Author Response to Review of

Tropospheric water vapor: A comprehensive high resolution data collection for the transnational Upper Rhine Graben region

Benjamin Fersch et al.
Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2022-57

RC: ***Reviewer Comment***, AR: *Author Response*, ☐ Manuscript text

Dear Referee,

we would like to thank you very much for taking the time to review our work. Your comments and constructive suggestions to our manuscript are highly appreciated.

1. General comments

RC: ***Water vapor is a crucial constituent of the atmosphere, not least because of its importance for severe weather events and climate change. The authors describe GNSS and InSAR datasets as input for assimilation in atmospheric models, along with the applied methods for merging. The datasets encompass the Upper Rhine Graben Region. The data are valuable and an interesting contribution for the scientific community.***

AR: *We are pleased having created a data set that is useful and valuable for the scientific community.*

RC: ***The article is not always easy to read, but I understand that this is due to the fact that different communities (GNSS, InSAR, WRF, ...) are coming here together for this joint work. Moreover, some abbreviations are not understandable at first reading. There is an appendix with the explanations, but it would be appreciated if more explanations are added in the text.***

AR: *We agree with the reviewer's impression that due to the multi-disciplinarity of the scientific content it is not easy to memorize the abbreviations and to understand the connections between the different parts. However, we think that the list of abbreviations given in the Appendix is the best way to have a central place for lookup. Otherwise, we would ask you for a detailed list of abbreviations that are unclear so that we can improve their definitions in the text.*

2. Specific comments

RC: ***line 12: What is meant with 2.5 mm global mean water equivalent? Average precipitable water? If yes, I would have expected a larger value.***

AR: *The reviewer is correct about the impression that 2.5 mm are too small. It's 25 mm or 2.5 cm. We corrected* *the typo. The numbers were by the way taken from the https://doi.org/10.1063/PT.3.2009 reference.*

RC: ***185: are applied***

AR: *Changed.*

RC: ***Equation 1: I suggest adding the gradient mapping function to grad(a,e)***

AR: *We have added the gradient mapping function as follows*

$$grad(a,e) = G_{NS} \cdot mf_G(e) \cdot \cos a + G_{EW} \cdot mf_G(e) \cdot \sin a \quad (1)$$
$$mf_G(e) = 1/(\sin e \cdot \tan e + 0.003) \quad (2)$$

RC: ***277: Where is Figure S4?***

AR: *Figure S4 is contained in the supplemental material (`https://doi.org/10.5194/essd-2022-57-supplement`). We have changed the reference as follows:*

> *The data was recorded by both satellites (A,B) along ascending orbit 88 between March 2015 and July 2019. All available datasets are visualized over time and indicating the along-track coverage in latitude in Figure S4 of the supplemental material.*

RC: ***Equation 3: is there a certain reason to use * instead of . ?***

AR: *It was a typo and has been changed.*

RC: ***466: derived***

AR: *Changed.*

RC: ***Figure captions 9 and 10, and others: please provide all the information in the figure caption, which is necessary to understand the figure.***

AR: *Thank you for pointing this out. We have updated the captions (now Figs. 8 and 9).*

> Fig. 8
> Collocated vs. reference GNSS estimated IWV at 5 validation stations. Continuous lines represent the reference GNSS station values and dashed lines the ones of collocated data. The colors mark the 5 different stations.

> Fig. 9
> Seasonal residuals between reference GNSS estimated and collocated IWV at 5 validation stations. The colors mark the 5 different stations.

RC: ***493: datasets***

AR: *Changed.*

RC: ***Equation A1, and other equations in the appendix: please add units***

AR: *Units were added to all equations in the Appendix.*

RC: ***Equations A8 and A9 denote the ZWD delay as a pure "wet" delay. On the other hand, A3 refer to a non-hydrostatic delay (not wet in the strict sense). Does this (small) difference cause any inconsistencies?***

AR: *We apologize, there was a typo in the submitted manuscript in Eq. (A9). The wet component of refractivity should be Davis et al. (1985):*

$$N_w = k_2' \frac{e}{T} + k_3 \frac{e}{T^2} \tag{3}$$

where k_2'*=22.1 K hPa*$^{-1}$ *and* k_3*=373900 K*^{2} *hPa*$^{-1}$*.*

As the ZWD in Eq. (A8) and (A9) just help to provide some information for the calculation of InSAR-derived ZWD, its influence is minimal.

Author Response to Review of

Tropospheric water vapor: A comprehensive high resolution data collection for the transnational Upper Rhine Graben region

Benjamin Fersch et al.
Earth Syst. Sci. Data Discuss., doi:10.5194/essd-2022-57

RC: ***Reviewer Comment***, AR: *Author Response*, ☐ Manuscript text

Dear Minyan Wang,

we would like to thank you very much for taking the time to review our work. Your comments and constructive suggestions to our manuscript are highly appreciated. Our answers to your comments are as follows:

RC: ***The authors introduce the process of how to get tropospheric IWV products based on occultation and satellite remote sensing synthetic aperture radar data, using data fusion and data assimilation methods in WRF, and describes the preliminarily evaluations of the quality in this paper.***

AR: *We employ propagation delays but not occultation for GNSS in this work.*

1. General comments

RC: ***1. In the introduction, it should be included the previous research on using similar methods and data to get the product. It is suggested to add.***

AR: *We have added the following paragraph to mention studies that combine either of the methods for determining trophospheric water vapor fields (L29-38):*

> Combining high temporal GNSS measurements with satellite products with low temporal but high spatial resolutions is obvious. Furumoto et al. (2003) applied GNSS water vapor measurements with radio acoustic soundings to improve water vapor profiles. Lindenbergh et al. (2008) combined Medium Resolution Imaging Spectrometer (MERIS) satellite data with GNSS data based on kriging techniques and Leontiev and Reuven (2018) used cloud fractions derived from Meteosat-10 to improve GNSS IWV interpolation. The assimilation of GNSS measurements in atmospheric models to reduce uncertainties of water vapor simulations is another promising approach which is widely used (see Wagner et al. (2022) for compilation) and also the assimilation of InSAR derived water vapor data can improve the spatial skill of precipitation forecasts (Mateus et al., 2021). Although the combination of single observational product types with local area atmospheric modeling is common, the rigorous fusion of multiple data sources with modeling has not been documented to our knowledge so far.

RC: ***2. At present, the contents before page 7 are far more than expectation. Many of them have nothing to do with the research itself, but are related to the fundamental data and methods. It is suggested that only the contents closely related to the study are kept. The structure of the paper needs to be greatly adjusted. The second section and the first section should be merged, and there is no need for subtitles. The second section is not an introduction to the data and research methods used in this paper, but is more like the expression in the master's or doctoral thesis technical document. It is not suitable in the scientific research paper. Pages 2 to 7 is suggested to be shortened to 3 pages at most. It is assumed that the reader is engaged in parts of research in this field. The third section should not be a description of the data set, but an introduction to the specific input data used in my research and the methods. You can also introduce the horizontal spacing and vertical resolution of observation points actually used in the research. It seems that the focus of input data and output products of this study is not prominent. The IWV product is in the form of grid data, right?***

AR: *We agree with the reviewer's impression that there is much information given before the actual data collection is addressed and described. The reason, why we decided to do this mini-review (Section 2) is the fact that this is a multi-disciplinary work where meteorology, geodesy, and photogrammetry come together. A reader may be rooted in one of these communities or come from a totally different discipline. Since the vocabulary and the approaches differ considerably among those groups we think it is valuable to present the state of the art of the different disciplines and to introduce their specific terms. We tried to keep this as concise as possible (about half a page per discipline).*

We have shortened Section 2 and updated the content description for the paper at the end of Section 1 which reads now as follows (L50–54):

> *In order to highlight the advantages and drawbacks of each approach and to elucidate the importance of comparing and combining different methods and disciplines, a brief introduction about methods and terminologies used by the different disciplines is provided in Chapter 2. Chapter 3 describes in detail the dataset and its creation process. Subsequently, we evaluate the dataset with independent observations in Chapter 4.*

and in Section 2, we have updated the statement to (L66-68):

> *In the following, also to make readers familiar with the terminologies used by the different disciplines, we provide a brief overview of the respective methods for tropospheric water vapor observation and modeling and highlight also how the strengths of the different data sources can be combined into something more valuable.*

Concerning the introduction of the dataset we agree that we could be more specific about the input data and output format. We have added a Table (Table 1) at the beginning of Section 3 to list the basic data products and their properties:

Data product	Temporal interval	Horizontal resolution	Vert. res.	1st raw product
ERA5	1h	31 km	~500 m	specific hum.
GNSS	1h	60 stations	-	ZTD
WRF	1h	2.1 km	~500 m	spec. hum.
InSAR	6–12 days	86 pts/km^2	–	ddSTDP
Tomography	1h, 6–12 days	2.1 km	~500 m	ZTD + ddSTD

Furthermore, we have changed the title of Section 3 to:

Tropospheric water vapor dataset

RC: ***3. The difficulty of this study is the treatment in the case of clear/cloudy circumstances, or precipitation, when the variation of water vapor in the lower atmosphere is more complex.***

AR: *From the perspective of assimilation in atmospheric models, the time-steps with higher water vapor variations are most promising to improve modeling results. If model simulations do not adequately represent these variations, assimilation may help to tie simulations closer to reality. That is also the main reason why we apply a high assimilation rate of 1 hour to include possible variations in our assimilation runs.*

RC: ***4. What are the thresholds of spatial and temporal matching during evaluation or colocation? The distance in space, and the time difference.***

AR: *We collocate GNSS ZTDs of 3 consecutive hours and always produce the ZTD or refractivty fields for the middle hour. In collocation we use all the GNSS stations available for the three hours. We do not have to do a spatial filtering since our algorithm will automatically weight more the closest GNSS stations to the point we will interpolate. As for the evaluation, for instance when we evaluate IWV shown in Figure 10, after the collocation of the GNSS ZTDs we interpolate exactly at the same time and location as the validation observations. Thus there is no temporal or spatial lag.*

RC: ***5. The evaluation results are insufficient. What is the variation of time series? What is the seasonal change from 2015 to 2019? And what is the difference with those in ERA5? In ERA5, GNSS bending angles and water vapor information from radiosonde in different height of the atmospheric are assimilated. InSAR data are not assimilated. Is IWV integration of GNSS assimilated? From the results, it is found that local water vapor field is more reasonable after assimilation of new observations like InSAR.***

AR: *The benefit of our regional atmospheric simulations is the increased spatial resolution as compared to ERA5 (2.1 km instead of 31 km). In this way we can include regional particularities much more detailed and resolve processes such as deep convection whereas parametrizations have to be applied in ERA5. In ERA5 however, many more variables are being assimilated. Regarding GNSS data, these are mainly profiles of GNSS radio occultation. We use ERA5 as driving data only at our domain boundaries. So, assimilation in regional atmospheric modeling is nevertheless beneficial to tie the simulations closer to reality. InSAR data is only available every 6–12 days. The impact of assimilating InSAR data is therefore comparably small for our overall results for each event. The combination of a better representation of processes in WRF due to smaller grid spacings and the assimilation of key variables such as water vapor seem to be the main reasons for a* *better model performance.*

Regarding the seasonal change from 2015 to 2019, this question is out of the scope of this work. We use data from 2015 to 2019 indeed but this is the case because the InSAR acquisitions processed happen to be for these years. However we focus on 4 different seasons, chosen so that there is also InSAR data available, since we would like to also advance the understanding of InSAR contributions to meteorology. Notice that the data for the 4 seasonal events are processed and evaluated separately. Only the InSAR acquisitions are processed for the overall time span. In the future, with larger temporal resolution of InSAR time series, one could have much more observations in one season and would not need to process InSAR acquisitions distributed over several years.

2. Specific comments

RC: ***1. The temporal and spatial range of the product should be stated in the abstract and summary. In the conclusion, we should summarize the above and data quality of this product, briefly explain the input data and methods.***

AR: *We have updated the information in the abstract as proposed. And added the following sentence to the conclusions (L459–464):*

> *It contains hourly 2-D fields of integrated water vapor (IWV) from the various disciplines and 3-D fields of water vapor density (WVD) for four multi-week, variable season periods between April 2016 and October 2018 at a spatial resolution of (2.1 km)$_2$. Zenith total delay (ZTD) from GNSS and collocation, and refractivities are provided as intermediate products. InSAR derived double differential slant total delay phases (ddSTDP) are available for March 2015 to July 2019. The original input data for this work were hourly time series from 66 GNSS stations, hourly ERA5 reanalysis fields from ECMWF, and hourly Sentinel-1A/B InSAR observations.*

RC: ***2. Abstract. Among should be changed, like one of. Guess should be changed, like estimated or other word. Add the physical quantity IWV in the abstract, which is very important. Add the time period from 2015 to 2019 (or others?), and the quantitative results like 0.98.***

AR: *We have changed the abstract as suggested.*

RC: ***3. L82. Delete the citation of the extreme weather events.***

AR: *We don't see an obvious reason, why we should skip Zhu et al. (2020) since it shows an application of GNSS data for extreme weather conditions. Therefore, we have decided to keep the reference.*

RC: ***4. L267 and 338, etc., no details like these are required.***

AR: *We think it is important to explain the structure of our dataset to the readers, since this publication is about describing and evaluating the data collection to potential users.*

RC: ***5. Figure 1 can be deleted. Please consider whether to delete section 3.5.***

AR: *We removed Fig. 1. For Section 3.5 we think that it needs to be kept since it explains the collocation dataset. However, we changed its title to:*

> ZTD, refractivity, and WVD based on collocation

RC: ***6. Figure 2 and Figure 4 (not 4S? Put Section 3.3 instead of section 3.5) should be placed in earlier pages.***

AR: *We moved the study region overview figure to the introduction (now Fig. 1). Figure 4 (now 3) is specifically linked to the description in Section 3.5 and therefore, we kept it there. Section 3.3 needs to be kept before 3.5 since the latter builds upon the data that is being described in Section 3.3.*

RC: ***7. L345. There is a logic problem. Generally, the citations should not be this section itself. It is section 3.***

AR: *We changed it to:*

> *For raw assimilation data we refer to the other datasets presented in this Section.*

RC: ***8. Figure 5. What are the first seven in the y vertical axis?***

AR: *As mentioned in L 431-432*

> *The mean values of the KGE over all 25 GNSS stations in the SAR area are presented on top, followed by the mean over the 5 validation stations*

the uppermost bar describes the performance of all 25 GNSS stations that are contained within the evaluation area (see Fig. 2). val_5 are the 5 selected validation stations (blue dots in Fig. 2) and subsequently the measures are given separately for these 5 stations. We have updated the captions of Fig. 2 (now Fig. 1)

> *Figure 1. Study area location and extent. WRF domain (650 x 670 km) and evaluation area (440 x 460 km, green area) with 56 GNSS stations for assimilation and tomography (blue squares), 5 GNSS stations for validation (cyan), 245 synop stations (black points) and the radiosonde station Idar-Oberstein (magenta). The red InSAR domain marks the core region where all datasets are available.*

and Fig. 5 (now Fig. 4)

> *Figure 4. Performance measures for IWV of PS-InSAR vs. GNSS for different subsets of the InSAR domain (see Fig. 1). stat_all describes the performance of all 25 GNSS stations in the InSAR domain, val_5 are the 5 selected validation stations (blue dots in Fig. 1) and the individual measures for these 5 stations, followed by the seasonal analysis of these 5 stations.*

RC: ***9. Figure 6. In the caption and text, = is inappropriate. The relationship between GSI and GNSS should be expressed in many words to prevent ambiguity. GSI is not referred to the NCEP assimilation system (Gridpoint Statistical Interpolation)?***

AR: *The captions of Fig. 6 (now Fig. 5) were updated as suggested. For the term GSI, the definition was missing in the text but is contained in the Appendix B. It refers to the combined assimilation of the GNSS, synptic, and InSAR data products. We have added a definition to the main text at L404-405:*

> Despite the already high accordances, slight improvements for all seasons are obtained by assimilation of tomography data (CA) as well as by the combined assimilation of GNSS, synoptic and InSAR data (GSI).

RC: ***10. Figure 10. More scales should be added in y axis. The longitude and latitude of Figures 4 and 11 are irregular, while those in Figure 2 are standard.***

AR: *We have increased the detail for the y-axes of Fig. 10 (now Fig. 09) and added the countries' frontiers to Fig. 11 (now Fig. 10) to make it better comparable to the areas defined in Fig. 2 (now Fig. 1).*