

Referee Report on ESSD-2026-242

TReAD: 40-year Taiwan Reanalysis Downscaling Dataset at 2 km

General Assessment

This manuscript presents the Taiwan Reanalysis Downscaling Dataset (TReAD), a 2-km, 40-year dynamically downscaled regional climate dataset produced with the WRF model. The dataset includes six Essential Climate Variables and is evaluated against station observations, gridded datasets and satellite-derived products. It addresses an important gap in long-term, high-resolution climate information for Taiwan, where complex topography limits the applicability of coarse-resolution reanalysis products.

The manuscript is generally well organised and presents a valuable dataset of broad interest to the climate, hydrology, ecology, disaster management and climate-impact research communities. The study is well within the scope of ESSD. The manuscript would benefit from clearer documentation and improved data accessibility, as well as a more transparent discussion of the dataset's limitations and recommended uses.

I recommend publication after minor revision, provided that these points are adequately addressed. Detailed comments are provided below.

Major Comments

1. Although the model configuration is broadly described, the manuscript does not provide enough detail for a third party to reproduce the simulations. Please provide further clarification on:

- A complete inventory of the external forcing, configuration and preprocessing inputs.
- A more detailed explanation of the nesting, nudging and blending methods, how they were implemented? and how the relevant parameters were specified?
- The preprocessing workflow used to generate the initial and lateral boundary conditions.

If configuration files, preprocessing scripts or workflow documentation can be made publicly available, this would substantially improve reproducibility and long-term usability. I encourage the authors to provide such material.

2. The Data Availability section should be expanded to provide clearer information for potential users. I recommend including a summary table of all variables, units, temporal and spatial resolutions, either in the main text or as a supplementary table. This would greatly improve the dataset's usability.

- The manuscript states that monthly products are available. Could the sub-daily outputs also be made available?

3. Several validation details require clarification

- How missing observational records were handled in the calculation of validation statistics? Were days with missing data excluded pairwise?
- Please clarify the elevation information used for the temperature adjustment and discuss the remaining representativeness mismatch between point observations and TReAD grid-cell means. This issue is particularly relevant in complex terrain, where horizontal and vertical sub-grid variability may not be fully addressed by a simple lapse-rate correction.
- Is the 6h spin-up in each re-initialisation sufficient for the land surface model to reach equilibrium, particularly for variables with longer memory such as soil moisture? Please provide a justification for retaining only the second 6 hours of each 12-hour simulation. In particular, please discuss whether the 6-hour adjustment period is expected to be adequate for variables influenced by land-surface memory. If this was not explicitly tested, this should be acknowledged as a potential limitation rather than necessarily requiring additional sensitivity experiments.

4. The manuscript would benefit from a dedicated subsection summarising dataset limitations and sources of uncertainty relevant to potential users.

Please discuss, where relevant:

- Limitations inherited from the ERA5 forcing.
- Uncertainty associated with model configuration and physical parameterisations.
- Potential limitations associated with the representation of land-use change.
- Uncertainty in the observational and gridded reference datasets used for validation, particularly in mountainous regions.
- Limitations associated with comparisons between point observations and grid-cell averages.

I also recommend adding concise user guidance addressing:

- Applications and regions for which TReAD is particularly well suited.
- Regions or conditions where users should have particular caution.
- Variables for which systematic biases remain and users may wish to undertake application-specific bias assessment or bias correction.
- Known limitations at high elevations and over small or remote islands.

The purpose should be to help future users understand the appropriate and inappropriate uses of the dataset.

Minor Comments

1. The manuscript states that 17 stations on the plain and 5 mountainous stations were used, giving a total of 22 stations. However, Table S1 appears to contain only 20 stations. Please verify the station inventory and ensure that the manuscript, figures, tables, supplementary material and validation statistics consistently report the same number of stations.

2. The dataset is described as covering 1980-2019 and most of the validation uses this period. However, some analyses and figure captions refer to 1980-2018, including the comparison of annual temperature time series and the spatial precipitation analysis (L255 and Fig.8). Please explain why 2019 is excluded from these/many analyses and ensure that the temporal coverage is stated consistently throughout the manuscript, figures, tables, captions and supplementary material.

3. The manuscript uses several datasets, including CWA observations, TCCIP data and Himawari-8-derived radiation data. Although the general sources are described, a concise summary table would make the validation framework considerably clearer. For each reference dataset, I suggest including: dataset name, variables used, spatial resolution, temporal resolution, period used in the present study, processing, interpolation, or elevation adjustment applied, purpose in the validation, relevant information on independence from ERA5/TReAD, where applicable.

4. The use of Typhoon Morakot as a precipitation case study is useful. However, please explain the rationale for selecting this event. For example, was it selected because of its severity, scientific importance, observational coverage or another reason? A broader multi-storm evaluation is not necessarily required. However, if TReAD has not been systematically evaluated for precipitation associated with typhoons having different tracks or intensities, this limitation should be acknowledged so that users do not overgeneralise the performance demonstrated for Morakot.

Suggested but non-essential additions

1. The current validation is broad and already includes several metrics and comparisons. Nevertheless, the quantitative description of dataset performance could be further strengthened by one or two additional diagnostics, such as:

- Mean Absolute Error
- Seasonal bias statistics, given the seasonal variations noted for temperature and radiation.
- Percentile-based error analyses, for example errors for extreme percentiles vs. moderate conditions.
- Probability distribution comparisons using quantile-quantile plots.

These are suggestions rather than mandatory requirements and I do not consider it necessary to include all of them.

2. If available, a brief estimate of the computational resources and storage requirements associated with producing the dataset would be useful for researchers interested in reproducing or extending the production workflow.

Editorial Comments

- The manuscript requires careful English-language revision. Several sentences are unnecessarily long or awkwardly phrased. A thorough language revision is recommended, particularly in the Introduction and Conclusion.
- Wind speed/velocity: Please clarify whether the manuscript refers to wind speed or wind velocity, as these terms have different scientific meanings. Use the appropriate term consistently throughout the manuscript, particularly in the Abstract and Section 3.5.
- Please ensure that all figures include descriptive titles and appropriate units for RMSE and other relevant metrics. Use consistent font sizes across all panels and remove redundant phrasing from the figure captions.
- Fig. 1: Increase the font size of the legends and axis labels for better readability.
- Fig. 4: Improve readability by standardizing the font size to match that used in Fig. 3.
- Figs. 6, 10, 16, 20, 24: Increase the line thickness in these figures for better visual readability.
- Fig. 7: Add units to the colorbar and include a descriptive title for the figure.
- Fig. 8: Verify and correct the time range shown (currently noted as 1980-2016) to align with the rest of the manuscript.
- Fig. 22: Remove the informal abbreviation "ie" from the figure or its label.
- Fig. 26: Change the y-axis label from "standardized deviations (Normalized)" to "Normalised standard deviation".

- L45: "Taiwan where its complexity of topography is high" for improved clarity. Rephrase, maybe to "Taiwan, where the topography is highly complex"
- L66-67: Shorten the sentence "Such topography makes regional-scale climate analysis and application studies more difficult" for conciseness.
- L73, L74, L77, L78: Avoid the repeated use of "dynamically downscaling" in these lines.
- L81: Clarify the meaning of "aforementioned logics"; consider replacing it with a more precise term such as "methods" or "approaches".
- L91: Specify the location of the two nested domains within Figure 2, or ensure they are clearly identifiable in the figure itself.
- L120 and L356: Define the abbreviation "UHI" at its first mention in L120 (where "heat island effect" appears), rather than introducing it later in L356.
- L164-165: "(see Table S1 for the details of the 22 stations including longitude, latitude and elevations)": this parenthetical is redundant because Table S1 already provides those details. Consider shortening it to "(Table S1)" or "(see Table S1)".
- L166: "It is worthy mentioning" change to "It is worth mentioning".
- L224-225: Correct the figure reference, the "mean biases" discussed here refer to Fig. 7, not Fig. 5.
- L231-232: "when evaluating ... separately (not shown here)". Could the authors include these results for completeness? Alternatively, this statement could be removed if the results are not essential to the discussion.
- L249: Remove the qualifier "relatively weak" when describing the overestimation, as this characterization is not supported by the data.
- L313-314: "to be better ... in 2009": Reformulate, make it shorter and clearer.
- L396-397: Rephrase the statement "since radiation ... weather systems" and add appropriate supporting references.
- L470-473: Rephrase the entire passage for improved clarity and conciseness.
- L480: "show varying results": change to "show varying performance".
- L582: "with cautions": change to "with caution".