation results show that MERRA2 and NOAA had the best performance, with MERRA2 better able to capture surplus and deficit soil moisture conditions. The study also revealed that precipitation forcing plays a vital role in improving RZSM simulations. Conversely, ERA5 overestimates RZSM, and SMOS exhibits a dry bias. From the time series variability, all datasets except SMOS can capture the temporal dynamics of RZSM well. Moreover, the accuracy of the RZSM products varies significantly in tropical climate zones. The newly developed dataset is better than a single RZSM product with five-fold cross-validation.

This study provides crucial insights for improving land surface and hydrological models in estimating RZSM, and proposes an approach to promoting the accuracy of RZSM estimation by maximizing the utilization of existing RZSM products, which is very helpful in drought management and agricultural production.

REFERENCES

- C. Albergel, C.R.E.M., T. Pellarin, J.-C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Pignatelli, E. Martin, 2008. From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. *Hydrology and Earth System Sciences*. 12, 1323-1337. <https://doi.org/10.5194/hess-12-1323-2008>.

Keywords: *Root zone soil moisture, validation, data fusion, seven products, random forest*