

Acoustic Emission Waveform Picking with Time Delay Neural Networks during Rock Deformation Laboratory Experiments

Thomas King^{*1}, Philip Benson², Luca De Siena³, and Sergio Vinciguerra¹

Abstract

We report a new method using a time delay neural network to transform acoustic emission (AE) waveforms into a time series of instantaneous frequency content and permutation entropy. This permits periods of noise to be distinguished from signals. The model is trained in sequential batches, using an automated process that steadily improves signal recognition as new data are added. The model was validated using AE data from rock deformation experiments, using Darley Dale sandstone in fully drained conditions at a confining pressure of 20 MPa (approximately 800 m simulated depth). The model is initially trained by manual picking of five high-amplitude waveforms randomly selected from the dataset (experiment). This is followed by semisupervised training on a subset of 300 waveforms.

Cite this article as King, T., P. Benson, L. De Siena, and S. Vinciguerra (2020). Acoustic Emission Waveform Picking with Time Delay Neural Networks during Rock Deformation Laboratory Experiments, *Seismol. Res. Lett.* **XX**, 1–10, doi: [10.1785/SRL20200188](https://doi.org/10.1785/SRL20200188).

Introduction

Acoustic emissions (AEs) are the transient elastic waves produced by the sudden redistribution (release) of stress when a material fractures. These signals are considered to represent the laboratory analog of natural earthquakes with wider applications spanning from nondestructive testing to vibration monitoring. In deformation experiments, AEs are related to the initiation and growth of fractures (e.g., Lockner *et al.*, 1992), matrix cracking (Scholey *et al.*, 2010), fluid flow through fractures (Benson *et al.*, 2019; Fazio *et al.*, 2019), and fiber breakage and various debonding processes (Bohse, 2000). Detection of these signals is a valuable asset as they provide a nondestructive and immediate feedback to dynamically evolving systems without the need for interference. However, a limitation of AE analysis is that results are often dependent on the successful discrimination of signals from the background noise. Because datasets are often very large, there is a need for automated picking tools that are sufficiently accurate and robust to handle waveforms in a medium that is continually evolving.

The simplest processing method to extract basic data uses stepwise calculation along the waveform time series of energy ratios (“signal-to-noise” ratio [SNR]). This type of analysis is often used to extract a discrete time series from a much longer (quasi-continuous) signal, with triggering being dependent on the arrival of large amplitude spikes (such as the first arrival of an earthquake). Although intuitive and less computationally demanding than other techniques, this method is prone to errors as the largest SNR may not refer to the first arrival of a genuine signal (Guoping *et al.*, 2004; Pomponi and Vinogradov, 2013). To reduce these false triggers, methods that

take advantage of the inherent characteristics of microseismic signals, that is, the frequency content, have been proposed (Zhang *et al.*, 2013). These methods use an automated approach via the Hilbert–Huang transform (HHT, Huang *et al.*, 1998) to identify low-frequency AE signals from high-frequency background noise. A further refinement of this approach combines the HHT with the Akaike information criterion (Akaike, 1973) to pick up the first arrival of seismic waves (Jia *et al.*, 2015).

Entropy-based techniques, which identify the incoherent nature of the noise signal by relating data to past and previous values, are also well documented. In medicine, Ródenas *et al.* (2015) reported the accurate detection of atrial fibrillation from electrocardiograms. A similar approach uses intrinsic mode functions (IMFs) to recognize focal electroencephalogram signals (Sharma *et al.*, 2015). Similar techniques have found their use at the field scale. For example, fluctuations in entropy were detected prior to the Mexican earthquake of 2017 (Ramírez-Rojas *et al.*, 2018). Entropy is highly sensitive to seismic amplitude, frequency, and phase changes (Jia *et al.*, 2019). It demonstrates immunity to noise under a wide range of environmental conditions due to a sensitivity to the coherency of a real signal, making them uniquely suited for seismological analyses. The technique can also be applied to active seismic

1. Department of Earth Sciences, University of Turin, Turin, Italy; 2. Rock Mechanics Laboratory, School of Earth and Environmental Sciences, University of Portsmouth, Portsmouth, United Kingdom; 3. Institute of Geosciences, Johannes Gutenberg University, Mainz, Germany

*Corresponding author: thomas.king@edu.unito.it

© Seismological Society of America

data (Zoukaneri and Porsani, 2015) by measuring the entropy of the instantaneous frequency content, increasing the level of information available for subsurface mapping.

Nonetheless, each of these techniques is limited by a user-defined critical value that defines when a time period is “signal” or “noise.” Often, this threshold is set based on the user experience and knowledge of the recording equipment. When tools are not calibrated, as is often the case in AE data (e.g., Høgsberg and Krenk, 2015), such an approach remains imprecise and will lead to further systematic errors regardless of the technique. Nevertheless, advances in machine learning and artificial intelligence tools are now ushering in a new era of data analysis that seeks to minimize biases enforced by user-defined parameters. Time delay neural networks (TDNNs) are automated signal classification techniques that are designed with the purpose of identifying patterns and trends in shift-invariant time-series data without explicitly knowing the beginning or end of a signal (Waibel *et al.*, 1995; Derakhshani and Schuckers, 2004; Peddinti *et al.*, 2015). Analogous to a 1D convolutional neural network (CNN), TDNNs are a form of recurrent neural network (RNN) that models the propagation characteristics of time-series data similar to entropy techniques. By constructing models of the key elements of audio, or elastic vibrations in the case of seismology, they can recognize different speakers when employed for speech recognition (Haffner and Waibel, 1992) and perform robustly in the presence of reverberations (Snyder *et al.*, 2015).

Here, we seek to take advantage of TDNNs by applying a simple noise or signal classification routine to AE data from a typical rock-deformation experiment using Darley Dale sandstone, in which the TDNN is used to define the onset of energy as the boundary between the two classifications. By iteratively training a neural network (NN) on a calculated time series of seismic envelopes, instantaneous frequency content, and estimated signal entropy, the method picks AE waveforms and automatically incorporates them into the evolving (updated) model, thus improving later picks. To validate our approach, AE source locations are solved using the pick times obtained from the trained model using a time difference of arrival method (e.g., Comanducci *et al.*, 2020). AE distribution and error parameters are then compared with those obtained from picking signals with a simple amplitude threshold (AT) method.

Data and Method

AE data

A 40 × 100 mm cylindrical sample of Darley Dale sandstone was deformed using a conventional triaxial deformation apparatus (Fazio, 2017) at a confining pressure of 20 MPa, representing approximately 800 m depth. To ensure the experiment collected an AE dataset that contains a large range of amplitudes (Nakamura *et al.*, 1972) and fracture mechanisms (Stanchits *et al.*, 2006), the sample was deformed at a deformation rate of 3.6 mm/hr until brittle failure following the

protocol of Fazio (2017). For a detailed overview of the AE data acquisition methodology, please refer to Appendix A.

Model parameters

A previous study proposed using only the waveforms as input data (Derakhshani and Schuckers, 2004). After extensive testing, that approach was found to perform less robustly with the high-noise conditions of the AEs measured here. Instead, we focus on the simple attributes of signal amplitude and frequency content as training parameters to classify the waveform. The former is modeled through the seismic envelope, whereas the latter is by instantaneous frequency content obtained through the HHT. A third parameter, permutation entropy (Unakafova and Keller, 2013), encompasses both aspects. A fourth time series, defined as the target model output, is used to classify the AE waveforms through binary separation of noise and signal data, assigned -1 and 1 , respectively. Signal is defined as the time period between the calculated onset of energy and the point at which energy drops below a prepick noise threshold. Because of uncertainties in the waveform content following the main arrival (i.e., reflections) the model is only trained on noise identified during the presignal period.

The seismic envelope is the most intuitive and useful parameter to apply as it represents the instantaneous energy of the signal (Fig. 1a). Typically, noise is represented by low-amplitude values and signal by high-amplitude values. The boundary between these two is often defined by a rapid increase in values as a signal arrives. However, in the presence of low-amplitude data, or strong scattering in which the envelope of the signal can become broadened, the increase becomes shallower or simply not present (De Siena *et al.*, 2016). To address this, the HHT provides a measure of instantaneous frequency content along the entire time series (Fig. 1b) that is independent of amplitude; therefore, any boundary between classifications is still preserved when energies are low. Noise is identified through a dominance of high-frequency energy, whereas the signal is identified through a consistent dominance of low-frequency waves. However, on its own, this method is likely to fail in the presence of high-frequency signals that are similar in content to the background.

The final training parameter, permutation entropy, describes the uncertainty and the degree of irregularity of a random time series (Fig. 1c). Ordinal patterns that represent the relationship between past and future values at each time-step of the waveform (i.e., an increasing trend) are calculated. A probability density function of the number of patterns within a moving window is then calculated, thus removing any dependence on the amplitude of the original waveform (Unakafova and Keller, 2013). The larger the value of permutation entropy, the higher the diversity of ordinal patterns is and the more complex the input data are. Thus, noise is characterized by high values and signal by low; however, the

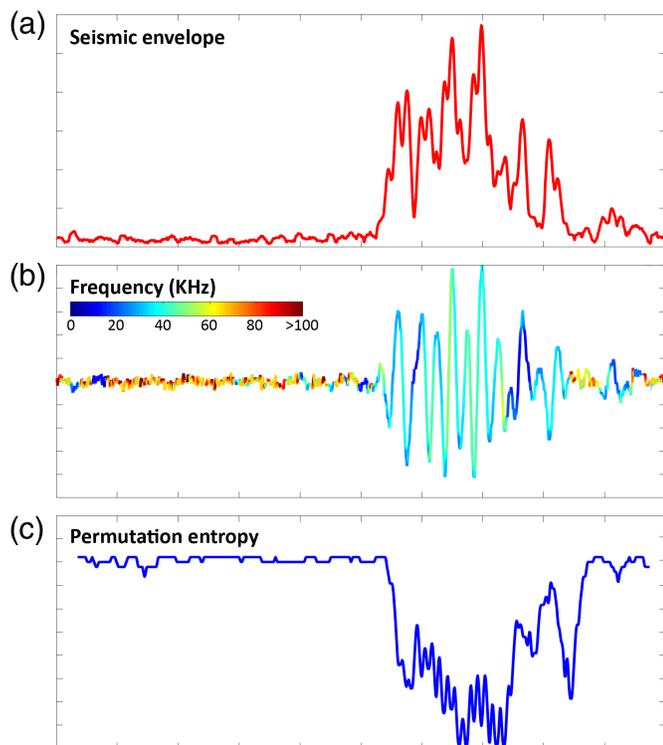


Figure 1. Input parameters used in model training. Data classifications, indicated previously, are noise and signal. (a) Seismic envelope. High values typically denote signal; however, low-amplitude data are poorly characterized. (b) Dominant frequency content of the acoustic emission (AE) waveform. Red and blue lines indicate high- and low-frequency content, respectively. Presignal noise is characterized by high frequency and the signal itself with low values. (c) Permutation entropy behaves similarly to seismic envelope, but the trend is opposite with low values identifying signal. Even at low amplitudes, permutation entropy detects coherency in frequency content. The color version of this figure is available only in the electronic edition.

boundary is slightly shifted to that of the seismic envelope. A more detailed overview on the training data is provided in [Appendix B](#).

Training routine

Similar to the previously described entropy method, a TDNN models the temporal trend (or trajectory) of the training data, classifying patterns in the data according to the target time series. A key difference of a TDNN from other multilayer NNs is that patterns are classified with shift invariance; a specific pattern may occur at any point in the time series. With each successive layer in the network, increasingly coarse trends in the training data are identified, therefore ensuring the model learns the most important features of the chosen classifications.

Five high-amplitude waveforms are randomly selected from an AE pool to create an initial model that is trained under supervised conditions. Their onsets are manually picked,

and the model iteratively trained on the concatenated input parameters (training dataset) after each waveform. At this stage, the model already has a relatively high degree of picking accuracy. Following this, the model undergoes semisupervised training. A waveform of any amplitude is randomly selected from the available pool. The input parameters are both calculated and simulated in the NN model generating an output.

Although easier than picking the signal directly from the waveform, it is still difficult to pick the signal onset from the model output due to the uncertain transition from noise to signal. With knowledge of the sample velocities, an effective approach was to define a window between the origin time (or first arrival time when picking subsequent waveforms in a single AE group) and the maximum possible arrival time for energy traversing the length of the sample. Within this window we calculate the instantaneous frequency content of the model output to characterize the rapid change in model values as a signal is detected. The onset of energy is then set when the frequencies and model output exceed 14 KHz and -0.95 , respectively.

Once the onset is identified, two picking quality ratios are calculated around this value: a short-term SNR of the original waveform and an SNR of the model output. After extensive testing, we set their values to a minimum of 2.0 for the former and between 0.3 and 0.9 for the latter as these identified a reliable onset when the training dataset was still small. When both ratios are met, the input parameters are added to the training data and the NN is updated. To improve computation speed, the model is trained in batches of 10 waveforms. Waveforms can be further downsampled to increase computation speed, although this will decrease data quality. Because this is an iterative process (Fig. 2), the model is steadily improved over time, which allows for the later inclusion of data that may not have been initially included. The model ceases training once the training dataset exceeds 300 waveforms as the quality of picking did not improve significantly after this stage.

To provide a robust comparison and best highlight the improvements offered by the proposed NN approach, AE data are also processed using a simple amplitude-threshold method to pick data. The root mean square (RMS) envelope of each waveform is calculated, and the pick time is chosen when the envelope exceeds 1.1 times the level of the background noise (using the same windowing approach as before).

Source location

Pick times are inverted for source location using a time difference of arrival (TDOA) method (Tobias, 1976). Although this method is only robust in the case of weak velocity anisotropy, the TDOA method is well established in microseismic source locations (e.g., Comanducci *et al.*, 2020) and works well for the near-isotropic Darley Dale sandstone as used here. These data are calculated by pairwise subtraction of the time of arrival values to each sensor from a single source. A fixed velocity is used in the calculation that is updated as time progresses (see

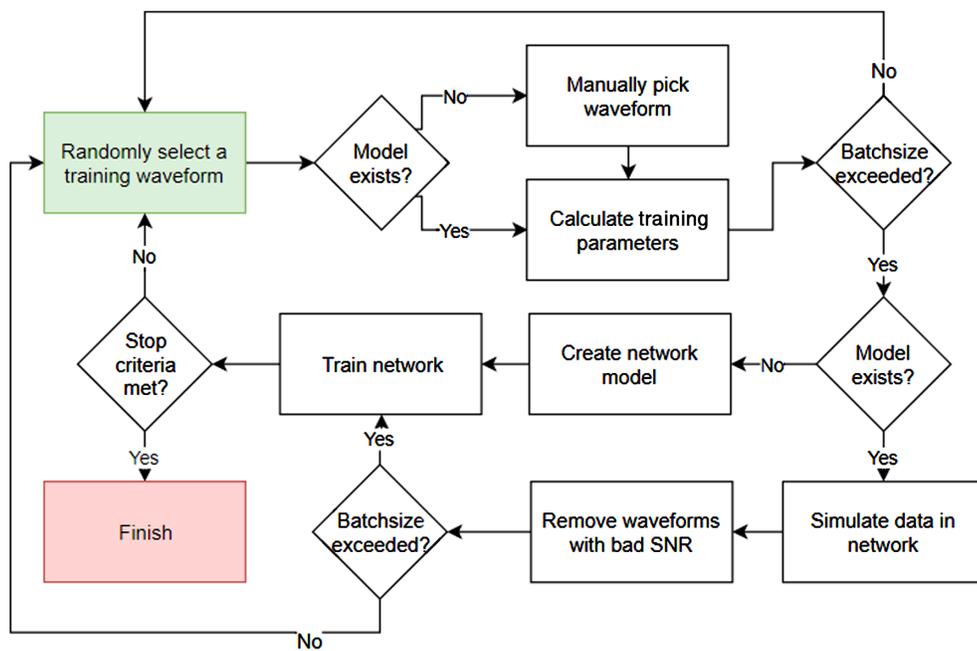


Figure 2. Flowchart detailing key elements of the training loop. SNR, signal-to-noise ratio. The color version of this figure is available only in the electronic edition.

Appendix A). The calculated TDOA values are determined through iterative estimation of the source location. The process iterates the location using the L2 norm of the location residual, comparing between calculated and TDOA values. It arrives to a local minimum at the best-determined source location. To be valid, sources are located with a minimum of six arrival times, although the accuracy greatly improves with the addition of more sensors.

Results and Discussion

At high amplitudes, both the NN and the AT methods performed equivalently. This shows that the seismic envelope is a key marker of an AE arrival. However, as amplitudes decrease and reach the level of the background noise, discrepancies in pick times become increasingly apparent (Fig. 3). Although it performed relatively well with impulsive arrivals, the AT method (Fig. 3, blue line) struggled with the strongly scattered or emergent arrivals that are characteristic of low-amplitude AE data. Such data are challenging for picking methods, as the first arrival of the *P* wave becomes more discriminatory due to the convolution of many wave modes that extend into the coda (e.g., Grosse and Ohtsu, 2008). The NN, on the other hand, was much more consistent in capturing the first arrival regardless of amplitude, wave mode, or waveform character. By modeling characteristic sequences of data, the TDNN is able to classify the patterns that uniquely define signal or noise in the time series. From the model output (Fig. 3, red line), the arrival time may then be obtained from a far simpler time series, even

for data that were not picked by the AT method. The windowing scheme was crucial to avoid picking late high-energy coda arrivals or even the coda from a prior AE in the prepick noise.

Although relevant in all forms of waveform analysis (e.g., seismic, microseismic), the occurrence of external (outside the sample) or electronic noise (often high amplitude, high frequency) with temporal and frequency characteristics similar to real events is a characteristic element of AE data processing. Waveform frequency data have already been shown to be a reliable parameter to characterize waveform data (96% accuracy), but they still lead to false-alarms in the prepick noise without additional constraints on arrival

time (Jiang and Zheng, 2020). Similar results are obtained when only amplitude data are used (Chen et al., 2019), but such an approach is likely to fail in high-noise environments. The method presented here is distinct from the previously mentioned studies in that it combines multiple input parameters to train the model (amplitude, frequency, and entropy). This provides a notable advantage over single feature methods as a more robust characterization of noise (rather than only signal) was obtained. However, the TDNN performed poorly in the presence of very high-amplitude data ($\text{SNR} > 800$). Because of the way the signal is decomposed during the HHT, the noise window is characterized by low-frequency information, thus resulting in errors during training. Because these data occurred very rarely, it was not considered detrimental to the overall model.

Sources that were located using both methodologies reveal the time-dependent evolution of the fault plane through the entire experiment (Fig. 4). Although events with high-location residuals are displayed here for completeness, the distribution of AE hypocenters is consistent with post-test visual inspection of the final fault (Fig. 4a, e.g., Lockner et al., 1992; Benson et al., 2010). Considering the evolution through time (Fig. 4b), we note that the event locations are diffuse throughout the sample during the early stages of deformation (windows 1, 2, and 3). As strain increases, during the period of strain-softening (window 4), the distribution of AE locations starts shaping a fault structure (strain localization). During crack coalescence and brittle failure (window 5), event

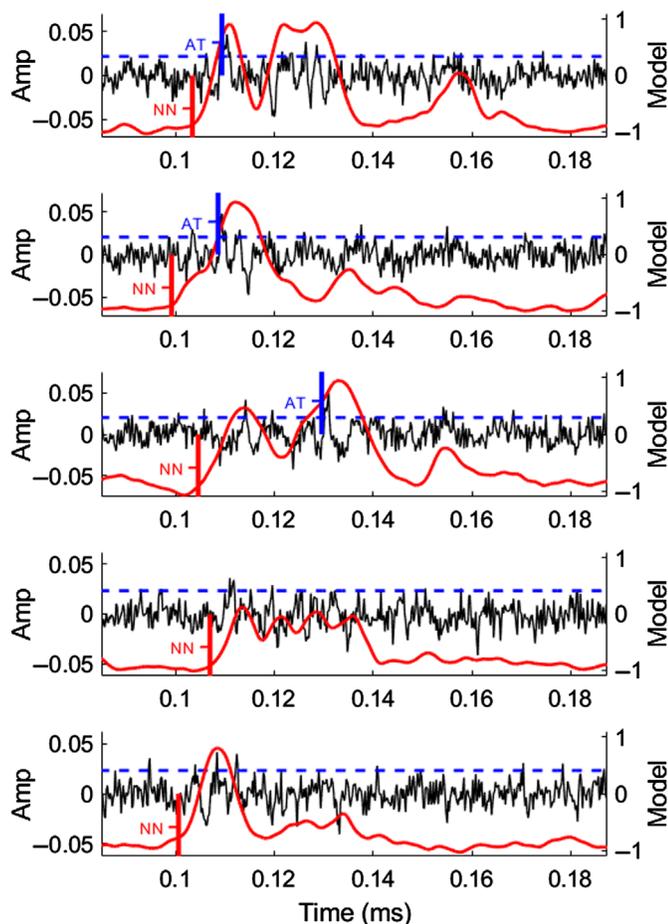


Figure 3. AE signal (black), neural network model output (NN, red), and amplitude threshold (AT, blue) highlight the discrepancies in picking quality between the two methods at low amplitudes. Strongly scattered or emergent waveforms are frequently miss picked due to late arrivals in the coda. The AT method also misses many low-amplitude arrivals that the NN is still able to detect. The color version of this figure is available only in the electronic edition.

locations are primarily located along the observed failure plane. Although the general distribution of events remains the same for both methods, inconsistencies in waveform picking (up to 0.1 ms) in Figure 3 resulted in differences in source locations that vary between 0.2 and 6 cm distance (Fig. 4, colored dots). The example events were randomly selected, and we did not observe a clear relationship between signal amplitudes and source location repeatability. A key component of the presented method is that the model is trained on data that came from the same dataset that is due to be picked. As AE datasets are typically very large, it is unfeasible to manually pick a dataset to truly validate our source locations. Nonetheless, a focusing of events toward a failure plane is much more evident when using picks from the NN compared with the AT method.

In terms of overall performance, TDNNs (or RNNs) are generally considered to be equivalent to the more powerful CNN for use in signal detection (Swapna *et al.*, 2018; Guo *et al.*, 2020; Yao *et al.*, 2020). Moreover, the number of training data required to create a well-performing model also remains relatively low for both approaches (e.g., Chen *et al.*, 2019). An important difference, however, is that the TDNN presented here outputs a transitional classification rather than the absolute output of a CNN. This provides a measure of uncertainty that is time dependent and thus physically much more appropriate for data generated by time-dependent processes such as waveform data and for the vast majority of prefailure processes.

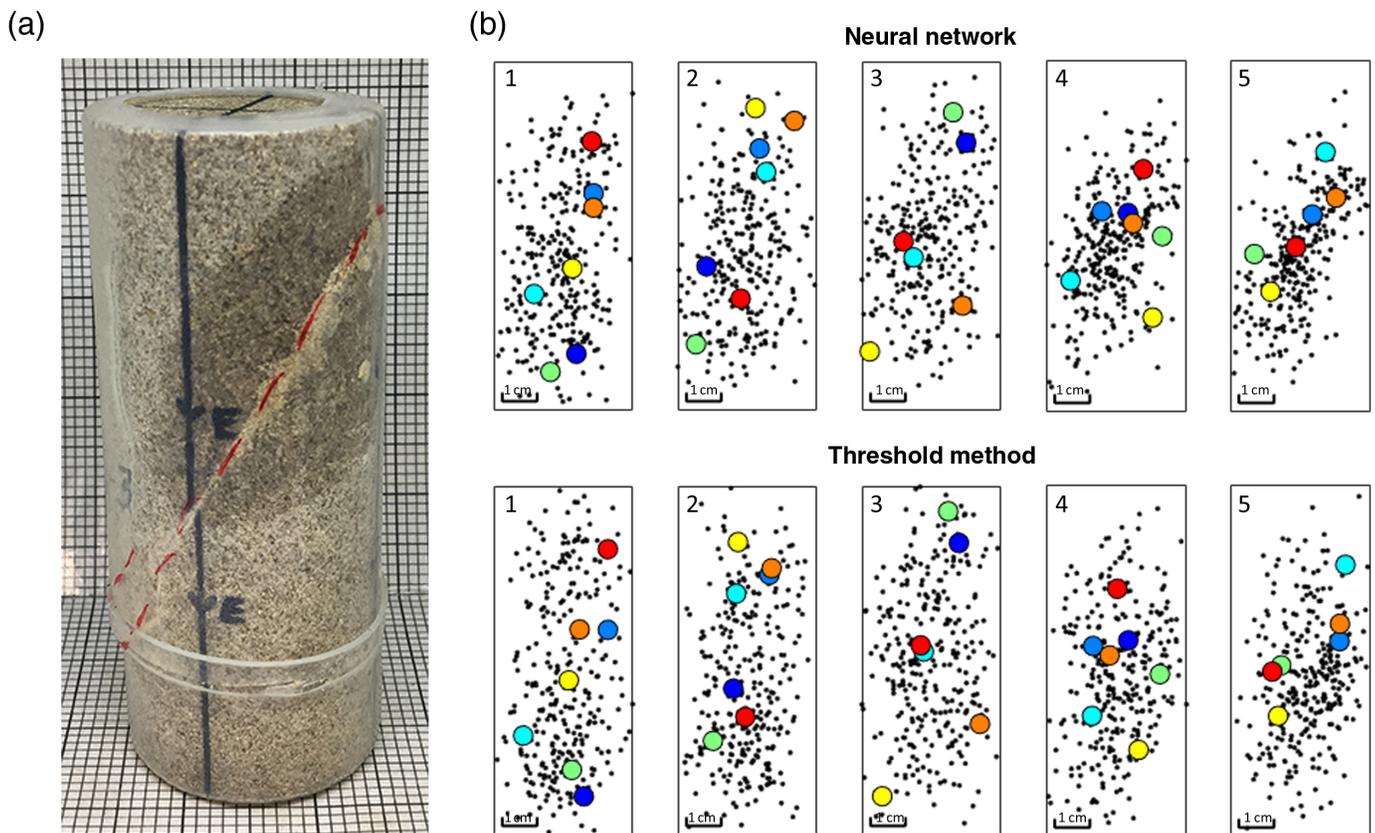
The semisupervised approach presented here may be considered advantageous due to the low-workload requirements when training the model on new datasets. To initiate a new training routine, only five high-amplitude waveforms need to be manually picked before the model can train itself. In addition, as the model improves, more low-amplitude data are incorporated into the model automatically, thus removing the requirement to manually pick difficult data. The choice of network, CNN or TDNN, is therefore tuned to the end-user requirements rather than on any standard performance metrics.

In any case, the models derived using the method proposed here may be applied to new datasets or modified as a triggering tool to use with continuous or real-time data. An important advantage of the TDNN over CNNs is that the input time series may be of any length, regardless of those used in the training routine. Rather than identifying individual elements (such as the first arrival), TDNNs characterize sequences of features to classify patterns. Thus, a continuous dataset may first be down-sampled to generate a quick approximation of arrival times (trigger), and the exact same model may then be used on the higher resolution recording to obtain a more accurate pick.

Conclusions

This study has presented a workflow and application of a machine learned waveform picking tool. A time delay NN is trained to recognize the onset of AE energy using instantaneous frequency, seismic envelope, and entropy measurements. Statistical results demonstrate the reliability of the method and highlight the potential of using multiple waveform characterization techniques to determine the arrival of acoustic energy.

We conclude that an automated process may be readily applied to AE datasets, as recorded here from a triaxial rock deformation experiment. Aside from the five waveforms used to initially train the model, no user intervention is required. A high degree of picking accuracy and the inclusion of low-amplitude data that may be missed by traditional single-parameter threshold methods result in datasets with a high-source location recovery rate and a reduction in the propagation of errors in further analysis of the data.



Data and Resources

Acoustic emission (AE) data are obtained at the Rock Mechanics Laboratory, University of Portsmouth. Raw data files, pick times, source locations, and errors are available at <https://zenodo.org/record/3958910>. Programming codes were developed in MATLAB version 2018a and are available at <https://github.com/thomaskingunito/programming>. All websites were last accessed in December 2020.

Acknowledgments

The authors thank Emily Butcher for sample preparation and Peter Ibemesi for assistance during the laboratory experiments. The authors would also like to thank Klaus Regenauer-Lieb and David Dempsey for their helpful comments during the development of the code.

References

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle, in *Second International Symposium on Information Theory*, B. N. Petrov and F. Csaki (Editors), Akademiai Kiado, Budapest, Hungary, 276–281.

Baud, P., and P. Meredith (1997). Damage accumulation during triaxial creep of Darley Dale sandstone from pore volumetry and acoustic emission, *Int. J. Rock Mech. Min. Sci.* **34**, nos. 3/4, 24-e1–24-e10.

Benson, P. M., D. C. Austria, S. Gehne, E. Butcher, C. E. Harnett, M. Fazio, P. Rowley, and R. Tomas (2019). Laboratory simulations of fluid-induced seismicity, hydraulic fracture, and fluid flow, *Geomech. Energy Environ.* **24**, 100169, doi: [10.1016/j.gete.2019.100169](https://doi.org/10.1016/j.gete.2019.100169).

Benson, P. M., B. D. Thompson, P. G. Meredith, S. Vinciguerra, and R. P. Young (2007). Imaging slow failure in triaxially deformed Etna

Figure 4. (a) Postfailure imaging of the sample highlights a singular failure plane propagating to the lower left. (b) AE sources that are located by both methodologies (black dots). Randomly chosen locations (corresponding colored dots) highlight the discrepancy in source location due to the different picking methodologies. The color version of this figure is available only in the electronic edition.

basalt using 3D acoustic-emission location and X-ray computed tomography, *Geophys. Res. Lett.* **34**, no. 3, doi: [10.1029/2006gl028721](https://doi.org/10.1029/2006gl028721).

Benson, P. M., S. Vinciguerra, P. G. Meredith, and R. P. Young (2010). Spatio-temporal evolution of volcano seismicity: A laboratory study, *Earth Planet. Sci. Lett.* **297**, nos. 1/2, 315–323.

Bohse, J. (2000). Acoustic emission characteristics of micro-failure processes in polymer blends and composites, *Compos. Sci. Tech.* **60**, no. 8, 1213–1226, doi: [10.1016/S0266-3538\(00\)00060-9](https://doi.org/10.1016/S0266-3538(00)00060-9).

Chen, Y., G. Zhang, M. Bai, S. Zu, Z. Guan, and M. Zhang (2019). Automatic waveform classification and arrival picking based on convolutional neural network, *Earth Space Sci.* **6**, no. 7, 1244–1261, doi: [10.1029/2018EA000466](https://doi.org/10.1029/2018EA000466).

Comanducci, L., M. Cobos, F. Antonacci, and A. Sarti (2020). Time difference of arrival estimation from frequency-sliding generalized cross-correlations using convolutional neural networks, *ICASSP 2020-2020 IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 4945–4949.

Derakhshani, R., and S. A. C. Schuckers (2004). Continuous time delay neural networks for detection of temporal patterns in signals, *2004 IEEE International Joint Conf. on Neural Networks*, Vol. 4,

- IEEE Cat. No. 04CH37541, 2723–2728, doi: [10.1109/IJCNN.2004.1381082](https://doi.org/10.1109/IJCNN.2004.1381082).
- De Siena, L., M. Calvet, K. J. Watson, A. R. T. Jonkers, and C. Thomas (2016). Seismic scattering and absorption mapping of debris flows, feeding paths, and tectonic units at Mount St. Helens volcano, *Earth Planet. Sci. Lett.* **442**, 21–31, doi: [10.1016/j.epsl.2016.02.026](https://doi.org/10.1016/j.epsl.2016.02.026).
- Fazio, M. (2017). Dynamic laboratory simulations of fluid-rock coupling with application to Volcano seismicity and unrest, *Ph.D. Thesis*, University of Portsmouth, School of Earth and Environmental Sciences, Portsmouth, England.
- Fazio, M., S. Alparone, P. M. Benson, A. Cannata, and S. Vinciguerra (2019). Genesis and mechanisms controlling tornillo seismo-volcanic events in volcanic areas, *Sci. Rep.* **9**, no. 1, 1–11.
- Gehne, S. (2018). A laboratory study of fluid-driven tensile fracturing in Anisotropic rocks, *Student Theses: Doctoral Theses*, University of Portsmouth, Portsmouth, England.
- Grosse, C. U., and M. Ohtsu (2008). *Acoustic Emission Testing*, Springer Science & Business Media, Berlin, Heidelberg, Germany.
- Guo, C., T. Zhu, Y. Gao, S. Wu, and J. Sun (2020). AENet: Automatic picking of P-wave first arrivals using deep learning, *IEEE Transactions on Geoscience and Remote Sensing*, 1–11, doi: [10.1109/TGRS.2020.3010541](https://doi.org/10.1109/TGRS.2020.3010541).
- Guoping, Z., W. Yanchun, and S. Rongliang (2004). An improved method for first arrival pickup using energy ratio, *Geophys. Prospect. Petrol.* **43**, no. 4, 345–347.
- Haffner, P., and A. Waibel (1992). Multi-state time delay networks for continuous speech recognition, in *Advances in Neural Information Processing Systems*, Morgan Kaufmann Publishers, Burlington, Massachusetts, 135–142.
- Heap, M. J., P. Baud, P. G. Meredith, A. F. Bell, and I. G. Main (2009). Time-dependent brittle creep in Darley Dale sandstone, *J. Geophys. Res.* **114**, no. B7, B07203, doi: [10.1029/2008JB006212](https://doi.org/10.1029/2008JB006212).
- Høgsberg, J., and S. Krenk (2015). Balanced calibration of resonant piezoelectric RL shunts with quasi-static background flexibility correction, *J. Sound Vib.* **341**, 16–30, doi: [10.1016/j.jsv.2014.12.006](https://doi.org/10.1016/j.jsv.2014.12.006).
- Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. Math. Phys. Sci.* **454**, no. 1971, 903–995.
- Jia, B., A. Sun, Y. Pan, H. Chen, and F. Liu (2019). Accurate method for picking up the first arrival time of microseismic signals based on entropy theory, *Advances in Civil Engineering*, 2019.
- Jia, R., Y. Tan, H. Sun, and Y. F. Hong (2015). Method of automatic detection on micro-seismic P-arrival time under low signal to noise ratio, *J. China Coal Soc.* **40**, no. 8, 1845–1852.
- Jiang, T., and J. Zheng (2020). Automatic phase picking from micro-seismic recordings using feature extraction and neural network, *IEEE Access* **8**, 58,271–58,278, doi: [10.1109/ACCESS.2020.2982184](https://doi.org/10.1109/ACCESS.2020.2982184).
- Lockner, D. A., J. D. Byerlee, V. Kukshenko, A. Ponomarev, and A. Sidorin (1992). Observations of quasistatic fault growth from acoustic emissions, *Int. Geophys.* **51**, 3–31.
- Nakamura, Y., C. Veach, and B. McCauley (1972). Amplitude distribution of acoustic emission signals, in *Acoustic Emission*, R. Liptai, D. Harris, and C. Tatro (Editors), ASTM International, West Conshohocken, Pennsylvania, 164–186, doi: [10.1520/STP35388S](https://doi.org/10.1520/STP35388S).
- Peddinti, V., G. Chen, V. Manohar, T. Ko, D. Povey, and S. Khudanpur (2015). Jhu aspire system: Robust lvcsv with tdnns, ivector adaptation and rnn-lms, *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, IEEE, 539–546.
- Pomponi, E., and A. Vinogradov (2013). A real-time approach to acoustic emission clustering, *Mech. Syst. Signal Process.* **40**, no. 2, 791–804.
- Ramírez-Rojas, A., E. Flores-Márquez, N. Sarlis, and P. Varotsos (2018). The complexity measures associated with the fluctuations of the entropy in natural time before the deadly México M8. 2 earthquake on 7 September 2017, *Entropy* **20**, no. 6, 477.
- Read, M. D., M. R. Ayling, P. G. Meredith, and S. A. F. Murrell (1995). Microcracking during triaxial deformation of porous rocks monitored by changes in rock physical properties, II. Pore volumetry and acoustic emission measurements on water-saturated rocks, *Tectonophysics* **245**, no. 3, 223–235, doi: [10.1016/0040-1951\(94\)00236-3](https://doi.org/10.1016/0040-1951(94)00236-3).
- Ródenas, J., M. García, R. Alcaraz, and J. Rieta (2015). Wavelet entropy automatically detects episodes of atrial fibrillation from single-lead electrocardiograms, *Entropy* **17**, no. 9, 6179–6199.
- Scholey, J. J., P. D. Wilcox, M. R. Wisnom, and M. I. Friswell (2010). Quantitative experimental measurements of matrix cracking and delamination using acoustic emission, *Compos. Part A: Appl. Sci. Manuf.* **41**, no. 5, 612–623, doi: [10.1016/j.compositesa.2010.01.008](https://doi.org/10.1016/j.compositesa.2010.01.008).
- Sharma, R., R. Pachori, and U. Acharya (2015). Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals, *Entropy* **17**, no. 2, 669–691.
- Snyder, D., D. Garcia-Romero, and D. Povey (2015). Time delay deep neural network-based universal background models for speaker recognition, *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, IEEE, 92–97.
- Stanchits, S., S. Vinciguerra, and G. Dresen (2006). Ultrasonic velocities, acoustic emission characteristics and crack damage of Basalt and Granite, *Pure Appl. Geophys.* **163**, no. 5, 975–994, doi: [10.1007/s00024-006-0059-5](https://doi.org/10.1007/s00024-006-0059-5).
- Swapna, F., K. P. Soman, and R. Vinayakumar (2018). Automated detection of cardiac arrhythmia using deep learning techniques, *Procedia Comput. Sci.* **132**, 1192–1201, doi: [10.1016/j.procs.2018.05.034](https://doi.org/10.1016/j.procs.2018.05.034).
- Tobias, A. (1976). Acoustic-emission source location in two dimensions by an array of three sensors, *Non-Destructive Test.* **9**, no. 1, 9–12, doi: [10.1016/0029-1021\(76\)90027-X](https://doi.org/10.1016/0029-1021(76)90027-X).
- Unakafova, V., and K. Keller (2013). Efficiently measuring complexity on the basis of real-world data, *Entropy* **15**, no. 10, 4392–4415.
- Waibel, A., T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang (1995). Phoneme recognition using time-delay neural networks, in *Backpropagation: Theory, Architectures and Applications*, 35–61.
- Yao, Z., Z. Wang, W. Liu, Y. Liu, and J. Pan (2020). Speech emotion recognition using fusion of three multi-task learning-based classifiers: HSF-DNN, MS-CNN and LLD-RNN. *Speech Comm.* **120**, 11–19, doi: [10.1016/j.specom.2020.03.005](https://doi.org/10.1016/j.specom.2020.03.005).
- Zhang, H. L., G. M. Zhu, and Y. H. Wang (2013). Automatic micro-seismic event detection and picking method, *Geophys. Geochem. Explor.* **37**, no. 2, 269–273.

Zhu, W., and T. Wong (1997). The transition from brittle faulting to cataclastic flow: Permeability evolution, *J. Geophys. Res.* **102**, no. B2, 3027–3041.

Zoukaneri, I., and M. J. Porsani (2015). A combined Wigner-Ville and maximum entropy method for high-resolution time-frequency analysis of seismic data, *Geophysics* **80**, no. 6, O1–O11.

Appendix A

Rock deformation experiment

Darley dale sandstone (DDS) is ubiquitous in rock mechanics testing. It is a brown–yellow, feldspathic sandstone with a modal composition of quartz (69%), feldspars (26%), clay (3%), and mica (2%) (Heap *et al.*, 2009). Previous studies report a connected porosity of $13.3\% \pm 0.8\%$ with grain sizes varying from 100 to 800 μm (Zhu and Wong, 1997). Pore sizes measured by a mercury porosimeter have a mean diameter of 10 μm but with significant microporosity (Read *et al.*, 1995). The unconfined compressive strength is 160 MPa (Baud and Meredith, 1997). At the scale analyzed here, no distinct layering or laminations were present. A cylindrical rock sample was cored using a diamond tipped hollow coring drill to prepare a 4 cm diameter sample that was then trimmed to 10 cm length with a diamond saw. End faces are accurately ground using a lathe fitted with a crosscutting diamond grinding disk with surfaces flat and parallel to within 0.01 mm.

Deformation was performed using a conventional triaxial deformation cell installed at the Rock Mechanics Laboratory, University of Portsmouth (Fazio, 2017). The sample presented here was deformed until brittle failure at a confining pressure of 20 MPa at a constant deformation rate of 3.6 mm/hr. Experimentation was performed under fully drained conditions to avoid any fluid-driven effects on acoustic emission (AE) frequency content (Benson *et al.*, 2010). These environmental conditions ensure that a high number of AEs is obtained and any time-dependent variations in the signal waveform are predominantly due to the scattering effects of microfractures, thus allowing for the sampling of a diverse range of deformation structure. Axial displacement is measured with a noncontact Eddy displacement system mounted to the apparatus. It comprises three sensors that accurately (submicron) measure the distance to a target steel plate attached to the driving piston. These readings are averaged and are used to set the target deformation rate via feedback to an axial stress intensifier.

For AE data acquisition, the protocol of Benson *et al.* (2007) was followed. The dry sample was positioned inside an engineered rubber jacket fitted with ports in which an array of twelve 1 MHz single-component piezoelectric transducers (PZTs, model PAC Nano30) was embedded (Fig. A1). These sensors have a relatively flat frequency response between 125 and 750 KHz. Sensor output is connected to preamplifiers set to 40 dB, focusing on data quality over quantity. An ITASCA-Image “Milne” recorder operates in a standard trigger model, downloading all 12 channels when any single channel passes a

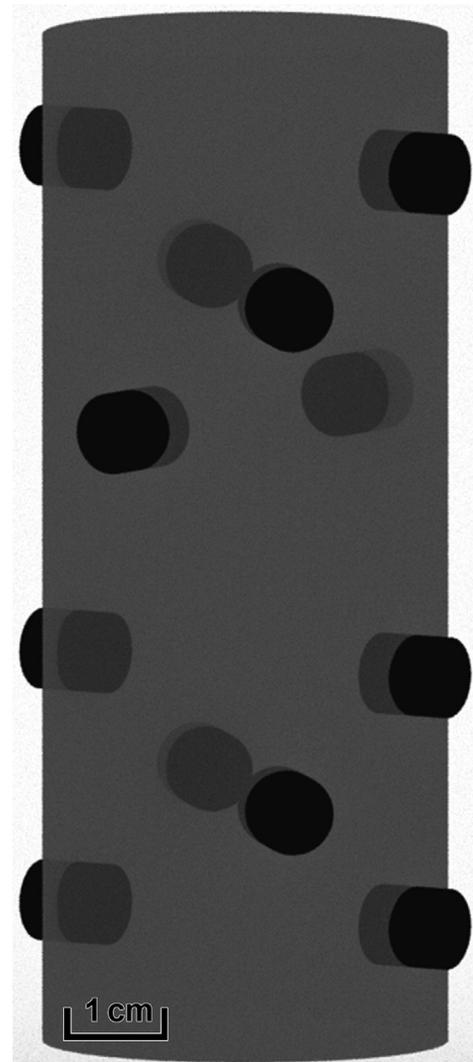


Figure A1. Sensor array.

set 100 mV threshold (e.g., Gehne, 2018). During experimentation, in addition to passive mode (recording of fracturing events), the sensors are also used in active mode for *P*-wave elastic velocity measurements to derive a velocity model for source locations (Fig. A2). Periodically, each PZT was triggered in sequence with a high voltage (200 V) pulse, with the energy recorded by the remaining 11 sensors in the array. SNR was further improved by pulsing each sensor multiple times (16) and stacking the received waveforms, with each survey of 12-sensors generating 144 nonunique ray paths and taking approximately 30 s to complete.

Mechanical data from the deformation testing reveal a typical response of DDS. A total of 24360 AEs was detected by the trigger recording system (Fig. A2, histogram 0.01% strain bins); however, it is unknown what percentage of these signals are noise related. Differential stress increases gradually (Fig. A2, black line), entering the elastic phase of deformation

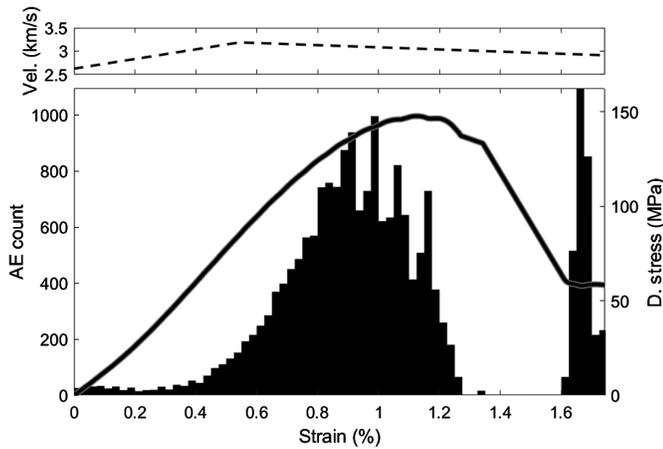


Figure A2. Velocity data, stress–strain curve, and raw AE count for Darley Dale sandstone deformed at 20 MPa confining pressure. Brittle failure occurs at approximately 1%, 2% strain.

at approximately 0.3% strain, coinciding with a rapid increase in the number of AEs recorded. During the period of strain-softening from 0.7% to 0.9% strain, the rate of AEs dramatically accelerates and peaks at ~ 1000 events/0.01% strain. As the sample goes through peak stress (145 MPa), several pulses in AE count correspond with minor drops in differential stress. However, the overall rate decreases with time. At $\sim 1.4\%$ strain the sample fails, with Figure 4a highlighting the formation of a fracture plane marked in red.

Appendix B

Model training parameters

Seismic envelope. The RMS envelope provides a scaled amplitude estimate of the AE trace. The envelope represents the instantaneous energy of the signal and is computed in a sliding window of 10 samples. Such a narrow window is selected to avoid smoothing of the low amplitudes at signal onset. In low-noise environments, the amplitude of the signal greatly exceeds that of the presignal noise and provides a good marker of the onset of energy. However, as noise increases, this onset becomes increasingly difficult to identify, resulting in a high degree of uncertainty.

Instantaneous frequency content. Intrinsic modal functions represent simple oscillatory modes of complex signals. Unlike harmonic signals, intrinsic mode function (IMF) can have variable frequency and amplitude content through time. Known as empirical mode decomposition (EMD), IMFs of individual AEs are obtained through a continuous screening process called “sifting” and must satisfy the following requirements: The number of extrema and the number of zero-crossings must either be equal or differ by a minimum of one, while the mean

of the upper and lower envelopes is zero. The procedure follows (Huang *et al.*, 1998):

1. The local extrema in the time series $X(t)$ are identified.
2. Local maxima are connected by a cubic spline line to produce an upper envelope.
3. This is repeated for the local minima to produce a lower envelope.

The difference between $X(t)$ and the mean of the two envelopes, m_1 , is the first component h_1 :

$$h_1 = X(t) - m_1.$$

Using h_1 as the new time series, this process is repeated k times until the standard deviation of h_{1k} , σ_k , is less than 0.8. This value was selected as it allows for a relatively small number of IMF components while still capturing signal complexity

$$\sigma_k = \sum_{t=0}^T \frac{|h_{k-1}(t) - h_k(t)|^2}{h_{k-1}^2(t)}.$$

Thus, h_{1k} is defined as the first IMF component of the data

$$c_1 = h_{1k}.$$

Typically, c_1 contains the highest frequency component of $X(t)$. Subsequent IMFs, which contain longer period data, are then calculated from the residual r_n in which

$$r_n = r_{n-1} - c_n.$$

The sifting process is stopped when r_n becomes a monotonic function from which no new IMF can be extracted. Consequently, the test data are decomposed into n empirical modes in which

$$X(t) = \sum_{j=1}^n c_j + r_n.$$

Once obtained, the Hilbert transform H , or instantaneous amplitude A_{inst} , of each IMF component is calculated as follows:

$$H(c_n)(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{c_n(\tau)}{t - \tau} d\tau,$$

in which τ is the width of the analysis window. The phase angle θ ,

$$\theta = \text{Im}(\log(H(c_n))),$$

is then used to compute the instantaneous frequency content f_{inst} as

$$f_{\text{inst}} = \frac{\Delta(\theta)/Ts}{2\pi},$$

in which Ts is the sampling period. To characterise the data for use in classification, each step of the vector $\text{freq}(t)$ is then defined as the $f_{\text{inst}}(t)$ of the highest $A_{\text{inst}}(t)$ of all IMF components in $X(t)$. Thus, freq is simply a high-resolution vector of the dominant frequency content through time.

Permutation entropy. The permutation entropy (PE) method describes the uncertainty and the degree of irregularity in a random series. To compute the empirical PE (refer to [Unakafova and Keller, 2013](#) for more details), ordinal patterns of the AEs are obtained with a delay of $\tau = 1$, indicating a distance of 1 between points in patterns and an order of

$d = 5$, meaning patterns contain six points ($d + 1$). With a high degree of overlap between points, it is possible to use all of the information about order relations between points in the AE waveform. The distribution of ordinal patterns is obtained by the simple enumeration of the type of pattern. The PE, computed in a sliding time window of 30 points to reflect the complexity in the time series, is calculated at time t as follows:

$$\text{PE}_d^\tau(t) = - \sum_{j=0}^{(d+1)!-1} \frac{q_j}{M} \ln \frac{q_j}{M} = \ln M - \frac{1}{M} \sum_{j=0}^{(d+1)!-1} q_j \ln q_j,$$

in which $q_j = \#\{k \in \{t, t-1, \dots, t-M+1\} \text{ has the ordinal pattern } j\}$ (with $0 \ln 0 := 0$) and M is defined as the number of ordinal patterns in the window.

Manuscript received 14 May 2020
Published online 30 December 2020