Response to the RC2

RC2: 'Comment on essd-2024-9'

The paper by Ma et al. focuses on generating in situ records relating to land-atmosphere interactions through an integrated observations network across the Tibetan Plateau. This work is immensely important for understanding the behavior of atmospheric boundary layer across various landscapes over the Tibetan Plateau, where site observations are notably scarce. Moreover, those measurements can be used for calibrating and assessing land surface models and remote sensing observations. The following comments warrant attention.

Response: Special thanks to you for these insightful comments. In the revised version of the manuscript, we have addressed essentially all the points raised. We hope that the modification made on the revised manuscript will cover the reviewer expectation. We appreciate the positive comments highlighting the contributions of our work. Our ongoing goals are to guarantee the accessibility and accuracy of these field observational data and to offer solid data support for the study of climate change and its environmental effects of the Tibetan Plateau. The revised contents are highlighted in blue in the following responses, corresponding changes are marked in red in the revised manuscript.

1. Abstract needs to be concise. The first two sentences had provided background information, please delete the sentence 'The TP is recognized … with diverse landscape'. Remove the content 'Scientific data sharing is critical for the TP … they bring about' into main text. Include more information about which kind of variable you are going to provide and temporal extent.

Response: It is really true as Reviewer suggested that the language in the Abstract is not concise enough. We have made a thorough modification to the abstract to improve it. The revised abstract is listed as follows. We gratefully appreciate for your advice.

Abstract: The climate on the Tibetan Plateau (TP) has experienced substantial changes in recent decades as a result of its susceptibility to global climate change. The changes observed across the TP are closely associated with regional land-atmosphere interactions. Current models and satellites struggle to accurately depict the interactions, critical field observations on land-atmosphere interactions here therefore provide necessitate independent validation data and fine-scale process insights for constraining reanalysis products, remote sensing retrievals, and land surface model parameterizations. Scientific data sharing is crucial for the TP since in-situ observations are rarely available in this harsh condition. However, field observations are currently dispersed among individuals or groups and have not yet been integrated for comprehensive analysis. This has prevented a better understanding of the interactions, the unprecedented changes they generate, and the substantial ecological and environmental consequences they bring about. In this study, we collaborated with different agencies and organizations to present a comprehensive dataset for hourly measurements of surface energy balance components, soil hydrothermal properties, and near-surface micrometeorological conditions spanning up to 17 years (2005-2021). This dataset, derived from 12 field stations covering a variety of typical TP landscapes, provides the most extensive in-situ observation data available for studying land-atmosphere interactions on the TP to date in terms of both spatial coverage and duration. Three categories of observations are provided in this dataset: meteorological gradient data (Met), soil hydrothermal data (Soil), and turbulent flux (Flux). To assure data quality, a set of rigorous data processing and quality control procedures are implemented for all observation elements (e.g., wind speed and direction at different height) in this dataset. The operational workflow and procedures are individually tailored to the varied types of elements at each station, including automated error screening, manual inspection, diagnostic checking, adjustments, and quality flagging. The hourly raw data series, the quality-assured data, and supplementary information including data integrity and the percentage of correct data on a monthly scale are provided via the National Tibetan Plateau Data Center (https://doi.org/10.11888/Atmos.tpdc.300977, Ma et al., 2023). The present dataset provides the benchmark constraints needed to evaluate and refine the land surface models, reanalysis products, and remote sensing retrievals. It can also be used to characterize fine-scale land-atmosphere interaction processes of the TP, as well as underlying influence mechanisms.

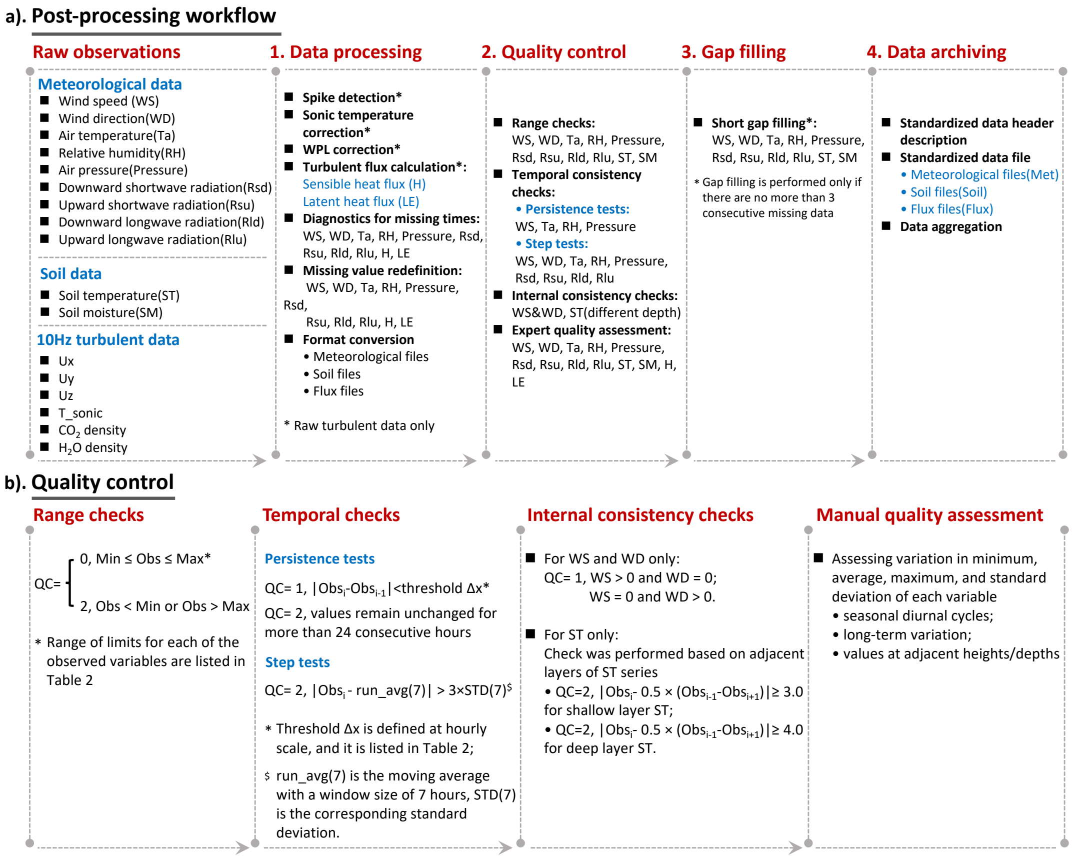
2. Section 2.1 and 2.2: Please provide a table in which each row represents one site and each column include one unique information. Then please provide the site name, location, climate, landscape type, installation of infrastructure, and measuring variables. If it is too large. It would be OK to provide two tables. One for basic information and another for introducing infrastructure installation, managing period, and measuring variables. Please provide as much details as you can for publishing a data paper.

Response: We agree with the reviewer in this regard, and we also believe that a data descriptor paper should include the table to facilitate the data users grasp the dataset as soon as possible. In fact, we have taken this into account when preparing the first draft of this manuscript, the Table 1 (provided after the references) presents not only the basic information about each station (e.g., latitude, longitude, elevation, landscape type), but also the observation infrastructure (sensor model, manufacture, height, units, and observing period of each variable). The design of this table was informed by several previous papers that were also published in the ESSD. We apologize for placing this table after the references because it was too long, so you may not have noticed it.

3. Section '2.3 Data post-processing workflow' needs further improvements.
(1) Figure 2: The information provide in this figure is a little bit general. It should be a summary of section 2.3.1 to 2.3.4. (i) We need to know the specific variables you are working on. (ii) Are you using those data processing approach for all variables? (iii) In the four modules, are you consistently applied these processing approaches to each variable and each site? I highly recommend that the author refer to previously published ESSD or other high-quality data papers and redesign the flowchart accordingly. I have provided the following paper for reference. Please note that there is no need to cite them.
Gebrechorkos, S. H., Peng, J., Dyer, E., Miralles, D. G., Vicente-Serrano, S. M., Funk, C., . . . Dadson, S. J. (2023). Global high-resolution drought indices for 1981–2022. Earth Syst. Sci. Data, 15(12), 5449-5466.
Pastorello, G., Trotta, C., Canfora, E. et al. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci Data 7, 225 (2020).
Beck, H. E., E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R.

F. Adler, 2019: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. Bull. Amer. Meteor. Soc., 100, 473–500.

Response: Thank you very much for the above suggestions. Following your suggestion, we have redesigned the Figure 2a (flowchart of data post-processing workflow), the new figure is listed as follows. Besides, the Figure 2b was added to clearly show relevant information used in the data quality control procedures.



(2) Section 2.3.1 to 2.3.4 require much more details: (i) Please list the relevant methods (equation, models, quantification metrics, etc) you used where are applicable. (ii) Definition of missing data should be quantified for each variable and each site. (iii) Provide a detailed description of the data header file format. Overall, this part is very important and much more details should be provided.

Response: (i) We have carefully addressed all the reviewer's concern about the description of the data quality control. The Figure 2b was added to list all the relevant information (equation, metrics, and threshold values) used in data quality control. Figure 2b is a further extension of the Figure 2a, and we think the revised Figure 2 can give the reader a systematic and in-depth understanding of the entire data post-processing process applied in this work.

(ii) We apologize for any misunderstanding you may have experienced. Instead of using NAN or -9999 to indicate missing data as they were in the raw data, we intended to use 9999.9. To prevent misunderstanding, we thus replace "Definition of missing data" to "Missing value assignment" in the Data control step in Figure 2. Thanks very much for your understanding.

(iii) Based on your suggestion, we have modified the description of the data header. In the revised manuscript, we described the naming format of data header in Section 2.3.4 Data archiving and Section 5 Data availability. The revised description is listed as follows:

- Section 2.3.4 (Line 349-351): During the archiving step, the header descriptions of the output files were first standardized to include information about the variable name, height/depth, and units. This information was expressed in the following format: variable_height/depth(units). The variable names are expressed as abbreviations which are listed in Appendix A.
- Section 5(Line 553-555): The data header of each CSV-formatted data file contains comprehensive information on the variable units and heights/depths, with naming format: variable_height/depth (units). The variable names are expressed as abbreviations. Appendix A provides a full list of the abbreviations for each variable.

4. Section 3 Data description: Much more details should be provided. Provide a table and listed all those variables this data set will provide. Indicate availability of each variable at a specific site. Provide unit for each variable and start date and end date (if applicable). The primary principle is assisting the data user quickly know how those valuable measurements fit their research.

Response: Thanks very much for pointing this out. We do understand your concern regarding the data description. Table 1 (apologize once again) summarized all the variables that included in the dataset, providing the necessary information of each variable to show the units of the variables, the heights and periods these variables observed, the models and manufactures these sensors used. Furthermore, the available period of each variable can be clearly observed from the Figure B1-B3 which were shown as Appendix. We think this information can help the data users identify which variables they need.

5. Section 4: it would be great if the authors can provide some application cases.

Response: This suggestion is highly appreciated. We searched for recent studies that directly using in-situ observations from the stations included in the current dataset. As examples of the application cases of the in-situ observations, we have selected a few representative examples of studies in the fields of fine-scale interactions analysis, model representation assessment, model development, remote sensing algorithm refinement, and key land surface parameter estimation. The updated content is listed as follows with newly added sentences highlighted in blue.

- **Fine-scale interactions analysis**
  Line 525-529: More specifically, targeted field campaigns across the vast grasslands and permafrost zones are indispensable for capturing the fine-scale interactions between the changing land surface and the overlying atmosphere. This is exemplified by the studies conducted by Li et al., (2015) and Wang et al., (2019), who investigated the lake-atmosphere interactions using in-situ observations from the Ngoring Lake and NAMORS station, respectively.
- **Model representation assessment**
  Line 529-532: It is possible to systematically verify model representations of hydrological and

thermal processes, as well as their interconnections, at various sites using this enhanced monitoring network. The work done by Liu et al. (2018), which evaluated the effectiveness of the WRF model in snowfall simulation using in-situ measurements, serves as an example of this.

- Model development

  Line 532-535: It will serve as pillars for improving model physics concerning cryospheric, hydrologic, and atmospheric processes in the intricate TP terrain. An example of this is the study done in 2013 by Chen et al., whereby a DEM-based radiation model was developed for an accurate estimation of instantaneous clear sky solar radiation using measurements from the QOMS station.

- Remote sensing algorithm refinement

  Line 535-538: Meanwhile, comprehensive field measurements are crucially needed to validate, calibrate, and refine remote sensing retrieval algorithms over the topographically complex terrain. For instance, Yuan et al., (2021) used in-situ measurements from this dataset to present an optimized canopy transpiration model and an improved technique for calculating soil evaporation with soil moisture and texture.

- Key land surface parameter estimation

  Line 538-541: Systemic biases in key land surface parameters in the reanalysis products can be decreased by incorporating synthesized ground-based datasets and revised satellite products through sophisticated data assimilation techniques. For instance, Qi et al., (2023) increased the accuracy of land surface temperature retrieval over the TP based on the in-situ data.