Estimation of seismic moment tensors using variational inference machine learning

Andreas Steinberg[1,*], Hannes Vasyura-Bathke[2,*], Peter Gaebler[1], Matthias Ohrnberger[2], Lars Ceranna[1]

[1]Federal Institute for Geosciences and Natural Resources (BGR), Hannover, Germany
[2]Institute for Earth and Environmental Sciences, University of Potsdam, Potsdam, Germany
[*]These authors contributed equally to this work.

Key Points:

- Ensemble of moment tensors of earthquakes are determined with Bayesian Neural Networks trained on synthetic waveforms
- Uncertainties in centroid location and time as well as uncertainties in Earth's structure are considered by variational inference
- Application to a subset of the 2019 Ridgecrest sequence and comparison to independent moment-tensor estimates shows robust performance

Corresponding author: Andreas Steinberg, `andreas.steinberg@bgr.de`

Abstract

We present an approach for estimating in near real-time full moment tensors of earthquakes and their parameter uncertainties based on short time windows of recorded seismic waveform data by considering deep learning of Bayesian Neural Networks. The individual neural networks are trained on synthetic seismic waveform data and corresponding known earthquake moment-tensor parameters. A monitoring volume has been pre-defined to form a three-dimensional grid of locations and to train a Bayesian neural network for each grid point. Variational inference on several of these networks allows us to consider several sources of error and how they affect the estimated full moment-tensor parameters and their uncertainties. In particular, we demonstrate how estimated parameter distributions are affected by uncertainties in the earthquake centroid location in space and time as well as in the assumed Earth structure model. We apply our approach on seismic waveform recordings of aftershocks of the Ridgecrest 2019 earthquake with moment magnitudes ranging from Mw 2.7 to Mw 5.5. Overall, good agreement has been achieved between inferred parameter ensembles and independently estimated parameters using classical methods. Our developed approach is fast and robust, and therefore, suitable for operational earthquake early warning systems.

1 Introduction

Robust and fast estimation of the source mechanism of earthquakes, i.e., the seismic moment tensor (MT), is important for many near-real time hazard assessments (earthquake early warning), and provides helpful information for evaluating appropriate measures and responses. Furthermore, hazard assessments and physics based aftershock probability calculations can be improved by using the inferred full seismic MT. Routine operational monitoring frameworks such as the United States Geological Survey (USGS) and GEOFON provide automatic centroid moment tensor (CMT) point-source solutions within minutes for moderate and large earthquakes (>Mw 4.5), usually in telseismic distances (Ekström, Nettles, & Dziewoński, 2012; Hanka & Kind, 1994). However, the MTs for smaller regional or local earthquakes, are often only analysed after manual inspection with delay times of up to days. The estimation of the full MT of smaller earthquakes (>Mw 3) can be important for detailed analysis of fore- and aftershock sequences, inference of local stress redistribution and especially, for seismicity monitoring in geotechnical applications (Cesca, Şen, & Dahm, 2014), where significant non double-couple (DC) components due to volumetric changes can be expected.

CMTs are usually estimated as solutions to an inverse problem by iterative comparison of synthetic and observed waveform data until a sufficient match is achieved. Forward modelling of synthetics is typically performed by assuming a point source and by considering a range of potential source model parameters and their combinations; whereas the uncertainties of the estimated parameters are quantified by considering data errors and theory errors which are introduced by the measurement and the assumptions in the inverse problem, respectively (Vasyura-Bathke et al., 2020). Uncertainties can be obtained through probabilistic approaches (Duputel, Rivera, Fukahata, & Kanamori, 2012; Kühn, Heimann, Isken, Ruigrok, & Dost, 2020; Stähler & Sigloch, 2014, 2016; Vackář, Burjánek, Gallovič, Zahradník, & Clinton, 2017; Vasyura-Bathke et al., 2020, e.g.), but these methods are computationally expensive and the estimation of CMT parameter densities can take tens of minutes to hours. Faster estimates would greatly increase the capabilities of earthquake early warning systems. Machine learning algorithms have been shown to be helpful and fast for seismic signal detection and localisation (Kriegerowski, Petersen, Vasyura-Bathke, & Ohrnberger, 2019; Smith, Ross, Azizzadenesheli, & Muir, 2021), phase picking (Mousavi, Ellsworth, Zhu, Chuang, & Beroza, 2020; Ross, Meier, & Hauksson, 2018) as well as initial characterization of the seismic source (e.g., Käufl, Valentine, O'Toole, & Trampert, 2014; van den Ende &

Ampuero, 2020). P-wave first-motion polarity can be used to determine the MT of earthquakes assuming a DC source, which has been shown to be fast and reliably to enhance MT catalogs using deep learning (Hara, Fukahata, & Iio, 2019; Ross et al., 2018; Uchide, 2020). Recently, deep learning has been used to train the so called Focal Mechanism Network (FMNet) to determine pure DC MTs based on full waveform synthetics Kuang, Yuan, and Zhang (2021). The FMNet has 16 trainable layers and was applied to four 2019 Ridgecrest earthquakes with magnitude larger than Mw 5.4. The network was trained on subjectively pre-defined Gaussian distributions as labels, describing the assumed distribution of the DC parameters strike, dip and rake.

Here, we present a machine learning framework employing several Bayesian Neuronal Networks (BNN) and using variational inference. Comprehensible consideration of errors are especially important for estimates obtained from unsupervised machine learning algorithms, as these are often treated and used as black boxes. Our BNNs are trained on synthetic waveforms with the aim to estimate MT parameters in near-real time considering errors in measurement and theory. We validate our approach on a subset of earthquakes from the aftershocks of the Californian Ridgecrest 2019-2020 sequence (Ross et al., 2019), as the Ridgecrest area is exceptionally well monitored with a dense station distribution, both in azimuth and distance (Fig. (1,a). The main shock of the 2019 Ridgecrest sequence was the Mw 7.1 2019-07-06 03:19:52 earthquake, preceded by several foreshocks of which the largest was the Mw 6.4 2019-07-04 17:33:49 earthquake. The following months several hundred aftershocks ¿Mw 3 were recorded (Ross et al., 2019). For the subset of earthquakes from the 2019 Ridgecrest sequence we investigate earthquakes with moment magnitudes M_W between 2.7 and 5.5. We compare our estimations with the moment tensors provided by the Southern California Earthquake Data Center (SCEDC).

2 Variational inference Neural Network estimation of Moment Tensors

Our main goal is to infer the radiation pattern and the orientation of the earthquake source. We train location specific neural networks for each point of a pre-defined grid of potential hypo-centers based on full sets of synthetic waveforms with associated source model parameters to be learned. We use a set of 41 broadband stations within a range up to 150 km around the center of our grid (Fig. 1,a). The grid (Fig. 1,b) extends horizontally 10.5 by 10.5 km, with a step size of 1.5 km. The vertical extent ranges from 2 km to 10 km depth, in 2 km steps.

As prior information our proposed framework needs a detection of an earthquake and the associated approximate source time. Furthermore, an approximate earthquake location can be considered. Nevertheless, it has already been demonstrated that detection and location of earthquakes are timely deliverable by other established machine learning based algorithms (Kriegerowski et al., 2019; Mousavi et al., 2020). Our approach does not estimate earthquake moment magnitudes and is indirectly limited to a range of magnitudes (e.g. between Mw 3 and 5) as the network training depends on signal processing parameters. The magnitude of earthquakes can be readily estimated in real time by other approaches (van den Ende & Ampuero, 2020).

2.1 Input

As input we use synthetic displacement waveform data calculated for a specific earthquake source and for all considered stations in E, N and Z components. Training on synthetic data has several advantages compared to training on recorded data sets. The procedure is applicable to regions with low seismicity, and furthermore, the use of synthetic waveforms allows exploring the full range of possible CMTs. Consequently, the training is not restricted by a biased set of catalog mechanisms from available

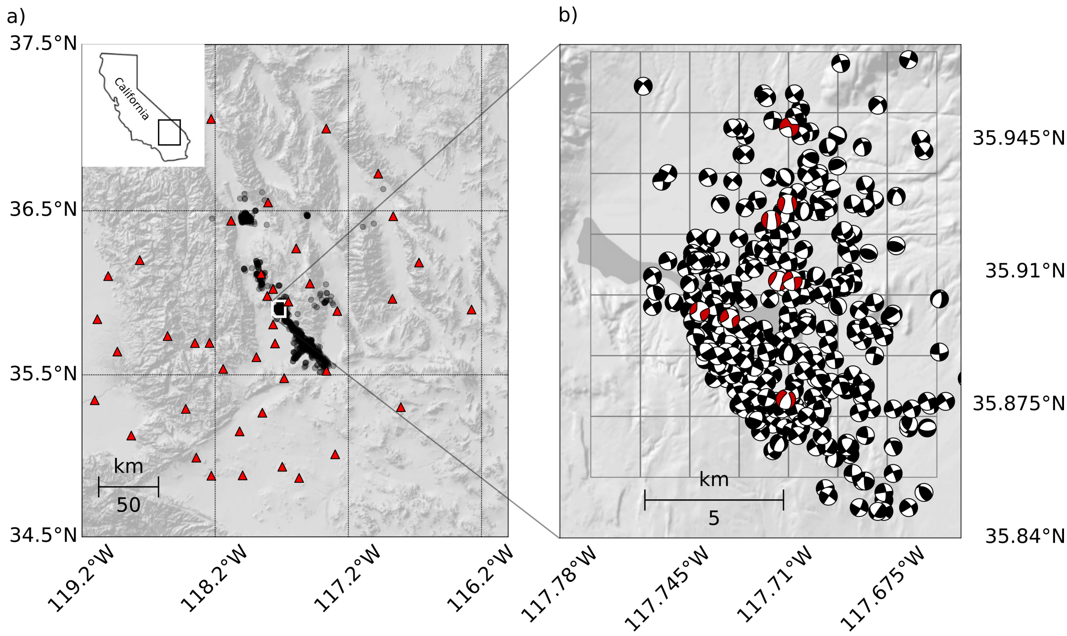

Figure 1: a) Region of interest, seismicity from 2019-07-04 to 2021-01-26 (black dots) and the station distribution (red triangles). Top-left inset shows the location of the map in California. The white rectangle shows the location of the study area. b) Zoom in to the study area. The black lines mark the grid of locations for which individual Neural Networks are trained. The focal mechanisms of earthquakes between July 2019 and December 2020 used for testing are plotted for full and double-couple CMTs in red and black, respectively. The indicated mechanisms are given as determined by SCEDC. Background in both a) and b) is a shaded relief of a digital elevation model.

observations, but it can be assured that the complete parameter space has been explored. For fast simulation of synthetic waveforms we use pre-calculated Green's functions (GF) stores from the Pyrocko software framework (Heimann et al., 2017, 2019). These GF stores are based on 1-D layered Earth structure models computed by using the reflectivity-type wavenumber integration method implemented in QSEIS (Wang, 1999). We calculate three different GF stores based on 1-D velocity profiles (Supp. Fig. S1): 1) for the entire Mojave Region used by the USGS and the SCEDC, 2) for the Coso Geothermal area (Wu & Lees, 1999) and 3) for a regional shallow velocity profile based on Crust2.0 (Bassin, Laske, & Masters, 2000).

We train our neural networks on pure synthetic waveforms without adding noise, because the characteristics of the noise would be learned as well by the networks. We filter the waveforms with a butterworth bandpass filter of fourth order between 0.8 and 2.4 Hz to avoid poor long-period response and weak long-period signals below the corner frequency of Mw 3.5 earthquakes (Aki & Richards, 2002). We assume a triangular source time function of fixed duration of 0.5 seconds, representative of earthquakes in the magnitude range 3-3.5 (Aki & Richards, 2002). Therefore, our trained networks are restricted to specific frequencies. This implies that our trained networks are only valid for a pre-defined magnitude range and that for studying earthquakes of different magnitudes, additional specific networks would need to be trained. For each source grid point location and the given 1-D Earth structure model we use the expected theoretical travel times to extract a snippet of waveform data 1 s before and 4 s after the theoretical first phase arrival. This also means that our extracted waveform snippets are relative in time and that they can be used for all possible centroid times in the training phase. To cut out real data, however, this means that the centroid time needs to be known.

We convert the extracted waveform snippets around the P-wave onset to form a 2D input image such that the rows represent the waveforms that are grouped first by channels (E, N, Z) and second by stations; the columns represent the samples over time. Finally, we normalize and re-scale the image by the absolute maximum amplitude of the full image such that all values fall between the closed interval of 0 and 1, where 0.5 indicates zero in the original waveform amplitudes as well as missing data. Due to this normalization all synthetics can be calculated for one single (but arbitrary) magnitude. The order of stations needs to be consistent for each image and must not change. Here, we chose an alphabetical order according to the station codes as arranging by azimuth or distance would be different for each considered source location and would cause artificial patterns which in turn would make efficient training of the networks difficult.

2.2 Labels

For each set of synthetic waveforms forming an input image we know the parameters of the underlying source. These are the output labels that our networks predict. The common MT parameterization with six independent components (Aki & Richards, 2002; Madariaga, 2007) seems a natural choice for describing the seismic source. However, a uniform sampling in this parameter space does not yield a uniform unique distribution of samples in moment-tensor space (Tape & Tape, 2015). Such a non-uniform and non-unique mapping would lead to bias in learned patterns for our networks. This problem can be solved by using spherical coordinates on the unit sphere of the fundamental lune description of the moment tensor(Tape & Tape, 2012b). Moreover, this parameterization allows for a uniform sampling of moment-tensors, with the advantage of only five independent parameters to describe the full spectrum of moment tensors. These five independent parameters (Tab. 1) are: κ as the strike-angle equivalent, σ as the rake-angle equivalent of the moment tensor slip angle, h as the dip-angle equivalent and the non-isotropic components v and w as the lune

Table 1: Lune parameter definitions and chosen discretization for constructing the training dataset.

Parameter	interpretation	min. value	max. value	step size
κ	strike angle	0	2π	0.1π
σ	rake angle	$-\frac{\pi}{2}$	$\frac{\pi}{2}$	0.2
h	dip angle	0	1	0.2
w	Lune latitude	$-\frac{3}{8}\pi$	$\frac{3}{8}\pi$	0.2
v	Lune co-longitude	$-\frac{1}{3}$	$\frac{1}{3}$	0.02

latitude and co-longitude, respectively (Tape & Tape, 2015). This parameterization of the MT clearly separates the radiation pattern from the source orientation. We choose a discretization of $0.1 \cdot \pi$ for κ, 0.2 for σ, h, w and 0.02 for v. This results to 171.600 synthetic waveform datasets that we use for training for each single location grid-point.

2.3 Network design

Instead of using deterministic network layers where scalar weights and biases are learned, we use their probabilistic expression with distributions of weights and biases. Each distribution is assumed to be Gaussian with mean μ and a standard deviation $\hat{\sigma}$ (e.g. Blundell, Cornebise, Kavukcuoglu, & Wierstra, 2015; Graves, 2011; Wen, Vicol, Ba, Tran, & Grosse, 2018). A neural network designed with such probabilistic layers (i.e., flipout layers) forms a Bayesian Neural Network (BNN) and can be considered as representing an ensemble of deterministic neural networks trained several times on the same input data. These BNNs allow to represent epistemic uncertainty in their inherent predictions due to limited training data and they yield a likelihood value to each drawn sample. Consequently, rather than predicting the same set of output labels given the same input data, repeated forward pass yields a distribution of output labels, i.e. uncertainties in lune parameters. This can vary for each individual BNN learned for the grid points, as the significance of specific seismic stations towards the source will vary.

Each single training iteration of a BNN consists of a forward pass and a back-propagation pass (Wen et al., 2018). In the forward pass a single sample is drawn from the output labels. During a backwards pass the gradients of the layer weights and bias distributions (i.e. means μ and standard-deviations $\hat{\sigma}$) are calculated with automatic differentiation and μ and $\hat{\sigma}$ are then updated to maximize an objective function depending on the input and output labels (Wen et al., 2018).

Our goal is to use a simple neuronal network architecture to avoid over-fitting and to allow for straightforward interpretation of the individual training steps. The network design (Fig. 2,a) is similar in rationale to Kriegerowski et al. (2019). We use three 2-D convolutional flipout hidden layers. The first two hidden layers are sensitive to the information over time only (Fig. 2,a). The first hidden layer has 8 filters and a 1 by 2 kernel and the second layer has 10 filters and a 1 by 30 kernel. The last 2-D hidden layer collects information over the station components with 12 filters and a 3 by 1 kernel. We use a dropout of 0.2 between convolutional flipout layers to robustly handle data errors and missing waveform data at particular stations. We downsample the output data of the convolutional flipout layers with a 2-D max pooling layer with a 3 by 4 kernel (example activations see Fig. 2,c) followed by flattening the data into a vector and feeding them into a fully connected dense flipout layer. The relatively simple network design allows for visual inspection of the activations in each layer (Supp. Fig S2). All convolutional flipout layers are activated using a

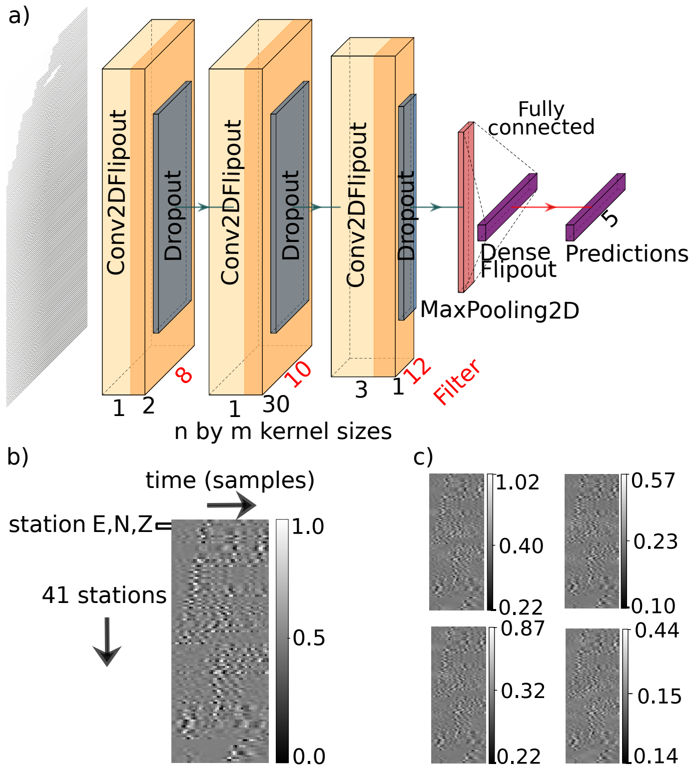

Figure 2: a) Design of an individual Bayesian Neural Network. b) example 2d normalized array input from synthetic waveforms. Blue arrows indicate hidden layers with RELU activation. c) exemplary activation of the pooling layer given the input of b)

Rectified Linear Unit function (RELU) (Glorot, Bordes, & Bengio, 2011). Finally, a non activated lambda distribution layer is used to hold the resulting distributions of predicted source parameters. As objective function (loss function) we use the negative log-likelihood and as optimizer the Adam algorithm (Kingma & Ba, 2014).

2.4 Variational inference from multiple BNN

The probabilistic output of the BNNs allows to combine inferences at several likely locations and centroid times of the earthquake's source. Each evaluation of a network with inputs yields a single prediction of the 5 MT parameters and the associated negative log-likelihood. The inferences from all these individual evaluations of networks can be combined and the source's errors in both centroid location and time can be propagated to uncertainties in MT parameter marginals through variational inference, yielding an ensemble of possible source mechanisms.

We consider an error in location within an ellipse around the assumed centroid location and evaluate the respective BNNs with the given input (Fig. 3). Note that waveform snippets are extracted differently from the waveform input according to theoretical arrival times at each receiver location (sec. 2.1).

Errors in centroid time result in shifts of the predicted theoretical arrival times and the extracted waveform snippets. We assume uniform distributed errors in timing and draw random samples within the timing errors and therefore, all BNNs at the considered grid points are evaluated several times. Consequently, we get different likelihoods to the differently extracted waveform snippets. In classical approaches in

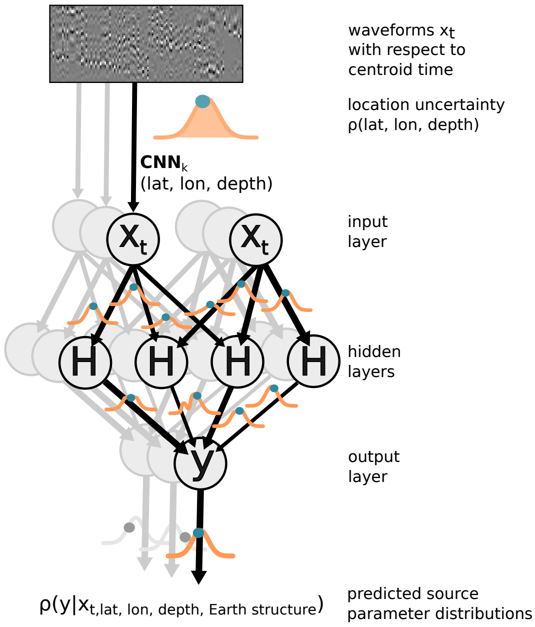

Figure 3: Training scheme, the normalized waveforms are timeshifted n-times within a timing error. The location uncertainty of an event determines the number of k BNN's that are used for prediction and contribute to the posterior probability density $\rho(\mathbf{m}|x)$.

seismology this corresponds roughly to shifting the waveforms to find the maximum correlation (e. g. Kühn et al., 2020).

Finally, errors in the Earth structure model can to be taken into account for robust inference on the estimated source mechanism (Vasyura-Bathke, Dettmer, Dutta, Mai, & Jónsson, 2021). The theory error from the choice of the 1-D Earth structure model can be included in our framework by training and evaluating BNNs on each grid-point for each Earth structure. This requires calculation of the full set of synthetic waveforms for different Earth structures; in our case, three structures (Supp. Fig S1). This results in total to over 100 Million waveform datasets on which the 588 BNNs (196 grid points times three Earth structures) are trained. The calculation of synthetic waveforms and the network training was done in parallel on several machines with a total of 128 CPUs over a period of three months. By using GPUs this time could be drastically reduced to a few days.

2.5 Moment tensor ensemble similarity

To asses the similarity between the predicted ensemble of MTs and a reference solution, e.g. from a catalog, we use the omega angle measure (Tape & Tape, 2012a). The omega angle has the advantage that focal mechanisms with opposite polarities are considered most dissimilar in contrast to other measures, e.g., the Kagan angle (Cesca et al., 2014; Tape & Tape, 2012a). The normalized omega angle distance d (Cesca et al., 2014; Tape & Tape, 2012a) between two moment tensors U_1 and U_2 with components I and J is calculated by:

$$d_\omega = \frac{1}{2}\left[1 - \frac{U_1 \cdot U_2}{||U_1||||U_2||}\right] = \frac{1}{2}\left[1 - \frac{\sum_{i,j=1}^{I,J} U_{1_{ij}} \cdot U_{2_{ij}}}{(\sum_{i,j=1}^{I,J} U_{1_{ij}}^2)^{\frac{1}{2}}(\sum_{i,j=1}^{I,J} U_{2_{ij}}^2)^{\frac{1}{2}}}\right] \quad (1)$$

It is defined between 0 and 1, for identical and opposite seismic radiation patterns between the two compared moment tensors, respectively. Note, that in order to calculate d_ω we need to convert our predicted MT ensemble from the Lune parameterization to the North-East-Down coordinate system (Aki & Richards, 2002).

3 Application to the Ridgecrest 2019 earthquake aftershock sequence

We train our networks for an area South of the Coso geothermal field (Fig. 1), which is known to host both induced and tectonic earthquakes (Monastero et al., 2005; Schoenball, Davatzes, & Glen, 2015). Significant non-DC components can be expected for earthquakes in this region (Ichinose, Anderson, Smith, & Zeng, 2003), potentially also for tectonic earthquakes, due to the influence of the geothermal reservoir. To test the performance of our framework we use recorded waveform data of the aftershocks that occurred between July 2019 and December 2020 to the Mw 7.1 Ridgecrest earthquake.

For these aftershocks, 8 full moment tensor solutions and 198 pure DC MT solutions (Fig. 1,b) are calculated (Hauksson & Unruh, 2007; Jordan & Maechling, 2003) and made publicly available by the SCEDC ((SCEDC), 2013). We compare the MT estimates of our approach to the moment tensors as determined independently by USGS and SCEDC (Hutton, Woessner, & Hauksson, 2010).

We download the waveform data for all events and for the 41 stations from the Southern California Seismic Network (California Institute Of Technology And United States Geological Survey Pasadena, 1926). Missing waveform data for any station and time period are replaced by zero values in the waveform data, which are then mapped to 0.5 values in the normalized input images. Measured waveform data are treated in the same way as our synthetic waveforms (sec. 2.1), i.e. data is restituted to ground displacement and down-sampled to match the Green's function sampling rate of 14 Hz.

For each aftershock we evaluate the BNNs for a total of 6000 samples. However, the number of activated BNNs depends on the uncertainties in centroid location and time as provided by the SCEDC catalog. The location uncertainty in horizontal and vertical position as given by the SCEDC is increased 10 times, as reported uncertainties are in the order of few hundreds of meters. The total ensemble of samples is then obtained by evaluating the activated BNNs equally.

3.1 Inferences for full moment tensors

We focus primarily on 8 aftershocks for, which a full moment tensor solution is available in the SCEDC catalog. We refer to these solutions as "reference" in the following.

We first evaluate only the waveform input with the BNN's trained using synthetics based on the *Mojave* Earth structure, which is the same as used by the SCEDC to determine their focal mechanisms (Supp. Fig. S1). Consequently, the reference and predicted MTs should be consistent in their epistemic uncertainty as the same Earth structure model and (mostly the same) dataset is used. We use the SCEDC catalog values for source position and centroid time. For the comparison we only consider uncertainty in centroid location. We find very good agreement of our predicted ensembles to most of the 8 reference moment tensors, with most of the omega angle distances d_ω being below 0.1 (Fig 4,a-h). Histograms of d_ω show their maximum mostly within the first few bins. Only, two ensembles of predicted moment tensors show small systematic errors (Fig 4,f and g). For those also the histograms of d_ω show their maxima at distances above zero.

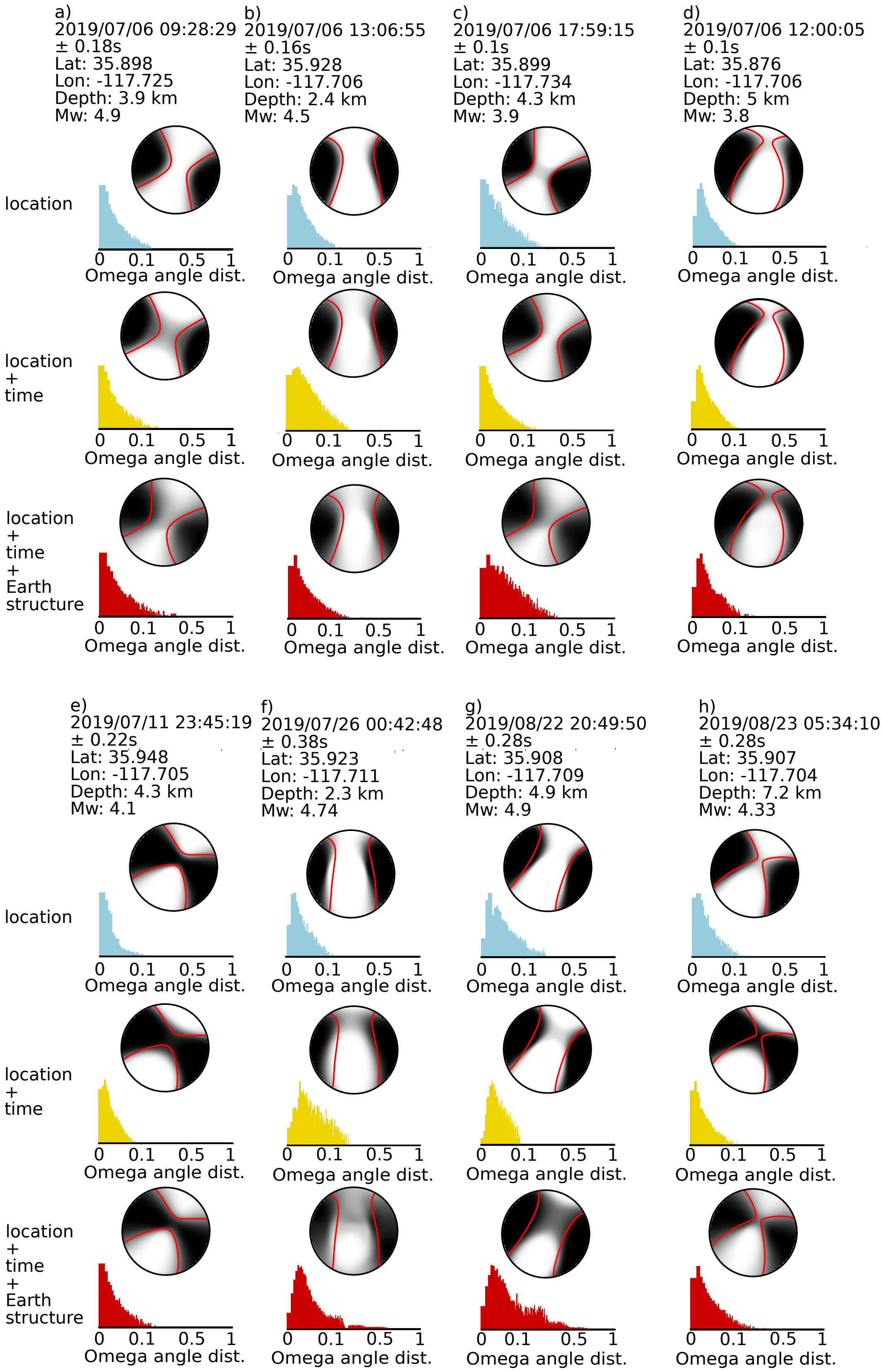

Figure 4: Inferred ensembles of full MTs considering different uncertainties. a) to h) show three fuzzy beachballs (BB), each based on 6000 MT predictions, where the reference MT (from SCEDC) is marked by red lines. The top BB is based on the predictions from BNNs trained on the *Mojave* Earth structure model and considering uncertainty in centroid location. The middle BB additionally includes uncertainty in the centroid time. The bottom BB shows the inferences of BNNs additionally considering inferences from all three Earth structure models. The normalized histogram of the omega angle distances d_ω between the reference MT and the ensemble of MT predictions is shown below each BB. Note that the x-axis for all subfigures scales quadratic.

In addition to uncertainty in centroid location we consider in the following uncertainites in the centroid time, which are also provided by the SCEDC catalog for each event. These uncertainties differ from earthquake to earthquake but they do not exceed 0.4 s for the considered aftershocks. When uncertainties in centroid times are considered the widths of some of the d_ω histograms increase for some ensembles of MT predictions (Fig 4,b,f) confirming the quality of the absolute centroid times of these aftershocks determined by SCEDC. However, it is worth mentioning that the widths of some of the d_ω histograms also decrease for some ensembles of MT predictions (Fig 4,g,h,c) suggesting biased absolute centroid times for those aftershocks in the catalogue.

Finally, in addition to uncertainties in centroid location and time we consider uncertainties in Earth structure. We evaluate the BNNs that have been trained on the synthetics from three different Earth structures (Supp. Fig. S1). The expected arrival times and thus extracted waveform snippets will be systematically different for each Earth structure. For some of the inferred MT ensembles the spread in d_ω histograms increases and some show values of up to 0.5 (Fig 4,f-g). For those events the *Mojave* structure model seems to be the most appropriate one and therefore uncertainties in Earth structure are overestimated. For other MT ensembles (Fig 4,a,d,e,h) the spread in d_ω histograms decreases or stays similar, meaning that the uncertainties in Earth structure are less crucial for those events. Nevertheless, the resulting ensembles of predicted MTs also comprise the solutions of considering only location uncertainty (Fig. 6) and, the maximum a-posterior (MAP) solution still shows good agreement between extracted waveform data snippets and synthetic waveforms calculated from the predicted source parameters (Fig. 5).

3.2 Inferences for double-couple moment-tensors

The SCEDC catalog also contains 198 pure double-couple focal mechanisms for events that occurred in the area of interest, which we refer to as reference in the following. Without visual inspection we let for the waveform data of each of those events our BNNs infer ensembles of 6000 MT solutions considering centroid location and time uncertainty. We compare the 198 reference focal mechanisms with our ensembles of MT parameter predictions from our framework by setting the predicted v and w values to zero, representing a pure double-couple source (Fig. 7,a). We also show d_ω between the reference mechanism and the predicted full seismic MT ensembles (Fig. 7,b). The additional degree of freedom of full MT solutions versus DC constrained solutions results in broadening and a slight shift of the histogram towards higher d_ω (Fig. 7,b).

With decreasing earthquake magnitude the spread of d_ω of the trained networks is increasing comparing the predicted ensembles of MT for the 198 earthquakes (Fig. 7,c). This spread is expected as the signal-to-noise ratio decreases with lower magnitude and larger d_ω values are expressions of an increase in uncertainty of the MT ensembles. However, the bulk part of d_ω shows distances below 0.1 and the predicted ensembles are in good agreement with the reference solutions across different magnitudes 2.7-4.5 (Fig. 7,d-g). We also notice a slight increase in the omega angle distances between reference and predicted source mechanisms for the largest of the 198 earthquakes. This might indicate a need for incorporating non-DC components in the source mechanism; whereas these components are missing in the catalogue descriptions.

4 Discussion

In general, we find a good agreement between the ensemble of predicted MTs and the independently determined and unseen moment tensor solutions from the SCEDC. Only a few predicted moment tensor ensembles show systematic differences (Fig 4,f and g), which could be due to several reasons, e. g. differences in the station configurations.

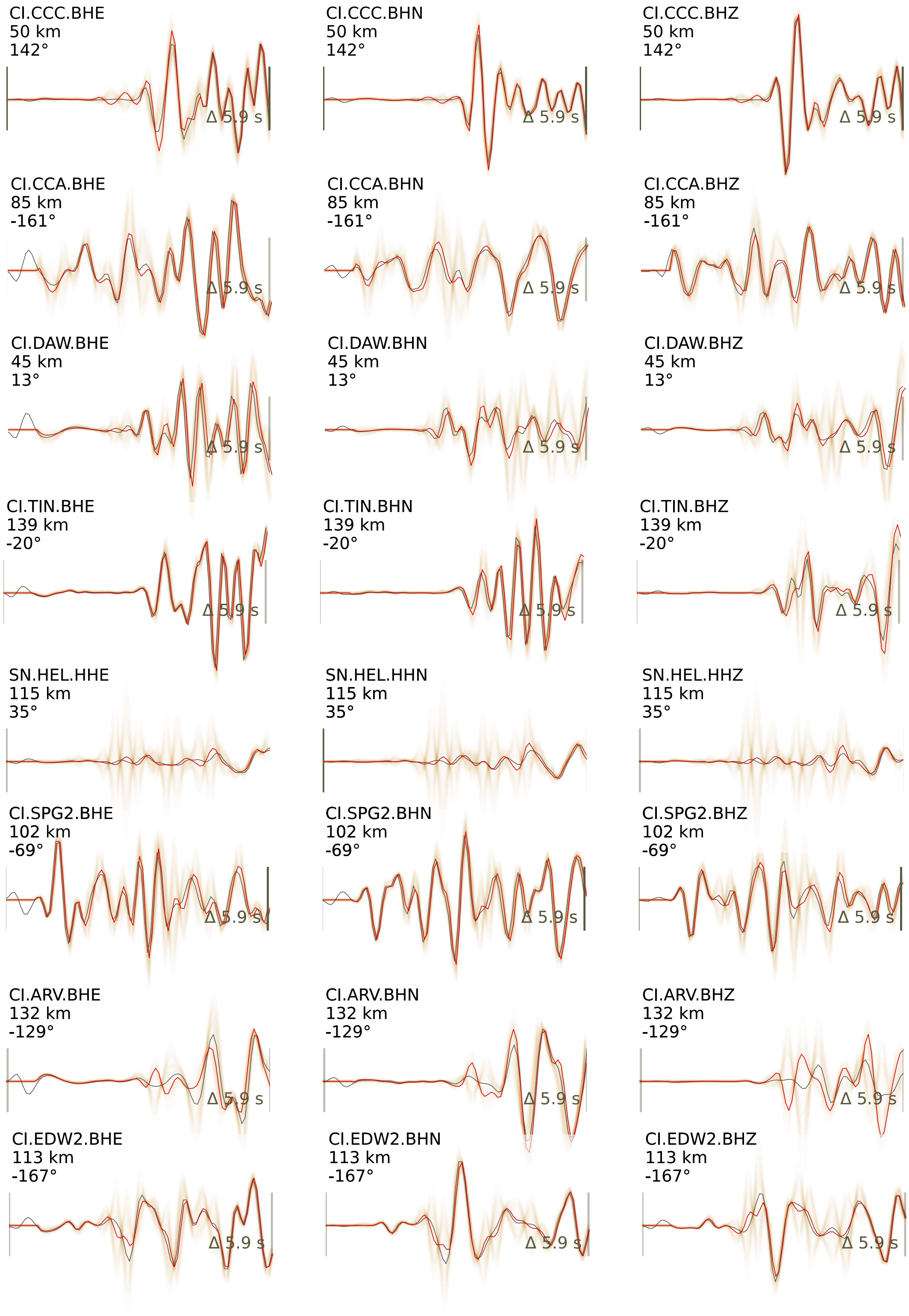

Figure 5: Exemplary waveform fits between observed waveforms (black) and synthetic waveforms based on the ensemble of estimated MT parameters (brown) with the MAP in red, for the Mw 4.1 earthquake on 2019/07/11 23:45:19. Note that the waveforms are displacements and normalized as described in section 2.1.

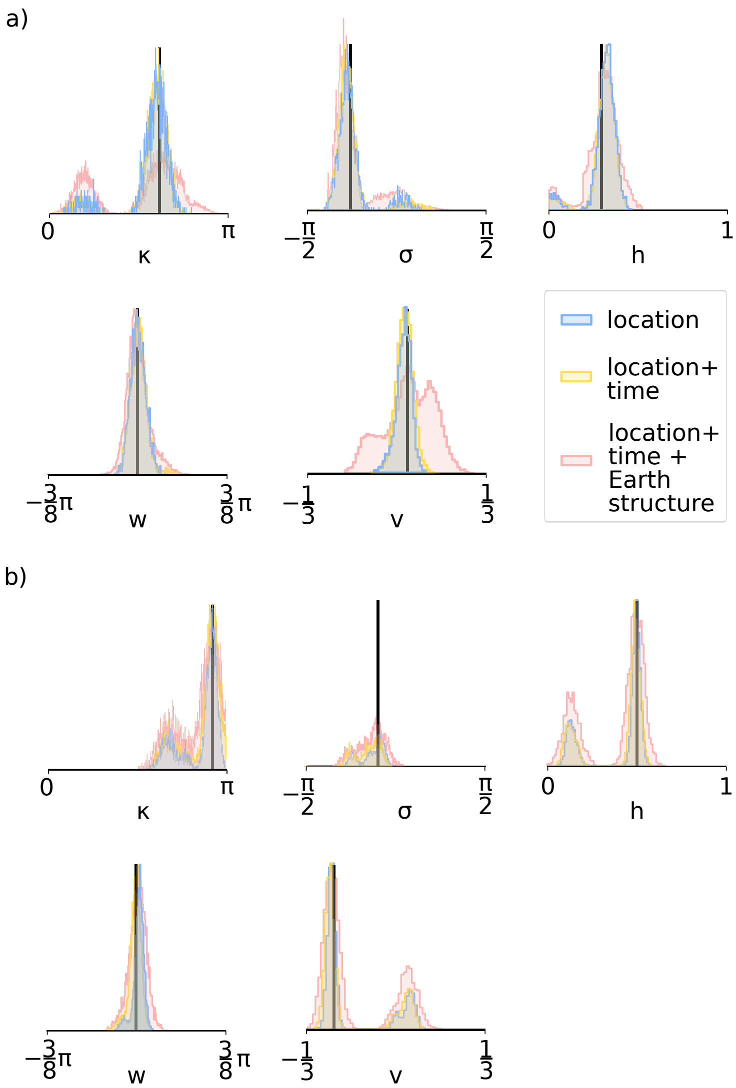

Figure 6: Normalized histograms of the ensemble of 6000 source parameter predictions for a) the Mw 3.8 earthquake on 2019/07/06 12:00:05 and b) the Mw 4.74 earthquake on 2019/07/26 00:42:48. Blue colors indicate the ensemble of predictions when using only the *Mojave* structure model and only considering error in location, yellow colors when using the *Mojave* structure model and considering error in location and timing and red colors the ensemble from all three considered Earth structure models and also considering errors in centroid location and timing.

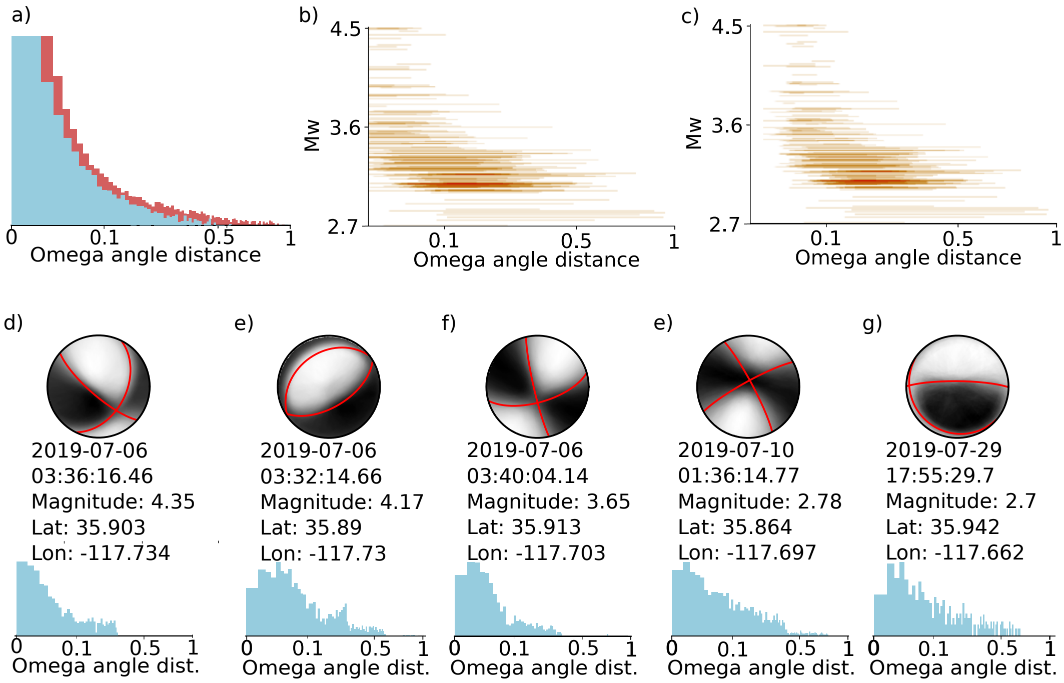

Figure 7: MT ensemble predictions for 198 pure double-couples focal mechanisms in the SCEDC catalog. a) omega angle distance as normalized histograms for v and w values set to zero in blue and in red with v and w left open, b) Density plot of lines drawn by the minimum and maximum omega angle distance d_ω for each earthquake magnitude as provided by the SCEDC catalog, with v and w set to zero, c) the same as in b) but with v and w left open, d) to g) show fuzzy beachballs, each based on 6000 moment tensor predictions, where the reference moment tensor (from SCEDC) is marked by red lines. Below: normalized histograms of the omega angle distances d_ω between the reference MT and the ensemble of MT predictions. Note that the x-axis for all subfigures scales quadratic.

Some of those systematic differences partly vanished by including also uncertainty in centroid time into the variational inference scheme (Fig 4, g). As we estimate the full seismic moment tensor the distribution and density of the non-DC components from the predicted ensemble can be inferred (Supp. Fig. S3 and S4). The main regions of high probability of solutions is consistent considering different sources of theory error. However, larger uncertainties for both the CLVD as well as the isotropic components, i.e. the lune v and w parameters, can be observed when additionally considering errors in Earth structure models (Supp. Fig. S3). It has been shown that an error in Earth's structure is often compensated by increased CLVD and isotropic components (Vasyura-Bathke et al., 2020).

We note that we evaluate the prediction accuracy of our framework by comparison with SCEDC cataloged moment tensors. These solutions, however, could potentially also be biased, deviating from the unknown "true" earthquake source. Variance reduction could be used to estimate the precision with respect to the real waveform data. We observe larger ω angle distances between the predicted MT ensemble and reference MTs when considering the inferences from several Earth structure models (Fig4,a-h). This is not unexpected, because the reference solutions are estimated with only one of the Earth structure models. However, it is also possible that the "true" unknown solution is better represented by our ensemble of predictions considering other Earth structure models. In regions with well known structure this approach likely overestimates the parameter uncertainties, but in regions with poorly known structure it might provide a more realistic representations of parameter uncertainties (Vasyura-Bathke et al., 2021).

The observation of a relation between spread of inferred parameter uncertainties with magnitude is a result of parameter selections before learning, such as filter and time window length, as well as decreasing signal-to-noise ratios for lower magnitudes. Our considered filter frequencies are optimal for earthquakes with magnitudes Mw 3 to 4, of which hundreds occurred during the 2019 Ridgecrest sequence (Ross et al., 2019). The station distribution around the Ridgecrest area and the good quality of the waveform data due to mostly remote station locations is exceptional and together with the statistically significant number of earthquakes this study area is bench-marking showcase to demonstrate the robustness and performance of our approach. It remains to be evaluated whether our approach performs equally well in areas with a sparse station network under worse noise conditions.

The novelty of our proposed framework lies in the estimation of ensembles of the full seismic MTs yielding uncertainties in parameter estimates based on seismic waveforms. A shortcoming in our approach is the current limited transferability of the trained BNNs to other study areas, unlike P-wave first motion polarity based approaches (e. g. Ross et al., 2018). We assume that under operational conditions on live incoming waveform data the prediction of the ensemble of full seismic MTs using the presented framework can be done a few tens of seconds after the earthquake, being almost near-realtime. Main factors that influence this response time are: 1) Our algorithm considers a waveform window of 5 s. 2) The safe restitution of the waveform data into displacement to avoid filter effects requires that at least several seconds of data are available (around 2 s for the chosen frequencies in the case study). 3) In its current form our approach requires the detection and location of an earthquake, which can be used to infer a centroid time and optionally, its uncertainty as prior knowledge. However, these can be delivered fast by other deep learning methods (Kriegerowski et al., 2019). 4) Finally, the evaluation of the waveform data by a single trained BNN takes a few hundred milliseconds and can be done in parallel for several BNNs at the same time. Hence, approaches based on P-wave first motion polarity only (Hara et al., 2019; Ross et al., 2018; Uchide, 2020) will likely outperform our proposed framework in terms of response time. Nevertheless, these time factors are not of importance for already cataloged data in a database, which can be searched fast by keeping the recorded waveform data in memory.

In principle, the presented method can be made independent of the particular station configuration at the expense of computational cost. This could be accomplished by calculating the synthetic waveforms for a distance-depth grid of locations and shifting the source and receiver relatively or by assuming a location grid of abstract receivers (van den Ende & Ampuero, 2020). The actual station locations can then be mapped to such an abstract receiver grid by interpolation or nearest neighbour. However, we do not expect that the framework could be made transferable to other regions, because of the characteristics of the assumed Earth structure models that are learned by the BNNs.

The choice of training a BNN for each considered grid point instead of training a single large neural network with waveforms from all possible locations, such as in Kuang et al. (2021), is a key point in our approach which allows us for estimating MT parameter uncertainties considering uncertainty in centroid time and location as well as uncertainty in Earth structure. Training a single large neural network with waveforms from all potential source locations would require to estimate additionally three location parameters (latitude, longitude and depth) as labels. This significantly increases the non-linearity of the problem and, consequently increases the required complexity of the neural network architecture, i.e. the number of trained filter weights and biases. In our view, a simple network architecture with few trainable parameters is favorable (Mignan & Broccardo, 2019) and, therefore, we chose to train multiple, but individually rather simple networks.

As a by-product of our approach it turns out that our BNNs also learned to be sensitive to the centroid location. Assuming that an earthquake occurred somewhere in the grid of BNNs, each BNN can be queried to return the log-likelihoods for the input data. The highest log-likelihoods should stem from BNNs learned for grid locations close to the true centroid location. We test this assumption for a Mw 3.9 earthquake included in the SCEDC catalog and indeed find a correlation of the log-likelihood values with distance to the centroid location (Fig. 8). As prior information only the centroid time and optionally its uncertainty is needed.

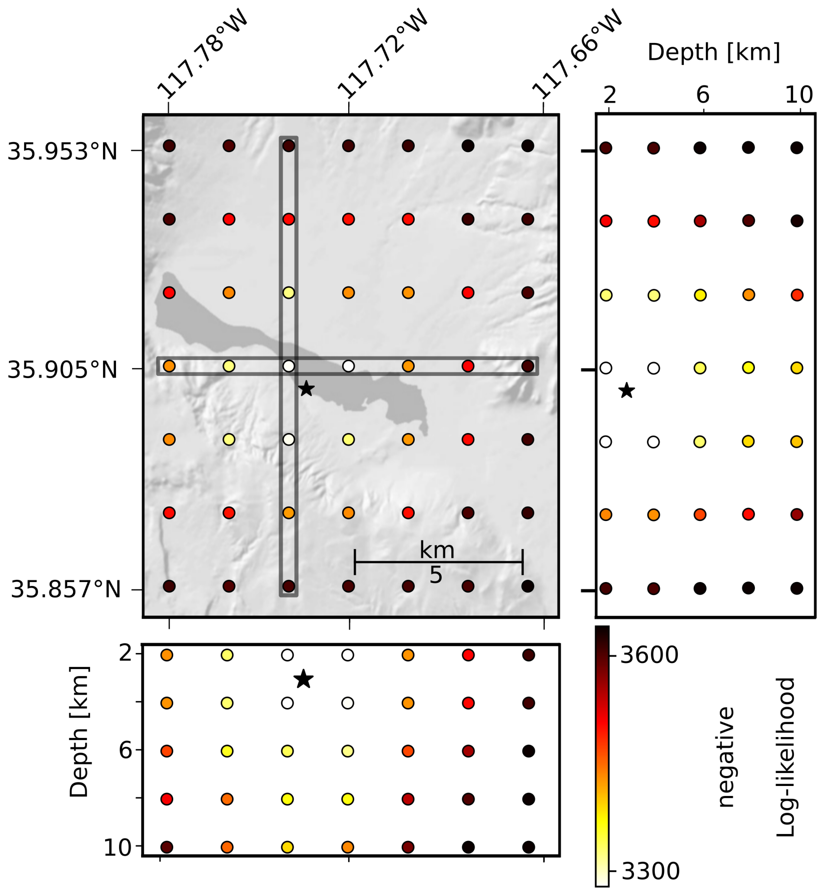

Figure 8: Earthquake centroid location inference. The grid points are colored by the negative log-likelihood values as inferred from evaluation of the BNNs for the real waveform data of the Mw 3.9 at 2019-07-06 17:59:15. The map view shows grid points at 4 km depth, whereas side views left and bottom show the grid-points at depth versus latitude and longitude along the profiles outlined with grey rectangles in the map view, respectively. The black star marks the centroid location as given by SCEDC for this earthquake.

5 Conclusions

We demonstrated that variational inference based on deep learning of Bayesian Neural Networks shows the capability to not only reproduce optimum parameter estimates of classical full moment tensor inversion, but it also yields uncertainties of the inferred MT parameters in near-real time. Our presented approach is flexible enough to optionally account for various cases of theory error that are well known to affect MT parameter estimates, i.e. errors in centroid location and time as well as errors in the assumed Earth structure.

The presented method has been successfully applied on local scale using field data of a subset of the Ridgecrest 2019 aftershock sequence, comprising 206 earthquakes with magnitudes Mw 2.7 to 5.5. The inferred ensembles of MT parameters have been compared to independently determined source mechanisms by the SCEDC.

One limitation of the presented approach is the non-transferable nature of the trained networks as they are trained for a specific Earth structure model, station setups, frequency filters and phase arrival time windows.

Our approach demonstrates the capabilities and the potential of machine learning for near-real time earthquake source mechanism estimation of small earthquakes with associated uncertainties. These are important information for hazard assessments and for providing other products to policy makers and public which are based on earthquake source analysis, e.g. shakemaps. The presented framework has the potential to be expanded upon and to be used in standardized automatic operational procedures.

Data availability statement

Data from regional seismometers are available via FDSN services from GEOFON and IRIS. The Caltech/USGS Southern California Seismic Network (SCSN) earthquake catalog, along with metadata and other ancillary data, such as moment tensors and focal mechanisms as been used and are available at `http://www.data.scec.org/index.html`. The Green's function stores used here are uploaded on Zenodo under DOI: 10.5281/zenodo.4643478 We make the code available and only use open-access waveform data for testing. We make the software available as jupyter notebook in the supplement and with pre-calculated example data as-well on Zenodo under DOI:10.5281/zenodo.4646666.

Acknowledgments

A.S was funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) and was supervised by Project Management Jülich (PtJ) (grand number 03EE4003A). H.V-B was partially supported by Geo.X, the Research Network for Geosciences in Berlin and Potsdam under the project number SO_087_GeoX.

We thank Nima Nooshiri for helpful discussions. We thank the colleagues at SCEDC for maintaining open data policies and open access catalogs. We use Keras (Chollet & Others, 2015) and TensorFlow probability (Abadi et al., 2016; Dillon et al., 2017) to build our network architecture.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... others (2016). Tensorflow: A system for large-scale machine learning. In *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)* (pp. 265–283).

Aki, K., & Richards, P. G. (2002). *Quantitative seismology.*

Bassin, C., Laske, G., & Masters, G. (2000). The current limits of resolution for surface wave tomography in North America. *EOS Trans. AGU*, *81*(F897).
Blundell, C., Cornebise, J., Kavukcuoglu, K., & Wierstra, D. (2015). Weight Uncertainty in Neural Networks. In *Proceedings of the 32nd international conference on machine learning (icml)* (Vol. 37, pp. 1613–1622).
California Institute Of Technology And United States Geological Survey Pasadena. (1926). *Southern california seismic network.* International Federation of Digital Seismograph Networks. Retrieved from `http://www.fdsn.org/doi/10.7914/SN/CI` doi: 10.7914/SN/CI
Cesca, S., Şen, A. T., & Dahm, T. (2014). Seismicity monitoring by cluster analysis of moment tensors. *Geophysical Journal International*, *196*(3), 1813–1826.
Chollet, F., & Others. (2015). *Keras.* \url{https://keras.io}.
Dillon, J. V., Langmore, I., Tran, D., Brevdo, E., Vasudevan, S., Moore, D., ... Saurous, R. A. (2017). Tensorflow distributions. *CoRR*, *abs/1711.10604*. Retrieved from `http://arxiv.org/abs/1711.10604`
Duputel, Z., Rivera, L., Fukahata, Y., & Kanamori, H. (2012). Uncertainty estimations for seismic source inversions. *Geophys, J. Int.*, *190*(2), 1243–1256. doi: 10.1111/j.1365-246X.2012.05554.x
Ekström, G., Nettles, M., & Dziewoński, A. M. (2012). The global {CMT} project 2004–2010: {C}entroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200*, 1–9.
Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 315–323).
Graves, A. (2011). Practical variational inference for neural networks. *Advances in Neural Information Processing Systems (NIPS)*, 2348–2356.
Hanka, W., & Kind, R. (1994). The {GEOFON} program. *Annals of Geophysics*, *37*(5).
Hara, S., Fukahata, Y., & Iio, Y. (2019). P-wave first-motion polarity determination of waveform data in western japan using deep learning. *Earth, Planets and Space*, *71*(1), 1–11.
Hauksson, E., & Unruh, J. (2007). Regional tectonics of the Coso geothermal area along the intracontinental plate boundary in central eastern California: Three-dimensional Vp and Vp/Vs models, spatial-temporal seismicity patterns, and seismogenic deformation. *Journal of Geophysical Research: Solid Earth*, *112*(B6).
Heimann, S., Kriegerowski, M., Isken, M., Cesca, S., Daout, S., Grigoli, F., ... Dahm, T. (2017). Pyrocko - An open-source seismology toolbox and library. *GFZ Data Services*, *v. 0.3*. Retrieved from `www.pyrocko.org` doi: http://doi.org/10.5880/GFZ.2.1.2017.001
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.
Hutton, K., Woessner, J., & Hauksson, E. (2010). Earthquake monitoring in southern California for seventy-seven years (1932–2008). *Bulletin of the Seismological Society of America*, *100*(2), 423–446.
Ichinose, G. A., Anderson, J. G., Smith, K. D., & Zeng, Y. (2003). Source parameters of eastern california and western nevada earthquakes from regional moment tensor inversion. *Bulletin of the Seismological Society of America*, *93*(1), 61–84.
Jordan, T. H., & Maechling, P. J. (2003). The SCEC community modeling environment: An information infrastructure for system-level earthquake science. *Seismological Research Letters*, *74*(3), 324–328.
Käufl, P., Valentine, A. P., O'Toole, T. B., & Trampert, J. (2014). A framework for

fast probabilistic centroid-moment-tensor determination—inversion of regional static displacement measurements. *Geophysical Journal International*, *196*(3), 1676–1693.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kriegerowski, M., Petersen, G. M., Vasyura-Bathke, H., & Ohrnberger, M. (2019). A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms. *Seismological Research Letters*, *90*(2A), 510–516.

Kuang, W., Yuan, C., & Zhang, J. (2021). Real-time determination of earthquake focal mechanism via deep learning. *Nature Communications*, *12*(1), 1432. Retrieved from `https://doi.org/10.1038/s41467-021-21670-x` 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. *Bulletin of the Seismological Society of America*, *110*(5), 2095–2111. doi: 10.1785/0120200099

Madariaga, R. (2007). Seismic source theory.

Mignan, A., & Broccardo, M. (2019). One neuron versus deep learning in aftershock prediction. *Nature*, *574*(7776), E1–E3.

Monastero, F. C., Katzenstein, A. M., Miller, J. S., Unruh, J. R., Adams, M. C., & Richards-Dinger, K. (2005). The Coso geothermal field: A nascent metamorphic core complex. *Geological Society of America Bulletin*, *117*(11-12), 1534–1553.

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. *Nature communications*, *11*(1), 1–12.

Ross, Z. E., Idini, B., Jia, Z., Stephenson, O. L., Zhong, M., Wang, X., ... Others (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. *Science*, *366*(6463), 346–351.

Ross, Z. E., Meier, M.-A., & 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.

(SCEDC), S. C. E. C. (2013). Southern california earthquake center. *Caltech. Dataset*.

Schoenball, M., Davatzes, N. C., & Glen, J. M. G. (2015). Differentiating induced and natural seismicity using space-time-magnitude statistics applied to the Coso Geothermal field. *Geophysical Research Letters*, *42*(15), 6221–6228.

Smith, J. D., Ross, Z. E., Azizzadenesheli, K., & Muir, J. B. (2021). Hyposvi: Hypocenter inversion with stein variational inference and physics informed neural networks. *arXiv*.

Stähler, S. C., & Sigloch, K. (2014). Fully probabilistic seismic source inversion – Part 1 : Efficient parameterisation. *Solid Earth*, *5*, 1055–1069. doi: 10.5194/se-5-1055-2014

Stähler, S. C., & Sigloch, K. (2016). Fully probabilistic seismic source inversion – Part 2 : Modelling errors and station covariances. *Solid Earth*(7), 1521–1536. doi: 10.5194/se-7-1521-2016

Tape, W., & Tape, C. (2012a). Angle between principal axis triples. *Geophysical Journal International*, *191*(2), 813–831.

Tape, W., & Tape, C. (2012b). A geometric setting for moment tensors. *Geophysical Journal International*, *190*(1), 476–498.

Tape, W., & Tape, C. (2015). A uniform parametrization of moment tensors. *Geophysical Journal International*, *202*(3), 2074–2081.

Uchide, T. (2020). Focal mechanisms of small earthquakes beneath the japanese islands based on first-motion polarities picked using deep learning. *Geophysical*

Journal International, *223*(3), 1658–1671.

Vackář, J., Burjánek, J., Gallovič, F., Zahradník, J., & Clinton, J. (2017). Bayesian ISOLA: New tool for automated centroid moment tensor inversion. *Geophysical Journal International*, *210*(2), 693–705. doi: 10.1093/gji/ggx158

van den Ende, M. P. A., & Ampuero, J.-P. (2020). Automated Seismic Source Characterization Using Deep Graph Neural Networks. *Geophysical Research Letters*, *47*(17), e2020GL088690.

Vasyura-Bathke, H., Dettmer, J., Dutta, R., Mai, P. M., & Jónsson, S. (2021, jan). Accounting for theory errors with empirical Bayesian noise models in nonlinear centroid moment tensor estimation. *Geophysical Journal International*. Retrieved from https://doi.org/10.1093/gji/ggab034 doi: 10.1093/gji/ggab034

Vasyura-Bathke, H., Dettmer, J., Steinberg, A., Heimann, S., Isken, M. P., Zielke, O., ... Jónsson, S. (2020, jan). The Bayesian Earthquake Analysis Tool. *Seismol. Res. Lett.*, *91*(2A), 1003–1018. Retrieved from https://doi.org/10.1785/0220190075 doi: 10.1785/0220190075

Wang, R. (1999). A simple orthonormalization method for stable and efficient computation of Green's functions. *Bulletin of the Seismological Society of America*, *89*(3), 733–741.

Wen, Y., Vicol, P., Ba, J., Tran, D., & Grosse, R. (2018). Flipout: Efficient pseudo-independent weight perturbations on mini-batches. *arXiv preprint arXiv:1803.04386*.

Wu, H., & Lees, J. M. (1999). Three-dimensional P and S wave velocity structures of the Coso geothermal area, California, from microseismic travel time data. *Journal of Geophysical Research: Solid Earth*, *104*(B6), 13217–13233.

6 Supplement