

## **General comments**

This study provides a high-resolution, long-term annual forest litterfall dataset across China, which could be a valuable contribution to forest carbon and nutrient cycling research. However, I have substantial concerns about the methodological framework, the reliability of the predictions, and the insufficient validation.

My main concern is the lack of a clear ecological basis linking the selected remote sensing and environmental variables to litterfall production.

The manuscript provides only a broad rationale for using climate and remote-sensing variables, but does not establish a sufficiently explicit ecological basis for the specific predictor set. Litterfall is influenced by vegetation biomass, tissue turnover, phenology, mortality, and disturbances. However, the authors use spectral indices such as NDWI, BSI, and NDBI without adequately explaining how they relate to these processes. Although these variables may indirectly reflect vegetation conditions, their relevance to litterfall production, and to its interannual variability, needs to be demonstrated.

The model seems to capture broad spatial correlations, for example, forests in warmer and wetter regions generally have higher productivity and litterfall. However, this does not mean the model can capture interannual changes caused by drought, extreme temperatures, tree mortality, or changes in leaf production and turnover. This is a major concern for a dataset intended to represent annual litterfall dynamics over 25 years.

Several methodological details are also unclear, particularly the definition and harmonization of ground observations, the selection of predictors, and the validation strategy. The proposed spatial matching approach appears to mainly address the mismatch between field locations and forest pixels. Its contribution to improving prediction accuracy for forest litterfall has not been clearly demonstrated.

This paper only compares the 30 m litterfall product with a 500 m MODIS-based results in few regions (Figure 3). The study lacks comparison with existing litterfall products (e.g., wang et al., 2026; 10.1016/j.rse.2026.115373) and independent validation, particularly the temporal variability.

Overall, although the dataset has potential value, I do not think the current methodology and validation are sufficient to support its reliability. I therefore do not recommend publication in its present form.

## **Specific comments**

1. L65–67. The review of existing remote-sensing-based litterfall mapping is insufficient. The authors should compare the ecological rationale, predictors, spatial and temporal resolution, and validation design of the present study with previous approaches, such as Hu and Huang (2019), Guo et al. (2025; Spatiotemporal prediction of forest litterfall in China by using multi-source data

and Transformer-CatBoost model), Zhao et al. (2022), and Wang et al. (2026; Remote sensing-based estimation of global litterfall dynamics in forest ecosystems and its potential impact on soil respiration).

2. Section 2.1. How many sites and observations were collected from CERN? What are their locations, forest types, and observation periods? What exactly was measured as litterfall production? leaves only, or total litterfall including branches and other components?
3. L100. Was GLC\_FCS30D used only to define the forest mask and summarize the resulting product, or was it included as a predictor in the Random Forest model? Was it used to distinguish different forest types, such as DNF, ENF, DBF, and EBF? These forests have different leaf turnover rates and litterfall dynamics. How were these differences considered in the model prediction and validation?
4. L120-124. The authors used seasonal precipitation and temperature variables from TerraClimate. Could this temporal aggregation obscure the effects of extreme drought, heat events, and seasonal phenology on litterfall? Were climate anomalies or extreme events considered?
5. L135-140. The manuscript states that multi-year averages without annual resolution were excluded, whereas annual observations and records containing more than six monthly measurements were kept. Please report how many records entered each category, how many annual values required missing-month imputation. Predictor variables should be temporally aligned with the exact observation window. The sensitivity of model performance and temporal variability to the cross-year imputation procedure should also be evaluated.
6. L 145–150 The outlier removal procedure, which excludes annual observations farthest from the site median, may also remove genuine extreme litterfall events caused by drought, storms, or other disturbances. How many observations were removed, and how did this affect the interannual variability?
7. L170–175. Please provide a complete description of all predictors, including their temporal resolution and aggregation methods. Were forest type and ecological region included? Were climatic predictors represented by annual values, seasonal means, or long-term averages?

The relationship between the selected Landsat spectral bands and indices (NDWI, BSI, NDBI, and NDVI) and litterfall production is unclear. What ecological processes are these variables supposed to represent? Without a clear explanation, the selection seems rather generic and could equally be applied to many other ecological variables.

8. Line 180. How were the training and testing samples divided? Were observations from the same sites included in both sets? Given the spatial autocorrelation and uneven distribution of field observations, random cross-validation may overestimate model performance. Spatially independent validation is needed. The authors should also evaluate whether the model can reproduce year-to-year variability at individual sites, rather than only spatial differences among sites.

*Ploton, P., Mortier, F., Réjou-Méchain, M., et al. (2020). Nature Communications, 11, 4540. Spatial validation reveals poor predictive performance of large-scale ecological mapping models*

9. Figure 4 mainly shows a broad latitudinal and forest-type gradient, which the authors themselves interpret in relation to forest productivity. It remains unclear whether the model provides the litterfall-specific information beyond these broad environmental and productivity gradients.
10. Figure 6 Wind speed is included in the analysis but is not clearly introduced in the Methods. What is its data source, and was it used as a model predictor or only in the subsequent analysis?