Reply to Referee #4' s comments

Title: A flux tower site attribute dataset intended for land surface modeling
No.: essd-2024-77

Shi et al. present a study on a flux tower site attribute dataset intended for land surface modeling. This dataset is very valuable for land surface modeling and beyond. Overall, the manuscript is well-written, I have some additional comments for consideration.

Thank you for your careful evaluation of this manuscript. We greatly appreciate your positive and constructive comments on our manuscript, which have significantly improved the quality of our manuscript.

All comments are addressed on a point-by-point basis below. The comments are laid out below in italicized font and specific concerns are numbered. Our response is given in normal font. The list of all related changes is given in blue text.

Comment 1 (L102): *"Picking years with a low gap-filled percentage for fluxes (latent and sensible heat) and vapor pressure deficit (VPD)." Why were these variables chosen for picking? Why not include precipitation, temperature, etc.? Please provide an explanation.*

Response1: Thank you for your questions. The quality screening in this study is based on the PLUMBER2 dataset. PLUMBER2 has already screened meteorological data, including the five key variables that have the largest influence on LSM simulations: incoming shortwave radiation, precipitation, air temperature, air humidity, and wind speed. Therefore, we did not conduct additional screening for variables such as precipitation and air temperature. We have added this information to the manuscript.

The reason for screening for VPD is that air humidity is calculated using VPD. However, the screening process of PLUMBER2 did not consider the gap-filled situation of VPD.

We selected latent heat and sensible heat for two main reasons: (1) latent and sensible heat are two of the most critical variables that need to be assessed in land-atmosphere exchange. Consequently, when the quality of latent and sensible heat is poor, we exclude all fluxes. (2) Lower quality observations for latent and sensible heat usually indicate reduced quality of other flux exchanges (e.g., carbon exchanges including GPP and respiration; friction velocities).

Origin (L177):

"The PLUMBER2 dataset got 170 sites by screening meteorological data."

Revised (L117):

"The PLUMBER2 dataset got 170 sites by screening meteorological data (including five key variables that have the largest influence on LSM simulations: incoming shortwave radiation, precipitation, air temperature, air humidity, and wind speed.)."

Comment 2: *One of our studies (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4732309) conducted land surface model evaluations at the site level using the PLUMBER2 datasets, which are very valuable. We took time to find seek additional model inputs, such as soil texture information, from the literature. This study is convenient for land surface modelers. Therefore, it would be even more valuable if it included more sites and longer period of data beyond the PLUMBER2 sites.*

Response 2: We sincerely appreciate your valuable suggestions and recognition. PLUMBER2 includes the major datasets available since the initial release of flux tower data. However, since flux tower datasets like FLUXNET2015 have not been released new versions, PLUMBER2 does not contain observational data updated in recent years. We acknowledge that these data are far from sufficient, but we believe this is a good starting point.

Based on your suggestions and those of other reviewers, we have included a call for the release of flux tower attribute data in the manuscript. The revisions are as follows:

Origin (L355):

"In land surface community, flux tower attribute data is currently not given enough attention."

Revised (L355):

"In land surface community, flux tower attribute data is currently not given enough attention. However, the site attribute data is almost as important as the flux tower observations themselves. We hope that future flux tower datasets will provide standardized site attributes."

Added (L442):

"We strongly advocate for the routine release of attribute data as part of flux tower data. Making such ancillary data more easily and routinely accessible would greatly increase the value and usability of the data."

Comment 3: *Another study of ours developed global 1km land surface parameters for earth system modeling (https://essd.copernicus.org/articles/16/2007/2024/essd-16-2007-2024.html), sharing some data sources with this study, such as PFTs classification. Combining this 1km data with the site-level study could enhance land surface modeling. For example, global 1km data could provide topography attributes for sites lacking this information. This could be discussed in the manuscript.*

Response 3: Thank you very much for your suggestions. We have considered using global topography data to supplement site-observed slope and aspect. However, given that most sites are located on flat terrain, as indicated by both site data and global data, and considering the potential for scale mismatch,

we believe that adding this data may offer limited benefit without a detailed assessment. Moreover, we observed certain discrepancies between global slope and site slope at some sites, as shown in the table below. Therefore, Therefore, we have not added global terrain data currently.

We appreciate your thoughtful comments and hope these replies satisfy you.

site	slope_global	slope_site
AT-Neu	15.7%	Flat
CH-Fru	5.2%	< 5%
CN-HaM	4.5%	Flat
DE-Bay	3.1%	10-15%
DE-Obe	3.3%	5-10%
FR-Pue	3.0%	Flat
IT-MBo	7.2%	Flat
US-Me4	9.4%	5-10%
US-NR1	4.9%	5-10%
US-SRG	3.6%	< 2%

Comment 4 (L92): *"Three global datasets were used to complement attribute data of sites lacking site-observed FVC, LAI, and soil texture." It's great to analyze the uncertainty by using global datasets to fill site data. How accurate are these global datasets? The authors could analyze the consistency between sites with complete attributes and the corresponding global datasets. This information would be valuable for readers to understand the uncertainties introduced by using global datasets.*

Response 4: We fully agree with your suggestion. Other reviewers have raised similar concerns requiring cross-checking of filled data to quantify their uncertainties.

We provide the quantification of uncertainties in the final dataset resulting from data filling. After careful consideration, the quantification of discrepancies between site data and filled data has been added to Sect. 3.2, illustrating the uncertainties of the filled data. The added information is as follows.

Add (Sect. 3.2):

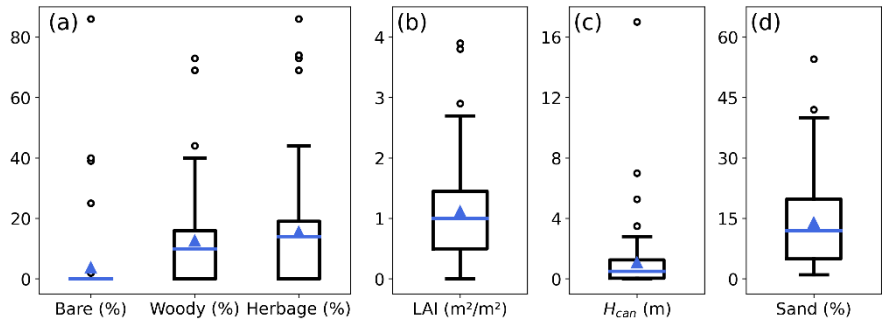

Figure 4. Quantification of discrepancies between site data and filled data for (a) PCT_PFT, (b) maximum LAI, (c) canopy height, and (d) the percentage of sand (at all sites for which both types of data sources are available). The 16 PFTs were divided into three main categories (bare soil, woody, and herbage) to be quantified separately quantification.

Add (L299):

Figure 4 quantifies the differences between site data and filled data at all sites for which both data sources are available, illustrating the inhomogeneities in the final dataset due to data filling. The differences in vegetation cover (including bare soil, woody, and herbaceous vegetation) generally fall within 20%, with a minority of sites exceeding 40%. The mean and median LAI differences are approximately 1 m^2/m^2. Canopy height deviations are primarily within 2 m, although a few sites exceed 4 m. Differences in sand content typically remain within 30%, with both mean and median differences below 15%. The quantification indicates that the filled data are generally reliable across most sites.

Comment 5 (L195): *It seems not all sites have elevation, slope, and aspect information from literature. How were these attributes obtained for sites lacking them? Seems not mentioned in the manuscript.*

Response 5: Thank you for your question. Yes, we did not find suitable global topographic data and only provided site-observed slope and aspect information in the attribute dataset.

Comment 6 (Table 2): *Should "LAI_default" be "LAI_max_default"? Similarly, for "LAI_site."*

Response 6: Thank you for your correction. This should be "LAI_max_default". We have revised the relevant description.

Comment 7 (L220): *"Resulting in six available sites. These sites were simulated to show the respective impact of different attributes in model results." This sentence is unclear. How many experiments were run for each site? This section could be written more clearly to make Section 3.3 easier to understand.*

Response 7: Thank you for your suggestion. We fully agree with you. We have provided a description of the experiments we conducted. The comparison before and after modification is as follows:

Origin (L219):

"These sites were simulated to show the respective impact of different attributes in model results. Table 2 provides an overview of the chosen sites along with their corresponding attribute information."

Revised (L219):

"Table 2 provides an overview of the selected sites along with their corresponding attribute information. Each site was simulated three times: 1) using site data for each attribute at each site, 2) using default data for each attribute at each site, and 3) using default data for the corresponding attribute at sites selected for each attribute separately, while maintaining site data for the remaining attributes. The comparison between simulations (1) and (3) aims to demonstrate the individual impact of each attribute, while the comparison between simulations (1) and (2) shows the combined impact of all four attributes."

We would like to thank you for your professional review work, constructive comments, and valuable

suggestions on our manuscript. We hope the correction made will meet with approval. These comments and suggestions have significantly improved the quality of our manuscript. Once again, thank you very much for the comments and suggestions.