Author Reply

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

Revision Date: 11 / 02/ 2016

Summary of Responses:

We thank anonymous 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**	**Page Reference (Origin)**	**Page Referred (New)**
1. **Fluency and precision of the text** Check the English please. Some parts of the manuscript are difficult to understand. A few examples: Page 1, lines 27-28 “As we all know, the more we know the characteristic and composition of crust which is an important part of lithosphere, the further we investigate deep earth”.Page 1, lines 36 What does “adjants” stands for? Page 1, lines 40-42 “in regions with good data coverage and uncomplicated structure but in regions with poor or no data coverage or complicated structure crustal thickness estimates are largely extrapolated” Those statements (and	We delete the sentence “As we all know, the more we know the characteristic and composition of crust which is an important part of lithosphere, the further we investigate deep earth”	Page 1, lines 27-28	
	“adjants” should be “adjacent regions”	Page 1, lines 36	Page 1, lines 34
	We rephrase “in regions with good data coverage and uncomplicated structure but in regions with poor or no data coverage or complicated structure crustal thicknessestimates are largely extrapolated”	Page 1, lines 40-42	Page 1, lines 38-40

several others in the manuscript) are not clear and would need some rephrasing. In addition, several typos and missing capitalization of letters in nouns should be checked carefully. Several sections of the text should be checked carefully. The consistency of the text is not always straightforward.	We check the English carefully and upload corrected manuscript by marked-up manuscript version with track changes in word showing the changes made		
2. **Bibliography** several aspects would require a more extensive referencing	We have extensive referencing added in references and marked-up manuscript version with track changes in word showing the changes made		
3. **Comments on the figures**			
(1) Table 1 provides many parameters but no unit is given.	Parameters in Table 1 have no units. Meanings of these parameters are illustrated in page 6,line 2~4.		
(2) Regarding the figures, I would suggest to merge Figures 1 and 2. Both describes the auto-encoder with one or two hidden layers, and could easily be merged. Similarly, Figures 3 and 4 could be merged. They both show crustal thickness from this study and from another model, used later for comparison.	In Figure 2, we use unsupervised learning techniques to pre-train at each layer and initialize parameters. After that, we adopt supervised learning techniques to train the whole network. However, Figure 1 illustrates how to use unsupervised learning techniques to initialize parameters, that is, we initialize the matrix W by reconstructing the input v of this network. Figure 3 describes our experimental results by using our method, while Figure 4 comes from Shapiro&Ritzwoller (2002). We can clearly compare the two type results by two figures.		
(3)In the text, several places are mentioned but are not located on any map, as the Wenchuan or Lushan earthquakes, the Longmenshan region, ...	We locate Wenchuan or Lushan earthquakes, the Longmen mountain and Sichuan-Yunnan block region in figure 3 and 4 in the revised version		
(4) in Figures 3 and 4, the caption doesn't mention that the blue lines are the boundaries of sedimentary basins.	We point out the blue lines are the boundaries of sedimentary basins in revised version		
4.Technical comments			

(1) the authors wrote (page 4, line 37) that the phase and group velocities of surface waves are not sensitive to similar depth layers. However, they extrapolates the phase velocities into group velocities, using the formula (4) , line 1-2, page 5. But the resulting periods for group and phase velocity dispersion curves are similar (10.0 - 30.0 mHz for group-velocity and 10.0 - 35.0 mHz for phase-velocity). This is not consistent with the period range described page 5, line 26, with Rayleigh phase velocities and group velocities (10 - 37.5 mhz). (Note that it should also be mHz and not mhz).	Our mistype period range described page 5, line 26		
(2) the authors only have constraints on Rayleigh-wave phase velocity (from Xie et al. 2013), that are derived for approximated group velocities, and joint inverted for crustal thickness. Why not using the Love-wave models in order to add different constraints to their inversion?	We aim to invert crustal thickness by using of dispersion data based on newly-developed neural network (that is deep learning neutral network). Simultaneously inverting Rayleigh-wave and Love-wave is our working in the future.		
(3)The authors used the phase-velocity model (from Xie et al. 2013) for periods between 33 and 100 s (10 - 30.0 mhz). But Rayleigh-wave phase velocities at periods of 33 s are mostly sensitive to depths of 30-80 km. In some regions (Sichuan Basin), the authors found some Moho	Where the stations are relatively sparse from Xie et al. 2013, the depth we attain is relatively shallow ,which demonstrate we need more observable data		

depths shallower than 30 km. Are those depths realistic?			
(4) Why don't the authors used periods of 8-40 s (25-125 mHz) as in Xie et al., (2013) to have additional constraints on the regions with shallow crust?	After tried many times we have found when we take higher frequencies as input variables the test errors of neural network are high		
(5) In the same way, Xie et al., (2013) used a grid of 0.5*0.5 _. Why do the authors downsample those maps to 2*2 _ (page 2, line 16)?	Meier (2007) invert surface wave data for global crustal thickness on a 2◦ × 2◦ grid globally using a neural network. This paper we still take a 0.5*0.5 grid as Xie et al.(2013) .		
5. **Comments on the method**			
(1)Another point that the author did not mention is how they invert the surface-waves velocities (phase and group) for Moho depth. They only mention (page 7, lines 21-25) that the dispersion curves are inverted for crustal thickness using 3 to 6 layers. Some additional information of the methodology seems needed. Did they used 1D, 2D or 3D sensitivity kernels? How did they defined the Moho discontinuity (velocity contrast, specific velocity, ...)?	When we can not show explicit functional relations , we take use of neural network to learn this function and show it by structure of neural network.		
	Our inversion based on neural network need not sensitivity kernels		
	Moho discontinuity is defined as velocity contrast.		
(2)The authors mention (page 7, line 35-36) "test error may be not the only criterion determining which neural network is best". So what are the other criteria that needed to be taken into account?	We also should take into account correlation coefficients as a criteria		