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 will scan the available literature for recently published work in the context of tropospheric water vapor and data fusion and update the introduction where possible.*

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 propose to reconsider Section 2 and check where it can be reasonably shortened. Further, we would specify more clearly in the last paragraph of the introduction our motivation to present this state of the art / review section. The readers familiar with the terms and the state of the art of the different disciplines can be directly pointed to section 3 with a respective statement.

Concerning the introduction of the dataset we agree that we could be more specific about the input data and output format and we would therefore update the general overview in the beginning of Section 3. The details about input data and processing are contained in the individual subsections but the general input products can be listed earlier as suggested. We will also think about another title for Section 3.

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 3–5 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 acqusitions 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 will update the abstract and conclusions accordingly.*

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 will consider your suggestions in the revised version.*

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

AR: *We would like to ask for the reason, why we should skip Zhu et al. 2020 since it shows an application of GNSS data for extreme weather conditions. Otherwise, we would opt for keeping 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 agree to remove Fig. 1. For Section 3.5 we think that it needs to be kept since it explains the collocation dataset. However, we propose to change it's 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 will check if there is any reasonable way to move those figures more towards the beginning.*

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

AR: *We will change that 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 see that the figure captions of Fig. 2 and 5 are not detailed enough and will change that for the revised version.

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: *We will change the figure captions accordingly. For the term GSI the definition is missing in the text but contained in the appendix B. It refers to the combined assimilation of the GNSS, synptic, and InSAR data products. We will add a definition to the main text.*

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 will increase the detail for the y-axes of Fig. 10. We are not sure what the reviewer exactly means with* standard *and* irregular *of Figs. 2, 4, and 11. Fig. 4 was designed to make the different layers optimally visible. We will check if we can add some more numbers to the coordinate axes.*