Anonymous Referee #2

Main Comments:

1. This study presents a highly valuable global GNSS climate data record derived from long-term GPS observations. The manuscript is well-structured, and the dataset is of substantial relevance to the meteorological, climate, and even geodetic communities. However, several aspects would benefit from clarification and technical refinement to enhance reproducibility and overall clarity. I would like to recommend the manuscript for publication after minor revisions addressing the following points

Response: We appreciate your thorough review and insightful comments. Our point-by-point responses to each comment are provided below.

Minor Comments:

1. The current title, "A comprehensive 22-year global GNSS climate data record from 5085 stations," may lead readers to assume that each of the 5085 stations have a continuous 22-year time series. To better reflect the actual structure of the dataset, I would like to suggest refining the title by including a phrase such as "covering up to 22 years" or "spanning up to 22 years". Similarly, the first sentence of the abstract and other related statements should be revised to accurately characterize the temporal extent and variability of the dataset across stations.

Response: We totally agree with your constructive suggestion, and the original title could indeed lead to unnecessary misleading and misunderstanding regarding the length of the dataset. To avoid this ambiguity, we have revised the title to:

"A global GNSS climate data record from 5085 stations spanning up to 22 years".

In addition, we have reviewed the whole manuscript to remove wording that might imply uniform 22-year coverage (such as in the Abstract, Introduction, and Section 2.1), and we now consistently state that the record lengths vary, using the phrase "***spanning up to a 22-year period***" throughout.

2. Line 11: The acronym GPAC is introduced in the manuscript without explanation. Please define this acronym upon its first use, whether it refers to your team, processing framework, or another entity, to ensure clarity for readers. Also, consider removing those unnecessary abbreviations in the abstract to maintain focus and readability.

Response: Many thanks for pointing these out.

First, GPAC actually represents our joint research centre. In the revised version, we spell out and define the acronym at first mention in the main text:

Lines 112-113

"*This reprocessing campaign, **led by the GNSS data processing for Positioning, Atmosphere, and Climate research centre (GPAC) and hereinafter referred to as "GPAC-Repro"**, adopted precise satellite orbit, xxxxxx*"

Regarding the Abstract, we have removed unnecessary abbreviations and refined these sentences:

Lines 10-12

"*This work presents a comprehensive global GNSS climate data record derived from 5085 stations, spanning up to a 22-year period 2000–2021. The dataset was generated using the state-of-the-art processing methodologies and precise products from the*

International GNSS Service (IGS) Repro-3 initiative."

Lines 13-16

"*A rigorous data screening and quality assessment framework was implemented, including formal error detection, offset identification, and extensive cross-validation with ERA5 reanalysis dataset, radiosonde profiles, and Very Long Baseline Interferometry measurements*"

3. Line 50: Given the importance of this methodological transition in the context of GNSS atmospheric monitoring, I suggest introducing a paragraph break. This will improve the logical flow by separating the historical context from the introduction of GNSS techniques.

Response: As per your suggestion, we have separated the paragraph into two. Thanks for helping us improve the quality of the manuscript.

4. Line 62: The current categorization of detection models may be misleading. AI-empowered methods, while modern, are still broadly encompassed within statistical methodologies. I recommend grouping the models into two categories of statistical and numerical and mentioning AI-empowered methods as a subcategory under statistical models for clarity and conceptual consistency.

Response: Thanks for this insightful advice. Basically, AI-empowered approaches can indeed be subsumed within statistical methodologies given their underlying mechanisms. We now revise the text to group detection models into two categories:

Lines 62-64

"*In recent years, the innovative utilisation of GNSS-derived ZTD and PWV estimates has spurred the development of* ***statistical (including artificial intelligence-empowered) and numerical approaches*** *for nowcasting and very short-range forecasting of weather extremes, such as heavy precipitation and tropical cyclones.*"

5. Line 145: The manuscript indicates the use of the Bernese GNSS Software (V5.2) for data processing, but no reference is provided. Please include an or a few appropriate citations to support reproducibility.

Response: Thanks for your reminder. A new reference has been added:

Lines 151-152

"*Advanced modelling and correction techniques were implemented using Bernese GNSS Software* ***Version 5.2 (Dach et al., 2015)****, incorporating the latest updates to enhance accuracy*"

References used here:

[1] Dach, R., Lutz, S., Walser, P., and Fridez, P.: Bernese GNSS software version 5.2. Astronomical Institute, University of Bern, 858, doi:10.7892/boris.72297, 2015

6. Table 1: Considering that the newer VMF3 has been released and is widely reported to offer improved accuracy over VMF1, please clarify the rationale for using VMF1 in this study.

Response: Thanks for raising this point. We fully acknowledge that the newer VMF3 generally improves the approximation of ray-traced slant delays, particularly at low elevations, relative to VMF1 (Landskron and Böhm, 2018). However, for this reprocessing campaign (initiated in 2017),

we intentionally adopted VMF1 to ensure homogeneity and reproducibility across 2000–2021 and to remain consistent with the IGS Repro3 inputs and the Bernese software V5.2. VMF1 remains widely used in high-accuracy geodetic processing and provides well-established global grids for the entire study period (Dach et al., 2021). In addition, we also noted in some previous studies that the benefits of VMF3 are nuanced, with mixed impacts on wet-mapping performance relevant to PWV retrieval (Feng et al., 2020).

To further refine this work, as mentioned in Section 7, our planned reprocessing campaign will incorporate multi-GNSS observations using Bernese V5.4 and will evaluate updated tropospheric models such as VMF3 to further improve the quality of the dataset.

Lines 695-697

> "*Given this, our ongoing research is conducting a new reprocessing campaign that will incorporate multi-GNSS observations using the latest Bernese V5.4* ***and updated tropospheric models like VMF3****, while managing inter-system and inter-frequency biases, harmonising antenna calibrations and metadata, and ensuring cross-system consistency.*"

References used here:

[1] Dach, R., Selmke, I., Villiger, A., Arnold, D., Prange, L., Schaer, S., Sidorov, D., Stebler, P., Jäggi, A., and Hugentobler, U.: Review of recent GNSS modelling improvements based on CODEs Repro3 contribution. Advances in space research, 68(3), 1263-1280, doi:10.1016/j.asr.2021.04.046, 2021.

[2] Feng, P., Li, F., Yan, J., Zhang, F., and Barriot, J. P.: Assessment of the accuracy of the Saastamoinen model and VMF1/VMF3 mapping functions with respect to ray-tracing from radiosonde data in the framework of GNSS meteorology. Remote Sensing, 12(20), 3337, 2020.

[3] Landskron, D., and Böhm, J.: VMF3/GPT3: refined discrete and empirical troposphere mapping functions. Journal of geodesy, 92(4), 349-360, 2018.

7. Table 1: The study relies on reprocessed orbit and clock products from the CODE analysis center. Given that several other IGS analysis centers also produce high-quality reprocessed solutions, it would strengthen the manuscript to briefly explain why CODE products were preferred for this work.

Response: In this work, the reasons for choosing the CODE products are as follows:

First, CODE products provide a complete and consistent suite of inputs for the entire study period, which avoids cross-centre inconsistencies and supports long-term homogeneity. In other words, to keep the dataset homogeneous, we used a single source in this release, despite other IGS Analysis Centres also producing high-quality reprocessed solutions.

Second, the CODE products are aligned with the Bernese software V5.2 and parameterisation, which reduces configuration uncertainty and improves reproducibility.

Third, these products are widely employed and also well-validated within the community, making it a reliable baseline for global climate-scale reprocessing.

8. How did you handle the station coordinates in your processing? Were the coordinates estimated simultaneously along with the tropospheric parameters and other parameters, or were a priori coordinates introduced and kept fixed during the GNSS data processing?

Response: Thanks for pointing this out. In this reprocessing, station coordinates were estimated simultaneously with the tropospheric parameters. A priori coordinates from the weekly combined solutions were taken as initial values and applied as tight constraints, which were not fixed during

the processing. Coordinates were treated as static over each 27-hour session, while receiver clocks were estimated epoch-wise.

9. Lines 177 and 190: Some units for physical constants are presented in a format inconsistent with SI or typographic standards. Please carefully review the entire manuscript to ensure uniformity in unit notation, including consistent use of spaces between values and units.

Response: Thanks for your kind reminder. We have gone through the whole manuscript, and all the issues occurred in the original manuscript have now been revised/refined.

10. Figures 7, 12 and 14: While the figures in the manuscript are generally informative and diverse, the legends and axis labels in Figures 7, 12, and 14 are too small and difficult to read. Please consider increasing font sizes and improving color contrasts to enhance readability, especially for printed versions.

Response: Thanks for your comment. In the revised version, almost all the figures, including Figs. 7 and 12, have been replotted or further refined to ensure readability. Regarding Fig. 14, as per the other reviewer's suggestion "*Everyone interested for sure knows the topography of Hawaii and the Andes*", we have removed it from the manuscript and updated the figure numbering throughout to reflect this change. Here are the updated figures 7 and 12:

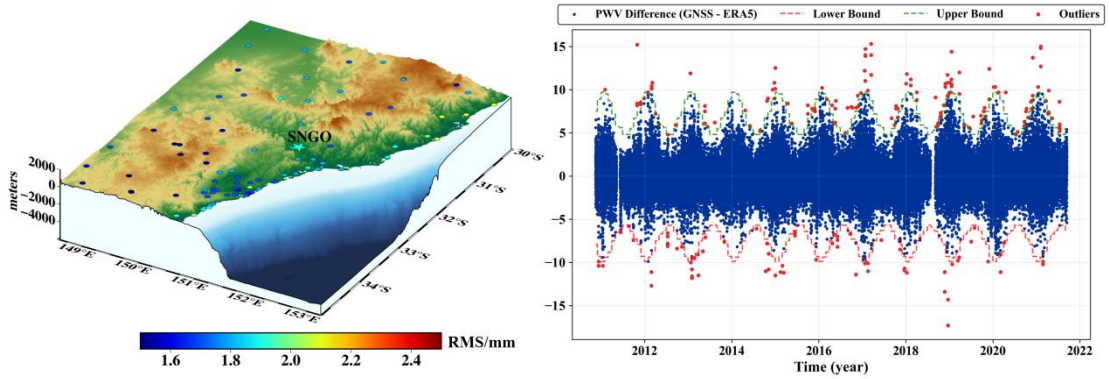

Figure 7. Identification of nearby stations for SNGO (a) and time series of PWV differences with threshold limits (b)

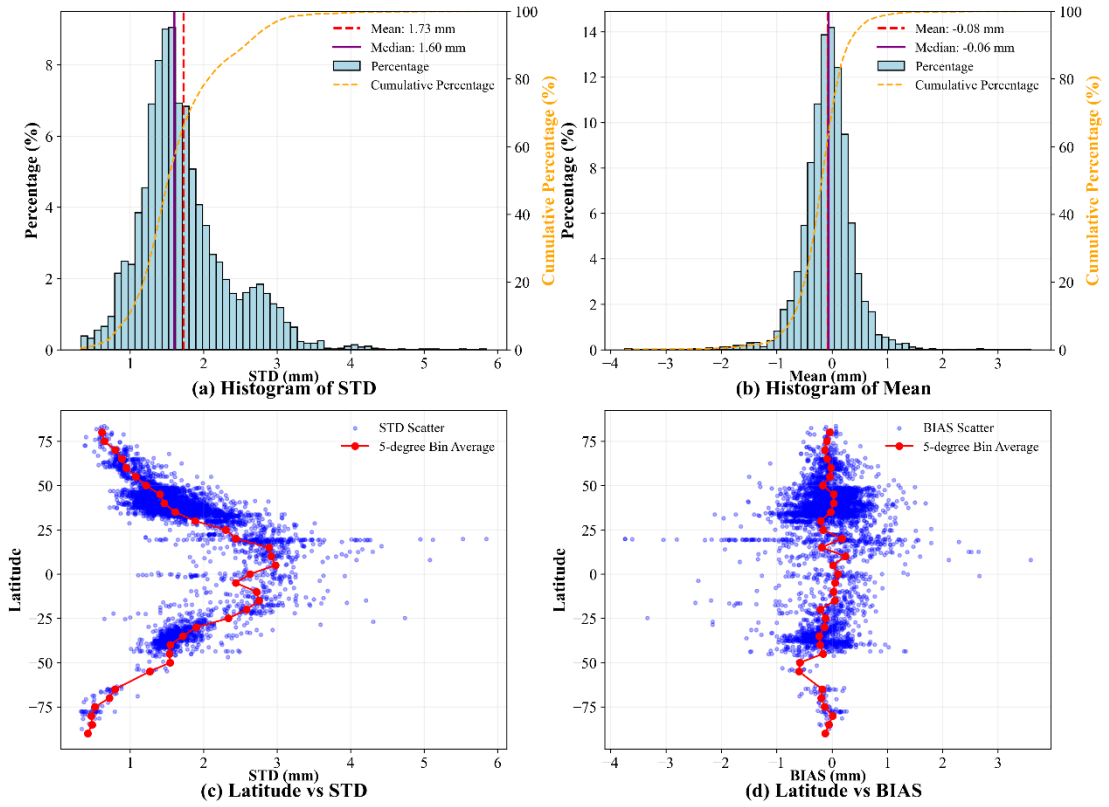

Figure 12. The mean and standard deviation of differences in PWV between GPS and ERA5

11. Line 521: When discussing monthly variations in PWV, it would be helpful to mention the number of hourly samples per month (e.g., range from A to B), as the number of days in a month varies and this affects statistical interpretations.

Response: Thanks for your valuable suggestion. We now quantify the completeness thresholds for both daily and monthly periods and add a new sentence in the revised version:

Lines 565-568

"*To minimise the impact of missing data on the analysis, we applied a strict inclusion criterion, i.e., only days with at least 21 hourly estimates and months with a minimum of 650 hourly samples were included in the calculation, corresponding to* ***at least 87.5 % (21 out of 24 hours) daily completeness and 87.4–96.7 % (650 of 672–744 hours) monthly completeness, nominally approximate 90 %.***"

12. Figure 19: The current subfigure titles in Figure 19 are somewhat generic, with panels (a) and (c) and panels (b) and (d) labeled identically. I would like to suggest updating the titles to more clearly differentiate between metrics and highlight their specific content.

Response: Thanks for your valuable suggestion, we have replotted this figure in the revise version:

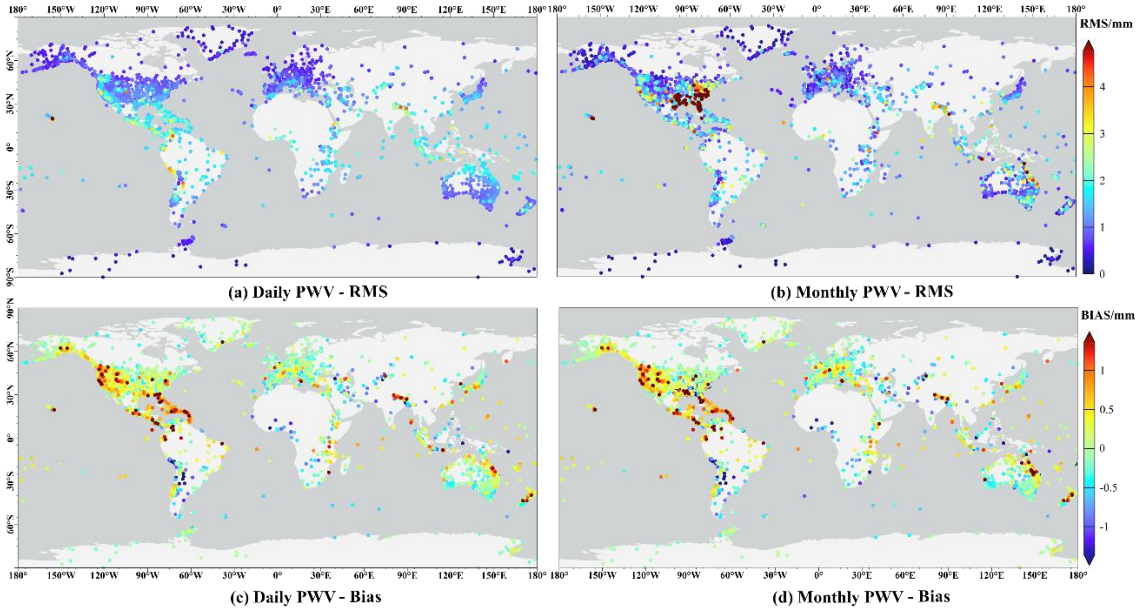

Figure 19. RMS (a, b) and Bias (c, d) statistics resulting from the comparison of daily and monthly mean PWV at all the stations against ERA5 dataset over the whole study period.

13. Line 624: As a gentle reminder, if the DOI for the dataset deposited in the PANGAEA data repository is available, please directly include it in the revised version. This will improve accessibility and ensure the completeness of your data publication.

Response: Thanks for your reminder. We have updated the Data Availability section in the revised version, in which the DOI of our submitted dataset has been provided:

Lines 673-675

"The global GNSS climate data record, including hourly ZTD and PWV estimates, ***described in this work is now available at: https://doi.org/10.1594/PANGAEA.982476 (Wang et al., 2025).*** *Additionally, the datasets have also been made accessible at: https://www.gnss.studio/Login, with its data download interface shown in Fig. 22."*