Solid Earth Discussions

Solid Earth Discuss., referee comment RC2
https://doi.org/10.5194/se-2021-136-RC2, 2022

Comment on se-2021-136

Anonymous Referee #2

Referee comment on "Common-mode signals and vertical velocities in the greater Alpine area from GNSS data" by Francesco Pintori et al., Solid Earth Discuss., https://doi.org/10.5194/se-2021-136-RC2, 2022

Pintori et al. use a version of the ICA method (called variational Bayesian ICA) to decompose vertical GPS position time series and hydrology/atmospheric predicted loading time series around the European Alps. They study the agreement between the ICs extracted from the GPS series and from the loading models for the period from 2010 to 2020. Their main conclusions are that 1) the vertical GPS series can be separated in a tectonic linear motion and variations caused by temperature and atmospheric/hydrology loading; and that 2) improved tectonic velocities are obtained by correcting the GPS series using ICs obtained from the GPS series themselves.

While the volume of work is of note, especially concerning the GPS data processing, I do not think the conclusions are supported by the data and methods used by the authors. It is reasonable to say that temperature variations, atmospheric pressure variations and hydrology load variations contribute to the variations observed in vertical GPS time series, especially at the annual period, as GPS positions react to these and many other phenomena together. A completely different thing is to say that the observed GPS variations of vertical position *are* originated or explained by these processes, as the authors repeatedly state in the manuscript. This is a clear misinterpretation of their analysis and I develop my reasoning in the paragraphs below.

Before that, and assuming conclusion 1 is right, it's very surprising that the authors do not try to remove the modeled loadings from the GPS series to test the impact on the estimated velocities. Instead, conclusion 2 is based on removing the GPS ICs from the GPS series, i.e., conclusions 1 and 2 are totally unrelated. The GPS ICs were obtained from GPS series that were previously detrended, explaining the small change of the estimated velocities from the filtered series. The ICA filtering also explains the reduction of the noise in the series and, therefore, of the estimated velocity uncertainty from the filtered series. Where I think this approach fails is that the raw series (used to estimate the velocity, the filtered velocity being very similar) and the filtered series (used to re-estimate the velocity uncertainty) are not consistent and therefore the velocity and its "improved" uncertainty are not consistent either. The authors could have tried a more aggressive filtering, like a band-pass filter leaving the trend and high-frequency noise only, or could have not consider colored noise in the velocity estimation (both ways are equivalent) and they will get even smaller velocity uncertainties. Unfortunately, this will not give any valuable information on the quality of the velocity and your ability to extrapolate it to understand tectonic physical processes. The only way to improve velocity estimates is to understand and reduce variability in the GPS series with proven corrections and models. If the white noise is more visible in the filtered series is probably because the GPS ICs absorb together a significant portion of the power-law noise that typically dominates the variance of the detrended GPS series, though this is not very clear from the IC PSDs in Fig. 3. Precisely, the power-law noise in the GPS series is only mentioned briefly and its influence on the GPS ICs and on the correlation with the loading ICs is not discussed at all.

With respect to the GPS ICs and their attribution of a geophysical origin, I enumerate below several points raising concerns on the authors' approach. Generally, many past publications have shown than GPS series and loading models do not see the same thing, except partly for the annual variation. Most of the variance in the loading model series is concentrated at the annual period. Compared to the PSD of the loading models, the GPS series contain a relatively higher variance at long periods with a distinct PSD slope and a PSD much richer in periodic artifacts at short periods. The authors briefly comment on the systematic errors that are present in the GPS series, but they do not try to make the GPS series more consistent with the model series. For instance, it is known the annual draconitic variation could significantly affect the comparison to the solar annual variation of the loading models. The results obtained by the authors are confusing (see points below) and do not refute findings from past publications, contrary to their claims to successfully separate geophysical signals from the GPS series. For instance, authors show no evidence that the HYDL series significantly explain variations in their GPS series. The GPS and NTAL annual seem to partly agree (see points below), so the authors introduce a thermal annual component in the discussion without providing strong evidence nor explanation of its spatial pattern. It is also probably worth mentioning that, if the GPS series were effectively explained by the combination of atmospheric/hydrology loading and temperature variations, as the authors claim, we should get the same GPS series out of the same GPS data when using different software, different strategies and different corrections. However, this is often not the case, especially when comparing global and regional GPS solutions.

Other general points:

- While I understand the objective of the ICA applied to the GPS series is to separate the variability into independent processes, I cannot understand the rationale for applying ICA to the NTAL and HYDL series. What are the independent processes to be separated in the atmospheric pressure loading or water loading? Even more confusing are the results from the comparison of a single GPS IC to a single NTAL/HYDL IC and the claim that the GPS series are explained by both. The ICA analysis is forcing the NTAL/HYDL series into non-gaussian independent components, even if they do not exist physically. This probably explains why the total NTAL annual is split across ICs with spatial patterns as orthogonal as possible. The same spatial patterns are found for the GPS series, probably because once the trend, offsets and annual are removed from the GPS series, what is left is a Gaussian or near Gaussian series with temporal & spatially correlated noise and also the above-mentioned systematic periodic errors. It may be that the easiest way for the ICA to force the separation of these residual series into ICs is by making their spatial patterns orthogonal (see another possible explanation in point 5 below). The authors' conclusion that GPS and loading see the same spatial patterns is therefore not very solid.

- The GPS and NTAL/HYDL series have different spatial samplings, which must complicate the interpretation of their comparison. Also related to the spatial sampling, it must be difficult to extract accurate NTAL values in the Alps due to the pressure model resolution and the short-scale changes in topographic gradient, making its comparison to the GPS series even less trustworthy. I suspect similar limitations exist when comparing GPS and HYDL model series in a mountain range.
- Each dataset used by the authors is decomposed in different numbers of ICs: 7 for GPS, although only 4 are discussed, and 3 for the model loadings. Then they compare the first 3 individual ICs and find weak correlations between them. The authors conclude on the origin of the individual GPS ICs based on their correlation to the individual loading ICs. However, this criterion is very weak, especially with correlation values around 0.6. As an example, similar (Pearson's) correlation values would be obtained between a pure sinusoidal and the same sinusoidal delayed almost pi/3, which is roughly two months if the sinusoidal has a period of one year. When subtracting one sinusoidal from the other, it is clear that we are not correcting much. The ratio of explained variance between the different ICs would have been more appealing, but, it is not clear that the individual ICs from different datasets correspond to the same fraction of the total signal (see point 1). So maybe the ICA method is not well adapted to this problem or should not be applied to the NTAL/HYDL series (see point 1). A band-pass filtered comparison of GPS and loading series would probably be more informative here. Also rather than filtering the GPS series, I think it would have been better if the authors had shown how the loading models change the variance of the GPS series, as it is done in many other publications. The loading would need to be computed at the station locations. It would have been even better to show how the GPS variance changes (not necessarily reducing) all along its power spectrum when correcting the loads.
- The authors are processing a regional network and aligning it to a global linear frame (IGb14) that does not include seasonal variations. The frame alignment of the daily solutions from regional networks acts as another CME-like filtering of the series, not discussed by the authors, but probably similar to the SFM method. The filtering is more efficient as the network size is smaller, but the authors do not provide enough information on this point. It is then difficult to interpret the common network-wide annual signal shown by the GPS IC1. I would expect the regional frame alignment would absorb part of this common GPS annual signal, making it difficult to compare to the loading model and also leaving an amplitude much smaller than the residual station-dependent annual signal that is probably captured by the IC4. However, the numbers in table 1 indicate the opposite, assuming the average "of the amplitude of the maximum displacement" is somehow related to the annual amplitude, which is not clear either. The annual variation is the most prominent signal in NTAL with amplitudes typically of a few mm, less than 1 cm at the center of large continental masses. So it's not clear what the authors mean with atmospheric loading amplitudes larger than 2 cm. It is also not mentioned which frame was used to create the loading series and whether they were detrended like the GPS series, especially the HYDL series.
- The 2nd and 3rd GPS ICs are particularly interesting. These represent daily E/W and N/S network tilts with a rather flat spectrum. The NTAL and HYDL show similar spatial tilts, but their physical meaning is dubious (see point 1) and their spectral content is completely different: mostly seasonal for NTAL and mostly interannual for HYDL. The origin of these network tilts is very likely not the same among the datasets, as stated by the authors. In addition, if the whole GPS network is truly moving like these two ICs and it is not an artifact of the ICA separation, I would first think of a problem with the reference frame alignment. As said in point 4, network-wide common mode signals, including daily tilts and annual up & downs, should be at least partly (if not totally)

absorbed by the frame alignment as these signals are not included in the linear reference frame and the network size is probably not large enough. Figure 7b must be wrong as there is no annual variation in the GPS IC2.