Reviewer #2: https://doi.org/10.5194/acp-2021-797-RC2

General Comments

The work presented in the manuscript gives an overall summary of applications of ground-based GPS observations in Europe of estimated time series of integrated water vapour (IWV), which to my knowledge is unique. It is broad in the sense that it deals with temporal scales from sub-daily to decades, while many previously published results often focus on one particular "signal", e.g. diurnal, annual, trends. As far as I can tell there are no new results in the manuscript, i.e. results that are different from what is already published. Three times it is stated that the results are "in line" with previously published results (lines 222, 323, and 399). Of course, it is also an important part of research to verify earlier findings, but if possible, I would appreciate if there was more emphasis on noted differences compared to earlier results. I am afraid I cannot help with the details. It is an impressive reference list and for me it is impossible to get a reasonably complete overall knowledge during the time allowed for the review.

Reply: Thank you very much for your affirmation.

Ground-based GPS is a unique technique to evaluate the quality of IWV from atmospheric reanalyses. The evaluation can provide information on how to improve their performances in retrieving IWV. However, most previous studies in Europe only evaluated the IWV from ERA-Interim produced by ECMWF, which has been superseded by ERA5 since 2019 August. The time length of those studies are also relatively short (<20 years).

To our knowledge, this is the first study which used 25 years of 1-hourly GPS IWV in Europe to evaluate the performances of the newly released ERA5 in modelling multiple temporal scale variations of IWV from intraday to decades. In addition to the ERA5 and ERA-Interim produced by ECMWF, this study also evaluated the IWV from four commonly used products developed by USA and Japan, which have rarely been evaluated in Europe. An advantage of this comprehensive evaluation is that it is capable to avoid impacts due to differences in reference GPS IWV data and evaluation methods, so that the comparisons on the performances of reanalyses are fair.

We believe the results, especially the evaluation of 1-hourly IWV, are very new and interesting to the community. This is because one of the most important advantages of ERA5 is its much higher temporal resolution compared to the other products (1-hourly v.s. 6-hourly), but its possible improvement has not been evaluated in Europe. GPS IWV is a unique data source to evaluate the 1-hourly ERA5 IWV, as it is famous for its high temporal resolution and high accuracy. Europe, especially its northern part, is characterised with unstable weather condition. Hence, it is a good study region to evaluate the performances of the reanalyses in modelling the high frequency variations of IWV. In this revised manuscript, we carried out more investigations on the intraday variations and diurnal cycles of IWV as suggested, such as the differences in inland and coast.

Moreover, Europe is known as the continent with the most significant warming speed. Evaluations of the long-term IWV trends from the atmospheric reanalyses with the 25 years of GPS IWV are also conducive to a better understanding of climate change.

Specific comments

1. L108: I do not understand the meaning of "integration rate of 95 %"? Can you explain what is being integrated?

Reply: We defined the integration rate of the daily IWV series at a GPS station as follows:

$$rate = \frac{N}{MJD_{last} - MJD_{first} + 1} \times 100\% \quad (1)$$

where N is the number of daily IWV estimates of the GPS station. MJD_{first} and MJD_{last} are the Modified Julian Dates of the first and last daily IWV estimates, respectively. We added this information to the revised manuscript.

2. L112: You report that the observations were weighted based on the elevation angle. Is it not important how the weighting was done (a weighting function including sine and cosine terms)?

Reply: Yes, we should provide more details on the GPS data processing. An elevation (e) dependent weighting function of $\sin e$ is adopted in the processing in addition to a cut-off elevation angle of 7°.

3. L192: It is mentioned that homogenisation was done as described by Yuan et al. (2021). I think such a process is critical and it deserves some more detail in your paper instead of having to go through the reference. For example, do you allow breaks to be inserted in the GPS IWV time series at a specific time epoch even if there has been no change noted in the log file for the hardware or the environment at the site?

Reply: As Reviewer #1 and Olivier Bock in their comment were questioning the homogenisation procedure used in the previous version of the manuscript, we decided to extend more on the homogenization and used a statistical changepoint detection tool (Wang, 2008) in combination with the available log file information. Moreover, the detection tool has been applied on the IWV monthly mean differences between GPS and all six reanalyses.

The procedure is as follows. For each GPS station, we first tested the changepoints by referring to its log file and identify them by using the detection tool developed by Wang (2008) based on the comparisons of the monthly mean IWV values of GPS and each reanalysis. A changepoint is accepted if it is reported by the tool in at least three GPS–reanalysis comparisons. Its amplitude is calculated as the average of those reported by the comparisons.

Regarding the undocumented changepoints, we determined them carefully. For each GPS station, we first identified the undocumented changepoints automatically reported by the detection tool based on the IWV monthly mean comparisons of GPS and each reanalysis. If similar changepoints within six months are reported by at least three GPS-reanalysis comparisons, they are considered as the changepoints from the GPS IWV series. Then, they are combined into one at the median month, and its amplitude is calculated as the average of those reported by the comparisons. By using all reanalyses in the changepoint detection tool, we hope to minimize the effect of these changepoints on the results here, although it cannot completely rule out that identical changepoints appear in different reanalyses by ingesting the same observational datasets through data assimilation. This limitation has been mentioned.

Reference

Wang, X. L.: Penalized Maximal F Test for Detecting Undocumented Mean Shift without Trend Change, Journal of Atmospheric and Oceanic Technology, 25, 368–384, https://doi.org/10.1175/2007JTECHA982.1, 2008.

4. L262: My interpretation is that you determine the amplitudes of the diurnal signal as the peak-to-peak value regardless of when the peaks occur. This makes me wonder if the results will be different if instead the phase and amplitud of the sine wave with a 24 h period is estimated, e.g, through the method of least squares. (In some studies also a semidiurnal term, a period of 12 h, is estimated.) It will be of interest if you comment on this, at least for a couple of sites in different climate zones?

Reply: Yes, there are harmonic analysis on the diurnal cycle of IWV (e.g., Steinke et al., 2019). We modelled the diurnal cycle with sine wave and compared to the peak-to-peak estimates as suggested. We also investigated the characteristics of several different climate zones. We are sure that this investigation will provide new results to the community.

Reference

Steinke, S., Wahl, S., and Crewell, S.: Benefit of high resolution COSMO reanalysis: The diurnal cycle of column-integrated water vapor over Germany, Meteorologische Zeitschrift, 165–177, https://doi.org/10.1127/metz/2019/0936, 2019.

5. L268: You find a correlation between the diurnal amplitude and the station height. Since station height (I guess) correlate with the site's distance to the ocean, another approach would be to correlate the amplitude with this distance. It is well known that the ocean (as long as there is no ice) acts like a low pass filter on daily variations in temperature and humidity.

Reply: This is a very interesting idea. We compared the impact of the difference between inland and coast on the diurnal amplitude of IWV as suggested.

6. L315: This whole section seems questionable if it is worth to be published? Do the GPS IWV data yield any new findings? Given the very high correlation between IWV from GPS and from the reanalyses, it seems as all the reported patterns, and their time dependences, will be seen by using reanalyses data only?

Reply: We agree with you that this part is out of the scope of this work, and thus it was removed. We will address the issues related to interannual variations of IWV in the future. The findings here are interesting, as very a few studies have investigated the teleconnections between the interannual variations of IWV in Europe and various climate indices. Furthermore, although the interannual patterns of IWV can be seen in the reanalyses, their performances can be validated by using GPS IWV measurements.

Technical Corrections

7. Line (L)1+: You use the American spelling of vapour, although ACP is a European journal?

Reply: Thank you for the suggestion. We used “vapour” and the style of English (UK) in the revised manuscript.

8. L97: ... IWV -using ... ?

Reply: Replaced with “by using”.

9. L17: 2%-18% --> 2 %–18 % (similar changes to be carried out many times in the manuscript)

Reply: Replaced hyphen with en dash.

10. L154: IWVs --> The IWV values ?

Reply: Replaced as suggested.

11. L157: reanalyses Compared --> reanalyses. Compared

Reply: Modified as suggested.

12. L203: IWVs are --> IWV for all sites and days are ?

Reply: Modified as “The daily mean IWV time series of each station is further aggregated into monthly mean IWV series…”.

13. L398: (29.5°E, 40.8°N), --> (29.5 °E, 40.8 °N), (see also L447-448)

Reply: Modified as suggested.

14. L444: 0-0,4 --> 0.0 – 0.4 ?

Reply: Replaced hyphen with en dash.

15. L446: 0,4-1 --> 0.4 – 1.0 ?

Reply: Replaced hyphen with en dash.

16. L480+: doi links are missing for almost all references and the established standard acronyms for journals are not used.

Reply: Added the doi links and used standard acronyms for journals.

17. Figure 2: The yellow colour is not ideal. I suggest to use cyan or magenta instead. You may also consider to use darker colours in Figures 5 and 8. Different colours in these figures are not really needed for clarity, although it may look nicer compared to have it all in black.

Reply: We used cyan to replace the yellow colour as suggested. We also used darker colours in Figure 5 and 8.