Segmentation of Common Conversion Point Stacks using Neural Networks

Lucas Sawade[1], Peter Makus[2], Ryan Manzuk[1], Stéphane Rondenay[3], Jessica C. E. Irving[4]

[1] Department of Geosciences, Princeton University [2] GFZ Potsdam [3] Department of Earth Science, University of Bergen [4] School of Earth Sciences, University of Bristol

Introduction

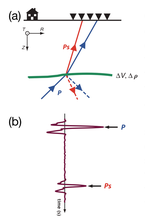

Figure 1: Cartoon showing the basic principle of teleseismic wave conversion below an array of stations. (a) The green line shows a perturbation in elastic properties, also referred to as a contrast in seismic impedance. In solid blue, the ray of an incoming, direct P-wave and in dashed blue its down-going reflection from the discontinuity in seismic properties. In solid red, the P-to-S converted wave and in dashed the downgoing S-wave reflected and converted at the discontinuity. (b) The corresponding arrivals recorded at stations of the array. Arrows pointing to the corresponding arrivals. Figure from Rondenay et al., 2017.

Converted teleseismic earthquake waves have been used for almost half a century to investigate Earth's sharp compositional discontinuities (Langston et al., 1977; Vinnik, 1977). During this time most of the focus rested on finding the exact location of the such discontinuities. However, different thermal and compositional differences can have a large effect on the transition itself. That is, width, intensity, and roughness could tell us more about thermal, chemical, and compositional structure of the upper mantle phase transitions. To assess this idea, we need to know the width of the subsurface discontinuities. Here, we explore the possibility of using a convolutional neural network to distinguish between discontinuity and surrounding mantle to analyze the discontinuities related to the mantle transition zone; i.e. the 410, which is the depth at which olivine transitions to wadsleyite, and the 660 which is the depth at which ringwoodite transitions to bridgmanite and ferropericlase. We are asking whether the discontinuities occupy different ranges in depth, do they have different characteristic mean amplitudes and how do these values change laterally? Do they vary at all? Or is the observed thickness solely controlled by filter width?

Dataset

Using PyGLImER (Makus et al., 2021), we first collected waveforms for a majority of the stations available in North America, rotated the waveforms to a P-Sv-Sh coordinate system, and computed receiver functions (RFs) using iterative time domain deconvolution with a Gaussian width parameter of 2.5 (Ligorría & Ammon, 1999). Then, we traced the P and S travel times using backwards ray-shooting in a 2-D plane through the GYPSUM 3D global P and S velocity model (Simmons et al., 2010), and computed the differential travel times. Using the travel times, we moveout corrected and stacked the traced receiver functions in circular bins spaced at 0.58° with radius 0.5° at their conversion points. An example slice through the resulting CCP volume is shown in Figure 2 on the left with accompanying map on the right.

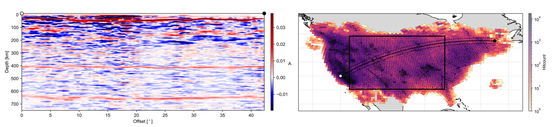

Figure 2: (Left) Vertical slice of through CCP Stack of radial receiver functions (RF). The colorbar indicates RF amplitude; where red corresponds to a downward increase in seismic impedance, often an increase in density and/or seismic velocity, and vice versa. The white and black dots at the top axis spine correspond to the dots indicated on the map on the right. (Right) Illumination map of P-S conversions at 410 km depth. The colorbar indicates the number of conversions that were collected in bins. Note that the map is interpolated and that the bins are not rectangular. The small circles show stations that provided RFs for the stack. The line across the continent indicates the path of the vertical slice. The vertical slice is interpolated using bins up to two bin distances away from the path; this buffer is indicated by the dashed line surrounding the path. The box indicates the data region of the CCP stack that we are using for the subsequent investigation.

CNN

Neural Networks are ubiquitous in our daily lives. They are used in, e.g., automated photo tagging on your mobile phone or autonomous cars to detect surrounding objects. Neural networks are generally modeled after the human brain and can be extremely powerful in detecting features in images after the user trains the network by giving it traced examples (the equivalent of teaching a human child to recognize objects with flash cards). The CCP dataset we have in our hands is nothing more than a stack of images in three dimensions with features (410 and 660) that are waiting to be detected. An example of feature detection in photos is shown in Figure 3.

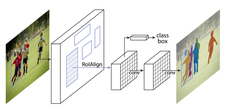

Figure 3: Example of a possible input and output of a Neural Network. In this case a more sophisticated approach "Mask R-CNN". Next to masks, a mask CNN also provides bounding boxes of individual features, which can be very powerful for overlapping objects. Figure from He et al., 2017.

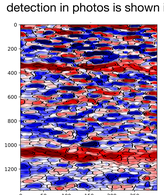

Figure 4: SLIC segmented CCP slice. The areas within black boundaries are selected for discontinuities and the remainder is assumed to be 'nothing'.

Dataset Labeling

We created a custom labeling software that makes use of 'super-pixels' that are found using the SLIC algorithm (Achanta et al. 2012) and are subsequently selected in a point-click fashion. We choose two labels: 'Discontinuity': 0 and 'Nothing': 1. In post-processing, the 410 and 660 discontinuity are separated using their depth ranges. Note that in Figure 4, the SLIC segmentation already does an incredible job at finding boundaries on a grayscale image — colormap is only for reference.

Training

We designed a simple two-layer convolutional neural network (CNN) that takes in three channel 101x101 pixel images and outputs a single class. We trained the CNN by taking 101x101 element x-, y-, and z-slices around a pixel in the CCP volume, which serve as the 3 channel inputs (see Figure 5). We use a class-balanced subset of 100,000 points from CCP volume to train the model. The training of the neural network is done on a single NVIDIA Tesla V100 and PyTorch. We used the Adam optimizer for this experiment with a weight decay value of 0.1, a learning rate of 0.0001, and cross-entropy loss. The training and validation evolutions are shown in Figure 6, left. The final filters of the first layer at the end of the training are shown in Figure 6, right.

CNN (cont.)

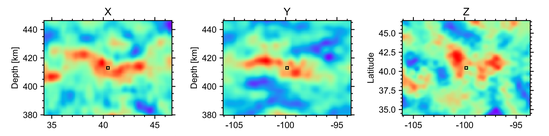

Figure 5: Example of three channel input for the CNN. Slices in each dimension (Lat, Lon, Z -> X, Y, Z). Here a slice around the 410 discontinuity. The black square indicates the location that is labeled by the CNN given these three input slices.

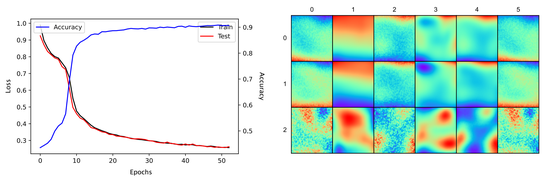

Figure 6: (Left) Model evolution. In black the training loss, and in red the loss associated with the held back data set. In blue the accuracy, i.e., fraction of correctly labeled data points. (Right) Final convolutional filters of the first layer in the neural network.

Results

Labeled Slices

The labeled volume is processed to remove spurious random classifications using common post processing routines such as opening, closing, remove small objects etc. The labeled discontinuities are also distinguished by their depth resulting in three masks: 410, 660, and 'nothing' (s. Figure 7).

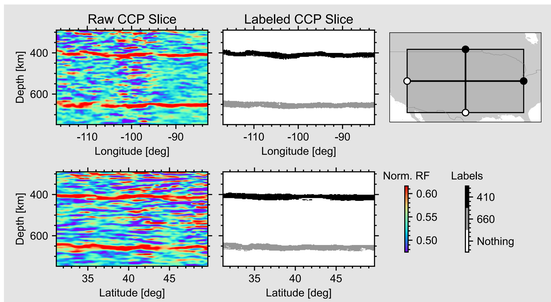

Figure 7: From Raw to labeled CCP slice. (Left column) Raw CCP slices with normalized amplitudes. Slice locations are indicated in the map in the top-right. (Center Column) Labeled CCP Slices. Colorbar for Norm RF and label are shown in the bottom right corner of the figure.

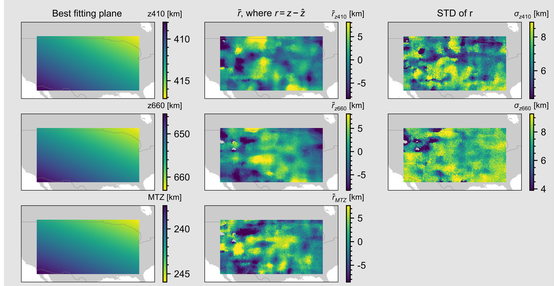

Figure 8: (Left column) Using the coordinates of the mask we can fit planes to the 410 and 660 discontinuity and compute corresponding MTZ thickness. (Center column) We subtract the fitted planes from the masked z coordinates to compute the mean vertical residual. In other words, positive values indicate the actual discontinuity is below the plane approximation and the standard deviation as for the horizontal coordinates. For the thickness, a positive value indicates that the actual MTZ is thicker than the ideal plane solution. (Right column) Standard deviations of the center columns' residuals.

Results (cont.)

Best fitting planar MTZ

In Figure 7, we fit a plane to the coordinates of the two mantle transition zone discontinuities within the masks and compute the vertical residuals and their distributions. According to planar fit, the MTZ, on average becomes thicker from south west to north east. The 410 and 660 both deepen northwestward. The 660 deepens at a higher rate than the 410 discontinuity leading. It is important to note that especially in the central part of the depth slices the plane approximation is not adequate to capture the strong relief of the discontinuities; a valley in both phase transitions stretches from north west to south within the plane. Both 410 and 660 are deeper in that region and the MTZ appears to be thicker. The "valley" in both discontinuities is clearly visibly in the 2D slice in Figure 9. While the ideal plane solutions for absolute depths of the 410 and 660 discontinuities look fairly different from trends shown in a previous study (Gao, 2014), the trend in transition zone thickness matches. This discrepancy in absolute depth is an effect of the updates to the upper mantle velocity structure from the GyPSuM model (Simmons et al., 2010). This effect is studied extensively in Burky et al., 2021 (see here: Dec 13 16-18.00, Board 0005, Poster Hall D-F).

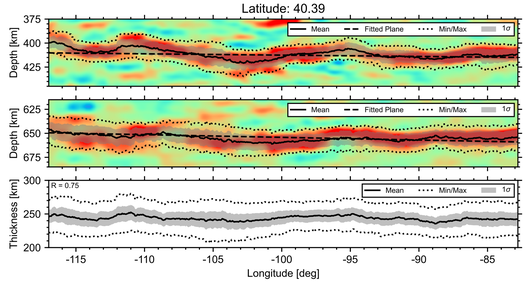

Figure 9: A latitude slice through the CCP volume at 40.39°N, with statistics of the masked depth values. (Top) 410 discontinuity, (Center) 660 discontinuity, (Bottom) Thickness of the mantle transition zone. The standard deviation is the combined standard deviation of 410, and 660, assuming that they are distributed independently (which is most likely not true). The R value in the thickness plot is the correlation coefficient between 410 and 660 mean depths.

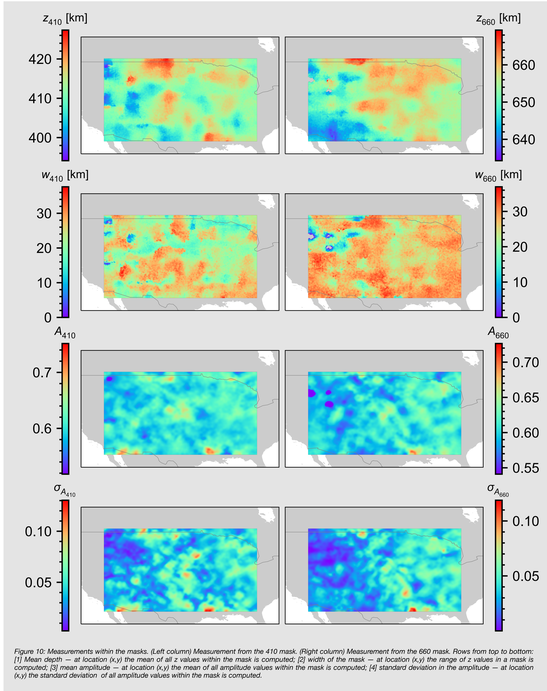

Figure 10: Measurements within the masks. (Left column) Measurement from the 410 mask. (Right column) Measurement from the 660 mask. Rows from top to bottom: [1] Mean depth — at location (x,y) the mean of all z values within the mask is computed; [2] width of the mask — at location (x,y) the range of z values in a mask is computed; [3] mean amplitude — at location (x,y) the mean of all amplitude values within the mask is computed; [4] standard deviation in the amplitude — at location (x,y) the standard deviation of all amplitude values within the mask is computed.

Results (cont.)

The widths of the discontinuities show small regional trends (Figure 10, second row). The 410 is on average thinner in the north east than the south west. A similar trend does not seem to exist for the 660 discontinuity, which only shows some anomalously thin locations in the north west. The average pulse width of the 410 is 23.43±4.42 km and slightly thicker for the 660 at 26.18±4.32 km. *Disclaimer*: These are *not* the actual thickness of the transitions, but rather the convolution of its width with the RF's low pass filter. This is by no means a direct measure of the transitions thickness but rather a first order tool to investigate lateral variations. The average amplitude (Figure 10, third row) in the 410 discontinuity does not show and significant regional trends, whereas the 660 seems to increase in mean amplitude from south west to north east. With the increase in amplitude, however, the standard deviation of the amplitude (Figure 10, fourth row) is increasing as well. The transition from low to high standard deviation seems to coincide with transition from shallow to deep 660. At this point, we are not certain whether this is a results of the mask that is found by the CNN or, indeed, the structural and compositional effect — further investigation is required.

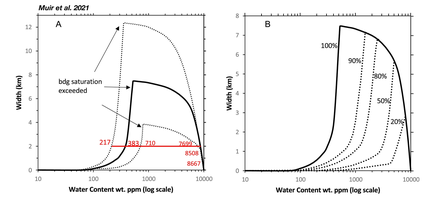

Figure 11: Effects of water content on the width of 660 discontinuity from Density Functional Theory. A wider transition could have implications on the water content in the transition zone. From Muir et al. 2021

Additionally, we find that the mean amplitude correlates positively with the depth of the 660 discontinuity (R = 0.88). Whether this is a regional effect or related to the path traveled or a result of the mask itself remains to be investigated.

Conclusions

Summary

- Using a simple convolutional neural network we detected the mantle transition zone discontinuities related to olivine phase transitions, and their width.
- We subsequently made measurements within the masks of the of the discontinuities finding
  - An average 410-depth of the 412.4±4.4 km
  - An average 660-depth of the 654.1±5.5 km
  - Both discontinuities show similar mean amplitude although the mean amplitude of the 410 is slightly higher with a larger spread:
    - 410 — 0.6052±0.0230 A;
    - 660 — 0.6047±0.0193 A.
  - We found a positive correlation between the depth of the 660 and its mean amplitude in the found mask correlate. At this point, we are not certain whether this is expected.
  - The standard deviation in amplitude measurement increases abruptly from south west to north east.

Outlook

More investigation is needed and a higher accuracy CNN required to make clear deductions. Plans for improvement:

- Pre-trained models for cross-learning the features present in the CCP slices
- Explore learning methods (not only 2D CNN, but R-CNN etc., 1D convolutional methods)
- Change RF processing techniques
- Instead of just RF slices additional data can be generated by processing the input RFs slices, say to 2D Fourier space etc.

References

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. and Süsstrunk, S. SLIC Superpixels Compared to State-of-the-art Superpixel Methods, *TPAMI*, May 2012.
- Burky, A. L., Irving, J. C.E, and Simons, F. J. (2021), Receiver Function Imaging of the Mantle Transition Zone beneath Eastern North America [DI15A-0005] presented at 2021 Fall Meeting, AGU, 6-17 Dec.
- Gao, S. S., and Liu, K. H. (2014), Mantle transition zone discontinuities beneath the contiguous United States, *J. Geophys. Res. Solid Earth*, 119, 6452– 6468.
- He, K., Gkioxari, G., Dollár, P. ,and Girshick, R. (2018). Mask R-CNN. *arXiv*. 1703.06870.
- Makus, P., Rondenay, S., Sawade, L., Ottemöller, L., and Halpaap, F. (2021). PyGLImER: A New Modular Software Suite to Image Crustal and Upper-Mantle Discontinuities Using a Global Database of Ps and Sp Receiver Functions. *Earth and Space Science Open Archive AID*.
- Muir, J. M.R., Zhang, F., and Brodholt, J. P. (2021). The effect of water on the post-spinel transition and evidence for extreme water contents at the bottom of the transition zone, *Earth and Planetary Science Letters*, Volume 565, 116909.
- Langston, C. A., and Burdick, L. J. (1977). Modeling crustal structure through the use of converted phases in teleseismic body-wave forms. *Bulletin of the Seismological Society of America*; 67 (3): 677–691.
- Ligorría, J. and Ammon, C. (1999). Iterative deconvolution and receiver function estimation. *Bulletin of Seismological Society of America*, 89, 1395-1400.
- Rondenay, S., K. Spieker, L. Sawade, F. Halpaap, and M. Farestveit (2017). GLImER: A New Global Database of Teleseismic Receiver Functions for Imaging Earth Structure. *Seismological Research Letters* 88(1), 39–48.
- Simmons, N. A., Forte, A. M., Boschi, L., and Grand, S. P. (2010). GyPSuM: A joint tomographic model of mantle density and seismic wave speeds. *Journal of Geophysical Research*, 115 (B12), B12310.
- Vinnik, L. P. (1977). Detection of waves converted from P to SV in the mantle. *Physics of the Earth and Planetary Interiors*, 15 (1), 39–45.

Abstract ID: S15D-0271 — CONTACT: lsawade@princeton.edu — GITHUB: lsawade