



**HAL**  
open science

# Automatic Identification of Mantle Seismic Phases Using a Convolutional Neural Network

J. Garcia, L. Waszek, Benoit Tauzin, N. Schmerr

► **To cite this version:**

J. Garcia, L. Waszek, Benoit Tauzin, N. Schmerr. Automatic Identification of Mantle Seismic Phases Using a Convolutional Neural Network. *Geophysical Research Letters*, 2021, 48 (18), 10.1029/2020GL091658 . hal-03350921

**HAL Id: hal-03350921**

**<https://hal.science/hal-03350921v1>**

Submitted on 12 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Geophysical Research Letters®

## RESEARCH LETTER

10.1029/2020GL091658

### Key Points:

- We train a one-dimensional (1D) Convolutional Neural Network to identify the arrival peaks of SS phases from a large data set of 58,567 handpicked waveforms
- The model is used to predict the arrival time of SS precursors relative to SS in stacked data and individual seismograms
- New maps of the 410 and 660-km discontinuities are generated using the model, and show excellent agreement to handpicked precursor maps

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

J. A. Garcia and L. Waszek,  
[jorgeagr97@gmail.com](mailto:jorgeagr97@gmail.com);  
[lauren.waszek@cantab.net](mailto:lauren.waszek@cantab.net)





### Citation:

Garcia, J. A., Waszek, L., Tauzin, B., & Schmerr, N. (2021). Automatic identification of mantle seismic phases using a Convolutional Neural Network. *Geophysical Research Letters*, 48, e2020GL091658. <https://doi.org/10.1029/2020GL091658>

Received 19 NOV 2020

Accepted 19 AUG 2021

## Automatic Identification of Mantle Seismic Phases Using a Convolutional Neural Network

J. A. Garcia<sup>1</sup> , L. Waszek<sup>1,2</sup> , B. Tauzin<sup>3,4</sup> , and N. Schmerr<sup>5</sup> 

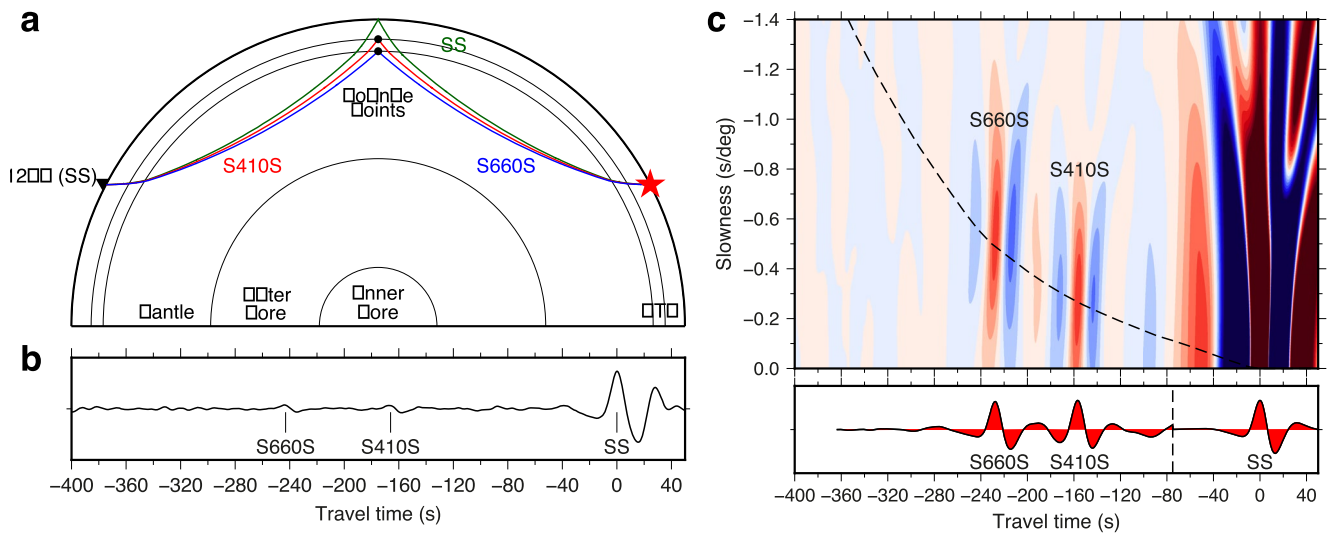
<sup>1</sup>Department of Physics, New Mexico State University, Las Cruces, NM, USA, <sup>2</sup>Physical Sciences, James Cook University, Douglas, QLD, Australia, <sup>3</sup>Laboratoire de Geologie de Lyon: Terre, Planètes, Environnement, Université de Lyon, Université Lyon 1 and Ecole Normale Supérieure de Lyon, UMR CNRS 5276, Villeurbanne, France, <sup>4</sup>Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia, <sup>5</sup>Department of Geology, University of Maryland, College Park, MD, USA

**Abstract** Typical seismic waveform data sets comprise hundreds of thousands to millions of records. Compilation is performed by time-consuming handpicking of phase arrival times, or signal processing algorithms such as cross-correlation. The latter generally underperform compared to handpicking. However, differences in picking methods creates variations in models and interpretation of Earth's structure. Here, we exploit the pattern recognition capabilities of Convolutional Neural Networks (CNN). Using a large handpicked data set, we train a CNN model to identify the seismic shear phase SS. This accelerates, automates, and makes consistent data compilation, a task usually completed by visual inspection and influenced by scientists' choices. The CNN model is employed to identify precursors to SS generated by mantle discontinuities. It identifies precursors in stacked and individual seismograms, producing new measurements of the mantle transition zone with quality comparable to handpicked data. This rapid acquisition of high-quality observations has implications for automation of future seismic tomography studies.

**Plain Language Summary** Energy from seismic events such as earthquakes propagates through the Earth as waves, which are detected as waveform signals at a global network of instruments. The travel times of these waves inform regarding the properties of the material they traverse, allowing seismologists to generate models of Earth's interior. Typical data sets of these measurements total hundreds of thousands. Compilation is performed by either manually handpicking the arrival times, or by signal processing algorithms. The latter generally underperform in comparison to handpicking, however inconsistencies between handpicked data sets create disagreement between models of Earth's structure. Here, we use machine learning methods to train a computational model to identify seismic waveforms. We use a large handpicked data set of shear waves "SS," which travel through the mantle and reflect once from Earth's surface. Precursors to SS are generated by discontinuities beneath Earth's surface, and their travel times relative to SS are linked to discontinuity depths. We train a model to identify SS signals, and find that it can also identify SS precursors. This capability to rapidly obtain new measurements with quality comparable to handpicked data has implications for automation of the development of new Earth models.

## 1. Introduction

Seismology is the major observational tool to map the structure and properties of Earth's interior. Global studies of the Earth benefit from hundreds of thousands of seismograms to make observations. The properties of seismic wave phase arrivals within seismograms (arrival time, amplitude, coda) provide measurements of Earth's velocity and attenuation structures. Although some studies use automated waveform processing to identify seismic phases (e.g., Chambers et al., 2005; Earle & Shearer, 1994; Houser et al., 2008), visual inspection of waveforms is commonly employed due to higher accuracy (e.g., Deuss, 2009; Flanagan & Shearer, 1998; Schmerr & Garnero, 2006; Waszek et al., 2018). However, handpicking is time-consuming, and susceptible to decisions and potential inconsistencies of the scientist. Different compilation methods produce a range of data sets resulting in variable geophysical models, as evidenced by variations in global mantle discontinuity topography maps obtained from the same data types (e.g., Deuss, 2009; Flanagan & Shearer, 1998; Huang et al., 2019; Schmerr & Garnero, 2006; Waszek et al., 2021).



**Figure 1.** (a) Ray paths of SS and its precursors, S410S and S660S. The red star denotes the location of the seismic event, and the black triangle a station to detect seismic waves. (b) An example of a high-quality seismogram showing the SS, S410S, and S660S arrivals. (c) Global vespagram stack for all data and cross-section through the theoretical relative precursor time and slowness. The precursor amplitudes have been magnified and normalized to the SS phase amplitude; magnification factor is typically around 30 (for the global stack here, it is 30.35).

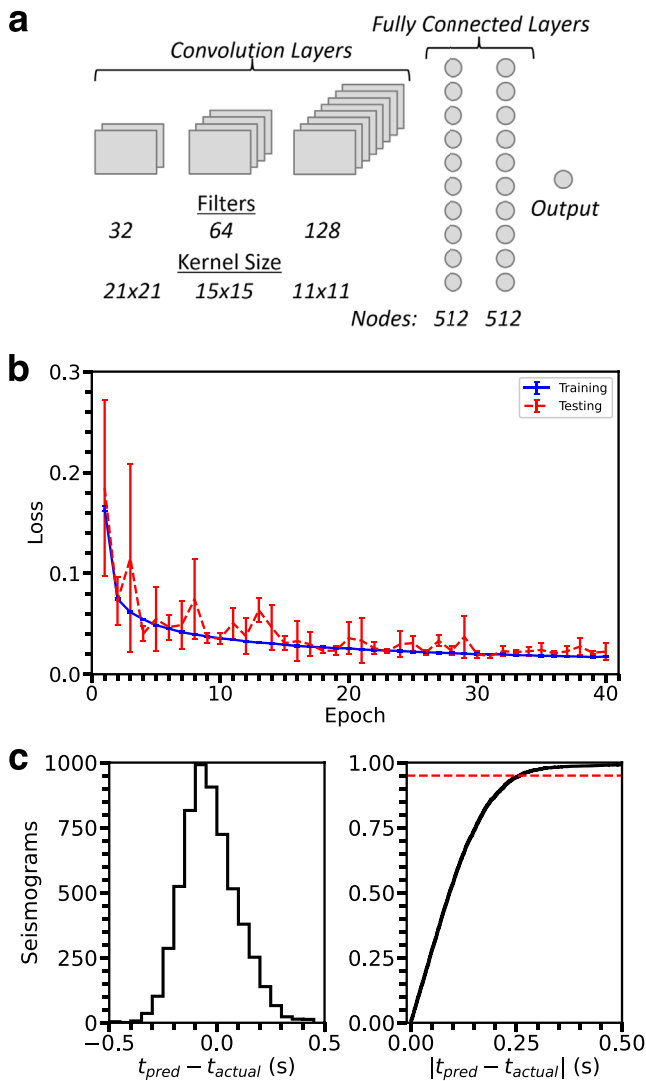
The task of picking seismic phases relies upon visual cues, in order to identify true detections amongst the background noise in seismograms, and distinguish the correct portion of the signal. Naturally, it is easier for the human eye to discern the correct peak associated with a particular seismic phase when trained to do so. This is the inspiration for the use of a Convolutional Neural Network (CNN) to perform this task, finding the necessary patterns through representation learning (LeCun et al., 2015). The use of deep learning trains a system that is capable of taking seismic data, identifying characterizing features of the waveforms, and producing an informed selection of arrival time and signal quality based on these signatures.

CNNs are the preferred deep learning algorithm for pattern recognition problems due to their ability to identify any set of objects given enough layers (Girshick et al., 2014; Krizhevsky et al., 2017; Simonyan & Zisserman, 2014). Within seismology, CNNs have proven capable of detecting and locating earthquakes (Perol et al., 2018), performing seismic arrival labeling (McBrearty et al., 2019), denoising data (Zhu et al., 2019), and picking arrival times of seismic phases (Ross et al., 2018; Zhu & Beroza, 2018).

Here, we apply CNNs to make new observations of seismic phases sensitive to mantle discontinuities. The two major global discontinuities at 410 and 660-km depth (“410,” “660”) bound the mantle transition zone (“MTZ”). They result from mineral phase transitions in olivine as pressure and temperature increase with depth (Ito & Takahashi, 1989; Katsura & Ito, 1989). Due to their opposing Clapeyron slopes, the discontinuities’ depths respond oppositely to temperature. In cold regions the 410 becomes shallower and the 660 becomes deeper; vice versa in hot regions. Consequently, their separation is a first order thermometer for the MTZ.

Mapping of mantle discontinuities globally has been achieved through measuring shear-wave reflections from underneath these boundaries (e.g., Deuss, 2009; Flanagan & Shearer, 1998; Houser et al., 2008; Huang et al., 2019; Shearer, 1993; Waszek et al., 2018). SS is a seismic shear wave phase with two legs in the mantle and one reflection from Earth’s surface (Figure 1a). Reflections from mantle discontinuities generate precursors to SS (“SdS,” where  $d$  is discontinuity depth), which arrive prior to the main phase. The SdS-SS travel time difference provide information regarding the discontinuity depth.

We use a CNN to train a model capable of identifying SS in seismograms. We use a large handpicked global data set of 58,567 SS data (Waszek et al., 2018), implementing a duplication and time-shift procedure to produce huge amounts of training data (316,262), described in detail below. Using the trained model, a scanning algorithm quantifies the quality of a phase signal within a waveform. We then employ the algorithm to output the arrival times and quality of SS precursors, in both stacked and individual seismograms.



**Figure 2.** (a) Diagram of the “RossNet” architecture used, from Ross et al. (2018). (b) Average loss of five models from the architecture used across epochs during the training stage (blue line) and testing stage (red line). These represent the error in fit of each model to the data. The error bars correspond to one standard deviation of the average loss of the models on the unseen testing set. The instance that results in the lowest overall loss is the set of weights used. (c) Histogram of prediction error (left) and cumulative histogram of absolute error (right) for the testing set; the red dashed line represents the 95th percentile of the data.

Maps of the depths of the 410 and 660 are generated, using the predictions to evaluate model performance. The study provides a new method to rapidly and automatically compile large high-quality seismic data sets and measurements, with implications for future work, particularly global tomography.

## 2. Seismic Data and Processing

Our study employs a large, handpicked data set of 58,567 SS waveforms (Waszek et al., 2018), aligned at the maximum peak (Figure 1b), and normalized to the SS amplitude. The seismograms are corrected for mantle and crustal structure using S40RTS (Ritsema et al., 2011) and Crust2.0 (Bassin et al., 2000), and filtered between 15–50 s. Our sampling rate is 0.1 Hz. A full description of processing methods is provided in Waszek et al. (2018, 2021).

Precursors to SS are typically too small in amplitude to identify on individual seismograms. Instead, the data are stacked in regional overlapping spherical caps partitioned by common reflection points, weighted by signal-to-noise ratio, where noise is the root-mean-square amplitude in the precursor window (–400 to –100 s). Vespagrams show stacked signals as a function of travel time and slowness relative to the main SS arrival (Figure 1c) (Davies et al., 1971). The cross-section taken along the dotted line is the predicted time and slowness of the precursors to SS in a standard reference model. Vespagrams are usually analyzed manually to measure the time and amplitude of the precursor signals. Here, we partition our data according to their bounce points in radii of 5°, 7.5°, 10°, and 15°. The different sizes account for heterogenous data coverage: smaller bin sizes in areas of higher data density provide greater resolution.

## 3. Model Training and Evaluation

The SS data set was divided into a training set of 90% of the seismograms, with the remaining 10% left as an unseen testing set to evaluate the model. For training, we used data uncorrected for crustal and mantle structure. Following the methodology of Ross et al. (2018), a small window (40 s) of the original waveform (500 s) was considered for the model input, starting at the theoretical onset time predicted by the 1D Earth model PREM (Dziewonski & Anderson, 1981). This smaller window represents the best return for minimizing computation time while maximizing pick accuracy when training the network (Figure S1). Additionally, it enabled us to augment the number of training records by creating variations of these segments, to obtain a more accurate model.

For each 40 s segment, we created five duplicate windows and applied a random time-shift across a uniform distribution of  $\pm 5$  s, increasing the training set by a factor of six. Although they are the same waveform shifted, to the CNN they appear as independent and different signals. This random shift allows the model to account for the variability of the location of the peak within a time window, thereby enhancing the spatial invariance of the model. For the testing set, only the original 40 s window was used.

We used the augmented data set to train a 1D CNN through the Keras library (Chollet et al., 2015), using the “RossNet” model architecture (Ross et al., 2018). The overall configuration of the layers is visualized in Figure 2a. The ReLU activation function (Nair & Hinton, 2010) was used in both the convolutional and fully connected layers. Model cost was evaluated with the Huber loss function (Huber, 1964), and the Adam

algorithm used for layer weight optimization (Kingma & Ba, 2014). In order to account for variations in model convergence due to random initialization of weights, we trained five different models for 40 epochs. The models were trained on two NVIDIA Tesla P100 graphics processing units (GPUs); each epoch took approximately three minutes to train.

Figure 2b shows the average and standard deviation of the loss over the models and epochs; this corresponds to the error in fit of the models to the data. Results are shown for the training and test data sets. Despite the variability of the errors in the testing data set, likely due to its comparatively small size, there is an overall trend of decreasing loss with increasing epoch. We use the best performing set of weights across all 200 instances for the remaining analysis. Summary statistics show an average prediction error on par with the sampling rate of the seismograms (0.1 s, Figure 2c). The cumulative histogram confirms that 95% of the model's prediction for the test data set are within 0.25 s of the picked arrival. This size of error is insignificant, since the picks are subsequently aligned to the maximum amplitude.

#### 4. Phase Waveform Quality

We desire not only the maximum arrival time of each waveform, but also the quality of the phase. Quality of the waveform during handpicking is normally judged visually in a qualitative manner. Here we propose a scheme to assign a quantitative description of the quality using the model.

The CNN model was constrained to accept only 40 s time windows as its input. Thus, we create a scanning algorithm that iteratively moves along the entire 500 s seismogram in 40 s windows to define the prediction quality through a statistical definition. The top three plots in Figure 3 provide an illustrative example of the scan algorithm. A 40 s window of data from time  $t$  to  $t + 40$  s is chosen and provided as input to the model to find the best matching shape to the ideal SS signal, giving an arrival time prediction for this window. The window from  $t + \Delta t$  to  $t + \Delta t + 40$  s is then analyzed; this process is repeated for the entire seismogram. The sliding window moves in steps of the seismograms' sampling rate, that is,  $\Delta t = 0.1$  s, requiring that both training and input data sets have the same sampling rate. As the scanning iterates, the SS arrival time will be identified consistently if enclosed in the windows. In some cases the model identifies phases outside of the window (Movie S1), as it recognizes the characteristics of the phase arrival, and subsequently predicts the location of the peak. If no discernible features are present, the best prediction time varies considerably as the scan iterates.

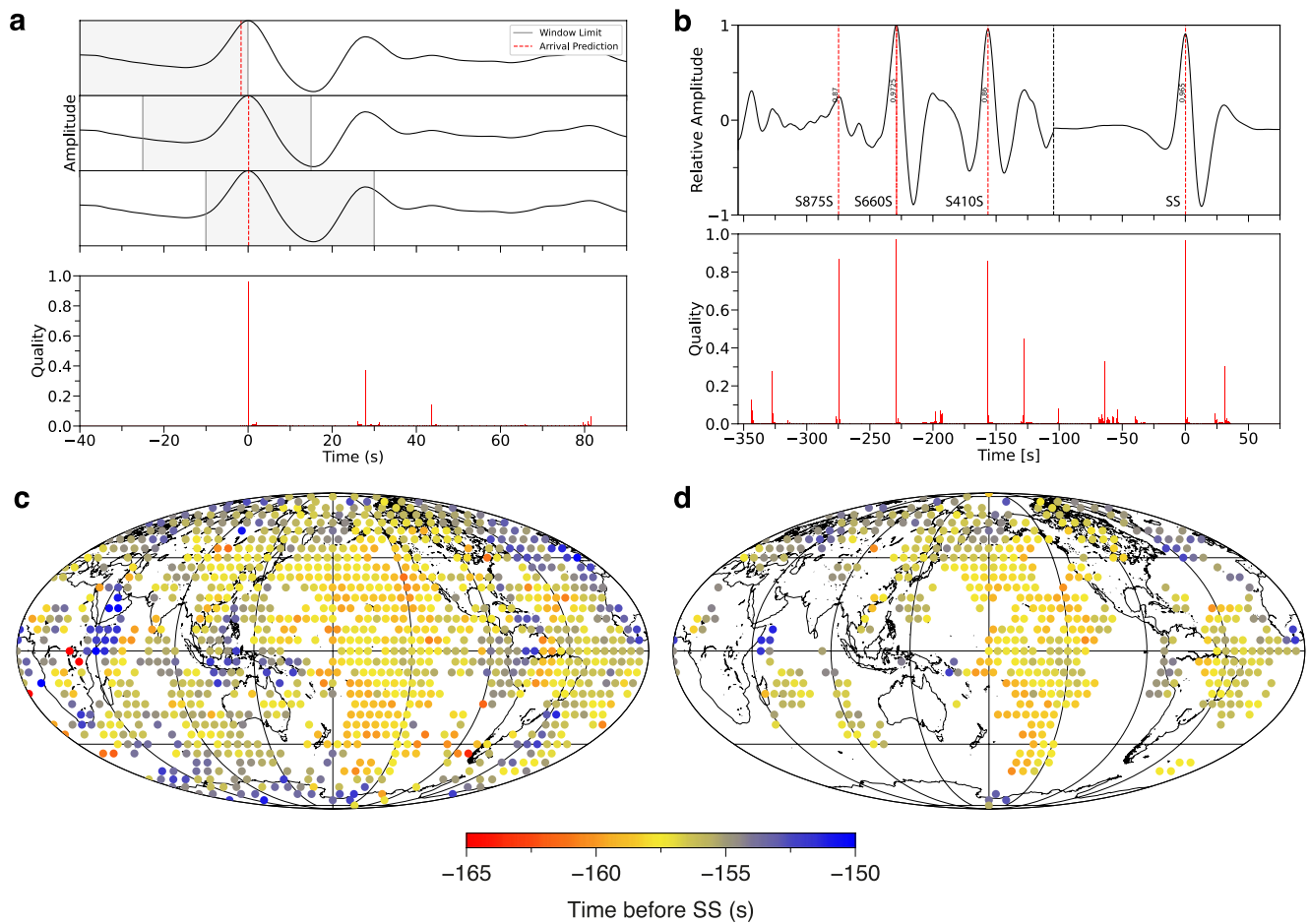
The obtained prediction times for a particular signal are not precisely the same throughout the scan. Due to slight differences in information within each window, the predicted time will vary by a value close to the sampling time of the data. We employ the DBSCAN algorithm (Ester et al., 1996) implementation in the Python scikit-learn library (Pedregosa et al., 2011) to perform density-based clustering of the predicted times. This way, a large amount of predictions that are close to each other form a tight cluster. Each prediction in a cluster is an approximate measure of a time  $\langle t \rangle$  of the signal, with a standard deviation corresponding to the error on the prediction  $\epsilon$ .

We use the quantity of predictions to define a quality measure for each signal. Let  $T$  be the window size used in the model, and  $\Delta t$  the sampling time of the seismogram. An ideal arrival will therefore appear  $T/\Delta t = 400$  times during the scanning process. The quality of a prediction  $q_{\text{pick}}$  is thus calculated:

$$q_{\text{pick}} = \frac{N_{\text{pred}} \Delta t}{T} \quad (1)$$

with  $N_{\text{pred}}$  the number of predictions within a cluster. We retain the prediction with the highest prediction frequency, or quality, as the SS "pick" for a particular seismogram. Correct identifications of SS result in higher quality of the main arrival compared to other features within the waveform (Figure 3a).

This scheme of defining a quality also allows us to determine the correct polarity of the SS signal. Since the model is only trained on seismograms with positive polarity SS signals, running the scanning window on a seismogram with a negative polarity SS peak results in inconsistent predictions with lower quality around the time of the SS arrival (Movie S2). In order to determine the polarity of an unknown seismogram, we employ the scanning algorithm on both the waveform and its inverse. For seismograms with an identifiable SS signal, the version with a positive polarity SS phase has the highest quality pick. Movie S2 also highlights



**Figure 3.** (a) Example of iterative prediction for a seismogram, at 0, 15, and 30 s, with the histogram of prediction rate. The portion of the waveform within the shaded area is used as input for the model. The red line is the predicted arrival for the given input. The signal is predicted consistently when enclosed by the window, and the true SS arrival is at the time of highest prediction rate. (b) Prediction of arrival times in a stacked cross-section and histogram. The four highest prediction times are marked on the cross-section. (c) S410S precursor arrival times for 5° bin stacked data picked using the deep learning model. The minimum prediction quality of picks retained is 60%. (d) Corresponding S410S arrival times measured using handpicking and visual quality checks.

the complexity in the characteristics of the SS signal onset, which is problematic for autopicking procedures; here, we remove some of these complications by training the code to identify the maximum amplitude.

## 5. Prediction of SS Precursors in Stacked Data

Precursors (“SdS”) may be approximated as lower amplitude versions of the main arrival with a similar shape. Thus, a model trained on the main arrival should be able to identify precursory signals in stacked waveforms due to their similarity, exploiting the pattern recognition capabilities of CNNs. We find that our scanning algorithm does indeed identify precursors as the highest quality predictions prior to the SS arrival (Figure 3b; Movie S3).

The handpicking quality criteria requires clear S410S and S660S signals in both the vespagrams and cross-sections, with no interfering phases or significant noise in the vespagram. The vespagrams are assigned qualities from “a” to “d.” The “a” vespagrams have no noise and clear precursors with waveforms very similar to SS, whereas “d” bins have much noise and the precursor shape is dissimilar to SS, and are not retained for analysis of any precursors other than S410S and S660S (see Waszek et al. (2018) for full methodology). Here, we use the CNN to obtain predictions of the S410S and S660S times for all of the stacks from the bins (corrected for 3D mantle and crustal structure). We retain picks with quality 0.6 or higher; following visual inspection, this is the lowest quality for which precursors could be identified (Figure S2).

The resulting maps of S410S arrival times for 5° bins show good agreement in the measurements from the CNN (Figure 3c) and handpicking (Figure 3d), with a correlation coefficient of 0.999. This indicates that, where both methods retain a bin, they measure the same relative arrival time for the precursor. This is true for both S410S and S660S picks in all bin sizes (Figure S3).

The CNN picks retain significantly more bins, which were removed by the handpicking quality procedure. The higher retention rates for the CNN is found for S410S and S660S measurements in all bin sizes (Figures S4–S11). This suggests that stricter visual quality procedures may remove useful information, meaning that the CNN can identify seismic signals in noisy data whereas handpicking cannot. Furthermore, the CNN model provides numerical measures of quality that the handpicking does not. Average quality of handpicked versus autopicked bins confirms that the bins removed by the CNN are indeed of lower quality than those retained by handpicking (Table S1). Furthermore, the average CNN quality also corresponds well to the handpicked quality, i.e., “a” quality bins have the highest CNN quality (Table S2). In order for the CNN method to retain the same number of bins as the handpicking, the minimum pick quality must be increased to as much as 0.86 for S660S in 5° bins (Table S3). This value drops as bin size increases, to 0.6125 for S410S in 15°, as the stacked signals become less similar to SS due to averaging over increasingly larger regions.

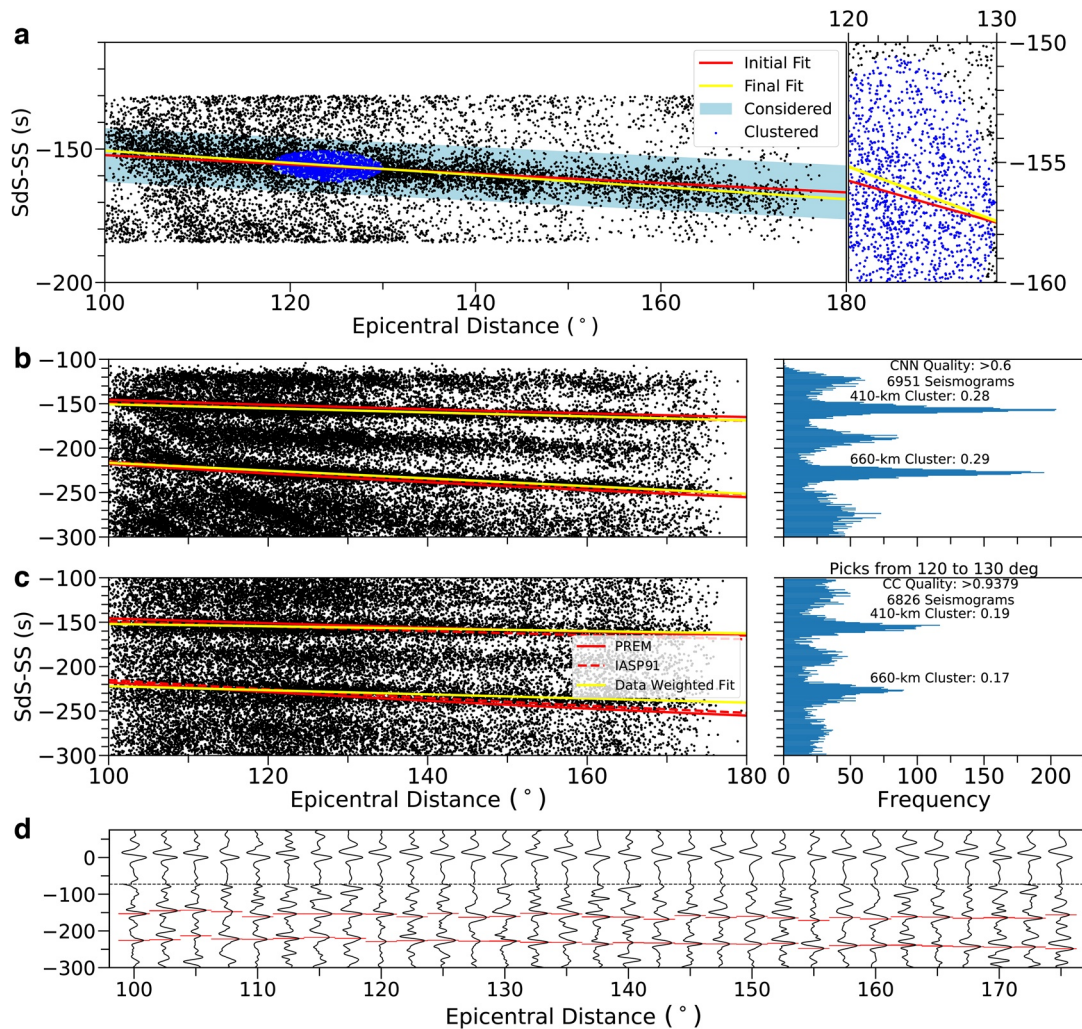
## 6. Prediction of SS Precursors in Individual Seismograms

Following the success of the CNN model for identifying precursors in the stacked data, we next apply it to precursors in individual seismograms. Normally, these can only be visually identified in the highest-quality waveforms due to their small amplitudes (e.g., Figure 1a). We scan the corrected data set, and consider the top 10 picks before the main arrival (Movie S4). Predictions with a quality below 0.6 are discarded, retaining a total of 38,985 measurements. This corresponds to multiple picks in some seismograms, and none in others. Examining the predictions as a function of epicentral distance (Figure 4a) reveals clusters corresponding to the 410 and 660, in addition to regional-scale discontinuities at 300-km and 520-km depth. The gaps with different slowness to the precursors (particularly between 100 and 120° distance) are interfering phases that the model does not pick, namely SdiffS660S which has a negative polarity, highlighting the CNN's ability to discard non-SdS signals.

The linear trends for both global discontinuities are calculated using the DBSCAN algorithm for density-based clustering, to determine statistically the predictions most likely to correspond to S410S and S660S based on density of picks. We select arrival time bounds of –185 to –135 s before the main arrival for S410S, and –250 to –200 s for S660S. These are selected to fully enclose the observed data trends, while excluding theoretical arrival times for other discontinuities, to ensure that the most dense cluster corresponds to robust picks. A linear fit, using the data identified by DBSCAN weighted by pick quality, is applied as an initial estimate for the trends. Subsequently, predictions within  $\pm 10$  s of this fit are considered to also be correct measurements for the discontinuity in question. The weighted linear model is then fit to this new set of data points (Figure 4a) and the picks updated to within  $\pm 10$  s of these new fits. The best fit for the 410 is  $t = -0.227\Delta - 127.925$ , and for the 660 is  $t = -0.437\Delta - 172.787$ , where  $t$  is time and  $\Delta$  is epicentral distance. Maps of the uncorrected and corrected relative travel time measurements are included in the Supplement (Figures S12–S14).

## 7. Discussion

The task of pattern recognition in seismology is not new. Cross-correlation has previously been used to generate SS data sets (Houser et al., 2008), measure precursor arrival times in stacked data for the mid-mantle (Waszek et al., 2018), and identify precursor signals in individual data (Shearer, 1991). It performs well when the two signals are noise and defect-free, but the majority of real data does not fulfill these criteria. Hence, a great deal of information is discarded since the automation procedure cannot discern between what might be defective data and hard-to-pick data. Setting the cross-correlation approach as our benchmark, we repeat the clustering analysis to identify 410 and 660 measurements from cross-correlation predictions, and compare to the CNN picks. A cutoff cross-correlation score of 0.9379 is required to obtain an equal number of precursor signals when using the cross-correlation method as compared to the CNN model



**Figure 4.** Predicted precursor relative arrival time as a function of epicentral distance for the individual seismograms, with theoretical (IASP91 and PREM) and fitted trends for the S410S and S660S measurements, comparing CNN and cross-correlation. (a) Visualization of the procedure for determining real measurements. By using density-based clustering, the most dense cluster consists of points that correspond to the observed trend, shown in blue. An initial linear fit using these points estimates the trend (red line). Points within an uncertainty cutoff are considered to be real measurements (light blue shading). A final linear fit is performed on this new set of points (yellow). Note the small difference between the initial and final linear models. (b) Predicted differential time as a function of epicentral distance for individual seismograms, with a histogram of picks between 120 and 130°, for the CNN models. Lines are predictions from IASP91 (Kennett & Engdahl, 1991) (red dotted), PREM (Dziewonski & Anderson, 1981) (red solid), and the best fit (yellow solid). Cluster quantities represent the proportion of picks in each cluster. (c) As in (b), for cross-correlation. (d) A random selection of seismograms and precursor picks for the CNN model. The width of the pick (red line) is proportional to  $2\sigma$  of the predicted arrival.

(i.e., 38,985 picks), significantly higher than the 0.6 typically used for automatic cross-correlation picking (Chambers et al., 2005). This result confirms the improved performance of the CNN model to identify signals more accurately than simple cross-correlation.

The histograms in Figures 4b and 4c are the number of predictions made between epicentral distances of 120 – 130° in time bins of 1 s. The two large Gaussian distributions correspond to predictions from the discontinuities, with the fraction of seismograms within the bin associated to that cluster shown. The CNN produces roughly twice as many predictions at this epicentral distance range, and identifies over 50% more precursors overall than using cross-correlation; for example, 410 picks are found for 28% of seismograms using the CNN compared to 19% from cross-correlation, demonstrating its greater predictive capabilities. This is further highlighted in the larger spread of discontinuity picks for the cross-correlation in Figure 4c compared to that for the CNN (Figure 4b).



Plotted in Figure 4d are a random selection of precursor picks from the CNN model with various epicentral distance and phase quality. These picks were considered by the clustering analysis to be true identification of precursors, and highlight the ability of the code to extract small signals from the noise. A corresponding examination of picks with a range of qualities confirms the marked improvement in waveform shape with increasing quality (Figure S2), and justifies our lower quality limit of 0.6. We note, however, that the CNN occasionally picks signals that appear to be sidelobes from negative amplitude interfering phases (Figure 4b). This is because the code picks the best-matching signal in a window regardless of shape, relying on the moving window to produce quality. A future goal is the implementation of a null output. In the meantime, the DBSCAN clustering analysis could be applied to remove interfering signals and their sidelobes. This would be particularly useful for mid-mantle precursors which have both positive and negative polarities (Waszek et al., 2018). The cross-correlation picks do not display the interfering negative signal gap, instead showing significant noise, highlighting its poorer performance.

SdS-SS differential travel time measurements from individual seismograms are a new type of measurement that is not yet widely used, primarily due to the difficulty in detection of the precursors. The measurements provide new, single path, unstacked observations of the MTZ discontinuities (Figures S12–S14), allowing for refinement of existing global and regional-scale seismic velocity models. For example, our preliminary global analysis presented here reveals that IASP91 (Kennett & Engdahl, 1991) provides a better fit to both the 410 and 660 (Figure 4). PREM uses 400 and 670 km for the discontinuity depths, and our measurements here are deeper and shallower than these values respectively. In consequence, the outputs and future developments from our algorithm represent a critical contribution to global seismology, in particular for tomography modeling efforts which require measurements from millions of seismograms. In theory, a model could be trained to identify any seismic phase, across various frequencies, representing a huge advance for compilation of data sets.

In addition to consistency of picking, and extraction of seismic signals from noise, the CNN technique provides a remarkable time saver in its capability to automatically process and pick seismic phases. Once a model is trained, the methods developed here allow for very rapid acquisition of new seismic data sets. The scanning algorithm picks a 140 s subset of a seismogram in approximately six seconds, which is similar to handpicking times, however the computer will continue to pick data constantly. Using a high performance computer, the scanning algorithm picked the entire data set of 58,567 signals in 10 h. In comparison, the same data set required several months for compilation via handpicking (Waszek et al., 2018). Naturally, the trained model will inherit any biases from the original data set, however the main advantage here is the consistency and reproducibility, as well as the overall benefits that an automated method provides. This method requires a significantly larger computational time than basic automatic algorithms (cross-correlation), but provides a much improved performance.

## 8. Conclusions

We have demonstrated the significant capabilities of CNNs for the task of picking seismic phases and exploiting the pattern recognition capabilities of these deep learning models. A trained model picks new data accurately and efficiently. It is able to identify other phases with similar features, and extract small-amplitude signals that typically appear masked by noise to the human eye. Thus, a CNN model trained on SS data can produce data sets of directly measured travel times for SS precursors in both stacked data and individual seismograms, the latter representing a new measurement to constrain Earth's upper mantle. Further consideration of deep learning models and potential applications to seismology could revolutionize the field by automatically picking waveforms as they become available. We encourage the use of and welcome contributions to our open-source Autopicker code.

## Data Availability Statement

A list of the data used, and all travel time measurements, is available at <https://doi.org/10.31905/DCPG-TRCH>. The Autopicker code is available at <https://doi.org/10.5281/zenodo.4827005>.

## Acknowledgments

The material is based upon work supported by NSF grants EAR-1661985 and EAR-1853662. J. A. Garcia was supported by a New Mexico Alliance for Minority Participation scholarship. L. Waszek is the recipient of a Discovery Early Career Research Award (project number DE170100329) funded by the Australian Research Council. L. Waszek acknowledges the Australian National University for hosting part of work performed in this project. B. Tauzin and this work are supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 793824. N. Schmerr and L. Waszek were supported by NSF grant EAR-1361325. Waveform data were obtained from the IRIS Data Management Center (NSF grant EAR-1063471).

## References

- Bassin, C., Laske, G., & Masters, G. (2000). The current limits of resolution for surface wave tomography in North America. *EOS Transactions of American Geophysical Union*, 81(F897).
- Chambers, K., Deuss, A., & Woodhouse, J. (2005). Reflectivity of the 410-km discontinuity from pp and ss precursors. *Journal of Geophysical Research*, 110(B02301). <https://doi.org/10.1029/2004jb003345>
- Chollet, F., et al. (2015). *Keras*. <https://keras.io>
- Davies, D., Kelly, E., & Filson, J. (1971). Vespa process for analysis of seismic signals. *Nature; Physical Science*, 232, 8–13. <https://doi.org/10.1038/physci232008a0>
- Deuss, A. (2009). Global observations of mantle discontinuities using ss and pp precursors. *Surveys in Geophysics*, 30(4), 301–326. <https://doi.org/10.1007/s10712-009-9078-y>
- Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference earth model. *Physics of the Earth and Planetary Interiors*, 25(4), 297–356. [https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7)
- Earle, P., & Shearer, P. (1994). Characterization of global seismograms using an automatic-picking algorithm. *Bulletin of the Seismological Society of America*, 84(2), 366–376.
- Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). *A density-based algorithm for discovering clusters in large spatial databases with noise*.
- Flanagan, M., & Shearer, P. (1998). Global mapping of topography on transition zone discontinuities by stacking ss precursors. *Journal of Geophysical Research*, 103(B2), 2673–2692. <https://doi.org/10.1029/97jb03212>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Ieee conference on computer vision and pattern recognition* (pp. 580–587). <https://doi.org/10.1109/cvpr.2014.81>
- Houser, C., Masters, G., Flanagan, M., & Shearer, P. (2008). Determination and analysis of long-wavelength transition zone structure using ss precursors. *Geophysical Journal International*, 174(1), 178–194. <https://doi.org/10.1111/j.1365-246x.2008.03719.x>
- Huang, Q., Schmerr, N., Waszek, L., & Beghein, C. (2019). Constraints on seismic anisotropy in the mantle transition zone from long-period ss precursors. *Journal of Geophysical Research: Solid Earth*, 124(7), 6779–6800. <https://doi.org/10.1029/2019JB017307>
- Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1), 73–101. <https://doi.org/10.1214/aoms/1177703732>
- Ito, E., & Takahashi, E. (1989). Postspinel transformations in the system  $Mg_2SiO_4-Fe_2SiO_4$  and some geophysical implications. *Journal of Geophysical Research*, 94(B8), 10637–10646. <https://doi.org/10.1029/jb094ib08p10637>
- Katsura, T., & Ito, E. (1989). The system  $Mg_2SiO_4-Fe_2SiO_4$  at high pressures and temperatures: Precise determination of stabilities of olivine, modified spinel, and spinel. *Journal of Geophysical Research*, 94(B11), 15663–15670. <https://doi.org/10.1029/JB094B11p15663>
- Kennett, B., & Engdahl, E. (1991). Travel times for global earthquake location and phase association. *Geophysical Journal International*, 105, 429–465. <https://doi.org/10.1111/j.1365-246x.1991.tb06724.x>
- Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/306538610.1145/3065386>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- McBrearty, I. W., Delorey, A. A., & Johnson, P. A. (2019). Pairwise Association of Seismic Arrivals with Convolutional Neural Networks. *Seismological Research Letters*, 90(2A), 503–509. <https://doi.org/10.1785/0220180326>
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *International Council of Machinery Lubrication*.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2). <https://doi.org/10.1126/sciadv.1700578>
- Ritsema, J., Deuss, A., van Heijst, H. J., & Woodhouse, J. H. (2011). S40rts: A degree-40 shear-velocity model for the mantle from new rayleigh wave dispersion, teleseismic traveltimes and normal-mode splitting function measurements. *Geophysical Journal International*, 184(3), 1223–1236. <https://doi.org/10.1111/j.1365-246X.2010.04884.x>
- Ross, Z. E., Meier, M., & Hauksson, E. (2018). P wave arrival picking and first-motion polarity determination with deep learning. *Journal of Geophysical Research: Solid Earth*, 123(6), 5120–5129. <https://doi.org/10.1029/2017JB015251>
- Schmerr, N., & Garnero, E. (2006). Investigation of upper mantle discontinuity structure beneath the central Pacific using ss precursors. *Journal of Geophysical Research*, 111(B8). <https://doi.org/10.1029/2005JB004197>
- Shearer, P. (1991). Constraints on upper mantle discontinuities from observations of long-period reflected and converted phases. *Journal of Geophysical Research*, 96(B11), 18147–18182. <https://doi.org/10.1029/91JB01592>
- Shearer, P. (1993). Global mapping of upper mantle reflectors from long-period SS precursors. *Geophysical Journal International*, 115(3), 878–904. <https://doi.org/10.1111/j.1365-246X.1993.tb01499.x>
- Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (Vol. 09).
- Waszek, L., Schmerr, N. C., & Ballmer, M. D. (2018). Global observations of reflectors in the mid-mantle with implications for mantle structure and dynamics. *Nature Communications*, 9(1), 385. <https://doi.org/10.1038/s41467-017-02709-4>
- Waszek, L., Tauzin, B., Schmerr, N. C., Ballmer, M. D., & Afonso, J. C. (2021). A poorly mixed mantle and its thermal state inferred from seismic waves. *In review*.
- Zhu, W., & Beroza, G. C. (2018). PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, 216(1), 261–273. <https://doi.org/10.1093/gji/ggy42310.1093/gji/ggy423>
- Zhu, W., Mousavi, S. M., & Beroza, G. C. (2019). Seismic signal denoising and decomposition using deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(11), 9476–9488. <https://doi.org/10.1109/tgrs.2019.2926772>