Review of manuscript acp-2020-77 “Trends of atmospheric water vapour in Switzerland from ground-based radiometry, FTIR, and GNSS data” by L. Bernet and co-authors.

General comments

This manuscript investigates the consistency of monthly means and linear trends of IWV from ground-based radiometry, FTIR, and GNSS data, and two modern reanalyses (ERA5 and MERRA-2) in Switzerland for the period 1995 to 2018. This study is of special interest to the climate community because it confronts several state of the art observational techniques and two of the currently best available reanalyses which assimilate a huge amount of observations (mainly from satellite remote sensors). Though the investigation is limited to the country of Switzerland, it shall shed some light on the uncertainties that generally impact both types of climate data and propose a methodology that may be replicated to other regions of the world. One specific source of uncertainty in trend analysis resides in the treatment of inhomogeneities in the data due to changes in the observing systems. In this respect the approach described by the authors with the GNSS data is original in the sense that it adjusts offsets (so-called jumps) in the time series that may be due to antenna and station changes. Though the proposed approach is interesting I think the way this issue is handled it is also the main weak point in the study.

The approach of fitting simultaneously biases (or changes in the mean) and a linear trend is an ill-posed problem which requires special care and validation because of collinearity. Indeed, it is well known that these parameters are highly correlated, e.g. successive downward changes in the mean can be compensated by an upward trend hence over-estimating (underestimating) an existing positive (negative) trend in the series. Interpreting the trend parameter alone may thus lead to erroneous conclusions. Though the authors notice in several places that the trends after bias correction increase, they do not consider it as problematic, e.g. at NEUC the GNSS trend increases from 0.33 to 0.74 mm/decade when 3 jumps are adjusted and at PAYE the trend increases from 0.32 to 1.14 mm/decade after bias correction (when 3 jumps are fitted). It is noticeable that the uncorrected trends are in very good agreement with the ERA5 trend at Bern (0.34 mm/decade). Inspecting the station information table cited by the authors, reveals that 6 out of the 7 stations listed in Table 2 are subject to 3 or more jumps. The trends at all these stations are relatively high (above 0.58 mm/decade). One exception is station EXWI which did not undergo a bias correction and which has the smallest trend (0.02 mm/decade). It may be pure coincidence, but this point needs to be checked. The large dispersion in the GNSS trend estimates over Switzerland (Fig. 8) compared to the relative flatness of the reanalysis trends (Fig. 10) is striking as well and may further support this problem.

I suggest that the authors further investigate the impact of the bias correction procedure on the GNSS trends and possibly reduce it. One critical aspect here is with the position of the change points. A misplaced change point, or a change point inserted in a series where there is no jump, is very likely to bias the trend estimate. So it is of prime importance to detect and correct only the true change points. It is not clear from the manuscript how the jumps are detected. Is it from the station position time series? The authors should be careful that though some jumps in the position time series may have a coincident jump in the ZTD or IWV time series, this is not always the case. The reason is that, e.g. in the case of an antenna change, a station height offset can be due to both a PCO and a PCV change while a ZTD offset is mainly due to a PCV change (as long as the position is estimated simultaneously with ZTD). Only when change points are properly detected can the method of estimating biases and trends together provide accurate trend estimates. In order to provide proper insight into the impact of the correction, it would be useful to show the trends estimated from corrected and uncorrected time series.

I recommend thus that the GNSS offset correction be improved or at least carefully assessed and the manuscript revised accordingly. The conclusions may change significantly. A number of additional specific comments are given below regarding other aspects of the work and the manuscript.

Specific comments

P4L99: could there be an effect of the changes of the TROWARA instrumentation in 2002 and 2004 on the IWV data and IWV trend estimated with the radiometer? Is there a bias adjustment similar to that applied to the GNSS data or any kind of recalibration?

Section 2.3: why did the author use the outdated Thayer constants and Bevis formula for the ZTD to IWV conversion? More elaborate and updated approaches are given in the COST GNSS4SWEC final report (see e.g. Chapter 6, Section 6.4.2). Note that the older refractivity constants may be responsible for a small bias in the GNSS IWV when comparing these data to other well calibrated observations (e.g. microwave radiometer and FTIR) and that ZHD and Bevis Tm computed from surface Ts observations and Ps may produce spurious diurnal and seasonal signals in the GNSS IWV estimates (due to spatial extrapolation errors and inaccuracies in the Bevis formula).

P7L200: what does this standard deviation represent exactly? Is it the variability of the IWV data at the nominal time sampling (in this case what is the sampling of the GNSS IWV data? 1 hour?) or is it a formal error?

Why is the standard deviation of TROWARA 2 or 3 times smaller than the GNSS standard deviation? (is it because of the time sampling?)

P8L221: explain how the jumps are detected (see also the general comments).

Fig. 2b: how does the corrected GNSS time series compare to the ERA5 data? Are the biases reduced?

P9L269: compared to the TROWARA trends which are of similar magnitude, the GNSS trends are not deemed significant because the standard deviation is 2 or 3 times larger. This prompts again for a clarification on the reason of the difference in standard deviation.

P9L271: this sentence is very speculative. Revise or remove.

P9L274-275: The shorter time period of EXWI is suggested but WAB1 has a much shorter period. Revise or remove.

P9L279-284: is the reliability of radiosonde data good enough to estimate trends? Similarly, the lidar trends are probably biased because they include clear sky measurements only. Moreover, the time period for the latter is much shorter. I think these results can hardly be inter-compared. Please remove.

P10L285-294: trend estimates computed for different periods must be compared with much care.

P10L295: why would trends be larger in summer? Do you have any clue for this?

Fig. 5: what is the benefit of using the mean GNSS trend here? how is this mean computed?

Add error bars on the monthly trend estimates to judge whether the variations are significant or not.

The GNSS trend for March is suspicious and needs to be checked.

P10L304: The fact that the GNSS stations are not located in the same reanalysis grid (ERA5 or MERRA-2?) does not sound like a good explanation for the differences. Please revise or remove.

P11L315: mention that the temperature trends explain part of it and that they are further discussed in section 4.2.

P11L325-328: contains some inaccuracies. Suggested correction:

"In case that the water vapour pressure e is smaller than e_s, the available water is in vapour phase, whereas for e > or = e_s it condenses. With increasing temperature, e_s increases, which leads for a given relative humidity (RH) to an increase of water vapour e. Changes in e_s can therefore directly be compared to changes in the amount of water vapour measured by e or IWV, assuming that the RH remains constant"

Give a reference to other studies concluding that RH remains nearly constant and be more specific about the conditions for this to hold, namely is this result valid at regional scale?

P12L347-355: I am not convinced of the impact of temperature inversions. Do you have further evidence of this? Instead, the effect of moisture transport is a better explanation (e.g. also discussed by Parracho et al., 2018).

P13L392: could the snow and icing problems effect the GNSS data used in this study?

P13L403: is it 3 or 4 stations?

P14L410: trends computed for different periods and regions must be compared with much care.

Fig. 13: these plots are weird. It seems that there is glitch in the data computation here. Please check carefully. Otherwise, a proper explanation of the MERRA-2 results should be given.

P15L444-445: I think it is not clear if it is the reanalyses results that are too smooth or the GNSS results that are too noisy. Can you clarify/discuss this point?

P15L462-468: The uncertainty of the reanalyses is not directly addressed along in this work. So, drawing doubt on these data in the final discussion sound rather awkward. This being said, it is true that the use of reanalyses for trend analysis has been debated but the quoted references suggesting that ERA5 and MERRA-2 are more accurate at estimating trends are too general to support such an assertion. Please revise or remove this paragraph.

Fig. 10 and 11: to ease the comparison with GNSS results, add the GNSS results onto the reanalysis results. The plots may be resized.

Table 1: add the number of jumps that are inserted in the trend analysis.