Review of manuscript se-2021-58

Dear Editor,

The manuscript entitled "***Imaging structure and geometry of slabs in the greater Alpine area - A P-wave traveltime tomography using AlpArray Seismic Network data*** " presents P -wave traveltime tomography using the data acquired from the temporary network of AlpArray and permanent broadband stations in the study area. The manuscript is well written and well organized. The network and the analysis of the data set are presented on a related paper in sufficient detail. I think the manuscript has the quality to be published in Solid Earth but at least a moderate revision would be necessary.

Dear reviewer,

Thank you very much for your detailed analysis of our manuscript. Your comments and suggestion will help us increase the quality of our work. In this document we try to respond to each point mentioned one by one.

Below are the details of my comments.

The number of stations used are quite large providing spatial coherence among the stations over the network. On the other hand, the number of events are relatively low (331) and some are lower magnitude events. What is the criteria for using Mw5.5 as threshold magnitude ? How many stations recorded the Mw5.5 events ? There is no information on the S/N of the low magnitude events. If the low magnitude events are recorded by fewer stations, would that create any bias as the average of traveltimes of each event is removed traveltime residuals?

We will write a new paragraph in which we describe the selection criteria for each event (and also for an onset on each station) in more detail. We chose a lower boundary of Mw5.5 because we experienced that the number of available picks for distances > 35° is not satisfactory anymore. However, one could try to lower the magnitude limit in a future work to possibly increase the number of events from the poorly covered azimuths especially when using waveform correlation for travel time determination.
We set a lower limit (100) for the number of onsets required for each event to avoid a bias of the array average. We rather experienced variations in the number of travel times per event related to the number of available stations due to delays in the deployment of the AlpArray stations.

The methodology is explained well and the parameters for the inversion are chosen appropriately. The misfit in Figure 12 drops fast during the first 3-4 iterations while stay almost flat for the rest. But the authors prefer to use the results after 12 iterations for very small improvements on the misfit. It would be valuable to present the results after fast drop in misfit (3 iterations) and after 12 iterations. Is there any overfitting the data by increasing the number of iterations ?

We chose to continue the inversion iterations until the roughness (smoothing norm) of the model saturates. This happens after 12 iterations. Thus, up to the 12^{th} iteration we can still get more detail into the model. After step 12 the smoothing norm stagnates and only the damping norm rises indicating that only the amplitude of the anomalies increases.

We also experienced the exact same behaviour in various synthetic tests, where the resulting model further approached the test model when continuing inversion iterations even if the misfit reduction was small.
There should not be any risk of overfitting of the data, as the final misfit of ~3.2 has not yet reached the limit of 1. Below this limit, we would start fitting the uncertainties of the travel time residuals into the model (overfitting).

The checkerboard tests are done but requires more, to present vertical and horizontal smearing in more detail. The authors state that the checkerboard anomalies are smeared at least 20 km at shallow depths below the crust. But they do not provide any information over rest of the domain. The initial checkerboard depth models should be provided together with the recovered patterns (Figure 10). Spike tests would be valuable to monitor smearing over the solution space.

We describe resolution capabilities of the model with depth when we evaluate the checkerboard results for each depth slice in section 4. We also show representative vertical slices though the recovered checkerboard (Fig. 10).
The used model is a combination of a classical checkerboard test and a spike test as there is unperturbed space in between the checkerboard tiles (which can be seen as kind of spikes). Therefore, we can monitor smearing effects quite well with this model and decided not to do additional spike tests.

The incorporation of the crustal model and upper mantle (<100km) into the inversion is a nice idea although more tests are needed to understand the influence of strong constraints on the velocity perturbations. The results after inversion should be compared to the initial model for depths < 100 km. Is there any bias on the transition from strongly constrained upper part (>100km) to unconstrained lower part ?

We will try to highlight the influences on the uppermost ~100 km when we update the vertical sections through the checkerboard to clarify the way we created the checkerboard model. There it also becomes visible in which way artifacts can occur within the weakly constrained parts of the crustal model.
We do not understand the last question regarding the bias on transition from strongly to unconstrained parts.

The authors compares their results to the previous teleseismic tomographic works. It appears that Zhao et al 2016 used lesser amount of data and attained similar resolutions. It would be nice if the authors display vertical cross sections of few profiles from the tomographic images of the previous works (e.g, Koulavov et al, 2009; Zhao et al 2016 ..) crossing the same structures.

We will add a comparison to other models that are digitally available in the same spatial domain.

The presentation of the 3D model in Figure 16 is the least satisfactory part of the paper. The Figure 16 does not make any impression neither as geology nor a velocity model. Depending on the level of velocity contours a different image with different size of low velocity holes and slab thicknesses would appear.

The main idea behind Fig. 16 is to demonstrate the complexity of the model results and to get an idea of the three-dimensionality of the different features. There is of course always some subjectivity involved, as what we see in the figure depends on the threshold value of the shown iso-surfaces. On the other hand, we learned that there is also a strong bias involved when 2-D slices through a 3-D model (or slabs) are presented, as perceived features such as a dipping direction or size will always depend on the exact profile positions and a natural slab does never follow only one direction along a single (or multiple selected) profiles from start to end as an idealized one would do. We experienced that a repositioning of a profile by less than ~50 km at only one end may significantly change the perceived geometric structures.

We should maybe clarify in the manuscript that the decision for a threshold value for the iso-surfaces is subjective and that the 3-D geometry changes for different values. Still, one should keep in mind that when trying to interpret idealized, discrete slab "boundaries" from a seismic velocity model, these are also often based on interpreting isolines (or iso-surfaces).

It appears that some of the research questions the authors posed in the introduction such as "how thick and how long are the descending slabs ? " remained unanswered.

We discussed the penetration depth of different slabs (and with that their lengths). We will try to answer questions regarding slab thickness in the discussion as well. However, it is difficult to make more than educated guesses on the slab thickness, as there are no discrete boundaries of a subducted (heated up) slab anymore, the further it is subducted, as its thermal signature blurs with time. Also, the discussed smearing effects, make such guesses difficult.

A geology based velocity model derived from the tomography can be used as synthetic test to better constrain the slab thicknesses, the extent of the low velocity zones.

The 3-D structure we show here is directly derived from the data and only depends on the chosen value of the iso-surfaces. For a geological interpretation of the velocity model (positive and negative anomalies) we want to refer here to the paper of Handy et al. (2021).

References

Handy, M., Schmid, S., Paffrath, M., Friederich, W., and the AlpArray Working Group: European tectosphere and slabs beneath the greater Alpine area – Interpretation of mantle structure in the Alps-Apennines-Pannonian region from teleseismic Vp studies, Solid Earth Discuss. [preprint], https://doi.org/10.5194/se-2021-49, in review, 2021.