Review Plomerova et al., SE-2021-56

The paper presents a teleseismic tomography of travel time residuals determined from records of the AlpArray backbone and the complementary experiment EASI that focuses on a narrow NS swath along the 13.3 E meridian. Most notable features in the derived model of P-wave velocity are two parallel positive anomalies dipping towards the north, one below the Alpine chain and one below the Bohemian Massif. The results provide new fuel for the long-lasting debate on slab polarity beneath the Alps and deserve publication after a substantial revision.

As a seismologist I have concentrated on aspects of data processing and inversion. There are several severe issues regarding the inversion procedure, the correction for crustal structure and the resolution tests that need to be dealt with. Moreover, data, data processing, uncertainties and crustal model need to be documented in more detail. Finally, the paper deserves a much better organization of thoughts and arguments in particular in the introduction and section 6.

Introduction:

The introduction is surprisingly short. I miss a proper description of the state of the art (what is known), of controversial issues (what is debated) and of how the manuscript contributes to answering some of these open questions. The introduction does not pose any research questions and thus entirely lacks a motivation for the work presented. I also miss a proper description of previous work done with EASI data and its relevance for the current work. At the end, a brief description of the contents of the paper would be welcome. Instead, the last paragraph rather resembles an abstract.

The meaning of the lines in the tectonic map in Fig 1 needs to be explained in the caption.

Data:

I assume that also data of the complementary EASI experiment are used. This should be made clear in the text.

Which procedures for checking data quality have been applied? The paper should be self-contained to some extent.

The description of the picking scheme and the event distributions should be put into main paper. Which trace is used as reference for cross-correlation and beam forming? Quantitative statements regarding error estimation of the picks and their probabilistic combinations should be made.

The description of the enhanced data set is confusing. Is the enhanced dataset just the original one with 201 events (in Figure S1 209 events) plus another 43 events within the 60 degree back-azimuth cone? Or is it a completely new selection from the overall available events containing only events within the 60 degree cone? This should be clarified in the manuscript. Regarding the figures S1, the latter seems to be the case.

There are two figures S1 in the supplement. Renumber to S1 and S2.

Information on crustal structure is taken from several sources. How is the 3D crustal model put together? How are transitions between subregions treated? How are crustal discontinuities treated, as transition regions or as real discontinuities? At least two horizontal sections through the 3D crustal velocity model should be shown. To which depth does the crustal model reach? Does it also contain parts of the uppermost mantle?

How is the correction for crustal structure done? One correction for all events assuming vertical incidence or event-wise corrections taking correct incidence angles into account?

The meaning of the end of the following sentence (line 100) "…. proper tomographic inversion in our target region to resolve structures, including the spatial limits of our images" is unclear? What are the spatial limits of the images? Do the authors want to say that resolution is also good at the model boundaries?

Estimated uncertainties of the travel time residuals should be documented in the manuscript. Show a histogram of the uncertainty distribution and give average and median values

Method:

It is stated that the matrix W_m in equation (4) provides horizontal smoothing. What about vertical smoothing? Is there also a damping term included in W_m , and if yes, what is the weight of smoothing relative to damping?

With 13 cells in vertical direction, I calculate a model depth of 13*30=390 km and not 435 km as stated in the text. Is there a specific reason for choosing 390 km (or 435 km) as the bottom of the model. Is this depth still warranted by intersection of rays given the length of the EASI profile?

Confining the inversion domain to a narrow stripe in NS-direction causes a problem for rays incident from easterly or westerly directions. They can carry significant travel time residuals but only pass through the uppermost parts of the inversion domain. Thus, the entire residual accumulated along the path through the mantle beneath the Alps needs to be explained by heterogeneities in the uppermost part of the model. This is probably impossible and definitely undesired. Basically, the travel time residual of the rays incident form the side could be much more easily explained by heterogeneities outside the model domain and hence do only bring little information for the uppermost model parts. A remedy for this problem could either be to discard the rays from the side or to extend the inversion domain in EW-direction to allow heterogeneities there. They will not be interpretable because of the lack of ray crossing but at least they inhibit a mapping of the entire travel time residual they carry into the uppermost parts of the model. This problem could be one reason for the rather low misfit reduction of 66 percent. Restricting ray incidence to a cone around the NS direction as done for data set 2 may mitigate this problem. What is the rationale behind choosing a 60 degree cone for data set 2?

What is the damping factor in Fig. S4? Is it epsilon**2 from equation (4)? Why do the authors plot data variance instead of misfit which is normalized to the data uncertainties. What is model variance? Is it the squared norm of the velocity perturbations? A definition of this quantity should be provided. The authors should provide a value of misfit normalized to the picking uncertainties to get an impression whether there is overfitting or maybe even severe underfitting. Regarding Fig. S4. I wonder that the authors get increasing data variance and decreasing model variance for decreasing damping factor. I would expect exactly the opposite. This strange behaviour needs to be explained by the authors.

Why do the authors only consider 2 iterations? Did they try more iterations and how does the inversion behave then?

Results:

In the "Data" section two dataset were presented. Which one was used to create the model shown in the figures? Fig. S5 which compares vertical sections obtained from the 2 data sets is only mentioned once. The issue is never discussed in the results or the interpretative section.

Regarding the vertical sections in fig. 3, I recognize significant perturbations in the gray-shaded upper 50 km of the model which appears to be the crustal domain. As the effect of the crust was subtracted from the travel-time residuals, why do the authors still allow perturbations in the upper 50 km? Theoretically, after crustal correction, the travel-time residuals should represent pure mantle structure and perturbations in the crust should be suppressed. How would the model change if perturbations in the upper 50 km were forced to zero? In particular, what would happen with the model between 50 km and 100 km depth, just beneath the crustal domain? The interpretation of a northward dip strongly depends on the velocity perturbations between 50 km and 100 km depth. Below the high velocity anomalies are rather vertical. Whether HV-EA is delaminated or not, also depends on the anomalies in this depth range. All the conclusions following in the paper about polarity flip, northward dip and detachment of the Eastern Alpine slab depend on this issue.

For a connection to crustal levels, one could plot the perturbations of the crustal model relative to a 1D-model of the crust into the vertical cross sections. But then, what to do with the anomalies that are already there?

What is the criterion for gray-shading in the model domain? Is it derived from the resolution matrix?

Resolution tests:

I would also like to see the results from a checkerboard test (with gaps) which nicely shows lateral and vertical smearing. In particular vertical smearing at shallow levels should be investigated because it may hide a detachment of the slab or falsely connect the high vp anomaly below 100 km with a shallower one further to the south giving the impression of a significant northward dip.

The remark that the polarity flip is more or less accepted is certainly a misconception. If I read Paffrath et al. (2021) correctly, they see a rather vertical, detached eastern Alpine slab and favour the interpretation of European provenance because its down-dip length can only be explained by the Tertiary shortening in the Eastern Alps accommodated by south-dipping subduction of European lithosphere. Basically, what is seen in the tomography is only the slab dip but tomography does not tell us the provenance of the slab. Even if it were dipping clearly northwards (in the tomographies by Mitterbauer (P4), Zhao and Paffrath, it is nearly vertical) it could still be overturned European lithosphere. Independent data are needed to decide this issue. The use of the term "polarity flip", however, already implies the interpretation of Adriatic provenance of the slab.

Resolution test 1 also mimics the real pattern quite well below 100 km depth while test 3 misses the increasing dip angle of the real pattern with depth. My remark regarding the treatment of crustal structure also applies to the resolution test. Since crustal corrections were subtracted before inversion of the real data, resolution test data should also be free of crustal contributions and anomalies in the crustal domain should be forced to zero.

Imaging the high-velocity perturbations in different tomography model:

2nd paragraph: If I read Paffrath et al (2021) correctly they do not postulate a polarity flip for the Central Alpine slab. They state a steep SE dip of the slab and also do not associate it with Adria. This should be corrected.

In general, I find this section not well structured. The polarity flip issue is discussed in several paragraphs interrupted by a discussion of the 2nd high velocity anomaly underneath the BM. There seems to be no real ordering of thoughts and arguments. This part should be streamlined and rewritten in a concise and non-repetitive way with clear order and structure of arguments.

I am strongly worried by the resolution tests shown in Fig. S7. For each one of the detached slab test models there are strong smearing artifacts reaching up to the surface. The artifacts are strongest in the crustal layer and they mimic an either northward or southward dipping continuous slab. I cannot see that this resolution test proves that a detachment can be resolved by the inversion. It rather suggests the opposite. I wonder how the result of these tests would look like if crustal perturbations were forced to zero. Possibly, all artificial positive anomalies between 50 km and 100 km could become much stronger pretending a dipping high velocity feature.