Responses to Referee #1

https://doi.org/10.5194/essd-2022-274-RC1

This paper is to present a high-quality global Integrated Water Vapour (IWV) dataset from 12,552 ground-based Global Positioning System (GPS) stations in 2020. Although the IWV data is useful, while it has no new methods or IWV improvement. Furthermore, it did not show new performances or achievements in weather prediction. Therefore it is against publication in the current form.

Reply: Thanks a lot for your affirmation on the usefulness of the dataset developed by this work. Our replies are shown with blue Arial font type. The related texts copied from the revised manuscript are shown with green Times New Roman font type.

Regarding your comment *of "it did not show new performances or achievements in weather prediction"*, it is noteworthy that this manuscript is a **data description paper** submitted to **Earth System Science Data (ESSD)** for the Special Issue of "**Analysis of atmospheric water vapour observations and their uncertainties for climate applications**". The aims and scopes of ESSD and the associated Special Issue are copied as follows:

Aims and scopes of ESSD

https://www.earth-system-science-data.net/about/aims_and_scope.html

*Earth System Science Data (ESSD) is an international, interdisciplinary journal for the publication of articles on **original research data (sets)**, furthering the reuse of high-quality data of benefit to Earth system sciences. The editors encourage submissions on **original data or data collections** which are of sufficient quality and have the potential to contribute to these aims.*

About this journal

*This journal aims to establish a new subject of publication: to **publish data** according to the conventional fashion of publishing articles, applying the established principles of quality assessment through peer review to data sets.*

Manuscript types of ESSD

https://www.earth-system-science-data.net/about/manuscript_types.html):

Data description papers

*These may **describe original research data, databases**, or combined datasets derived from them.*

ESSD understands original research data as data generated by observation of the Earth system. In most cases this excludes reprocessing, postprocessing, or reanalysis of model outputs (such as climate model outputs, environmental modelling outputs, output of classification algorithms, and so forth). For the most part, ESSD does not publish model-based forecasts or projections. Articles may describe observational data extended or verified with model output or other combinations of model output and observational data produced for specific purposes or specific regions where observations remain too sparse or otherwise too limited. In all cases users must discover a unique accessible verifiable data product.

Although examples of data outcomes may prove necessary to demonstrate data quality, extensive interpretations of data – i.e. detailed analysis as an author might report in a research article – remain outside the scope of this data journal. ESSD data descriptions should instead highlight and emphasize the quality, usability, and accessibility of the dataset, database, or other data product and should describe extensive carefully prepared metadata and file structures at the data repository.

Articles in this category should not focus on instrumentation, methodology, data extraction, or data treatment except when that information helps quantify uncertainties or otherwise facilitates validation of data presented.

Please carefully read the ESSD data policy and editorial if unsure what kind of data are within the scope of the journal.

Aim of the special issue

https://essd.copernicus.org/articles/special_issue10_1118.html

The absorption of energy by water vapour keeps the average temperature on Earth at a level that makes life possible. However, the average temperature increased in recent decades due to increased greenhouse gas emissions. This anthropogenic increase is amplified by the water vapour feedback. Water vapour is further an important component of the water cycle and indirectly of the energy cycle. The role of water vapour in these cycles in a changing climate needs to be fully understood and quantified. Thus, ***high-quality observations of water vapour*** *and the characterisation and validation of associated uncertainties are of high importance for climate applications.* ***Topics covered around atmospheric water vapour include, among others, retrieval development, data records, uncertainty characterisation and validation, inter-comparisons, climate applications, and model validation.*** *This special issue originates from the GEWEX Water Vapor Assessment (G-VAP, http://gewex-vap.org) and is open for all submissions within scope, not just from G-VAP.*

According to the scopes and aims of ESSD and the Special Issue, this work aims to overcome the limitations of previous GPS IWV datasets in quantity and quality, rather than achievements in weather prediction.

Regarding your comment of *"Although the IWV data is useful, while it has no new methods or IWV improvement"*, we cannot agree with it, because the dataset did show achievements in quantity as well as in quality.

This work provided an enhanced global dataset with 12,552 GPS stations through the collaboration with Nevada Geodetic Laboratory (NGL), USA, whereas most existing GPS Integrated Water Vapour (IWV) datasets only include several hundred stations. Moreover, compared to NGL's operational GPS IWV product, the quality of the GPS IWV dataset is enhanced by employing accurate meteorological information from ERA5 for the GPS IWV retrieval with a significantly higher spatiotemporal resolution. A dedicated data screening algorithm is also implemented.

The status, challenges, and the aims of this work are stated in Section 1 Introduction and copied as follows:

However, the insufficient spatial densities and coverages of current available GPS IWV datasets severely limit their capabilities in sensing Earth system processes over various spatial scales. Although approximately ten thousand of continuous dual-frequency GPS stations are currently in operation around the world, most existing regional and global GPS IWV datasets only include several hundred continuous stations or campaign measurements spanning less than one year, such as IGS (Wang et al., 2007; Heise 2009), EUREF (Pacione et al., 2017), HyMeX (Bock et al., 2016), EUREC4A (Bock et al., 2021), and GURN (Fersch et al., 2022). Combining their solutions for an integrated analysis is possible but it could be impacted by systematic differences due to various software packages and strategies used in their GPS data processing as well as approaches adopted for the conversion of IWV. Nevada Geodetic Laboratory (NGL) has taken a big step forward by integrating global GPS observations at more than 10,000 stations and providing their tropospheric ZTD and IWV estimates as well as other products (Blewitt et al., 2018). Although NGL processes the GPS observations uniformly by employing many state-of-the-art models and strategies, the accuracy of its operational IWV product is limited by the inaccurate auxiliary meteorological information used in IWV retrieval. An enhancement of NGL's operational GPS IWV product will provide more accurate IWV estimates and thus it can be more valuable for many applications, such as analysis of extreme weather and diurnal variations of IWV, and assessment of other IWV retrieval techniques.

The evaluations by using *in-situ* reference radiosonde and synoptic measurements did show the superiority of the enhanced GPS IWV product developed in this work compared to NGL's operational product. Details of the evaluations can be found in Section 5 Evaluation of enGPS versus opGPS. A short summary in the Abstract is copied as follows:

The enhancement is reached by employing accurate meteorological information from the fifth generation of European ReAnalysis (ERA5) for the GPS IWV retrieval with a significantly higher spatiotemporal resolution. A dedicated data screening algorithm is also implemented. The GPS IWV dataset has a good agreement with in-situ radiosonde observations at 182 collocated stations worldwide. The IWV biases are within ±3.0 kg m^{-2} with a Mean Absolute Bias (MAB) value of 0.69 kg m^{-2}. The standard deviations (SD) of IWV differences are no larger than 3.4 kg m^{-2}. In addition, the enhanced IWV product shows substantial improvements compared to NGL's operational version, and it is thus recommended for high-accuracy applications, such as research of extreme weather events and diurnal variations of IWV, and intercomparisons with other IWV retrieval techniques. Taking the radiosonde-derived IWV as reference, the MAB and SD of IWV differences are reduced by 19.5% and 6.2% on average, respectively. The number of unrealistic negative GPS IWV estimates are also substantially reduced by 92.4% owing to the accurate Zenith Hydrostatic Delay (ZHD) derived by ERA5.

The dataset developed by this paper has wide potentials for various applications. For example, in this manuscript, we carried out intercomparisons of the GPS IWV with six different radiosonde types, which can be useful for the evaluations on possible wet/dry biases in specific radiosonde types. Particularly, the IWV intercomparisons between GPS and the radiosonde types of RS41 and Meisei were rarely reported as far as we know, and hence these results are valuable.

In addition, a potential application of this dataset is suggested by Frank Fell from the community comment (https://doi.org/10.5194/essd-2022-274-CC1):

> *We are currently establishing a total column water vapour data record from Sentinel-3 MWR observations covering the global ice-free global ocean.* ***The dataset presented by Yuan et al. has a potential to serve as ground-truth and therefore is of significant interest to our work, especially since it contains many GPS coastal stations.***

Therefore, we believe that this work fits scope of ESSD, especially the aim of this Special Issue. The dataset has been enhanced in not only quantity but also in quality, and thus it is beneficial to the community. With the free sharing of the dataset, the "*achievements in weather prediction*" and many other potential applications can be addressed not only by us but also by other international researchers in the future. We hope you can kindly reconsider this submission as a data description paper.