Journal of Geophysical Research: Solid Earth

Supporting Information for

A Little Data goes a Long Way: Automating Seismic Phase Arrival Picking at Nabro Volcano with Transfer Learning

Sacha Lapins[1], Berhe Goitom[1], J-Michael Kendall[2], Maximilian J. Werner[1], Katharine V. Cashman[1] and James O. S. Hammond[3]

[1]School of Earth Sciences, University of Bristol, UK; [2]Department of Earth Sciences, University of Oxford, UK; [3]Department of Earth and Planetary Sciences, Birkbeck, University of London, UK.

Corresponding author: Sacha Lapins (sacha.lapins@bristol.ac.uk)

Contents of this file

Introduction

In this supporting information document, we provide full details of our U-GPD model architecture (Fig S1), cross-entropy loss function used for model training (Text S7), velocity model used to associate detected phase arrivals and locate events (Table S8), and comparisons of U-GPD model performance on test data using varying levels of data augmentation (Fig S2). We also provide plotted locations using the U-GPD model, base GPD model (with two threshold values) and existing manual catalogue (Figs S3-S6).

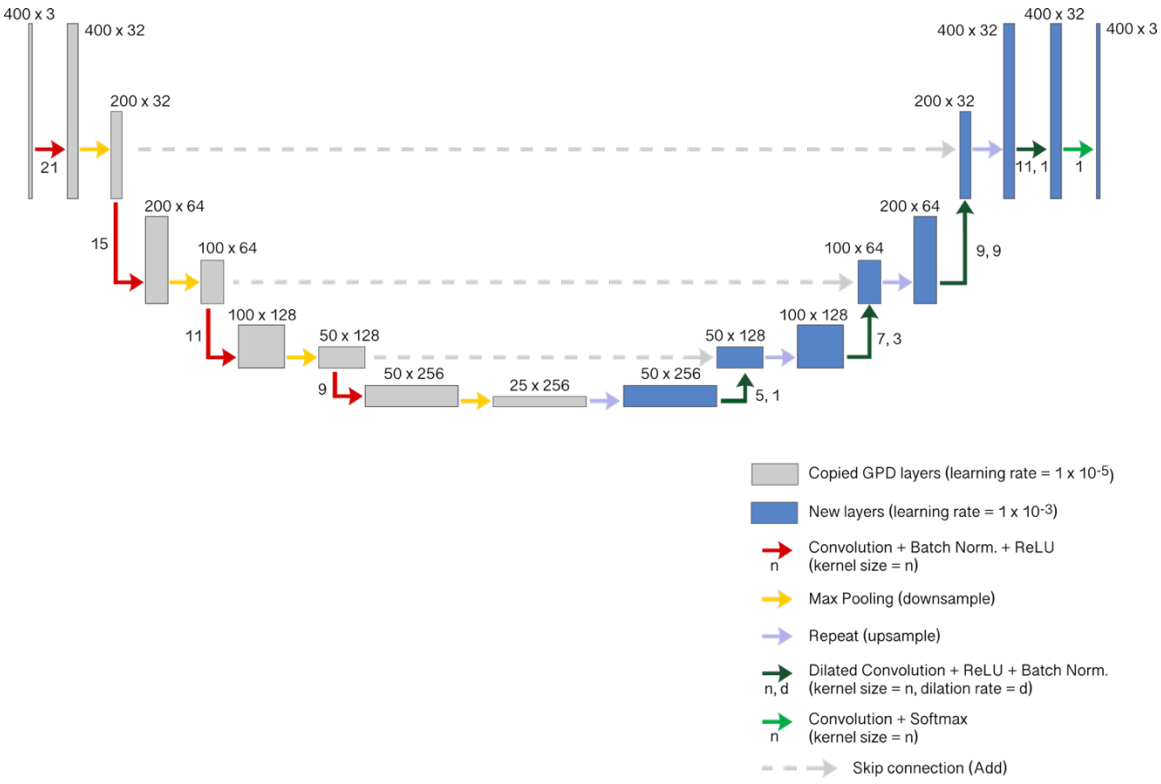

Figure S1. U-GPD model architecture.

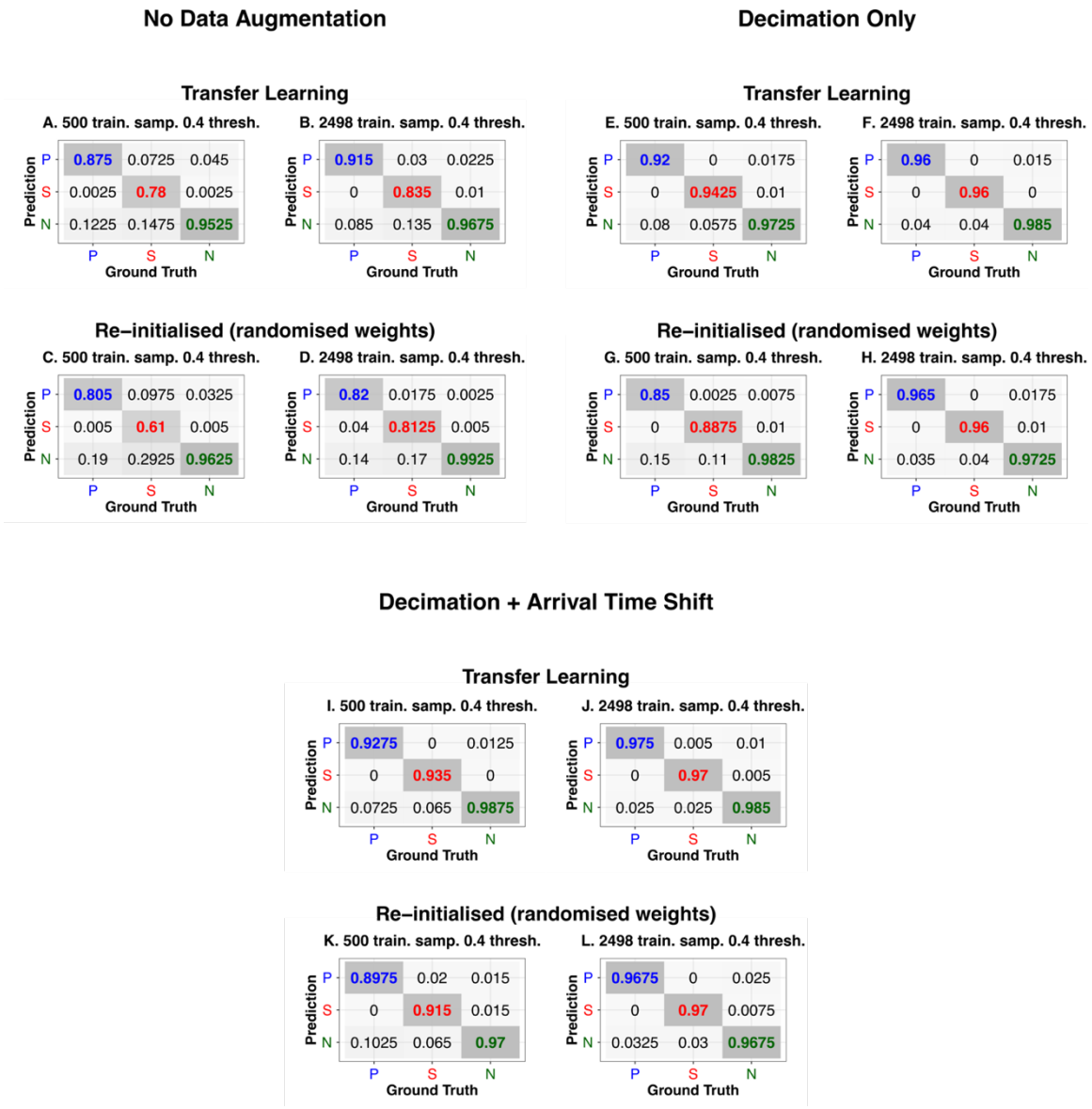

Figure S2. Results of U-GPD model, trained with different levels of data augmentation, applied to our test dataset.

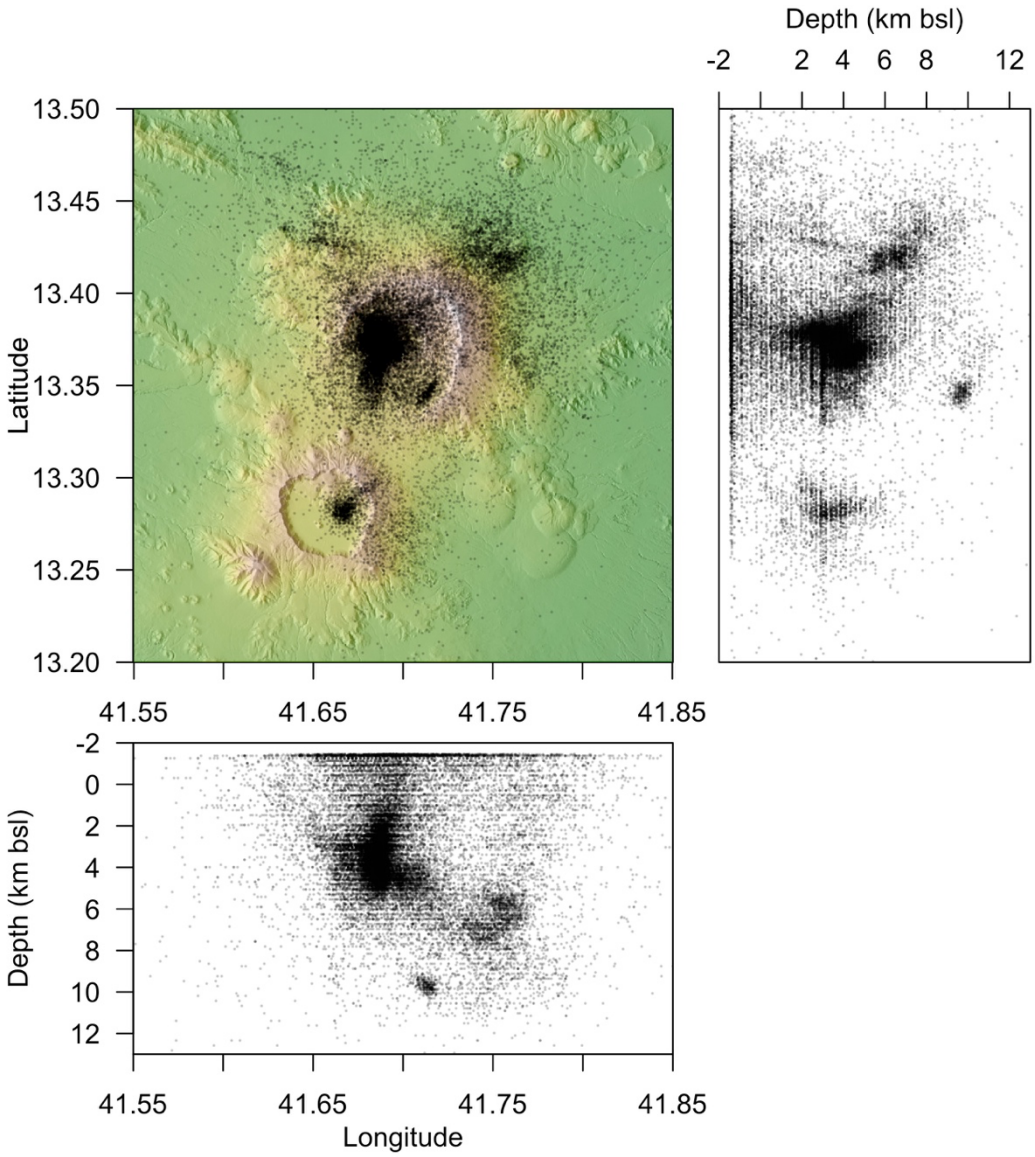

Figure S3. Event locations using P- and S-wave phase arrivals detected by U-GPD model with 0.4 threshold (no. of events = 33,950).

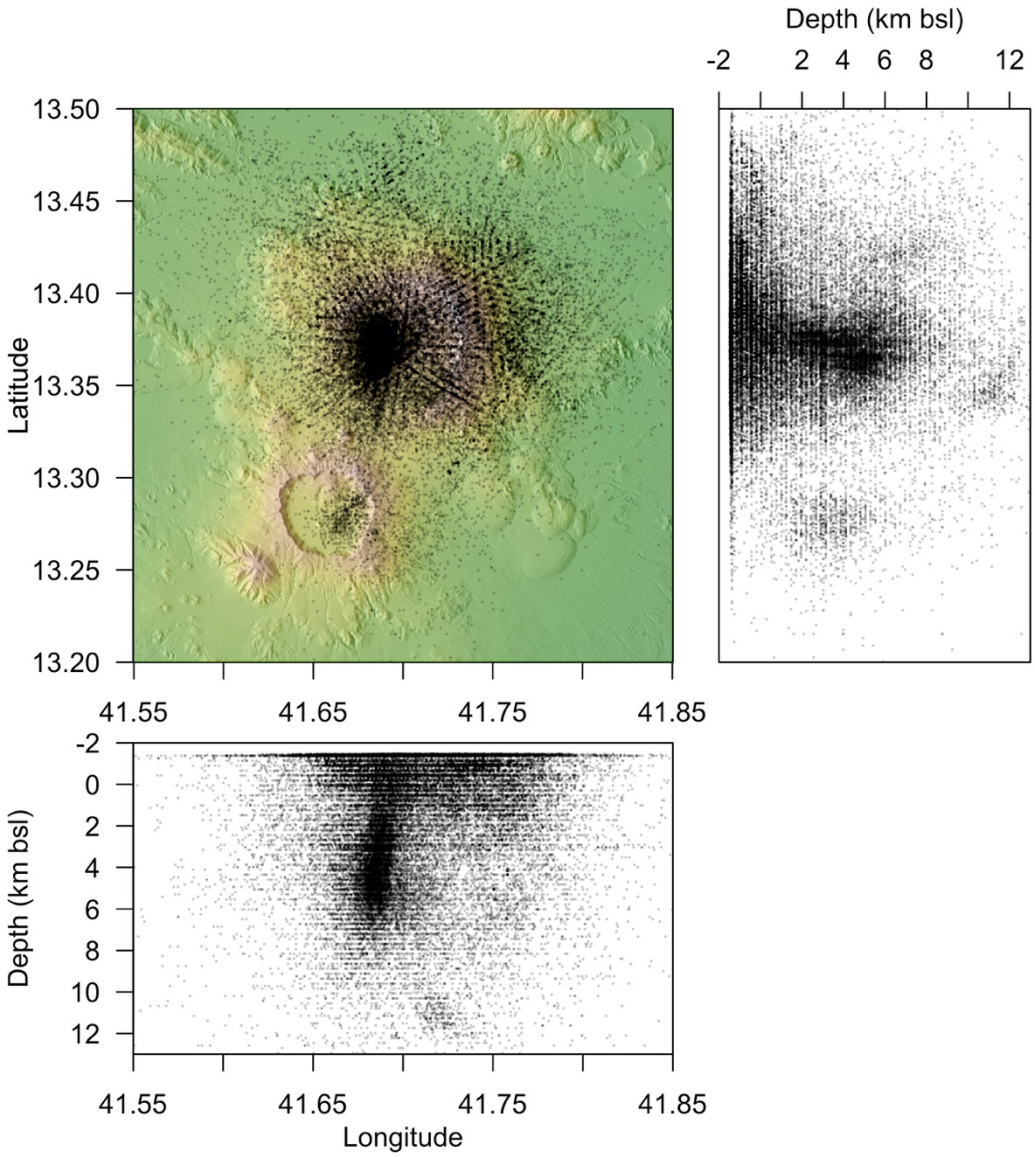

Figure S4. Event locations using P- and S-wave phase arrivals detected by original GPD model with 0.9 threshold (no. of events = 41,007).

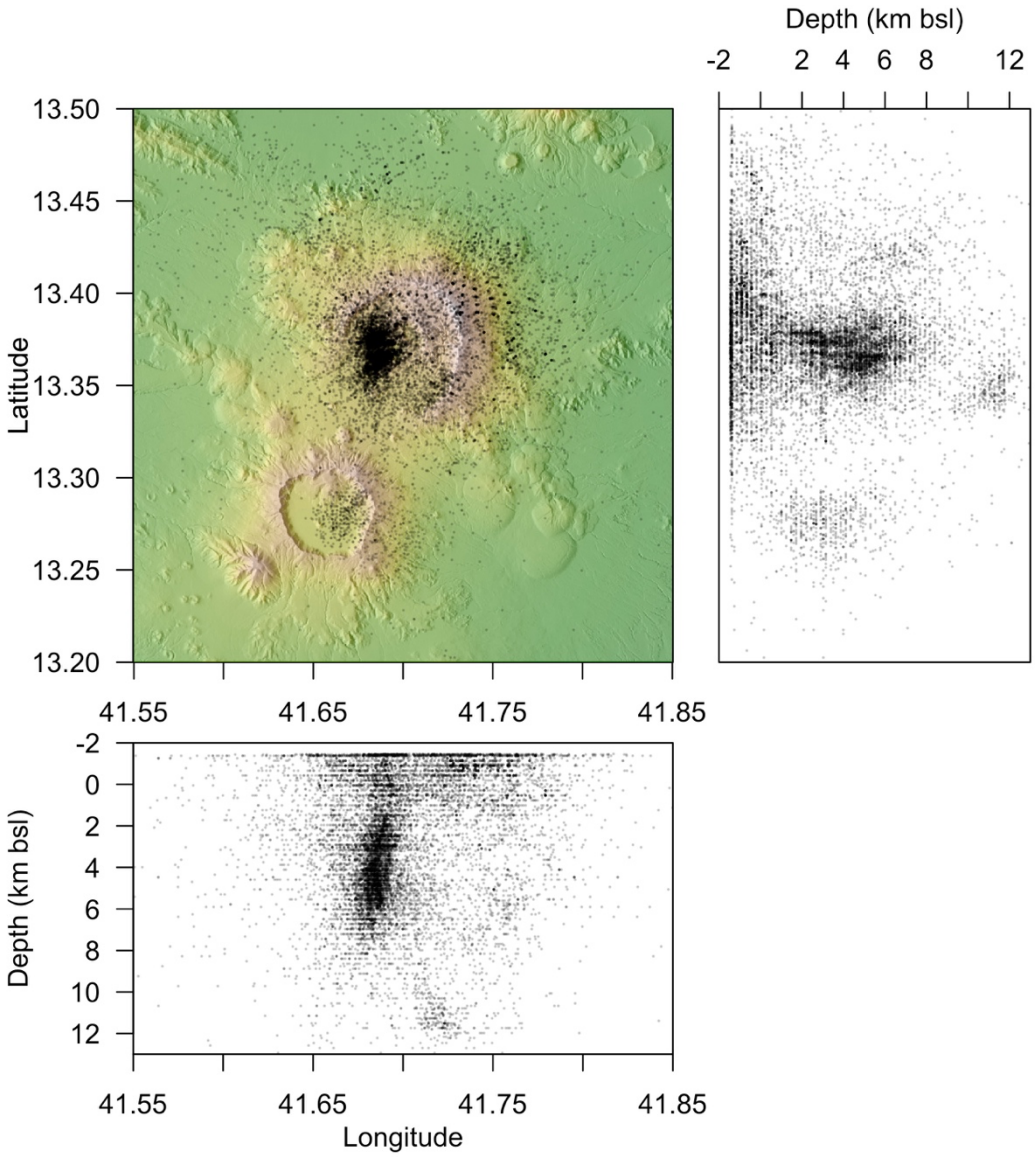

Figure S5. Event locations using P- and S-wave phase arrivals detected by original GPD model with 0.99 threshold (no. of events = 13,319).

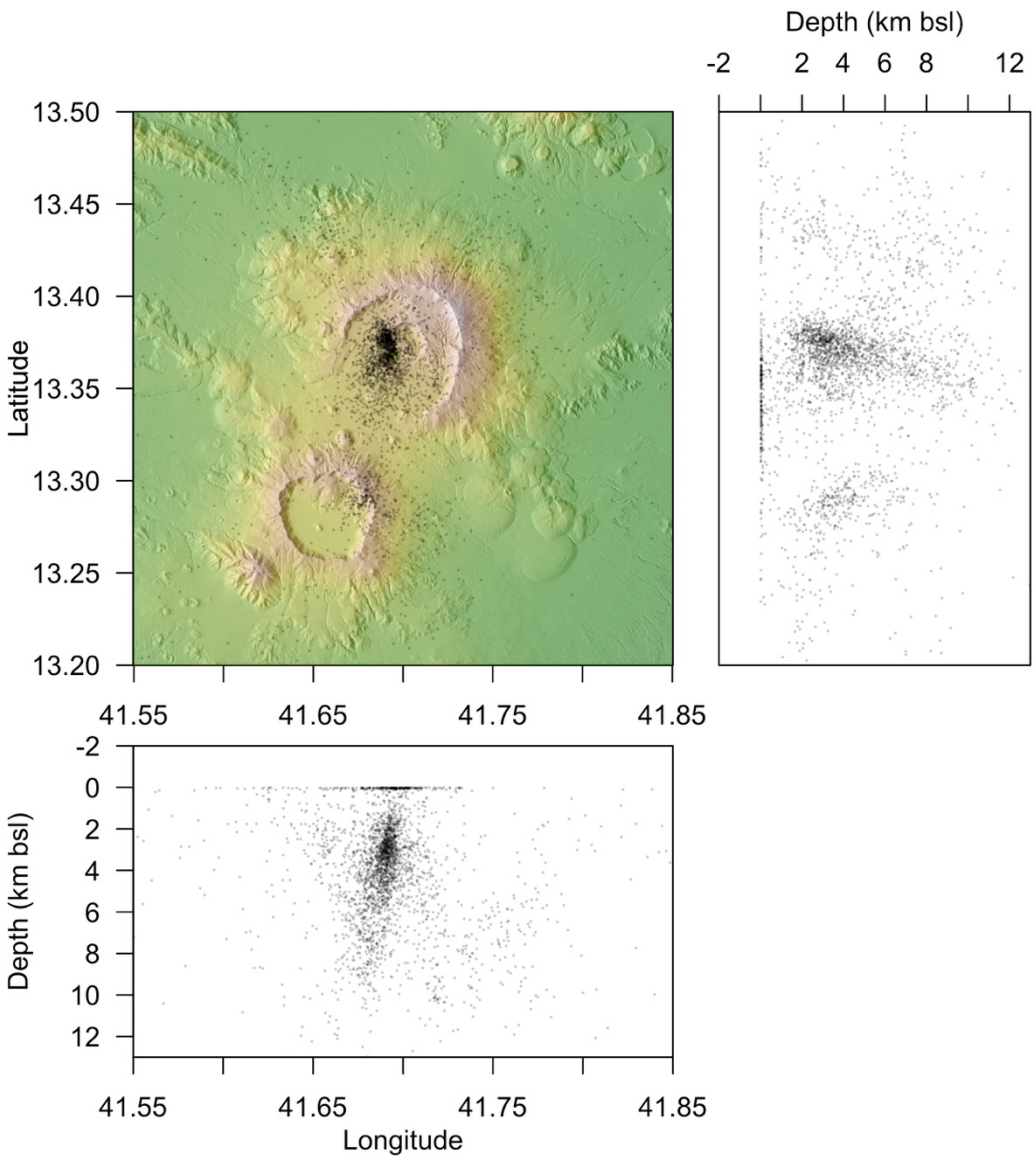

Figure S6. Event locations using P- and S-wave phase arrivals from original manual catalogue (no. of events = 2,984).

Text S7.

Supervised deep learning models are optimized through minimizing a loss function that measures the distance between the model's prediction and ground truth labels. To train our U-GPD transfer learning model, we use a cross entropy loss function, defined as:

$$\boldsymbol{L(p,q) = -\sum_{i=1}^{3}\sum_{x} p_i(x) \log q_i(x)}$$

where $\boldsymbol{i}$ is class/channel number (P, S or N), $\boldsymbol{x}$ are the datapoints in our waveform, $\boldsymbol{p_i(x)}$ is the ground truth probability distribution for $\boldsymbol{x}$ and $\boldsymbol{q_i(x)}$ is the model's predicted distribution for $\boldsymbol{x}$. The value of this loss function decreases as the model predictions converge towards the ground truth labels.

Table S8. P-wave 1D velocity model used to associate model phase arrival picks and locate events (linear gradient between layers, assumed Vp/Vs ratio of 1.76 for S-wave velocity model).

Depth for top of layer (km b.s.l.)	Vp (km / s)
-1.5	4.10
3.0	6.10
8.0	6.80
25.0	7.40