

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

Sacha Lapins¹, Berhe Goitom¹, J-Michael Kendall², Maximilian J. Werner¹, Katharine V. Cashman¹ and James O. S. Hammond³

¹School of Earth Sciences, University of Bristol, UK; ²Department of Earth Sciences, University of Oxford, UK;

³Department of Earth and Planetary Sciences, Birkbeck, University of London, UK.

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

Contents of this file

Figure S1

Figure S2

Text S3

Table S4

Introduction

In this supporting information document, we provide full details of our U-GPD model architecture, cross-entropy loss function used for model training, velocity model used to associate detected phase arrivals and locate events, and comparisons of U-GPD model performance on test data using varying levels of data augmentation.

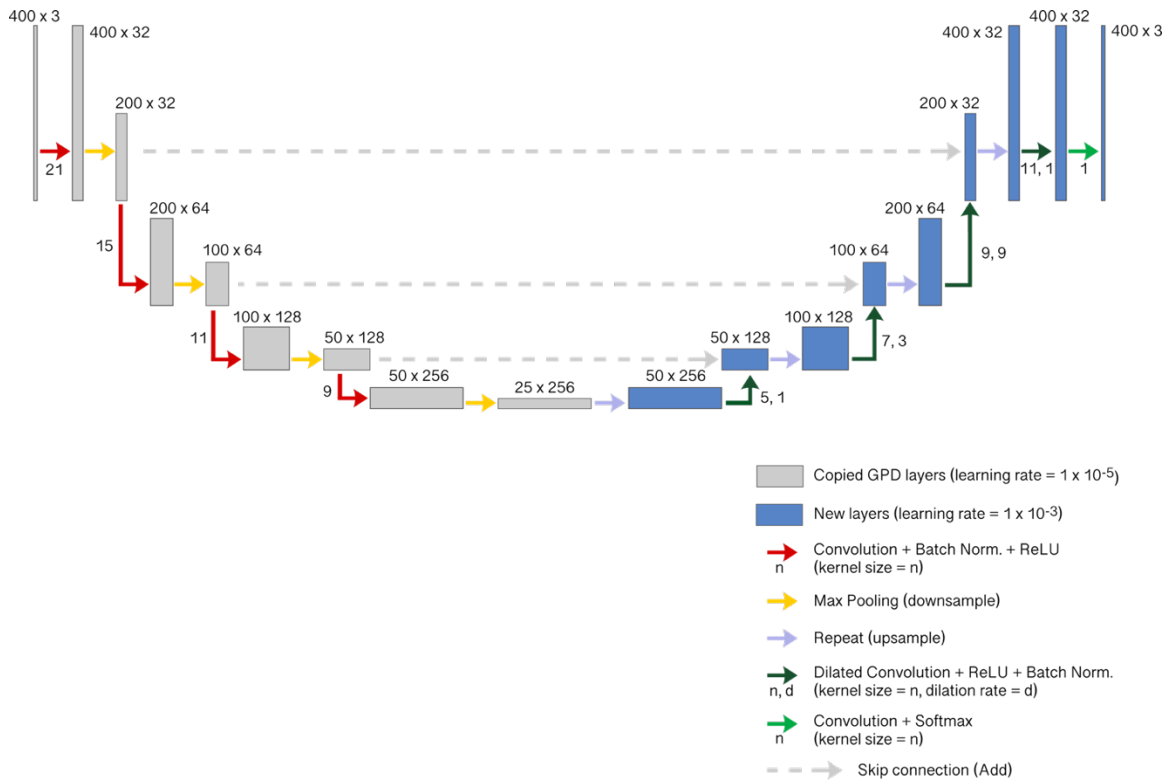
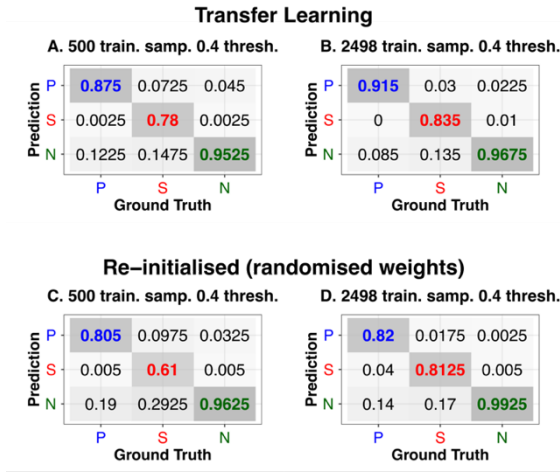
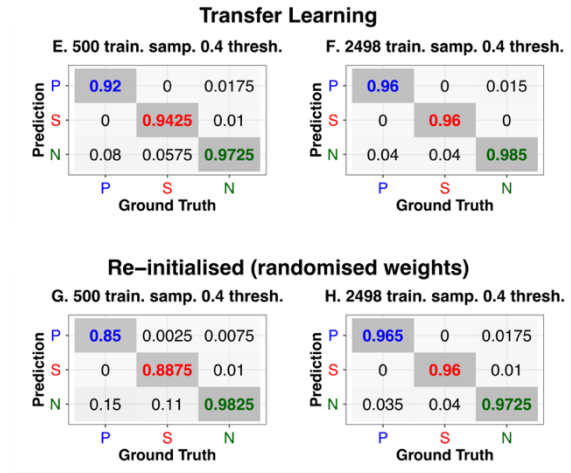


Figure S1. U-GPD model architecture.

No Data Augmentation



Decimation Only



Decimation + Arrival Time Shift

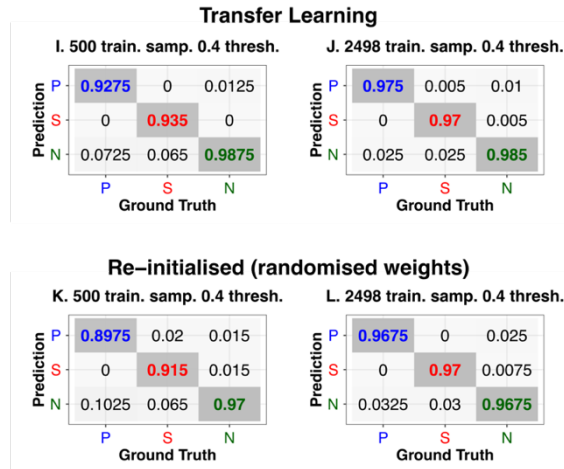


Figure S2. Results of U-GPD model, trained with different levels of data augmentation, applied to our test dataset.

Text S3.

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:

$$L(\mathbf{p}, \mathbf{q}) = - \sum_{i=1}^3 \sum_x \mathbf{p}_i(\mathbf{x}) \log \mathbf{q}_i(\mathbf{x})$$

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

Table S4. 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