A multibranch, multitarget neural network for rapid point-source inversion in a microseismic environment: examples from the Hengill Geothermal Field, Iceland

Nima Nooshiri [1], Christopher J. Bean [1], Torsten Dahm [2,3], Francesco Grigoli [4], Sigríður Kristjánsdóttir,[5] Anne Obermann [6] and Stefan Wiemer [6]

[1] *Geophysics Section, Dublin Institute for Advanced Studies, Dublin Ireland. E-mail: nima@cp.dias.ie*
[2] *GFZ German Research Centre for Geosciences, Potsdam, Germany*
[3] *Institute of Geosciences, University of Potsdam, Potsdam, Germany*
[4] *Department of Earth Sciences, University of Pisa, Pisa, Italy*
[5] *Iceland GeoSurvey, Reykjavik, Iceland*
[6] *Swiss Seismological Service, ETH-Zurich, Zurich, Switzerland*

SUMMARY

Despite advanced seismological techniques, automatic source characterization for microseismic earthquakes remains difficult and challenging since current inversion and modelling of high-frequency signals are complex and time consuming. For real-time applications such as induced seismicity monitoring, the application of standard methods is often not fast enough for true complete real-time information on seismic sources. In this paper, we present an alternative approach based on recent advances in deep learning for rapid source-parameter estimation of microseismic earthquakes. The seismic inversion is represented in compact form by two convolutional neural networks, with individual feature extraction, and a fully connected neural network, for feature aggregation, to simultaneously obtain full moment tensor and spatial location of microseismic sources. Specifically, a multibranch neural network algorithm is trained to encapsulate the information about the relationship between seismic waveforms and underlying point-source mechanisms and locations. The learning-based model allows rapid inversion (within a fraction of second) once input data are available. A key advantage of the algorithm is that it can be trained using synthetic seismic data only, so it is directly applicable to scenarios where there are insufficient real data for training. Moreover, we find that the method is robust with respect to perturbations such as observational noise and data incompleteness (missing stations). We apply the new approach on synthesized and example recorded small magnitude ($M \le 1.6$) earthquakes at the Hellisheiði geothermal field in the Hengill area, Iceland. For the examined events, the model achieves excellent performance and shows very good agreement with the inverted solutions determined through standard methodology. In this study, we seek to demonstrate that this approach is viable for microseismicity real-time estimation of source parameters and can be integrated into advanced decision-support tools for controlling induced seismicity.

Key words: Neural networks, fuzzy logic; Computational seismology; Induced seismicity; Earthquake source observations.

1 INTRODUCTION

Deep learning is now widely used in the automation of many scientific and non-scientific operations that, only a few years ago, would have been considered impossible without human intervention. Tasks such as speech, face and object recognition (e.g. Google Images and Google Assistant), recommending the best video to watch to hundreds of millions of users every day (e.g. YouTube and Netflix) or defeating the best human players at the ancient game of Go (Silver *et al.* 2017) can now be easily performed by increasingly more powerful machines. Deep Learning techniques have started to become popular also within the seismological community (Marone 2018; Bergen *et al.* 2019) with

applications ranging from rapid earthquake characterization (Lomax *et al.* 2019; Jozinović *et al.* 2020; van den Ende & Ampuero 2020), seismic-phase picking (Ross *et al.* 2018; Zhu & Beroza 2019; Woollam *et al.* 2020; Mousavi *et al.* 2020), seismic-signal classification (Kong *et al.* 2019; Meier *et al.* 2019), earthquake-magnitude estimation (Mousavi & Beroza 2020; Münchmeyer *et al.* 2021), seismic-event detection and localization (Perol *et al.* 2018; Kriegerowski *et al.* 2018; Mosher & Audet 2020), fault-plane determination (Uchide 2020; Kuang *et al.* 2021) and ground-motion modelling (Esfahani *et al.* 2021), to mention but a few.

   Among the many seismological applications that would greatly benefit from the use of deep learning techniques, induced seismicity monitoring and characterization is probably one of the highest ranked since it deals with massive seismicity data sets that are too large to be manually analysed. The recent cases of felt earthquakes associated with underground industrial operations such as Basel (Majer *et al.* 2007; Deichmann *et al.* 2014), Castor (Cesca *et al.* 2014) and Pohang (Grigoli *et al.* 2018; Ellsworth *et al.* 2019), to cite a few, highlighted the limitation of the current tools to characterize induced seismicity and the need for new paradigms for managing the risk that these activities pose for the local communities living close to such industrial sites (Edwards *et al.* 2015; Mignan *et al.* 2015; Lee *et al.* 2019; Schultz *et al.* 2021). New risk-based induced seismicity management tools like the Adaptive Traffic-Light System (ATLS) overcome the limitations of the traditional systems that have no capability to forecast the seismicity produced by a specific injection protocol, thus proving ineffective in preventing felt induced earthquakes (Wiemer *et al.* 2015; Mignan *et al.* 2017). On the other hand, the ATLS combines seismic data and geomechanical models in real-time to forecast induced seismicity and calculates risk scenarios, providing a framework for effective decision making. However, this requires robust and rapid information on microseismic sources obtained in (quasi-) real-time and in a highly automated way. While this may be achievable for seismic-event locations, the real-time determination of seismic moment tensors for weak earthquakes is extremely challenging both from a computational and methodological point of view. In the hydraulic stimulation of a geothermal well on Geldinganes, Iceland (Broccardo *et al.* 2020), for example, the Bayesian full-waveform moment tensors estimated took many hours to be calculated (Hofmann *et al.* 2021) and therefore this information could not be used within the ATLS system, reducing its performance.

   Real-time analysis of microseismicity is extremely important not only for managing the risk related to induced earthquakes but also for the geomechanical characterization of the reservoir (e.g. cap rock integrity) and last but not least to optimize hydraulic stimulation operations performed during the creation of Enhanced Geothermal Systems (EGS). For instance, precise locations of induced seismicity yield important information on the geometry of the fractures (Phillips *et al.* 2002; Rutledge *et al.* 2004), and their source mechanisms are fundamental for understanding the state of the subsurface stress-field (Bohnhoff *et al.* 2004; Song *et al.* 2014) and, in particular, to monitor the stress changes within the reservoir as a consequence of fluid injection/withdrawal operations (Zoback 2010; Terakawa *et al.* 2020). However, source moment tensor determination of microseismic events still remains challenging, since they require the modelling of high frequency waveforms and reliable source locations. The combination of computational demand and the degree of human intervention required means that some important source information is not available in real-time.

   A promising alternative to the current methods for rapid source-parameter estimation is to use supervised deep learning (SDL) and build a predictive model from existing labelled data, that is seismic waveforms (as features) accompanied by their corresponding source parameters (as labels). Such data-driven models, once trained, can compute source parameters of new microseismic events almost instantaneously. However, SDL is only applicable where there is enough good quality labelled data that captures the entirety of possible scenarios. In many SDL applications in seismological fields, such a data set is hard to collect and time consuming to label (i.e. to obtain manual solutions with low uncertainties). It may also be that at the early stages of an engineering project, previous potential data are simply not available (e.g. no previous induced events).

   The aims of this paper are twofold: firstly, to develop a SDL model that can be used for rapid inversion of microseismic source parameters; and secondly, to demonstrate the use of synthetically generated data for the purpose of training SDL algorithms on such tasks. In our approach, the inverse problem is represented in compact form by parallel convolutional neural networks which yields the full moment tensors and spatial locations of seismic events. In essence, we train a neural network (NN) that can be evaluated rapidly and with moderate computational requirements, yielding seismic moment tensor and location results within a small fraction of a second once seismic waveforms are available. This makes our approach suitable for automated and real-time processing like microseismicity source characterization in a geothermal context. Another attractive feature of our approach is that the NN is trained entirely on synthesized data (this is known as 'sim-to-real' in the deep-learning community; Tobin *et al.* 2017; Peng *et al.* 2018; James *et al.* 2019). Our training set captures the physics of wave propagation with sufficient accuracy (see Section 2.3), consists of random sources uniformly distributed in location and moment-tensor space, and reflects important perturbations presented in real-world data, such as observational noise and missing stations. Therefore, it is directly applicable to regions for which there is insufficient real data including temporary seismic networks and hydraulic stimulation experiments, for example. In this study, we seek to demonstrate that our approach can provide moment tensor in near real-time and can be viably integrated into decision-support frameworks like ATLS. We also discuss the robustness of synthetic data results and suggest important considerations for ensuring good NN generalization when training with synthetic data. As a demonstration test, we apply the method to real data from three small magnitude earthquakes that occurred at the geothermal field system in the Hengill area, Iceland. This is the demonstration site in the EU-GEOTHERMICA project COSEISMIQ (COntrol SEISmicity and Manage Induced earthQuakes; http://www.coseismiq.ethz.ch), the core objective of which is to implement innovative tools for real-time control, decision support and risk management of induced seismicity following geothermal operations.

2 METHODOLOGY

2.1 The inverse problem

In science and engineering inverse problems, the goal is—broadly speaking—to estimate physical parameters that characterize a model, **m**, given some set of data, **d**. In these problems, where the physics is adequately understood, the parameter estimation is approached by evaluating a forward operator G that relates **m** and **d** such that $\mathbf{d} = G(\mathbf{m})$ minimizing a residual norm. In our case, G can be a wave-propagation code that calculates synthetic seismograms given a source–receiver geometry and earthquake rupture model. For a real-time application like induced seismicity monitoring, this approach remains challenging because it requires repeated evaluation of the often computationally expensive forward operation $G(\mathbf{m})$. Moreover, because of the low amplitude and high frequency content of the recorded microseismic signals, routine inversion procedures can become unstable and manual parameter tuning is often required, which can be both cumbersome and time consuming. In a practical setting where **m** is required with urgency, this may be a limiting factor.

A second class of algorithms is based on a direct approximation of the inverse mapping $G^{-1}(\mathbf{d})$. This can be achieved via means of a predictive model that is calibrated using a data set of past and/or possible future observations. These data-driven models, once trained, are computationally cheap at evaluation time and thus offer great potential for rapid real-time information on seismic source parameters in a monitoring context. Our approach belongs to this group of methods. In this paper, we introduce a deep learning-based inversion algorithm that forms a regression model to estimate seismic-source parameters rapidly. In other words, a NN algorithm is trained to encapsulate the information about the relationship between seismic observations and underlying point-source models. The learning-based inversion model is calibrated using a synthetic data set giving us full control over all aspects of the data environment and the noise model to help better understand the capabilities of the approach.

2.2 Supervised deep-learning model

The deep learning method used in this study is a supervised learning technique, which uses a set of labelled data for building a predictive model that relates observable seismic waveforms with microseismic source parameters. Specifically, we utilize convolutional neural networks (ConvNets; LeCun *et al.* 2015) to achieve this. This is a class of deep NNs most commonly applied to analysing visual imagery. ConvNets apply repeated convolutional and pooling (down-sampling) operations organized in sequential layers to the input images, resulting in a set of learnable filters that automatically engineer the appropriate features for invariant pattern recognition. The core idea is then to treat the seismic inverse problem as an image recognition task. In this study, we present a multibranch, multitarget regression ConvNet to estimate the source parameters of microseismic earthquakes recorded by a seismic network (Fig. 1). The NN is designed as a multibranch model that includes two parallel ConvNets for earthquake moment tensor and location estimation, respectively. Convolution operations within the two ConvNets are performed on different types of descriptors (input images) for feature extraction. The extracted high-level features are then flattened and fused as an input of a subsequent fully connected network. The idea of designing a multibranch NN in this paper has been motivated by the application of a combination of different input data in Grond software tool (Heimann *et al.* 2018; Kühn *et al.* 2020), that is a robust inversion algorithm for earthquake source characterization. Kühn *et al.* (2020) showed that it is beneficial to combine time-domain waveforms with other input data types and perform joint inversion of waveform attributes to improve the stability of solutions.

As input data, the moment-tensor branch of the model takes in three-component seismic waveforms, that are recorded by a dense station network, in the form of a RGB colour model. In computer image-recognition problems, the input of a colour image is represented by the intensity values of the image in three primary colours of red, green and blue. In a similar fashion, we map the three-component seismic data (two horizontal and one vertical) to the three primary colours, where seismic amplitudes represent pixel values. This has the advantage that the rich information content on the source type and rupture mechanism conveyed by the waveform shape is preserved. The resultant input data, denoted by $\mathbf{d}_{\mathcal{M}} \in \mathbb{R}^{3 \times N_r \times N_t}$, form a 3-D numerical array of seismic-waveform amplitudes the size of which is the number of seismometer components (3) by receivers (N_r) by time samples (N_t). For the hypocentre-location branch, however, we take a different approach to provide it with a descriptor that is more appropriate for waveform-based earthquake localization. Instead of using time-domain seismograms, we perform a signal transformation for each individual record and construct characteristic functions (CFs) based on kurtosis, the fourth moment of local amplitude distribution (Langet *et al.* 2014; Poiata *et al.* 2016). Transient changes in the non-stationary statistics of the signal, like phase arrivals associated with seismic source, generates sharp increases in the values of the kurtosis function, making it an efficient CF to highlight the onset of direct phase arrivals (Saragiotis *et al.* 2002). Moreover, a comparative study conducted by Beskardes *et al.* (2018) showed that, compared to other CFs such as short-term average over long-term average (STA/LTA) and envelope, kurtosis CF is robust in the presence of all types of noise and provides significantly better results under conditions of low signal-to-noise ratio (SNR). We obtain kurtosis waveforms using a recursive scheme (Poiata *et al.* 2016) and then stack them along channels to create a waveform stack for each station. The resultant data, denoted by $\mathbf{d}_{\mathcal{X}} \in \mathbb{R}^{N_r \times N_t}$, span a 2-D numerical array of stacked kurtosis-waveform amplitudes and corresponds closely to single-channel grey-scale images in visual imagery algorithms.

As mentioned earlier, this study implements a supervised deep-learning methodology that infers the relationship between seismic observations (input features) and underlying point-source models (desired outputs). Accordingly, the input features for training are labelled
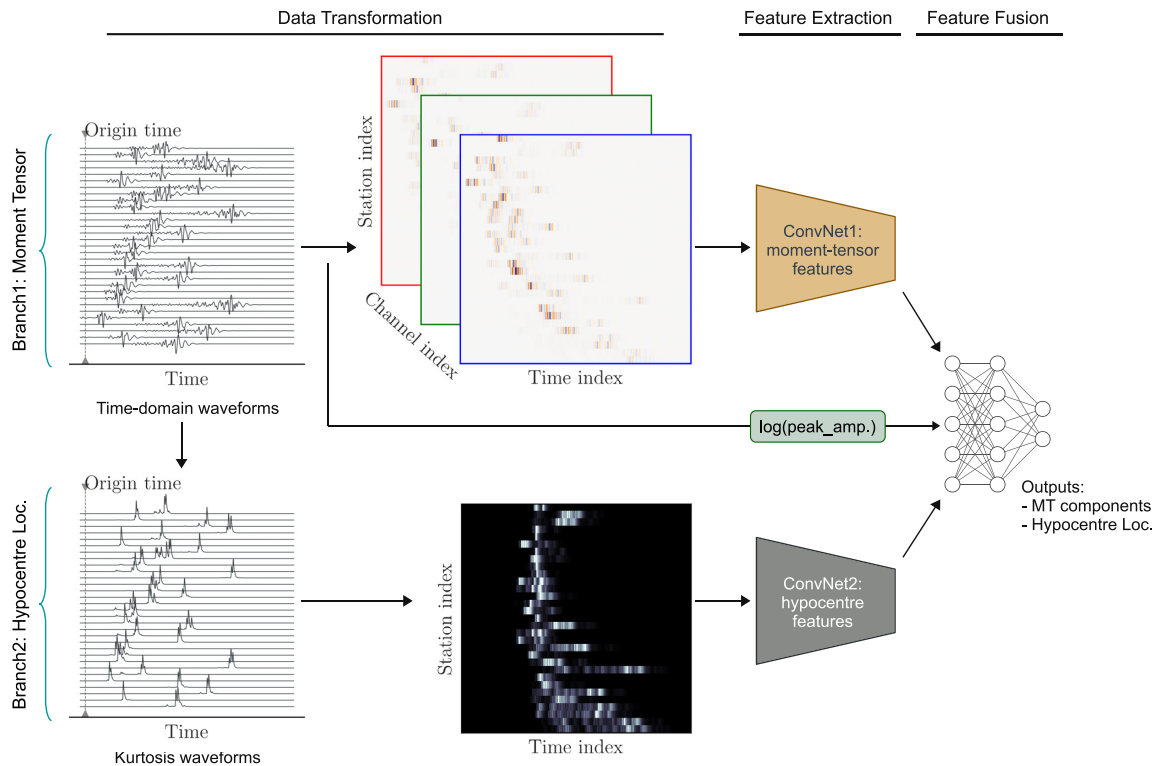
Figure 1. Schematic diagram of the overall work-flow, showing the inputs, feature extraction, feature fusion and the outputs. For simplicity, neural-network layers and their sizes are not shown. The network is designed as a multibranch model that carries out parallel ConvNets. Top branch takes in three-component seismic waveforms and is trained to extract moment-tensor features. Bottom branch is fed by stacked kurtosis waveforms and learns hypocentre-location features during the training step. The extracted features by each branch are then flattened and, combined with the logarithm of the peak amplitude, are used as inputs to the fully-connected network. For the exact number of layers and the size of each layer, please refer to Table 1.

with the seismic-source parameters, that are moment-tensor parameters, $\mathbf{m}_{\mathcal{M}} \in \mathbb{R}^6$, and hypocentre spatial location, $\mathbf{m}_{\mathcal{X}} \in \mathbb{R}^3$. The output of the NN is organized into a vector as:

$$\mathbf{m} = [\mathbf{m}_{\mathcal{M}}, \mathbf{m}_{\mathcal{X}}]^T \in \mathbb{R}^9. \tag{1}$$

The task addressed here is then a multitarget regression, that is a type of structured output where the output space consists of multiple quantitative values. Our training data set therefore consists of many ($\{\mathbf{d}_{\mathcal{M}}, \mathbf{d}_{\mathcal{X}}\}$, $\mathbf{m}$) samples, taken from an appropriate data set of seismic waveforms and corresponding point sources. We will discuss how these samples are constructed in the following sections.

2.3 Building a training data set

Our approach builds a training data set by calculating synthetic velocity seismograms from a population of randomly generated micro-seismic sources. The NN learns the essential characteristics of the seismic wave field radiated by a point source from this synthetic data, so that this knowledge can be transferred to real data. This approach is helpful in scenarios with little data availability of sufficient quality, variety and abundance for building a training data set. This includes, for instance, regions with a low seismicity rate, data sets that do not capture the entirety of possible source mechanisms, or short duration temporary seismic networks. The generation of a high-quality training data set based on forward modelling reduces or even removes the dependence on manually labelled training data.

In our approach, the seismic sources are uniformly distributed in the space of all moment tensors of unit norm, $\hat{\mathbb{M}}$, and located throughout a particular monitoring volume, $\hat{\mathbb{V}}$, which must be determined in advance. As introduced by Tape & Tape (2015), the notion of a set of moment tensors being uniformly distributed is that equal volumes in the moment-tensor space $\hat{\mathbb{M}}$ contain approximately equal numbers of moment tensors (in other words, equal volumes are equally likely). In this paper, we apply the method described by Stähler & Sigloch (2014) to generate a uniform population of unit moment tensors. By definition, a unit 5-sphere, $\mathbb{S}^5$, is the set of points $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_6) \in \mathbb{R}^6$ with $\|\hat{\mathbf{y}}\| = 1$. Stähler & Sigloch (2014) identified $\hat{y}_i$ with the components of unit moment tensor and used the transformation method of Tashiro
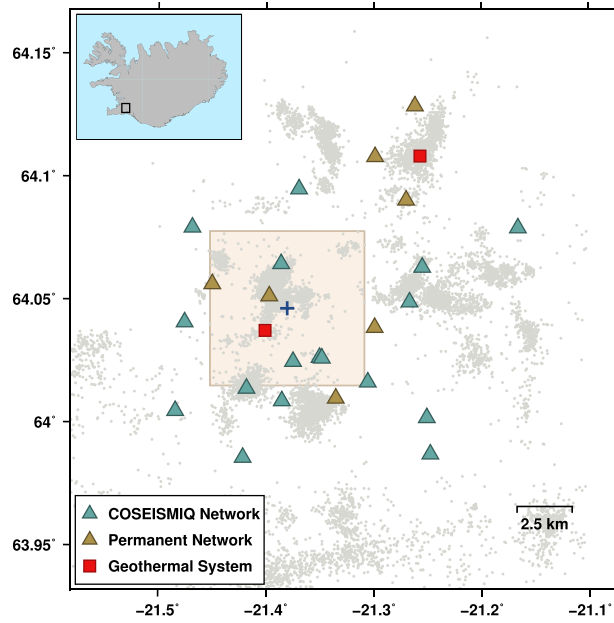
Figure 2. Map view of the station configuration used and the source region defined (shaded rectangle area) in the Hengill geothermal area, southwest Iceland. The source region spans a surface area of 50 km^2 and extends from 0 to 7 km in depth. The triangles show the positions of seismic stations and the '+' symbol indicates the local reference location. The scattered dots denote the seismicity locations between December 2018 and January 2021 provided by the Iceland GeoSurvey (ÍSOR) bulletin data.

(1977) to generate uniform random points on the surface of the $\mathbb{S}^5$ as follows:

$$\hat{\mathbf{y}} = (\hat{y_1}, \ldots, \hat{y_6}) = \begin{pmatrix} \sqrt{x_1 \sqrt{x_2}}\, cos(2\pi x_3) \\ \sqrt{x_1 \sqrt{x_2}}\, sin(2\pi x_3) \\ \sqrt{(1-x_1)\sqrt{x_2}}\, cos(2\pi x_4) \\ \sqrt{(1-x_1)\sqrt{x_2}}\, sin(2\pi x_4) \\ \sqrt{1-\sqrt{x_2}}\, cos(2\pi x_5) \\ \sqrt{1-\sqrt{x_2}}\, sin(2\pi x_5) \end{pmatrix} \in \mathbb{S}^5, \tag{2}$$

where x_i is drawn from a uniform distribution on the interval [0, 1]. The function f defined by

$$f(\hat{\mathbf{y}}) = f(\hat{y_1}, \ldots, \hat{y_6}) = \begin{pmatrix} \hat{y_1} & \frac{1}{\sqrt{2}}\hat{y_4} & \frac{1}{\sqrt{2}}\hat{y_6} \\ \frac{1}{\sqrt{2}}\hat{y_4} & \hat{y_2} & \frac{1}{\sqrt{2}}\hat{y_5} \\ \frac{1}{\sqrt{2}}\hat{y_6} & \frac{1}{\sqrt{2}}\hat{y_5} & \hat{y_3} \end{pmatrix}, \tag{3}$$

is then applied as a bijective map between $\mathbb{S}^5$ and the space of 3×3 symmetric matrices that preserves norms (Silver & Jordan 1982; Tape & Tape 2015; Ford *et al.* 2020). In other words, the function f maps $\mathbb{S}^5$ onto the space $\hat{\mathbb{M}}$ of unit moment tensors, and thus $\mathbb{S}^5$ is essentially $\hat{\mathbb{M}}$ (i.e. they are isomorphic). Uniformly distributed moment tensors generated in this way have uniformly distributed orientations but their source types become sparse toward the boundary of the fundamental Lune source-type plot (see fig. 10 of Tape & Tape 2019). This characterization is anticipated for an appropriate population of synthetic moment tensors with homogeneous probability (Ford *et al.* 2010; Chapman & Leaney 2012; Tape & Tape 2015). Magnitudes of the random moment-tensor sources can be drawn from the Gutenberg–Richter (GR) model (Gutenberg & Richter 1944) fitted to the frequency–magnitude distribution observed in a real seismicity catalogue, from the area of interest if it is available, or following a well established law, if not. Details on parameter estimation for the GR model and the earthquake catalogue used in this study are given in Section 2.4.

As monitoring volume, $\hat{\mathbb{V}}$, we choose a rectangular region in the Hengill geothermal area in southwest Iceland, that is the demonstration site in the project COSEISMIQ. The selected region also hosts the two largest geothermal power plants in Iceland, the Nesjavellir and the Hellisheiði. The latter is also the third-largest geothermal power station in the world. In addition the study region covers an area where significant seismicity, both natural and induced, has been seen according to an existing local earthquake catalogue (Fig. 2). The volume extends from surface to 7 km in depth, and spans a surface area of ∼50 km^2. For our demonstration test, we generated a total of 25 000 seismic sources uniformly distributed in the space of $\hat{\mathbb{M}}$ and volume of $\hat{\mathbb{V}}$ defined as above. It should be noted that, once trained on a certain monitoring volume, the NN is specific to that region. Therefore, parameter estimation for seismic sources that locate outside the prior spatial range is likely to be pointless because the trained NN would be required to extrapolate to regions that are not seen previously. Although we have not addressed this in the present study, our method has the potential to be extended to larger seismically active regions for monitoring
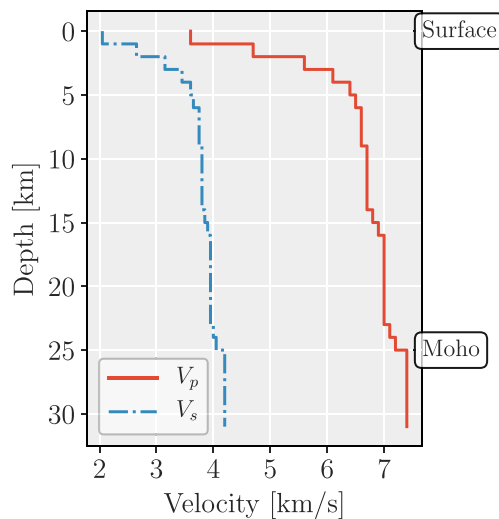
Figure 3. Layered *P*- and *S*-wave velocity models (V_p and V_s, respectively) used to calculate synthetic waveforms (modified from Tryggvason *et al.* 2002).

applications. This can be achieved by increasing the size of the generated training data set and the time duration of the modelled seismic waveforms when the size of the monitoring volume increases.

Having chosen the prior distributions of the source mechanisms and locations, we then calculate a set of three-component velocity seismograms for every source recorded at a network of receivers. We use a subset (24 stations) of a dense seismic network deployed in the Hengill geothermal area consisting of combined permanent stations of Reykjavík Energy (OR), that are operated by the Iceland GeoSurvey (ÍSOR), and temporary short period and broadband sensors installed within the framework of the COSEISMIQ project. Fig. 2 shows an overview of the station configuration and source region. Seismic-waveform modelling is performed through the use of pre-computed data bases of Green's functions based on a reflectivity-type wavenumber integration method (QSEIS code, Wang 1999) provided by the Pyrocko-GF framework (Heimann *et al.* 2019). We assume a half-sinusoid source–time rate function of duration of 30 ms. The synthetic waveforms contain body-wave phases and include near- and far-field wave components. We simulate the wave propagation with a sampling rate of 200 Hz and considering a crustal *P*- and *S*-wave velocity model of southwest Iceland derived by Tryggvason *et al.* (2002). The layered velocity model is shown in Fig. 3.

2.4 Ambient-noise modelling and adding procedure

In realistic measurement conditions, noise is a constant feature in seismic data that interferes with the waveforms radiated from seismic sources and contaminates the original signals. Particularly in microseismic monitoring, noise commonly obscures microseismic events and so makes the determination of accurate source parameters challenging. Therefore, to provide a more realistic training set, we develop a work-flow to incorporate realistic noise within the synthetic seismic data.

Here, we apply a statistical modelling method to obtain a more accurate characterization of real seismic noise and use field noise recorded on the COSEISMIQ seismic network in the Hengill area. The modelling technique is based on the statistical covariance method, in which the background noise is considered as a multivariate Gaussian random field (Gouveia & Scales 1998). Such a multivariate distribution is described by its covariance matrix and mean vector, that are derived to be similar to that of the observed noise. To compute the noise covariance matrix and mean vector, N_e time segments each of length N_t samples are selected from time windows of recorded noise, within which no earthquake-generated signals have been recorded. These segments of noise data are assumed to be realizations of the probability distribution that characterizes the ambient noise. The data are then concatenated to form a noise data matrix, $\mathbf{A}$, of size N_t by N_e (Fig. 4). The noise mean vector, $\boldsymbol{\mu} \in \mathbb{R}^{N_t}$, is computed across the N_e traces and removed from each column of the matrix $\mathbf{A}$, resulting in the reduced matrix $\hat{\mathbf{A}}$. The noise covariance matrix, $\mathbf{C}$, is calculated using (Priestley 1981)

$$\mathbf{C} = \frac{\hat{\mathbf{A}}\hat{\mathbf{A}}^T}{N_e}, \tag{4}$$

where $\hat{\mathbf{A}}^T$ denotes the transpose of the matrix $\hat{\mathbf{A}}$. Knowing the mean and covariance matrix of noise—as a multivariate Gaussian distribution—we can readily generate noise records that are realizations of the observed noise as follows (Aster *et al.* 2019):

(i)   Compute the Cholesky decomposition $\mathbf{C} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is a lower-triangular matrix.
(ii)   Generate a vector of Gaussian white noise, $\mathbf{w}$, with zero mean and unit standard deviation.
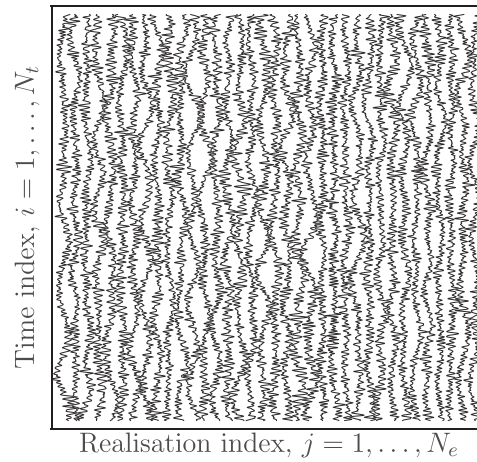(iii)   Let $\mathbf{z} = \mathbf{L}\mathbf{w} + \boldsymbol{\mu}$.

Figure 4. Noise data matrix constructed from N_e time segments of ambient noise in the seismic traces. These time segments are assumed to be realizations of the multivariate Gaussian distribution that characterizes the noise field over a finite time window, N_t.
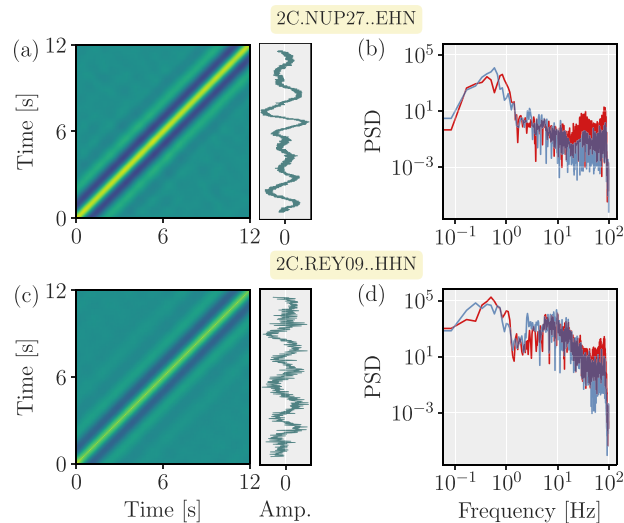

Figure 5. (a) Covariance matrix (left-hand panel) and mean vector (right-hand panel) calculated from noise traces recorded at example station 2C.NUP27 (with short-period response band). They represent the statistical properties of noise as a multivariate Gaussian field at this station. (b) Power spectral density (PSD) quality check. Red curve is the PSD of a synthetic noise trace created by drawing a random realization from the mean and covariance matrix shown in (a). Blue curve shows the PSD of an observed noise record. (c, d) Same as (a) and (b) but for example high broad-band station 2C.REY09. The sampling rate of synthetic and observed noise data in both examples is 200 Hz, and the interstation distance in this case is ∼6.6 km.

This procedure produces a modelled noise record, **z**, with the same statistics as the observed noise field. The steps (ii) and (iii) above can be repeated to generate as many noise records as the number of synthetic waveforms available. Figs 5(a) and (c) illustrate the mean and covariance matrix calculated for two example stations deployed in the COSEISMIQ seismic network in the Hengill area. Essentially, each covariance matrix has a main diagonal band with a width showing the degree of temporal correlation of the ambient noise. Comparison of power spectral densities (PSDs) of observed and synthetic noise depicts a good agreement between them and indicates that noise traces modelled by the statistical approach have characteristics close to the field measurements (Figs 5b and d).

To obtain a good approximation of the statistics of the observed noise, we choose a minimum number of noise segments, N_e, of 2400. In cases where this criterion is not met—for example, due to data gaps/incompleteness—we alternatively use the convolution-based modelling method, where a randomly chosen time segment of recorded noise is convolved with a random Gaussian noise to create a modelled noise trace. The number of stations for which we used the convolution-based noise modelling varied between one and four (out of 24 stations) depending on the time period for which we analysed the noise records.

In this study, we assume that the level of noise varies with time across the seismic network—mostly due to general weather conditions and seasonal changes. Therefore, the observed noise data are first divided into half-month intervals. Then, for each subset, the statistical noise
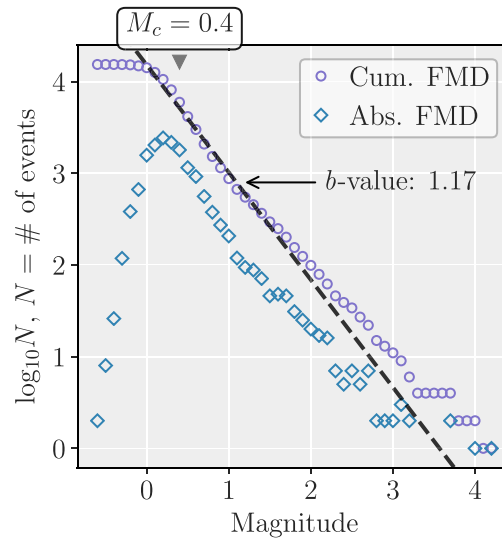
Figure 6. Analysis of frequency–magnitude distribution (FMD) of seismicity in the Hengill area between December 2018 and January 2021. The magnitudes of earthquakes have been provided by the Iceland GeoSurvey (ÍSOR) bulletin data. Diamonds and circles are the absolute and cumulative FMDs, respectively, and the dashed line is the estimated Gutenberg–Richter law. The inference for b-value and cut-off magnitude, M_c, are indicated in the figure. Magnitude bin-size of 0.1 was used to estimate the b-value and M_c.

modelling described above is performed for every station in the seismic network independently. The derived statistical noise models are thus functions of station location and half-month of the year. The origin times of the seismic events in our synthesized training set are uniformly distributed in time. For each event, we select the noise models corresponding to the event origin time for all available stations and, finally, add noise to the synthetic waveforms for each station separately.

A realistic earthquake-magnitude distribution is of considerable importance when incorporating noise into the synthetic seismic data because the SNR tends to decrease as seismic events become smaller in magnitude. We consider the GR magnitude distribution of seismic activity in the Hengill area to draw the magnitudes of the events for the training data set. We fit the GR model (Gutenberg & Richter 1944) to the observed frequency–magnitude distribution of seismicity in the Hengill area between December 2018 and January 2021 provided by the Iceland GeoSurvey (ÍSOR) bulletin data. The cut-off magnitude, M_c, is computed by using the maximum-curvature technique of Wiemer & Wyss (2000) with an upward adjustment of 0.2 magnitude units suggested by Woessner & Wiemer (2005). The b-value is estimated by using the maximum-likelihood method of Aki (1965). The estimated distribution and corresponding values are shown in Fig. 6. It should be noted that, because of our choice of the GR model to draw magnitudes, rather than a uniform magnitude distribution, the trained NN would be indirectly limited to small magnitude range ($M \leq 2$). Note that, in this study, seismic data from a very local seismic network is simulated and a realistic noise honouring the statistics and the SNRs of field noise recordings is added to build the training data set. By choosing uniform magnitudes as relatively large, one would end up with a training set in which the overall SNR is unrealistically high since there will be too many clean, almost noise-free seismic records in the synthesized data set. This can lead to NN predictions that are biased in different ways. Therefore, we suggest to train additional specific NNs for selective magnitude ranges.

2.5 Parametrization of point sources

Seismic sources in our training data set are determined by a point in 3-D space and a second-order tensor—the moment tensor. The spatial location of each source, given by $\mathbf{m}_\chi = (\Delta x, \Delta y, \Delta z)$, is expressed by Cartesian offsets from a reference point on the Earth's surface in east, north and depth directions. We choose the centre of the surface area spanned by our pre-defined monitoring volume as the reference location and transformed all geographical locations to Cartesian coordinates relative to this local origin (Fig. 2). To parametrize the moment tensor, rather than using the six components as independent parameters directly, we follow the approach of Kikuchi & Kanamori (1991) and express the full moment tensor in the form of linear combination of a set of basis tensors, $\mathbf{M}^l$:

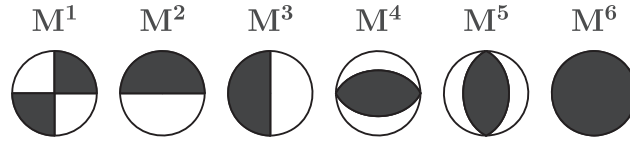$$\mathbf{M} = \sum_{l=1}^{6} a_l \mathbf{M}^l, \tag{5}$$

Figure 7. Six basis mechanisms, $\mathbf{M}^l$, used to parametrize full moment tensors (Křížová *et al.* 2013; Mustać & Tkalčić 2016). An arbitrary moment tensor is represented as a linear combination of $\mathbf{M}^l$.

where a_l are the scalar coefficients. In this study, we used the following set of six elementary moment tensors as basis tensors (Křížová *et al.* 2013; Mustać & Tkalčić 2016):

$$\mathbf{M}^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathbf{M}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \mathbf{M}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

$$\mathbf{M}^4 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{M}^5 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{M}^6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(6)

The mechanism diagrams for these basis tensors are shown in Fig. 7. The deviatoric part of the full moment–tensor solution is determined by the first five tensors, while the isotropic part is represented by the sixth tensor. The six independent components of the full moment tensor are given by the a_l coefficients:

$$\mathbf{M} = \begin{pmatrix} -a_4 + a_6 & a_1 & a_2 \\ a_1 & -a_5 + a_6 & -a_3 \\ a_2 & -a_3 & a_4 + a_5 + a_6 \end{pmatrix}.$$

(7)

The problem is then recast into determining a_l coefficients, so that the source mechanism of an example event in our training data set is represented by $\mathbf{m}_{\mathcal{M}} = (a_1, \ldots, a_6)$. One advantage of this parametrization is that it provides a unique relation between a_l coefficients and the full moment tensor. Moreover, any arbitrary moment tensor can be represented by a linear combination of the basis tensors and specific solutions, such as pure deviatoric or double-couple, can be obtained through different subgroups of $\mathbf{M}^l$. Therefore, our NN model is required to learn the *contribution level* of different basis source mechanisms to the full moment tensor. Furthermore, numerically speaking, the prior distribution of the coefficients a_l are unimodal and symmetric around zero and this large degree of symmetry in the final layer is welcome and desired by the NN regression model. Our numerical tests showed that this makes the learning process in our NN model faster and the network converges to a regression behaviour with higher accuracy after sufficient training iterations.

3 NN TRAINING AND INFERENCE

3.1 Data pre-processing and preparation

Several pre-processing steps are required to use the generated data set for NN training and evaluation. In essence, three-component velocity seismograms at all receivers are bandpass filtered, downsampled to the Nyquist sampling rate, windowed and tapered. A seismic record section of T s is cut for every event starting 0.2 s before the first break point of the event at the very first recording station that detects the event. To account for uncertainty in the starting time of the waveform cut, that arises from detection-time uncertainty, in the training data set we used each seismic-record section three times with different random shifts between -0.1 and $+0.1$ s applied to the starting time. This data augmentation also expanded the training data set by a factor of three. To avoid scaling issues, each 3-D raster is normalized by the maximum absolute (peak) amplitude over all records, resulting in pixel intensities in the range of $[-1, 1]$. The resultant time-domain data are directly used as inputs to the moment–tensor branch of our deep-learning model. To feed the hypocentre-location branch, however, they are transformed into kurtosis CFs and stacked along channels. Each waveform stack is normalized by division through the maximum, so that the grey-scale range is $[0, 1]$. Since data normalizations mentioned above discard information on event magnitude and hence the scalar seismic moment, we therefore store the logarithm of the peak amplitude—the normalization factor—for each example and it is provided as a further input to the fully connected network (Fig. 1). It should be noted that channels and stations can be in any order as long as the order is retained through the entire data set. In our data set, the channels and stations are sorted alphabetically and by latitude, respectively. In this study, we prepare the waveforms for our application by bandpass filtering between 1.5 and 6 Hz, downsampling from 200 sample per seconds (sps) to 50 sps, and cutting time windows of $T = 8$ s lengths for each event. We also randomly split our data into training and test sets, where the training batch contained 80 per cent of the data. The test set is used as a *hold-out set*; that is, a subset of the data is held back from the training of the model, and then this hold-out set is used to check the model generalization performance.

Table 1. NN architecture used for the training process. Each pooling step is to downsample a feature map by taking the maximum value over windows of size 2×2. The kernel size specifies the height and width of the 2-D convolution windows.

Block	Layer	Stage	No. of filters	Kernel size
Feature extraction	1	C→B→P	32	1×21
(parallel ConvNets)	2	C→B→P	32	1×11
	3	C→B→P	64	3×3
	4	C→B→P	64	3×3
	5	C→B→P	128	3×3
Feature fusion	1	F	512	-
	2	F	10	-

Note. C, convolution; B, batch normalization; P, pooling; F, fully connected.

3.2 NN architecture and hyper-parameters

Table 1 summarizes the topology of the NN model used for the training process. We design the moment-tensor and location branches—parallel ConvNets—with identical architectures. In each branch, the first two convolutional layers are comprised of 1-D kernels that slide over the time dimension, so these layers extract features that are relevant over time and recognize the temporal patterns in input data. The following three convolutional layers in each branch consist of 2-D kernels, that are designed such that they extract features over seismic stations. The number of kernels in our NN increases as the NN gets deeper. As we move forward in the layers, the patterns get more complex, hence there are more combinations of patterns to capture. By increasing the number of kernels in subsequent layers, we enable the NN to capture as many combinations as possible.

To fine-tune the model parameters, we use the training data available after the split (i.e. 80 per cent of the total data) and split it again into training/validation parts with another 80/20 ratio. When tuning the hyperparameters, we fit the model to the training split and evaluate its performance on the validation part. The validation part is also held back from the training of the model, but is instead used to obtain an unbiased evaluation of the model fit on the training part when selecting model parameters. After all of the training experiments are concluded (i.e. optimal model parameters are chosen), we fit (train) the final model on the *aggregate* of the training and validation parts (i.e. 80 per cent of our total data). The training is performed using the stochastic optimisation algorithm Adam (Kingma & Ba 2015), that is based on adaptive moment estimation, with a learning rate of 3×10^{-4}. By varying the learning rate over the range of 10^{-4} and 10^{-3}, we found that the chosen value produced the best results when the model performance was assessed against the hold-out validation set. Network regularization is provided through batch normalization (Ioffe & Szegedy 2015) applied to all of the convolutional layers. During the training phase, we perform 150 epochs of learning, where an epoch is one iteration over the training set. We select the model at epoch 150 as the final model. The training process took about 50 min using a single NVIDIA Tesla P100 graphics processing unit (GPU, 4.7 TFLOPS at double precision).

3.3 Training with station dropout

One important characteristic of the method presented in this study is that it requires a predetermined seismic network configuration, so the number of stations to be included in each training sample is fixed. In fact, the geometry of the seismic network is learned implicitly by the NN during the training phase. This makes the proposed method suitable for monitoring and real-time applications. However, in a realistic seismic-monitoring scenario, it may well be the case that a fraction of stations comprising the seismic network are temporarily non-operational, for instance due to extreme weather conditions in winter (particularly in our case study in Iceland), failure of power supply, or water damage. This leads to gaps in the fixed data structure. To address this issue and to further improve the generalization performance and robustness of our NN model, we apply a technique that is similar to that of Kriegerowski *et al.* (2018) and can be referred to as *station dropout*. During the training phase, a fraction of stations in each example is randomly removed (i.e. zeroed out) at every iteration with probability p using samples drawn from a Bernoulli distribution (Fig. 8). This implies that every station has probability p of being temporarily unavailable, meaning that it will be entirely removed during the current training step, but it may be available during the next step. The effect of this random dropout is that the NN cannot rely on data from any given set of seismic stations, since some of them might be unavailable at any time during training. Therefore, the NN is forced to learn more general patterns from the data leading to a more robustly trained network that generalizes better. Moreover, Jozinović *et al.* (2020) showed that this strategy is valid as their ConvNet model was capable of predicting intensity measurements at stations with missing data. In this study, we consider a dropout probability of 0.1 for the seismic stations used. In the realm of practical deep learning, dropout is a popular and highly successful technique for regularizing deep NNs to avoid overfitting (Hinton *et al.* 2012; Srivastava *et al.* 2014). Deep-learning practitioners often implement dropout on hidden layers, while here it is applied to the input layer.

3.4 Uncertainty propagation by Monte Carlo station dropout

The NN model trained for source parameter estimation can be seen as a deterministic function that only returns point predictions. In other words, the deep learning-based inversion algorithm does not incorporate uncertainty information into its predictions by default. Considering that noise is an ever-present obstacle in all seismic data recordings, it becomes necessary to quantify uncertainty in predictions made by
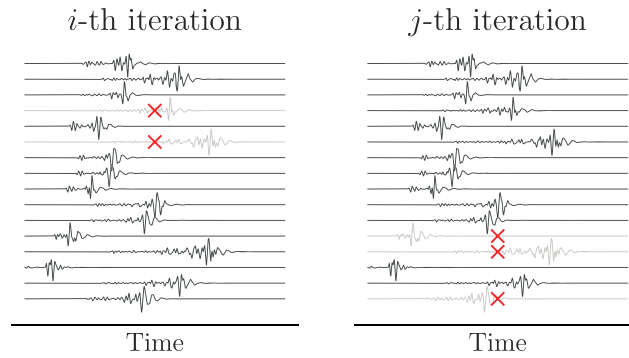
Figure 8. Schematic illustration of random station dropout. Some of the stations in each example are randomly dropped out (by setting waveform amplitudes to zero) with a given probability *p* using samples that are independently drawn from a Bernoulli distribution. The randomization is applied at each training iteration.
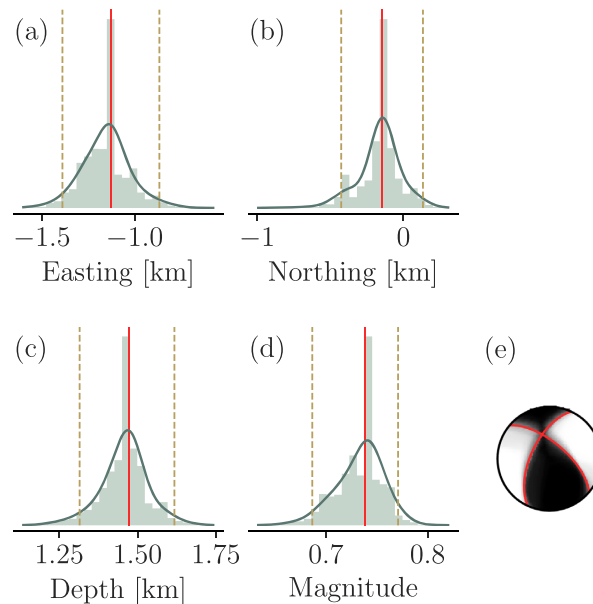

Figure 9. Uncertainty propagation by Monte Carlo station dropout for an example event. (a)–(d) Distributions and corresponding Gaussian kernel density estimates of 2500 random solutions for hypocentre location and seismic magnitude. For each distribution, the vertical solid line indicates the median value and the vertical dashed lines show the shortest interval that contains 95 per cent of the distribution density (also referred to as 95 per cent credible interval). (e) Fuzzy moment tensor illustrating the solution uncertainty (the smaller the variability of ensemble solutions, the clearer the separation between regions of compressional and dilatational motions). The projection shows ensemble of 2500 mechanisms (radiation patterns). Nodal planes of the median solution are indicated in red.

the trained model since the actual outcome can be influenced by noisy observations. A measure of prediction uncertainty provides valuable information that can be used in systems that make decisions. However, for non-linear and algorithmic techniques, such as deep NNs, there is typically no analytical way to propagate uncertainties in the data to uncertainties in the estimated (predicted) parameters. In such cases, Monte Carlo methods for uncertainty propagation can be applied, in which the parameter-estimation problem is re-solved many times for independent noise realisations (Aster *et al.* 2019). The resulting ensemble of solutions is then binned and statistically examined to quantify uncertainties in the estimated parameters arising from uncertainties in the data. The Monte Carlo analysis clearly yields empirical uncertainty estimates.

In this paper, to derive predictive uncertainty along with point parameter estimates of our trained model, we simply apply the station dropout at inference time—the process of running the trained model on new data—and generate model predictions in a stochastic way. By running multiple forward passes through the trained model with a different set of dropped stations each time, we obtain an ensemble of results, whose distributions can be used to present point estimates and corresponding uncertainties (Fig. 9). We refer to this analysis as *Monte Carlo station dropout*. A similar approach was used by Mousavi *et al.* (2020), where they implemented a channel dropout after every layer of their NN and used it during both training and prediction. It should be pointed out that, by randomly dropping (i.e. zeroing out) some stations in an example, we effectively add an impulse type of noise (more specifically, pepper noise) to the input image of the ConvNet (Boncelet 2009). This is a type of image noise that is caused by dead or damaged pixels (i.e. pixels that contain no information about the true image). Station dropout clearly has a similar effect and results in input images with missing pixels, hence it is considered as noise on inputs.
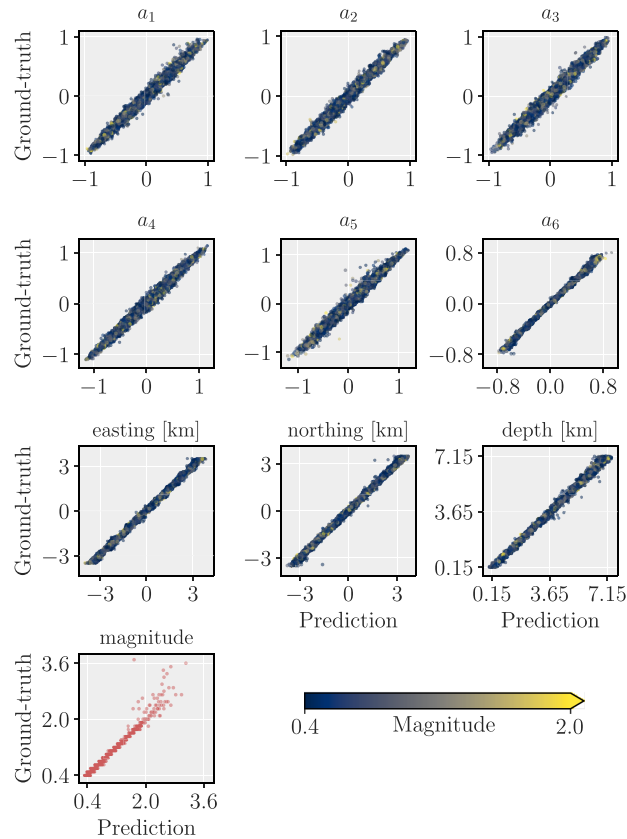
Figure 10. Scatter plots of the NN predictions against ground-truth values for 5000 microseismic events in the synthetic test set. Symbols are colour-coded by the magnitudes of seismic events.

4 RESULTS AND DISCUSSION

4.1 Evaluation of the NN predictive performance on synthetic test data

Once the NN is trained, we evaluate its predictive performance on the hold-out test set consisting of 5000 synthetically generated examples. This data set is similar to unknown data, because the model has not seen it before. Fig. 10 shows the predicted location coordinates and moment-tensor coefficients obtained from the trained NN versus the ground-truth values. The linear correlations for all parameters reveal that the trained NN has precisely learned the essential characteristics of seismic waves excited by a moment-tensor source, so it is capable of providing full information on the spatial location as well as the mechanisms of microseismic sources. In order to more easily interpret Fig. 10, we plot the distributions of differences between ground-truth and NN-predicted source parameters of the events in the synthetic test set. Location differences in the direction of the three Cartesian coordinate system axes are shown in Figs 11(a)–(c). It can be seen that there is no systematic trend in predicted localizations as indicated by the high degree of symmetry of the distributions.

To measure the differences between ground-truth, M, and predicted, $\hat{M}$, moment tensors, we calculate the distances between them in the space of 3×3 matrices (Tape & Tape 2019). That is, M and $\hat{M}$ are treated as elements of $\mathbb{R}^9$ and the normalized distance between them is then related to angular distance $\chi = \angle(M, \hat{M})$ and calculated by:

$$d(M, \hat{M}) = \sin \frac{\chi}{2} = \left[\frac{1}{2} (1 - \cos \chi) \right]^{\frac{1}{2}} = \left[\frac{1}{2} \left(1 - \frac{M \cdot \hat{M}}{\|M\| \|\hat{M}\|} \right) \right]^{\frac{1}{2}}, \tag{8}$$

where $M \cdot \hat{M}$ is the Euclidean inner product of M and $\hat{M}$ in $\mathbb{R}^9$ and $\|.\|$ denotes the matrix norm. Distances calculated by using eq. (8) range between 0 and 1, that correspond to identical and opposite seismic radiation patterns between the two compared moment tensors, respectively. The histogram of distances between true and predicted moment-tensors is shown in Fig. 11(d). From this, we find that 95 per cent of the distances are below 0.1, denoting a high degree of similarity between predicted and true moment-tensors in both source type and orientation. Visual comparison of resulting source mechanisms corresponding to a random batch of outputs is given in Fig. 12. Corresponding focal-mechanism plots match almost perfectly well, indicating that the NN is able to satisfactory invert for seismic moment tensor.
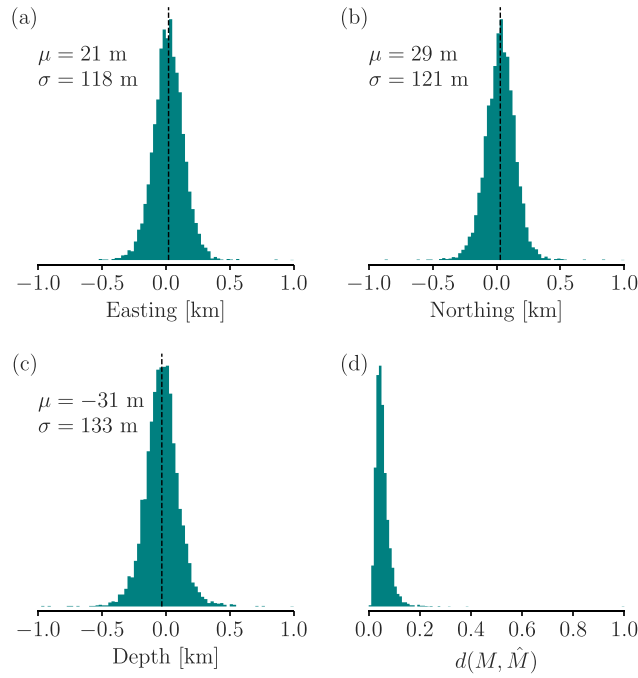
Figure 11. (a)–(c) Differences between ground-truth and NN-predicted locations for 5000 events in the test set in easting, northing, and vertical directions, respectively. Mean values, μ, and standard deviations, σ, of the underlying distributions are indicated on each subplot. (d) Distances between ground-truth, M, and NN-predicted, $\hat{M}$, moment tensors for 5000 events in the test set.
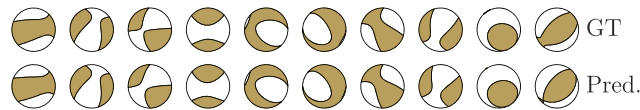

Figure 12. Focal-mechanism plots for 10 randomly chosen seismic events in the hold-out test set: (top panel) ground-truth (GT) mechanisms for the microseismic events generated using forward modelling, and (bottom panel) mechanisms for the same events predicted by the NN.

Table 2. Manually estimated source parameters for three microseismic events in the Hengill geothermal field, southwest Iceland, for which manual moment-tensor solutions are available.

ID	Date	Time	Lat. (°)	Lon. (°)	Depth (km)	Mag.
1	2018-12-23	18:52:49.438	64.04279	−21.39984	2.14	1.21
2	2019-01-14	15:27:20.821	64.05262	−21.39498	2.85	1.25
3	2019-01-29	19:16:41.969	64.05427	−21.39013	2.09	1.60

4.2 Application to example recorded events

Next, we invert waveform data from three small-magnitude microseismic events in the Hengill geothermal area that occurred inside our monitoring volume. We chose these events because 'manual' moment-tensor solutions are already available for them (Tables 2 and 3; the manual moment-tensor solutions obtained by using the inversion tool *Grond* (Heimann *et al.* 2018; Kühn *et al.* 2020) and provided by C. Rossi and S. Cesca as personal communication, 2020). Moreover, these events have been recorded by all but two seismic stations used for training, providing a chance to test the trained NN for the case where data from some stations are not available. We processed data for the same events with our trained NN to estimate the point-source parameters in a real-time fashion (at this stage, we can presume that these events have already been identified by an existing automatic detection algorithm). Then, following the instrument response corrections, the data are pre-processed in the manner explained in Section 3.1 and fed into the trained NN. The NN solutions are summarized and shown in Table 3. The comparison shows that the NN results are a very good match to the manual inversion solutions. Differences in the hypocentre locations are just in the order of a few hundred meters in both horizontal position and depth. For all three solutions, we find that the NN has predicted shallower depths. This could be related to differences in the velocity models used by the two methods to calculate the Green's functions. Regarding the focal mechanisms, the solutions predicted by the NN show very similar patterns in both source type and orientation to the manual inversion solutions. Overall, the performances of the NN is very consistent with standard techniques, but with the big advantage of providing rapid solutions. While those manual solutions were obtained in a time frame of tens of minutes per event, our trained NN is able to produce similar results within a fraction of a second. Moreover, time required to obtain manual solutions is an estimate only after having

Table 3. Focal mechanism comparisons and differences in hypocentre locations between NN predictions and manually inverted solutions (the manual moment-tensor solutions obtained by using the inversion tool *Grond* (Heimann *et al.* 2018; Kühn *et al.* 2020) and provided by C. Rossi and S. Cesca as a personal communication, 2020). Location differences are given as offset positions of the NN-based location relative to the manual location. The depth column indicates the median depth value and its 95 per cent credible interval (CI) predicted by the NN.

	Location differences (m)			Depth (km)		Focal mechanisms	
ID	Δlon.	Δlat.	Δdepth	Median	95% CI	Neural Net.	Manual
1	−238	21	−572	1.57	[1.38, 1.70]		
2	115	241	−702	2.15	[1.92, 2.34]		
3	−583	261	−530	1.56	[1.38, 1.71]		
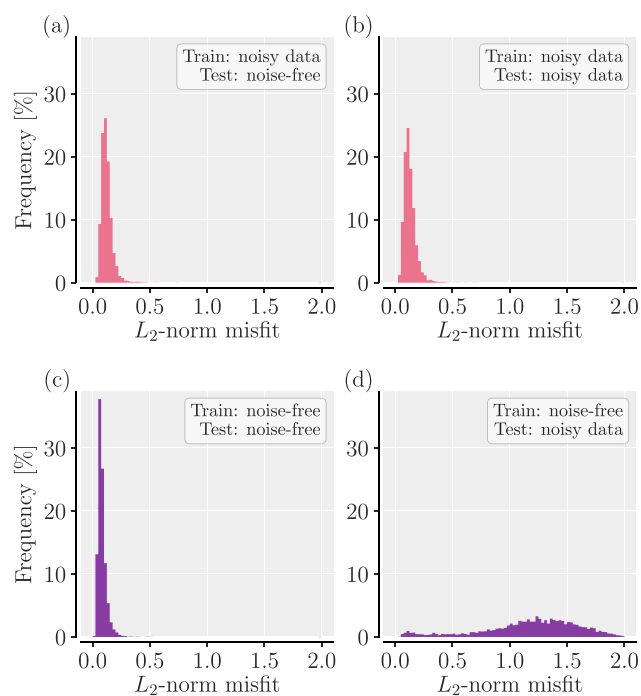
Figure 13. Histograms of test errors for (a,b) the NN trained on synthetic seismic data with added noise and (c,d) the NN trained on exact noise-free synthetic seismic data. In (a) and (c), the NNs are tested against noise-free synthetic seismic data, and in (b) and (d) they are tested on synthetic seismic data with added noise. Each test set contains 5000 seismic events. The errors are measured by the L_2-norm misfit (Euclidean distance) between vectors of ground-truth values and the NN predictions.

tuned the input parameters and found the best settings, whereas when applying the NN setting adjustments and manual pre-processing are not required and the solutions are obtained automatically. These features offer the potential for automatic and rapid real-time information on microseismic sources in a deep geothermal context, which renders the system suitable for seismic monitoring applications. It should be noted that, currently, a database of microseismic earthquakes in the Hengill geothermal field that has manual moment-tensor solutions is not available and the above is just a preliminary demonstration of application of the method on real data. We are currently constructing a 'manual' catalogue of stable moment-tensor solutions for seismic events in this region and will present the inversion result comparisons with the NN approach for a larger data set in a follow-on study.

4.3 The importance of adding noise to the training data

When training with synthetic seismic data, an important question is to investigate the implications of adding noise to the training data set for the generalization performance of the trained NN. We wish to derive a NN approximation to the inverse mapping $G^{-1}(\mathbf{d})$ that performs well on previously unobserved data, not necessarily those on which it was trained. To ensure that the NN has not picked up and learned the
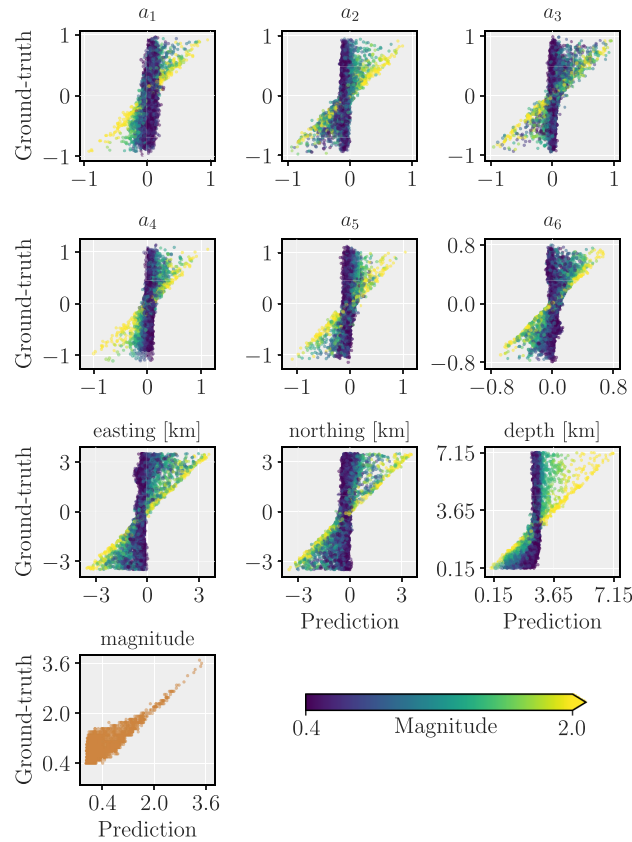
Figure 14. Scatter plots of ground-truth values against NN predictions for 5000 examples in the synthetic test set. Here, the NN is trained on exact noise-free synthetic seismic data and is applied on seismic data with added noise. Symbols are colour-coded by the magnitudes of seismic events.

seismic noise added to the training data as patterns, we test our trained model on both exact noise-free and noisy synthetic seismic data. The histograms of test errors are shown in Figs 13(a) and (b). Here we measure the test errors as the L_2-norm differences (Euclidean distances) between the ground-truth values and the NN predictions. Our NN, trained in the presence of noise, has a remarkably high noise tolerance and its inverse-mapping structure has low sensitivity to the presence of noise in seismic traces. It should be pointed out that the noise is drawn from a statistical model conditioned by field data, so the NN does not encounter precisely the same noise data in training and test steps.

Additionally, in order to assess the importance and effects of adding noise to the synthetic seismic data on the generalization performance of the NN, we consider a second NN trained on exact noise-free seismic data and examine its performance using similar hold-out test sets mentioned above. Fig. 13(c) demonstrates that the predictions of this network for the noise-free synthetic data is very good as indicated by small errors. However, when it is applied to synthetic seismic data with added noise, its generalization performance becomes very poor (Fig. 13d). This indicates that a NN trained on noise-free synthetic seismic data approximates the exact inverse mapping well but performs badly on seismic data contaminated by noise. In fact, such a model falsely associates 'field' noise as coming from variations in the microseismic source parameters leading to erroneous predictions (Fig. 14). Another interesting feature can be observed by looking at the magnitudes of seismic events when such a NN is applied on the noisy data. It can be seen that, as the seismic events become smaller in magnitude, the predictions become biased and the model performance deteriorates (Fig. 14). We thus conclude that without incorporating noise, a NN trained on noise-free synthetic seismic data may not be able to provide stable results when inverting for source parameters of small earthquakes that are within the microseismic range ($M \leq 2$) and its application will be limited to moderate and large seismic events with high SNRs.

5 CONCLUSION

We have demonstrated how a supervised deep-learning model can be used to rapidly invert for point-source parameters. The deep-learning model consists of two convolutional NN branches with individual feature extraction, where one branch distils information relevant to the seismic moment tensor from time-domain seismic data, while the other branch learns hypocentre-location features from kurtosis-waveform characteristic functions. Combining extracted high-level features, the proposed model simultaneously yields full moment tensor and spatial location of seismic sources. Once the NN model is trained, a single inversion can be executed within a fraction of a second.

The NN can be trained using synthetic seismic waveform data only and can therefore be easily applied to cases that lack large amounts of labelled training data, such as temporary seismic networks and down-hole monitoring arrays used for hydraulic stimulation experiments or regions with low seismic activity. Within our parametrization, we are able to obtain well-sampled uniformly distributed seismic sources, in

both moment-tensor and location space, making the presented work-flow (from building a training set to training the model) computationally efficient. A key finding when using synthetic data as a training set for the NN is to incorporate noise that honours the expected field conditions for real seismic data.

The trained NN is specific to a pre-defined monitoring volume and also a certain frequency range. However, the method can be easily extended to larger regions for monitoring applications, that can be achieved by increasing the size of the generated training set and the time duration of the modelled seismic waveforms. Since the frequency band of the trained NN is not adjustable, our recommendation is to create a bank of NNs each trained on a different frequency range.

The method is robust with respect to perturbations such as observational noise and missing data. It has been shown to achieve excellent performance on the test set and a very good agreement with the inverted locations and moment tensors of example real events. This approach offers great potential for rapid real-time information on seismic sources in a deep geothermal context and can be viably used for microseismic monitoring tasks in general.

REFERENCES

Aki, K., 1965. Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits, *Bull. Earthq. Res. Inst., Tokyo Univ.,* **43,** 237–239.

Aster, R.C., Borchers, B. & Thurber, C.H., 2019. *Parameter Estimation and Inverse Problems,* 3rd edn, Elsevier.

Bergen, K.J., Johnson, P.A., de Hoop, M.V. & Beroza, G.C., 2019. Machine learning for data-driven discovery in solid earth geoscience, *Science,* **363**(6433), doi:10.1126/science.aau0323.

Beskardes, G.D. *et al.*, 2018. A comparison of earthquake backprojection imaging methods for dense local arrays, *J. geophys. Int.,* **212**(3), 1986–2002.

Bohnhoff, M., Baisch, S. & Harjes, H.-P., 2004. Fault mechanisms of induced seismicity at the superdeep German Continental Deep Drilling Program (KTB) borehole and their relation to fault structure and stress field, *J. geophys. Res.,* **109**(B2), doi:10.1029/2003JB002528.

Boncelet, C., 2009. Image noise models, in *The Essential Guide to Image Processing,* pp. 143–167, ed. Bovik, A.C., Academic Press.

Broccardo, M. *et al.*, 2020. Induced seismicity risk analysis of the hydraulic stimulation of a geothermal well on Geldinganes, Iceland, *Nat. Hazards Earth Syst. Sci.,* **20**(6), 1573–1593.

Cesca, S., Grigoli, F., Heimann, S., González, Á., Buforn, E., Maghsoudi, S., Blanch, E. & Dahm, T., 2014. The 2013 September–October seismic sequence offshore Spain: a case of seismicity triggered by gas injection?, *J. geophys. Int.,* **198**(2), 941–953.

Chapman, C.H. & Leaney, W.S., 2012. A new moment-tensor decomposition for seismic events in anisotropic media, *J. geophys. Int.,* **188**(1), 343–370.

Deichmann, N., Kraft, T. & Evans, K.F., 2014. Identification of faults activated during the stimulation of the Basel geothermal project from

cluster analysis and focal mechanisms of the larger magnitude events, *Geothermics,* **52,** 84–97.

Edwards, B., Kraft, T., Cauzzi, C., Kästli, P. & Wiemer, S., 2015. Seismic monitoring and analysis of deep geothermal projects in St Gallen and Basel, Switzerland, *J. geophys. Int.,* **201**(2), 1022–1039.

Ellsworth, W.L., Giardini, D., Townend, J., Ge, S. & Shimamoto, T., 2019. Triggering of the Pohang, Korea, earthquake (Mw 5.5) by enhanced geothermal system stimulation, *Seismol. Res. Lett.,* **90**(5), 1844–1858.

Esfahani, R.D.D., Vogel, K., Cotton, F., Ohrnberger, M., Scherbaum, F. & Kriegerowski, M., 2021. Exploring the dimensionality of ground-motion data by applying autoencoder techniques, *Bull. seism. Soc. Am.,* **111**(3), 1563–1576.

Ford, S.R., Dreger, D.S. & Walter, W.R., 2010. Network sensitivity solutions for regional moment-tensor inversions, *Bull. seism. Soc. Am.,* **100**(5A), 1962–1970.

Ford, S.R., Kraft, G.D. & Ichinose, G.A., 2020. Seismic moment tensor event screening, *J. geophys. Int.,* **221**(1), 77–88.

Gouveia, W.P. & Scales, J.A., 1998. Bayesian seismic waveform inversion: parameter estimation and uncertainty analysis, *J. geophys. Res.,* **103**(B2), 2759–2779.

Grigoli, F. *et al.*, 2018. The November 2017 Mw 5.5 Pohang earthquake: a possible case of induced seismicity in South Korea, *Science,* **360**(6392), 1003–1006.

Gutenberg, B. & Richter, C.F., 1944. Frequency of earthquakes in California, *Bull. seism. Soc. Am.,* **34**(4), 185–188.

Harris, C.R. *et al.*, 2020. Array programming with NumPy, *Nature,* **585,** 357–362.

Heimann, S. *et al.*, 2017. Pyrocko—an open-source seismology toolbox and library, GFZ Data Services, doi:10.5880/GFZ.2.1.2017.001.

Heimann, S., Isken, M., Kühn, D., Sudhaus, H., Steinberg, A., Vasyura-Bathke, H., Daout, S., Cesca, S. & Dahm, T., 2018. Grond—a probabilistic earthquake source inversion framework, GFZ Data Services, doi:10.5880/GFZ.2.1.2018.003.

Heimann, S., Vasyura-Bathke, H., Sudhaus, H., Isken, M.P., Kriegerowski, M., Steinberg, A. & Dahm, T., 2019. A Python framework for efficient use of pre-computed Green's functions in seismological and other physical forward and inverse source problems, *Solid Earth,* **10**(6), 1921–1935 doi:10.5194/se-10-1921-2019.

Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R.R., 2012. Improving neural networks by preventing co-adaptation of feature detectors, *arXiv preprint arXiv:1207.0580,* available at https://arxiv.org/abs/1207.0580

Hofmann, H. *et al.*, 2021. Soft stimulation treatment of geothermal well RV-43 to meet the growing heat demand of Reykjavik, *Geothermics,* **96,** doi:10.1016/j.geothermics.2021.102146.

Hoyer, S. & Hamman, J., 2017. Xarray: N-D labeled arrays and datasets in Python, *J. Open Res. Software,* **5**(1), doi:10.5334/jors.148.

Hunter, J.D., 2007. Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.,* **9**(3), 90–95.

Ioffe, S. & Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on Machine Learning,* Vol. **37,** pp. 448–456.

James, S. *et al.*, 2019. Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks, in *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 12 619–12 629.

Jozinović, D., Lomax, A., Štajduhar, I. & Michelini, A., 2020. Rapid prediction of earthquake ground shaking intensity using raw waveform data and a convolutional neural network, *J. geophys. Int.,* **222**(2), 1379–1389.

Kikuchi, M. & Kanamori, H., 1991. Inversion of complex body waves–III, *Bull. seism. Soc. Am.,* **81**(6), 2335–2350.

Kingma, D.P. & Ba, J., 2015. Adam: a method for stochastic optimization, in *3rd International Conference on Learning Representations (ICLR),* available at https://arxiv.org/abs/1412.6980.

Kong, Q., Inbal, A., Allen, R.M., Lv, Q. & Puder, A., 2019. Machine learning aspects of the MyShake global smartphone seismic network, *Seismol. Res. Lett.,* **90**(2A), 546–552.

Kriegerowski, M., Petersen, G.M., Vasyura-Bathke, H. & Ohrnberger, M., 2018. A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms, *Seismol. Res. Lett.,* **90**(2A), 510–516.

Křížová, D., Zahradník, J. & Kiratzi, A., 2013. Resolvability of isotropic component in regional seismic moment tensor inversion, *Bull. seism. Soc. Am.,* **103**(4), 2460–2473.

Kuang, W., Yuan, C. & Zhang, J., 2021. Real-time determination of earthquake focal mechanism via deep learning, *Nat. Commun.,* **12**(1), 1432, doi:10.1038/s41467-021-21670-x.

Kühn, D., Heimann, S., Isken, M.P., Ruigrok, E. & Dost, B., 2020. Probabilistic moment tensor inversion for hydrocarbon-induced seismicity in the Groningen Gas Field, The Netherlands, Part 1: testing, *Bull. seism. Soc. Am.,* **110**(5), 2095–2111.

Langet, N., Maggi, A., Michelini, A. & Brenguier, F., 2014. Continuous Kurtosis-based migration for seismic event detection and location, with application to Piton de la Fournaise Volcano, La Réunion, *Bull. seism. Soc. Am.,* **104**(1), 229–246.

LeCun, Y., Bengio, Y. & Hinton, G., 2015. Deep learning, *Nature,* **521**(7553), 436–444.

Lee, K.-K. *et al.*, 2019. Managing injection-induced seismic risks, *Science,* **364**(6442), 730–732.

Lomax, A., Michelini, A. & Jozinović, D., 2019. An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network, *Seismol. Res. Lett.,* **90**(2A), 517–529.

Majer, E.L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B. & Asanuma, H., 2007. Induced seismicity associated with enhanced geothermal systems, *Geothermics,* **36**(3), 185–222.

Marone, C., 2018. Training machines in earthly ways, *Nat. Geosci.,* **11**(5), 301–302.

Meier, M.-A. *et al.*, 2019. Reliable real-time seismic signal/noise discrimination with machine learning, *J. geophys. Res.,* **124**(1), 788–800.

Mignan, A., Landtwing, D., Kästli, P., Mena, B. & Wiemer, S., 2015. Induced seismicity risk analysis of the 2006 Basel, Switzerland, enhanced geothermal system project: influence of uncertainties on risk mitigation, *Geothermics,* **53,** 133–146.

Mignan, A., Broccardo, M., Wiemer, S. & Giardini, D., 2017. Induced seismicity closed-form traffic light system for actuarial decision-making during deep fluid injections, *Scient. Rep.,* **7**(1), 13607, doi:10.1038/s41598-017-13585-9.

Mosher, S.G. & Audet, P., 2020. Automatic detection and location of seismic events from time-delay projection mapping and neural network classification, *J. geophys. Res.,* **125**(10), doi:10.1029/2020JB019426.

Mousavi, S.M. & Beroza, G.C., 2020. A machine-learning approach for earthquake magnitude estimation, *Geophys. Res. Lett.,* **47**(1), doi:10.1029/2019GL085976.

Mousavi, S.M., Ellsworth, W.L., Zhu, W., Chuang, L.Y. & Beroza, G.C., 2020. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking, *Nat. Commun.,* **11**(1), doi:10.1038/s41467-020-17591-w.

Münchmeyer, J., Bindi, D., Leser, U. & Tilmann, F., 2021. Earthquake magnitude and location estimation from real time seismic waveforms with a transformer network, *J. geophys. Int.,* **226**(2), 1086–1104.

Mustać, M. & Tkalčić, H., 2016. Point source moment tensor inversion through a Bayesian hierarchical model, *J. geophys. Int.,* **204**(1), 311–323.

Paszke, A. *et al.*, 2019. PyTorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems 32,* pp. 8024–8035, eds Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché Buc, F., Fox, E. & Garnett, R., Curran Associates, Inc.

Peng, X.B., Andrychowicz, M., Zaremba, W. & Abbeel, P., 2018. Sim-to-real transfer of robotic control with dynamics randomization, in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA),* doi:10.1109/icra.2018.8460528.

Perol, T., Gharbi, M. & Denolle, M., 2018. Convolutional neural network for earthquake detection and location, *Sci. Adv.,* **4**(2), doi:10.1126/sciadv.1700578.

Phillips, W., Rutledge, J., House, L. & Fehler, M., 2002. Induced microearthquake patterns in hydrocarbon and geothermal reservoirs: six case studies, *Pure appl. Geophys.,* **159**(1), 345–369.

Poiata, N., Satriano, C., Vilotte, J.-P., Bernard, P. & Obara, K., 2016. Multiband array detection and location of seismic sources recorded by dense seismic networks, *J. geophys. Int.,* **205**(3), 1548–1573.

Priestley, M.B., 1981. *Spectral Analysis and Time Series,* Academic Press.

Ross, Z.E., Meier, M.-A. & Hauksson, E., 2018. P wave arrival picking and first-motion polarity determination with deep learning, *J. geophys. Res.,* **123**(6), 5120–5129.

Rutledge, J.T., Phillips, W.S. & Mayerhofer, M.J., 2004. Faulting induced by forced fluid injection and fluid flow forced by faulting: an interpretation of hydraulic-fracture microseismicity, Carthage Cotton Valley Gas Field, Texas, *Bull. seism. Soc. Am.,* **94**(5), 1817–1830.

Saragiotis, C., Hadjileontiadis, L. & Panas, S., 2002. PAI-S/K: a robust automatic seismic P phase arrival identification scheme, *IEEE Trans. Geosci. Remote Sens.,* **40**(6), 1395–1404.

Schultz, R., Beroza, G.C. & Ellsworth, W.L., 2021. A risk-based approach for managing hydraulic fracturing-induced seismicity, *Science,* **372**(6541), 504–507.

Silver, D. *et al.*, 2017. Mastering the game of Go without human knowledge, *Nature,* **550**(7676), 354–359.

Silver, P.G. & Jordan, T.H., 1982. Optimal estimation of scalar seismic moment, *J. geophys. Int.,* **70**(3), 755–787.

Song, F., Warpinski, N.R. & Toksöz, M.N., 2014. Full-waveform based microseismic source mechanism studies in the Barnett Shale: linking microseismicity to reservoir geomechanics, *Geophysics,* **79**(2), KS13–KS30.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.,* **15**(56), 1929–1958.

Stähler, S.C. & Sigloch, K., 2014. Fully probabilistic seismic source inversion – Part 1: efficient parameterisation, *Solid Earth,* **5**(2), 1055–1069.

Tape, W. & Tape, C., 2015. A uniform parametrization of moment tensors, *J. geophys. Int.,* **202**(3), 2074–2081.

Tape, W. & Tape, C., 2019. The eigenvalue lune as a window on moment tensors, *J. geophys. Int.,* **216**(1), 19–33.

Tashiro, Y., 1977. On methods for generating uniform random points on the surface of a sphere, *Ann. Inst. Stat. Math.,* **29**(1), 295–300.

Terakawa, T., Seo, W., Kim, K.-H. & Ree, J.-H., 2020. Three-dimensional pore fluid pressures in source region of 2017 Pohang earthquake inferred from earthquake focal mechanisms, *Geophys. Res. Lett.,* **47**(9), doi:10.1029/2019GL085964.

Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W. & Abbeel, P., 2017. Domain randomization for transferring deep neural networks from simulation to the real world, in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* pp. 23–30, doi:10.1109/IROS.2017.8202133.

Tryggvason, A., Rögnvaldsson, S.T. & Flóvenz, Ó.G., 2002. Three-dimensional imaging of the P- and S-wave velocity structure and earthquake locations beneath Southwest Iceland, *J. geophys. Int.,* **151**(3), 848–866.

Uchide, T., 2020. Focal mechanisms of small earthquakes beneath the Japanese Islands based on first-motion polarities picked using deep learning, *J. geophys. Int.,* **223**(3), 1658–1671.

van den Ende, M. P.A. & Ampuero, J.-P., 2020. Automated seismic source characterization using deep graph neural networks, *Geophys. Res. Lett.,* **47**(17), doi:10.1029/2020GL088690.

Virtanen, P. *et al.*, 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python, *Nat. Methods,* **17,** 261–272.

Wang, R., 1999. A simple orthonormalization method for stable and efficient computation of Green's functions, *Bull. seism. Soc. Am.,* **89**(3), 733–741.

Wessel, P., Smith, W.H.F., Scharroo, R., Luis, J. & Wobbe, F., 2013. Generic mapping tools: improved version released, *EOS, Trans. Am. Geophys. Un.,* **94**(45), 409–410.

Wiemer, S. & Wyss, M., 2000. Minimum magnitude of completeness in earthquake catalogs: examples from Alaska, the western United States, and Japan, *Bull. seism. Soc. Am.,* **90**(4), 859–869.

Wiemer, S., Kraft, T. & Landtwing, D., 2015. Seismic risk, in *Energy from the Earth: Deep Geothermal as a Resource for the Future?,* Vol. **62**, pp. 263–295, eds Hirschberg, S., Wiemer, S. & Burgherr, P., vdf Hochschulverlag.

Woessner, J. & Wiemer, S., 2005. Assessing the quality of earthquake catalogues: estimating the magnitude of completeness and its uncertainty, *Bull. seism. Soc. Am.,* **95**(2), 684–698.

Woollam, J., Rietbrock, A., Leitloff, J. & Hinz, S., 2020. HEX: Hyperbolic Event eXtractor, a seismic phase associator for highly active seismic regions, *Seismol. Res. Lett.,* **91**(5), 2769–2778.

Zhu, W. & Beroza, G.C., 2019. PhaseNet: a deep-neural-network-based seismic arrival-time picking method, *J. geophys. Int.,* **216**(1), 261–273.

Zoback, M.D., 2010. *Reservoir Geomechanics,* Cambridge Univ. Press.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1 Training and validation average losses (also called learning curves) for 150 epochs of learning, where an epoch is a complete pass (iteration) over the training set. As the number of epoch increases, the training loss smoothly decreases and stabilizes. Meanwhile, the validation loss follows a similar decreasing pattern. After 150 iterations, both training and validation losses become stable. Both converged loss functions demonstrate that the NN is well trained.

Figure S2 Testing with only time-domain waveform branch (red) and after adding the kurtosis branch (blue): (a-c) differences between ground-truth and NN-predicted locations of 5000 test events in easting, northing and vertical directions, respectively. (d) Distances between ground-truth and NN-predicted moment tensors for the aforementioned test events. After adding the kurtosis branch, the location standard deviation errors reduced by 43, 35 and 34 percent in the east–west, north–south and vertical directions, respectively. Kurtosis waveforms can especially sharpen *P*- and *S*-phase arrivals, thus help constrain locations.

Figure S3 Testing with exact noise records (red) and modelled noise (blue). In the former, we used observed noise records and in the latter we used modelled noise to add to the synthetic waveforms in the test data set, and applied the trained NN on both test sets. The results do not show any biased predictions, confirming the noise-modelling procedure. Panel labelling is as in Fig. S2.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.