Author Reply

Authors Name: Xian-QiongCheng , Qi-He Liu, Ping-Ping Li , Yuan Liu
Paper Name: Inverting Rayleigh surface wave velocities for crustal thickness in eastern Tibet and the western Yangtze craton based on deep learning neural networks

Revision Date: 5 /20/ 2018

Summary of Responses:

We thank the referee for his working for this paper, who has given many good suggestions, which we are incorporated in this revised work.

Below are the responses of work we have done.

Comments and Suggestions	Response
1. Inclusion of noise in training data The authors do not mention whether or not the synthetic training data contain noise. A neural network trained on noise-free synthetic data will perform very poorly on real data containing noise (e.g. Meier et al., 2007, figure 8b). If noise was included in the training data, the authors should describe this. Otherwise, it should be included.	We have trained our neural network on synthetic data with random noise and we have stated in the revision.
2. **Conversion from group velocity to phase velocity** The authors calculate group velocity from a published phase velocity map using the standard formula (their equation 4). However, including both phase and group velocity will only add new information if the phase and group velocity are measured independently (as is commonly the case). Therefore it is misleading to include the calculated group velocity in this paper. The group velocity data should be removed from the study or replaced by group velocity data measured independently. (Generally phase velocity is more sensitive to deeper structure so it is easier to infer deep structure from phase velocity measurements.)	We do not adopt the calculated group velocity and retrain our neural network on phase velocity only in the revision
3. **Benefit of deep neural network versus shallow neural network** A deep neural network is one with more than one hidden layer, whereas a shallow neural network has just one hidden layer. The additional complication of using a deep neural network is justified if the mapping has a hierarchical structure. For example, in image processing, it is common to move from the more elementary aspects of the input data (e.g. the values of the individual pixels) to intermediate parts (such as the distribution of edges) and finally to the most abstract aspects (such as the subject of the image). While it is undoubtedly true that the Earth has a hierarchical structure, ranging from individual grains to entire continents, the authors do not demonstrate that the dispersion data	We retrain our neural network and find that more hidden layers can get more lower test errors than shallow neural network does , which can be demonstrated in table 1

contain sufficiently complicated information to justify a deep neural network. The paper does not currently demonstrate that a deep neural network offers any improvement over a shallow neural network, such as that used by Meier et al.(2007). A comparison should be given.	
4. **Non-unique solutions** The authors focus on the non-linearity of the inverse problem, but they do not mention that it is also non-unique. Conventional optimisation of a neural network can lead to meaningless outputs for a non-unique mapping, as shown in figure 3b of Meier et al. (2007). Ideally, the method should be changed to solve for a probability distribution, for example using histogram or median networks (Devilee et al., 1999) or a mixture density network (Meier et al., 2007). Otherwise, the authors should attempt to quantify the range of non-uniqueness, or at least mention it in their discussion.	we have not considered about the uncertainty of crustal thickness which should be revealed by deep mixture density network in a probabilistic manner in our future work
5.**Unattributed quotations** Some explanatory sections are taken verbatim from other work, for example the paragraph beginning at 3.2:19 is identical to the second paragraph of section 3 of de Wit et al. (2014). These sections should be attributed, and either paraphrased or written in quotation marks.	In the revision we re-write in quotation marks on identical to paragraph of Wit et al. (2014)
6. **Meaning of 'data-driven'** It is misleading to say that the method 'data-driven' (e.g. lines 1:9–11). The inversion is model-driven; it is trained using a large number of synthetic data which are generated using a known forward mapping (in this case, the calculation of dispersion by normal mode summation). The role of the neural network is to approximate the inverse relation apparent in the synthetic dataset. The description 'data-driven' is appropriate when the forward mapping is not known (or not used). An example would be speech processing, where the meaning of a word cannot be calculated from its audio waveform.	Our manuscript aims at inverse problem, so meaning of data-driven in the manuscript is that we have no idea of inverse relationship , although the forward mapping is known. That is, we have no model describing how to infer crustal thickness from phase velocity. So we think this belongs to data-driven problem.
7. Lateral resolution of crust thickness Figures 7 and 8 show a comparison of the crust thickness model in this study with the crust thickness model in Xie et al. (2013). Although the two models are based on the same data, the result in this study appears to resolve much finer features. The authors should explain how this higher resolution is achieved and whether it is justified.	In the discussion we talk out our result resolve much finer features than other models, and these finer features revealed by our result is consistence with Wang et.al(2010) who attained the crustal thickness estimated by the H-k stacking method based on the broad band tele-seismic data. We think this higher resolution is achieved as deep sSAE works very well in learning useful high-level feature for better representation of input raw data.
3. Corrections to the writing There are some errors in the writing, but I have not listed them in detail, in the expectation that the body of the text will change	We check the English sentence by sentence and upload revised manuscript