Geophysical Journal International

Geophys. J. Int. (2025) **242,** 1–13
Advance Access publication 2025 June 20
Research Paper

https://doi.org/10.1093/gji/ggaf155

Forecasting induced seismicity in enhanced geothermal systems using machine learning: challenges and opportunities

Sadegh Karimpouli [1], Grzegorz Kwiatek [1,2] Patricia Martínez-Garzón [1,3] Danu Caus [4,5,6] Lei Wang [1] Georg Dresen[1] and Marco Bohnhoff [1,7]

[1] *GFZ Helmholtz Centre for Geosciences,* 14473 *Potsdam, Germany. E-mail: sadegh.karimpouli@gfz.de*
[2] *GMuG Gesellschaft für Materialprüfung und Geophysik mbH, Dieselstraße* 9, *D*-61231 *Bad Nauheim, Germany*
[3] *RWTH University of Aachen,* 52062 *Achen, Germany.*
[4] *DKRZ German Climate Computing Centre,* 20146 *Hamburg, Germany*
[5] *Helmholtz Centre Hereon,* 21502 *Geesthacht, Germany*
[6] *Helmholtz AI,* 85764 *Neuherberg, Germany*
[7] *Department of Earth Sciences, Free University Berlin,* 14195 *Berlin, Germany*

Accepted 2025 April 27. Received 2025 March 27; in original form 2025 January 6

SUMMARY

Induced seismicity poses a significant challenge to the safe and sustainable development of Enhanced Geothermal Systems (EGS). This study explores the application of machine learning (ML) for forecasting cumulative seismic moment (CSM) of induced seismic events to evaluate reservoir stability in response to fluid injections. Using data from the Cooper Basin (Australia), the St1 Helsinki geothermal project (Finland), and a controlled laboratory injection experiment, we evaluate ML models that integrate catalogue and operational features with various frameworks. Results indicate that feature-rich models outperform simpler ones in complex seismic environments like the Cooper Basin and laboratory cases, where seismicity is promoted by earthquake interaction and fault reactivation. However, in scenarios like St1 Helsinki, with minimal event clustering, additional features offer limited predictive benefits. While ML models are promising, several challenges impede reliable forecasting, including data scarcity from operational wells, the extrapolation demands of cumulative output (i.e. CSM) and the difficulty of predicting abrupt CSM increases for large seismic events. Enhancing model robustness requires synthetic data augmentation and improved feature selection capable of capturing diverse reservoir dynamics. These advancements may enable more accurate near real-time forecasts of problematic induced seismic events, informing operational decisions to mitigate seismic risks while maximizing energy extraction, and hence offering a pathway for broader adoption of ML in renewable energy development and management.

Key words: Machine learning; Earthquake interaction, forecasting, and prediction; Induced seismicity.

1 INTRODUCTION

Safe usage of renewable energy resources is crucial for mitigating greenhouse gas emissions and thus global warming. Deep geothermal energy production plays an important role in this aspect and provides a sustainable and reliable base-load heat and power supply. Complementary to using heat from shallow depth ranges, 'Enhanced Geothermal Systems' (EGS) aim at extracting heat from deeper reservoirs, typically at around 4 km depth in crystalline basement or sedimentary basins, at temperatures reaching or exceeding 100 °C. Developing EGS involves creating fluid pathways in low-permeability rock formations through hydraulic stimulations, which are typically accompanied by mostly small-scale induced seismicity (Majer *et al.* 2012; Ellsworth 2013). Induced seismic events large enough to be felt can reduce public acceptance and may in some cases lead to structural damage to buildings, exceptionally even resulting in personal injuries. For example, in the Basel Deep Heat Mining project in Switzerland, deep geothermal field development triggered over 10 000 earthquakes, some of which reached magnitudes up to M 3.4 (Deichmann & Giardini 2009; Giardini 2009). These larger events were widely felt by the population, resulted in 9 million USD insurance claims and ultimately caused abandonment of the project. More recently, the occurrence of a M_w 5.5 earthquake in 2017 near Pohang, South Korea, left 82 people injured and damaged approximately 2000 houses. This event has been linked to a nearby deep geothermal project under development (Park *et al.* 2020).

Mitigating induced seismicity in EGS projects is essential for ensuring a safe and efficient geothermal energy extraction (Mukuhira *et al.* 2023; D. Li *et al.* 2024). For example, Kwiatek *et al.* (2019)

successfully applied a near real-time 'adaptive stimulation concept' during the currently deepest geothermal stimulation in urban Helsinki, Finland (St1 Deep Heat project). Based on earlier findings that the seismic energy release was stable and proportional to hydraulic energy input, they mitigated the injection scheme. This allowed keeping the maximum magnitude below the threshold (M2) given by local authorities. It was found that the maximum magnitude evolution in time follows predictions from theoretical models (McGarr 2014; van der Elst *et al.* 2016; Galis *et al.* 2017). For example, if a conservative model of McGarr (2014) is used, the expected maximum seismic energy release is bounded by the total injected fluid volume. In other words, the maximum earthquake size (seismic deformation) is constrained by the total elastic strain energy stored due to fluid injection. Galis *et al.* (2017) developed a slightly different fracture mechanics-based scaling relation between the largest magnitude of self-arrested earthquakes and the injected fluid volume. To this end, they used the size of ruptures induced by localized pore-pressure perturbations and propagating on pre-stressed faults.

Applying these models, Kwiatek *et al.* (2019, 2022) and Bentz *et al.* (2020) suggested that deviation from seismic moment release that exceeds predictions from above cited models indicates an unstable seismic evolution and should trigger warning and reassessment of injection operations. Primarily this requires a high-resolution seismic network continuously monitoring the balance between seismic and hydraulic energy, whereas the forecasting of large events itself is based on the premise that the deviation from a stable, pressure-controlled system can be detected in ample time. Kwiatek *et al.* (2019) pointed out that monitoring a single seismic parameter and hydraulic data alone may not be enough to handle the potentially complex processes in the reservoir. Martínez-Garzón *et al.* (2020) summarized observational factors from reservoir stimulations that would point at a high potential for run-away earthquakes to occur. These factors include a notable drop in *b*-value, deformation localization along main seismogenic faults, high and/or increasing seismic injection efficiency (the ratio of radiated to hydraulic energy), dominance of shear faulting, abundant earthquake triggering and earthquake interaction. However, not all of these factors are necessarily observed in every case. For instance, a drop in the *b*-value is evident in Basel (Ritz *et al.* 2022) and the Geysers EGS (Martínez-Garzón *et al.* 2014), but it remains unclear in Pohang (Woo *et al.* 2019). Martínez-Garzón *et al.* (2020) concluded that any warning system should incorporate further parameters that may potentially indicate undesirable system changes, among them *b*-value, clustering and triggering properties or focal mechanism similarity.

Accurate forecasting of large induced earthquakes is vital for a warning system in an EGS project. However, some large earthquakes—runaway events—do not conform to the theoretical models of maximum expected magnitude and stable, pressure-controlled seismic moment release (Ripperger *et al.* 2007; Galis *et al.* 2017; Norbeck & Horne 2018; Kwiatek *et al.* 2019, 2022). Monitoring the stability of the geothermal system in near real-time and using solely the seismic energy release seems to be too limited to capture operational and subsurface complexities (Kwiatek *et al.* 2022, 2024). This calls for incorporating additional parameters that characterize the potential changes in the structural inventory of the reservoir or changes in stress conditions, as well as more advanced forecasting models (Kwiatek *et al.* 2022). Machine Learning (ML) algorithms are now widely used as powerful data-driven tools for forecasting problematic events in seismology (Rouet-Leduc *et al.* 2017; Karimpouli *et al.* 2023; Saad *et al.* 2023; Karimpouli, Kwiatek, Ben-Zion *et al.* 2024). In the EGS applications, ML is applied as aid in routine processing of seismic data (Leong & Zhu 2024; Okamoto *et al.* 2024). However, ML can potentially also be used in forecasting problems, mostly on a regional scale (Holtzman *et al.* 2018; Limbeck *et al.* 2021; Mehrabifard & Eberhardt 2021; Qin *et al.* 2022; Hui *et al.* 2023). Hincks *et al.* (2018) used seismicity data from Oklahoma from 2009 to 2016 to predict annual moment release using a developed Bayesian Network. They found that the injection depth of fluids relative to the crystalline basement most strongly correlates with seismic moment release (the largest induced events typically occur in the basement). Wozniakowska & Eaton (2020) estimated the seismogenic activation potential of each well using a Logistic Regression ML model on a data set from 6466 multistage horizontal hydraulic fracturing wells drilled into the Montney Formation over a large region in western Canada. They found that injection depth and well distance to the Cordilleran thrust belt are the most important parameters.

This study uses ML to forecast cumulative seismic moments from a future volume increment of injected fluid known *a priori*. This allows us to compare such a prediction with the maximum magnitude models cited above to decide if the seismic moment evolution remains stable. To this end, we use the seismicity catalogue and fluid injection data from one laboratory experiment and two case studies of stimulated reservoirs that differ with respect to the geological and structural environment (Cooper Basin, Australia and St1 Helsinki, Finland). Using two ML models and applying catalogue/operational features, we will discuss challenges and opportunities of applying ML for induced seismicity monitoring in EGS.

2 METHODOLOGY

Suppose an injection well operates within an ongoing EGS. For operational decision-makers, an effective tool would be one that provides reliable forecasts of upcoming earthquakes potentially induced by subsequent fluid injections and/or extractions. However, the primary challenge lies in the lack and resolution of sufficient data from such operating wells to train an ML-based earthquake predictor (Saad *et al.* 2023; Karimpouli, Kwiatek, Ben-Zion *et al.* 2024). A common approach is to use data from other cases. However, Kwiatek *et al.* (2024) demonstrated that even two wells within the same EGS, located as close as 500 m apart (e.g. St1 Helsinki), exhibit significantly different seismic responses. This suggests that, even data from a nearby well within the same EGS, may not be suitable for training a predictor for the current well (let alone another system). As a result, the problem must be redefined to ensure that the predictor can be trained on limited data while still providing an effective decision-making tool for operational use.

Bentz *et al.* (2020) analysed ten geothermal injection operations and found that, during a stable injection phase of an EGS, cumulative seismic moment (CSM) scales with cumulative fluid volume injected (CVI) (see Fig. 2 therein). Based on trends in the magnitude evolution with respect to injected fluid volume, they identified two types of geothermal reservoirs: those that result in stable, self-arrested earthquake ruptures during fluid injection, and thus follow the expected trends in seismic moment evolution in time, and those that are potentially unstable, and do not conform with the existing models. In logarithmic scale, theoretical seismic energy released in response to a total volume of fluid injection show slopes ranging from 1 (Hallo *et al.* 2014; McGarr 2014; Z. Li *et al.* 2022) to 1.5 (van der Elst *et al.* 2016; Galis *et al.* 2017) (Fig. 1). Accordingly, an

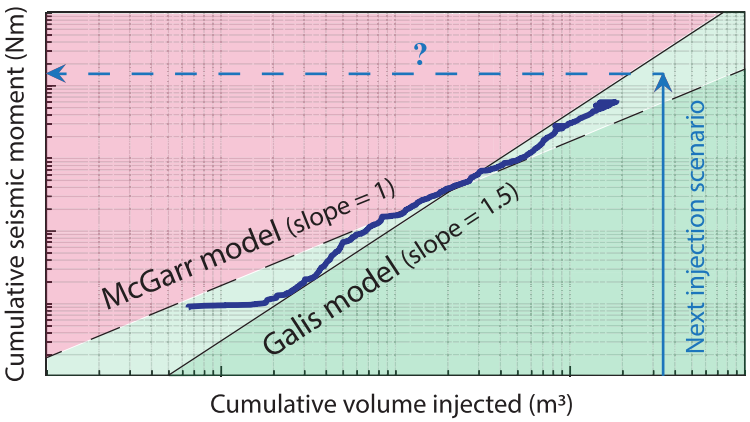

Figure 1. Schematic illustration of cumulative seismic moment (CSM) plot to cumulative volume injection (CVI) in logarithmic scale showing trend of the moment evolution (dark-blue curve). The slopes of theoretical (McGarr 2014 and Galis *et al.* 2017) models vary between 1 and 1.5 (dashed and solid line) defining 'stable' and 'unstable' seismic energy release, shown as green and red areas, respectively. The offset of these models depends on reservoir and fluid properties. We aim at predicting CSM (dashed arrow) based on the next injection scenario (solid arrow), allowing us to evaluate if the reservoir will remain stable or become unstable in the next phase of injection.

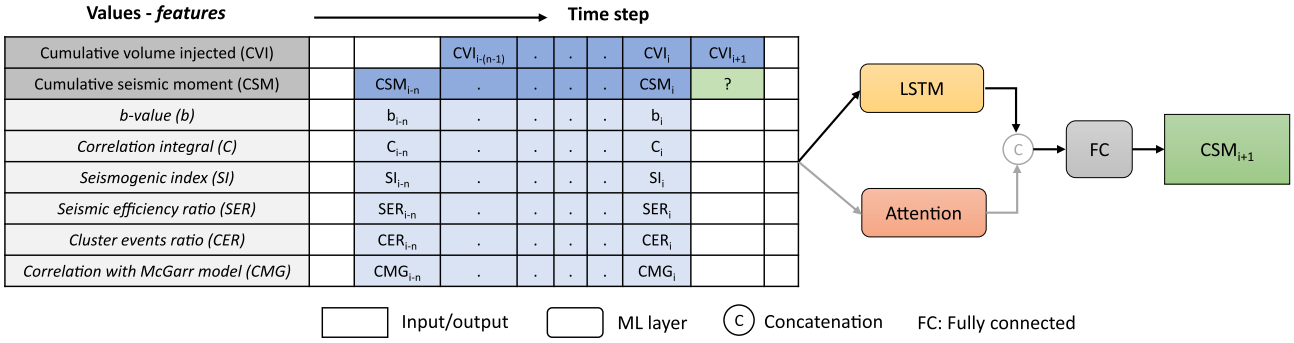

Figure 2. Inputs, output and ML layers used to construct three individual models in this study (see the text).

expected CSM for the next CVI may be located below these models or between them (green area in Fig. 1). We define a reservoir to be in a stable phase if its seismic moment evolution remains below or within limits given by the theoretical models having a slope of 1–1.5 (green area in Fig. 1). If this trend exceeds theoretical predictions of self-arrested events (i.e. moves from green to red area in Fig. 1), we define the reservoir to be in an unstable phase, with an enhanced potential for runaway earthquakes. Therefore, we define our problem as *forecasting the cumulative seismic moment for the next step in the injection scenario*.

2.1 Input/output data and features

Consequently, the input data are injection volume and the corresponding induced seismicity catalogue. For injection data, the volume of fluid, which will be injected in the next step, is also known. For catalogue data, we use cumulative seismic moment (CSM) as well as other catalogue/operational features such as:

Gutenberg–Richter (GR) b-value (b): The GR frequency–magnitude distribution describes the relationship between the earthquake magnitude (M) and number of earthquakes (N) below this magnitude:

$$\log_{10}(N_M) = a - bM, \tag{1}$$

where a and b are constant values (known as a- and b-value). A decreasing trend of b-value has been proposed to reflect accumulation of stress and damage and the dominance of larger fault segments preparing to fail (Amitrano 2003; Scholz 2015).

Correlation integral (C): The correlation integral is a measure of closeness and localization of events, providing insights into the stress accumulation and release processes, describing the change in event number (N) with separation distance (r) as:

$$c = \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=0}^{n} \sum_{j=0}^{n} H\left(r - |x_i - x_j|\right), \tag{2}$$

where n indicates the amount of data in the analysis window, x the hypocentre coordinates, and H the Heaviside step function:

$$H = \begin{cases} 1 & x>0 \\ 0 & x\leq 0 \end{cases} \tag{3}$$

Seismogenic index (SI or Σ): Assuming a Gutenberg–Richter distribution, Seismogenic Index is defined as (Shapiro *et al.* 2010):

$$\Sigma = \log_{10} N - \log_{10} Q + bM, \tag{4}$$

where N is the cumulative number of induced seismic events larger than M and Q is the cumulative injected fluid volume. The b parameter is the aforementioned GR b-value. SI determines the expected level of seismic activity in each EGS and is independent of the injection time.

Seismic efficiency ratio (SER): depends on the characterization of the fractured zone and the injected or removed fluid volume (Hallo *et al.* 2012, 2014). It accounts for the partitioning of the elastic strain energy release ($\sum M_o$) into seismic processes with

respect to the theoretical total seismic moment ($\sum M_{\mathrm{TOT}}$) such as McGarr model (McGarr 2014):

$$\mathrm{SER} = \Sigma M_o / \Sigma M_{\mathrm{TOT}} \tag{5}$$

Clustered events ratio (CER): Following Zaliapin *et al.* (2008), we use the nearest-neighbour approach to classify events as background or clustered. The distance between i and j events is defined as (Baiesi & Paczuski 2005):

$$\eta_{ij} = \begin{cases} t_{ij}\left(r_{ij}\right)^{d_f} 10^{-bM_i} & t_{ij} > 0 \\ \infty & t_{ij} \leq 0 \end{cases}, \tag{6}$$

where t_{ij} and r_{ij} are the time and space distance of two events, d_f is the fractal dimension of event epicentres, b is the GR b-value, and M is the magnitude. Based on a threshold distance η_0 background and clustered events are detected (Zaliapin & Ben-Zion 2013). CER is defined as:

$$\mathrm{CER} = N_{\mathrm{CE}} / \left(N_{\mathrm{CE}} + N_{\mathrm{BE}}\right), \tag{7}$$

where N_{CE} and N_{BE} are the number of background and clustered events. It reveals seismicity deviation from background seismicity (following a Poissonian distribution) to clustered seismicity, signifying a complex interplay of stress accumulation and fault instability (Martínez-Garzón *et al.* 2018; Picozzi *et al.* 2023).

Correlation with McGarr model (CMG): quantifies the deviation from seismic moment evolution trend with respect to the theoretical model of McGarr. McGarr (2014) introduced a simple model, which forecasts the next largest magnitude as:

$$\Sigma M_0 = \frac{2\mu\left(3\lambda + 2G\right)}{3}\Delta V, \tag{8}$$

where λ and G are Lame's elastic parameters of the geothermal reservoir, G is the modulus of rigidity, μ is the friction coefficient and ΔV is the cumulative fluid injection. When $\lambda = G$ and $\mu = 0.6$, the model simplifies to:

$$\Sigma M_0 = 2G\Delta V. \tag{9}$$

The CMG computes the correlation between the observed CSM and the ones obtained from the McGarr model. A high correlation means the trend follows a stable state of the reservoir.

2.2 Machine learning models

Fig. 2 illustrates the input data and ML models that we use in this study. Input data are selected as time-series with a length of n time steps up to the current time step i. The output is the cumulative seismic moment (CSM) in the next time step $i + 1$ for a known fluid injection volume (i.e. CVI_{i+1}), which is the independent external feedback parameter. The ML models generally consist of an LSTM (Long Short-Term Memory) network, paired with an Attention mechanism and followed by a fully connected (FC) layer.

The LSTM (Hochreiter & Schmidhuber 1997) is a type of neural network stemming from Recurrent Neural Networks (RNN) (Elman 1990) designed to capture features of sequence data. These features can be helpful for various downstream tasks, such as prediction and forecasting of future variables. Unlike typical multilayer perceptrons (MLP, Rumelhart *et al.* 1986), LSTMs have a context memory layer that is helpful in remembering historical variability and correlations within the input stream. LSTMs appeared as an improvement to the original RNNs, from the need to solve the problem of vanishing and exploding gradients, which came as a side-effect of having the memory context layer. These problems were to some extent solved by using various the so-called 'gates' within the original RNN architecture, resulting in the subsequent LSTM architecture being more stable to train, among other things. Gates are essentially mathematical operators designed to improve the memory range of the LSTM, but in practice the LSTM often performs well at relatively short-term learning, struggling to capture very long-term dependencies. To mitigate this issue, the ML community proposed the concept of Attention (Bahdanau 2014). Attention mechanisms were significantly popularized within the transformer architecture in a natural language processing context (Vaswani *et al.* 2017). The attention mechanism has the ability to expand the memory range of a prediction task in practice, by placing emphasis on certain input samples that seem more important for predicting, even though these input samples might be encountered further away in the past. This is a form of empirical 'selective' memory, based on weighing and prioritizing long-term inputs, as opposed to the more short-term view of the LSTM. So, from one perspective, by combining the hidden memory embedding vector of the LSTM (providing a more 'tactical' view on the recent data) with the Attention layer (adding a more 'strategic' view on the long-term data), we hope to get the best of both approaches, resulting in improved overall predictions. From another perspective however, fusing the LSTM and Attention via a layer with learnable weights FC has the potential to balance other strong points and shortcomings of these methodologies in certain scenarios, such as the ability of the LSTM to capture smoother/cumulative average trends underlying the data, and the ability of the Attention layer to be more reactive to the immediate, sharp, stepwise behaviour of the input. To sum up, LSTM and Attention can excel at different aspects while analysing the input data, hence we fuse them via a learnable layer FC to take advantage of their complementary nature across different input scenarios.

In this study, we use three different models as follows:

(a). **First model (LSTM):** Inspired by theoretical models (e.g. McGarr 2014), we start with only two time-series of cumulative fluid volume injected (CVI) and cumulative seismic moment (CSM) as input data and apply a many-to-one LSTM model to predict CSM in the next step (Fig. 2). The LSTM layer contains 64 units and we impose a dropout rate of 10 per cent on it. Following is a dense layer with 1 output channel to predict the variable of interest. We use the '*Adam*' optimizer (Kingma & Ba 2014), Mean Absolute Error (*MAE*) as the loss function and a batch size of 128 to train the model for 200 epochs. We use a linear activation function throughout our models.

(b). **Second model (LSTM+Att+FC)**: Due to the stepwise behaviour of the input data, we concatenate an Attention layer with the LSTM layer (Fig. 2) as a memory embedding. This way, we use LSTM with the main intention to empirically capture short-term history, although theoretically it can also capture longer term, cumulative, smooth trends. Additionally, we use a flattened attention vector to capture and emphasize important features from the longer term history, coping in parallel with the variable, stepwise nature of our data. The fully connected layer/FC acts as a fusion layer, intended to learn from both input streams: LSTM and Attention. The inputs, outputs and parameters are similar to the first model (i.e. LSTM); however, for a more expressive fusion function, we use two FC layers in the last part, having eight neurons each, and a single projection channel for the output.

c. **Third model (LSTM+Att+FC+Feat)**: Laboratory experiments have revealed that catalogue features contain valuable information about the stress evolution (Karimpouli, Kwiatek, Martínez-Garzón *et al.* 2024) and could be used for large labquake forecasting (Karimpouli, Kwiatek, Ben-Zion *et al.* 2024). Therefore, in

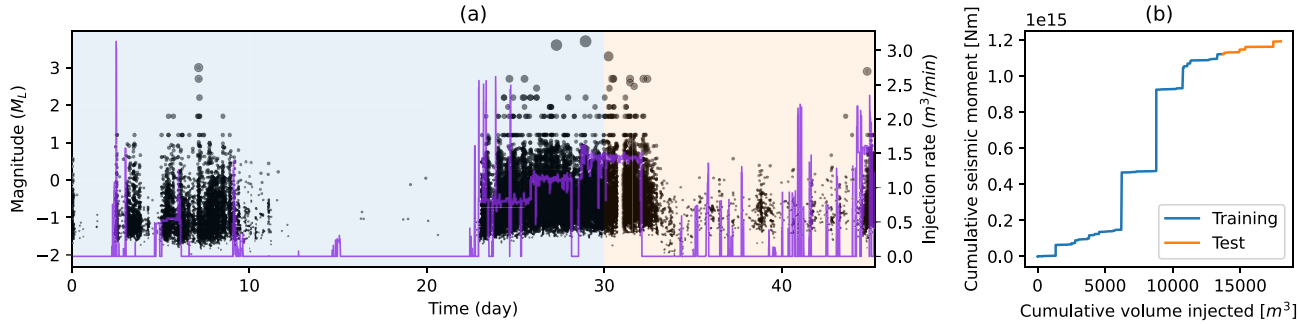

Figure 3. Cooper basin data set used in this study. (a) Magnitude of induced events (black dots) and injection rate (purple curve) along the time measured from the start of the injection. (b) Cumulative seismic moment versus cumulative volume injected. The first 65 per cent of data (blue area in (a) and blue curve in (b)) is used as the training data, and the following 35 per cent (orange area in (a) and orange curve in (b)) is used as both validation and test data.

this model, we increase the inputs by adding catalogue/operational features (Fig. 2). These features reflect the background seismicity (e.g. SI) and response of the subsurface structure (e.g. b, SER and CMG) and may prepare some signs of the preparatory phase before large earthquakes (e.g. C and CER). The output model architecture and model parameters are the same as in the previous model (i.e. LSTM+Att+FC).

3 CASE STUDIES

3.1 Cooper Basin, Australia

The Cooper Basin, located in northeast of South Australia, is a geothermal area known for its high-temperature hot rock geothermal resources (Baisch *et al.* 2006). This basin contains thick layers of Precambrian granite, which host significant heat resources, making it ideal for EGS. As one of the primary geothermal developments in the Cooper Basin, the well Habanero#1, drilled in 2003, encountered high temperatures around 250 °C at depths exceeding 4200 m. During stimulation of well Habanero#1, over 20 000 m^3 of water was injected to create or enhance fractures in the granite, leading to > 23 000 seismic events with $-2.0 < M_L < 3.7$ and completeness magnitude of -0.7 (see Data availability). The seismicity revealed a nearly planar structure 2×1.5 km^2 in size being activated, with an apparent thickness of 150–200 m (Baisch *et al.* 2006, 2009). Event magnitudes and injection rates are shown in Fig. 3(a). As indicated in Fig. 3(b), we use 65 per cent (∼ 30 d) of all data as the training set and the remaining data as both evaluation and test data. Bentz *et al.* (2020) noted that the Cooper Basin stimulation contained some episodes of potentially unstable seismic response, as the seismic injection efficiency was continuously increasing during the stimulation campaign.

The cumulative volume injected (CVI) and cumulative seismic moment (CSM) are computed in temporal intervals of every 30 min. This means that, in each step, all data from the start of injection and up to the current step are used for computing the cumulative inputs and to extract the features listed in Section 2. After input/feature computation, sequences of time-series are selected with 12 h length and 30 min steps, leading to time-series with a length of 24. Therefore, depending on the model, input size is either 2×24 (LSTM and LSTM + Att + FC) or 8×24 (LSTM + Att + FC + Feat), where 2 and 8, respectively, represent the number of input features, while 24 is the sequence length. Overall, 2183 sequences are produced, which are subsequently split into 1418 training sets and 765 validation/test subsets. It should be noted that only for the third model (LSTM + Att + FC + Feat), in which eight features are used, we first remove 200 initial sequences and then split them. This is due to the small number of events in the early steps, where feature values fluctuate strongly.

3.2 St1 Helsinki

The St1 Deep Heat geothermal project in Finland is aimed at providing renewable district heating using deep geothermal energy. Located in Espoo, near Helsinki, the project involves drilling two wells down to about 6400 m to access the heat from crystalline rock formations. This heat is then intended to be transferred to a district heating network, supporting sustainable heating solutions for the local area. In 2018, stimulation of the OTN-3 well involved injecting a total of 18 160 m^3 of fresh water into deep crystalline rocks over a period of 49 d, from June to July. The injection process utilized variable well-head pressures between 60 and 90 MPa and flow rates between 400 and 800 $l\,min^{-1}$, allowing for controlled fracturing while minimizing seismic impact. To monitor and manage seismicity, a robust seismic monitoring system was set up, using a 24-station borehole seismometer network around the well. The seismic catalogue contains 55 707 detected events, with $-1.0 < M_W < 1.7$ and a completeness magnitude of -0.5, all originating near the stimulated volume (Kwiatek *et al.* 2019). Fig. 4 shows the data set and data division into training and validation/test sets in this study. Kwiatek *et al.* (2019, 2022) concluded that the 2018 stimulation was an example of a reservoir displaying a stable seismic response.

We use the same parametrization of the models as in the case of Cooper Basin and forecast the seismic moment release in time windows of 30 min. Overall, we obtain 2767 sequences, which are then split into 1936 and 831 sequences for the training and validation/test data, respectively.

3.3 Experimental data

Wang *et al.* (2020a, c) conducted laboratory fluid-injection induced fault reactivation experiments on reservoir sandstone samples at upper crustal stress conditions in the configuration of triaxial compression setup. The cylindrical rock samples (50 mm × 100 mm in diameter and length) were cut by a prefabricated sawcut fracture oriented at 30° with respect to the cylinder axis. To record Acoustic Emission (AE) events (laboratory earthquakes) associated with fault slip, 16 piezoelectric transducers (PZTs) were mounted to the sample surface before experiments. At a constant confining pressure of 35 MPa and pore pressure of 5 MPa, the samples with sawcut

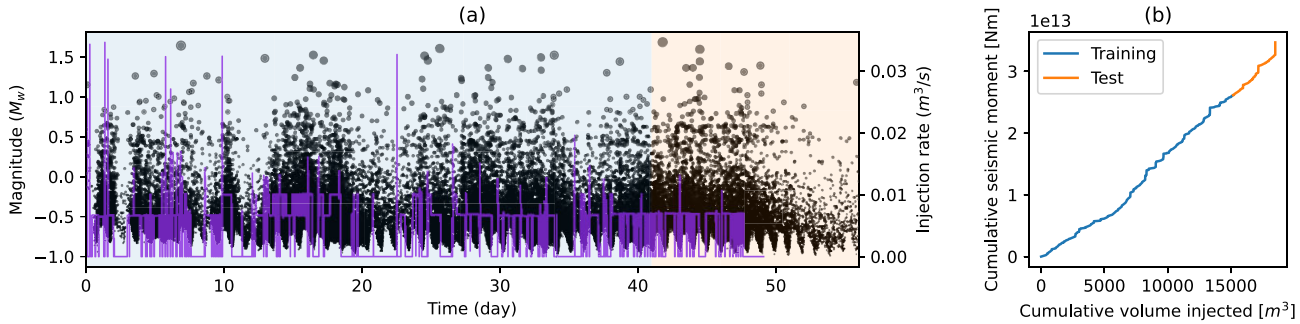

Figure 4. Similar to Fig. 3, but for the St1 Helsinki data set.

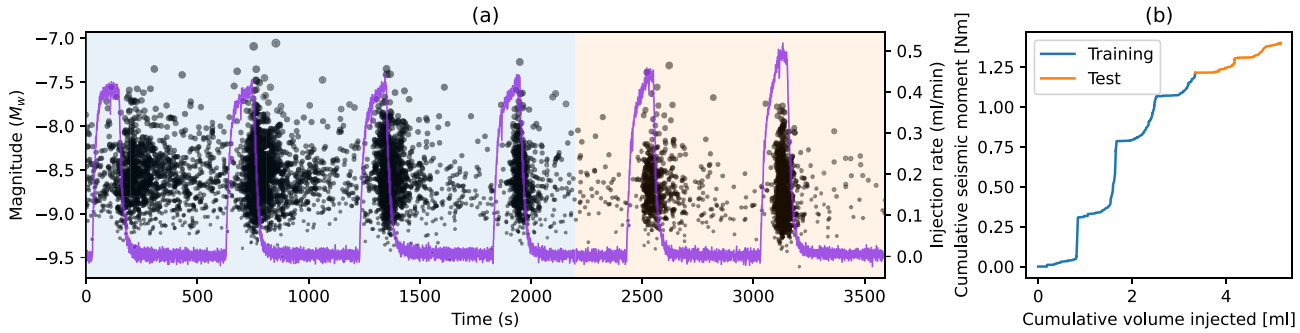

Figure 5. Similar to Fig. 3, but for the experimental data set.

faults were first axially loaded up to the critical stress state of the fault. Then, the position of the axial loading piston was fixed, and distilled water was injected from the bottom end of the sample to induce slip at undrained conditions. Using a syringe pump, fluid pressure was increased stepwise from 5 to 29 MPa with a pressurization rate of 2 MPa min^{-1}. The whole fluid injection protocol was divided into six stages, and in each stage a duration of 10 min was composed of a ramp phase lasting for 2 min (4 MPa fluid pressure increment) and a subsequent plateau phase at constant pore pressure for 8 min.

The six-stage fluid injection caused episodic slow stick-slip events with a peak slip rate of up to 4 $\mu m\,s^{-1}$. Injection-induced fault slip was accompanied by a total of 3 983 AE events with hypocentres distributed across the fault surface (Wang *et al.* 2020b). The moment magnitudes of AE events ranged from M_w −9 to M_w −7, and their temporal evolution during fluid injection is shown in Fig. 5(a). Fig. 5(b) presents the relation between cumulative seismic moments of AEs and cumulative fluid injection volume in the laboratory experiment.

For the laboratory data, note that here time and volume are in seconds and milliliters, respectively. However, all procedures for input features computation and model training are similar. We first compute all features with 1 s steps. Then, all sequences are produced for 24 steps corresponding to 24 s. However, the input array size for the ML models remains similar to the other cases. Finally, 3476 sequences are obtained, which are split into 2259 and 1217 as the training and validation/test data, respectively.

4 RESULTS

Figs 6–8 illustrate the results of forecasting the seismic moment release of the test data employing all three models (Fig. 2), as well as the feature variation and importance for all three cases. For each sequence of the test data, we assume that all data are known based on the catalogue and injection rate and the prediction is generated only for the next time step (i.e. next 30 min).

The obtained results show that although forecasting by LSTM (and LSTM + Att + FC in the case of experimental data) is very close to the observed values in all analysed case studies (Figs 6, 7, 8a, c), none of the ML models could forecast all larger earthquakes (indicated by sharp jumps in observed cumulative seismic moment (CSM) values in Figs 6, 7, 8b, c). Instead, we generally see an underestimation of CSM, and the CSM forecasts are delayed with respect to the actual observed data, which is not acceptable. This means that a generally low Root Mean Square Error (RMSE) obtained between the predicted and observed CSM does not guarantee that a model produces reliable forecasts. Kwiatek *et al.* (2024) proposed the best performing model to be the one that produces minimal, yet positive deviation from the observed values. This means that a slight overestimation of the next value is preferred, whereas underestimation is heavily penalized. In the Cooper Basin and in the laboratory fluid injection experiment, we observe that adding more features led to a slight overestimation of the actually observed seismic moment release (see third ML model in Figs 6, 8a, c). However, like the previous models, the third model (LSTM + Att + FC + Feat) is not sensitive to the rapid changes in CSM release either. Interestingly, the introduction of additional features seems to help the third model improve predictions, as long as the injection volume increases, even if the observed CSMs do not. For example, horizontal intervals in the observed CSM in Figs 6, 8(a) show absence or very small release of seismicity. However, in these intervals, fluid is still injected, storing potential energy in the reservoir and preparing conditions for future large earthquakes. Unlike the two other models, LSTM + Att + FC + Feat forecasts CSM with ascending values. This means that using additional features, the model learnt to overestimate even if real CSM is constant, while for example cumulative fluid volume injected (CVI) increases, SER decreases and so on (Figs 6, 8e, f). Unfortunately, in the case of St1 Helsinki, the CSM release remains underestimated even with additional features.

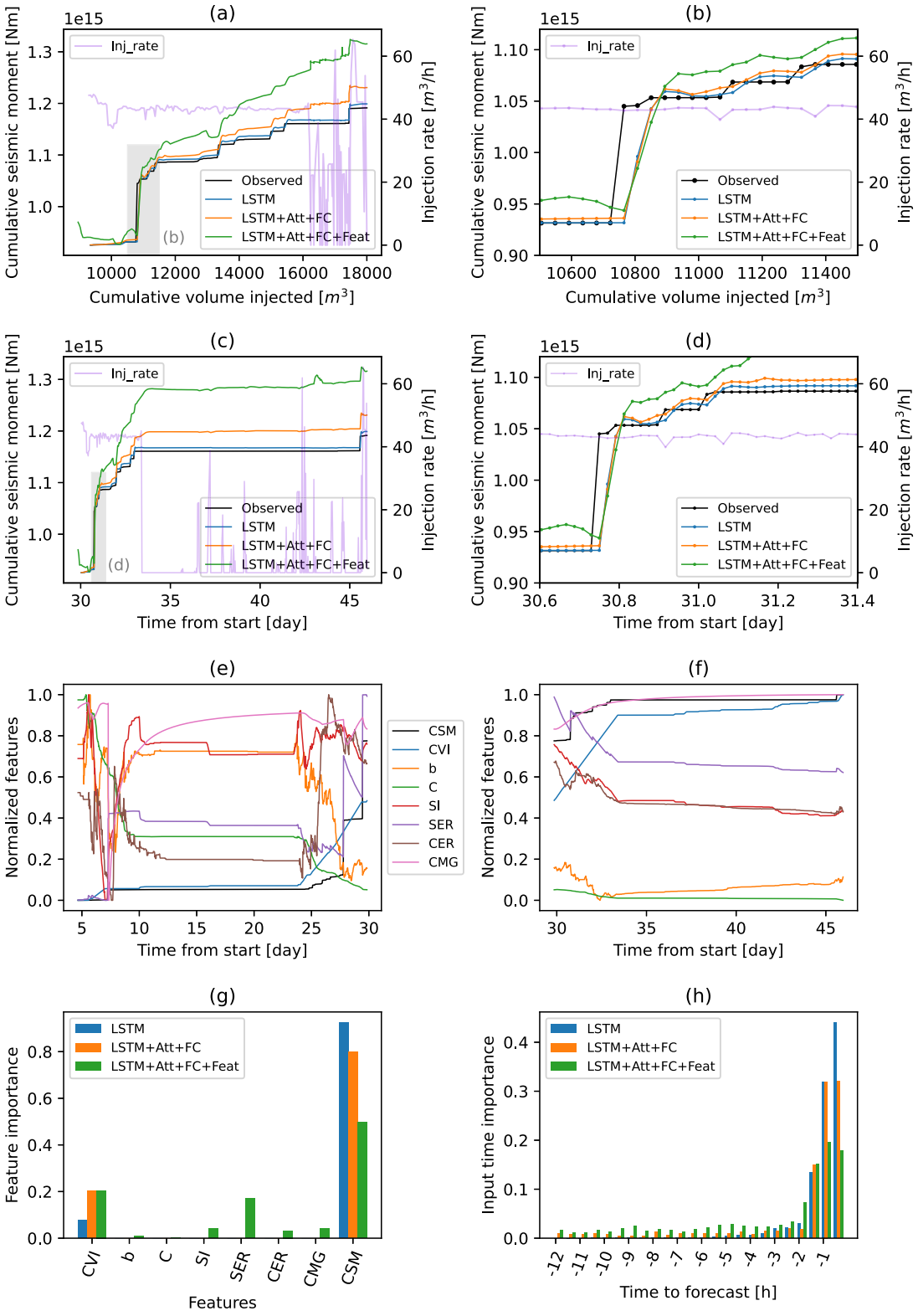

Figure 6. Forecasting cumulative seismic moment (CSM) release for the next 30 min for Cooper basin, Australia, using three different models: LSTM (blue), LSTM + Att + FC (orange) and LSTM + Att + FC + Feat (green) (see section 3.1). In all plots, the observed values are in black. (a) Temporal evolution of CSM with cumulative volume injected (CVI). The shadowed rectangle is the zoom area plotted in (b). (c) Temporal evolution of CSM with time. The shadow rectangle is the zoom area plotted in (d). Injection rate is plotted as the right-axis in (a–d). (e–f) Plots of features with time for training and test phases. Features' values are normalized for a better comparison. (g) Importance weights for each feature in every model. Note that in the first two models (i.e. LSTM and LSTM + Att + FC), only CSM and CVI are inputs. (h) Importance of time steps within the input time-series.

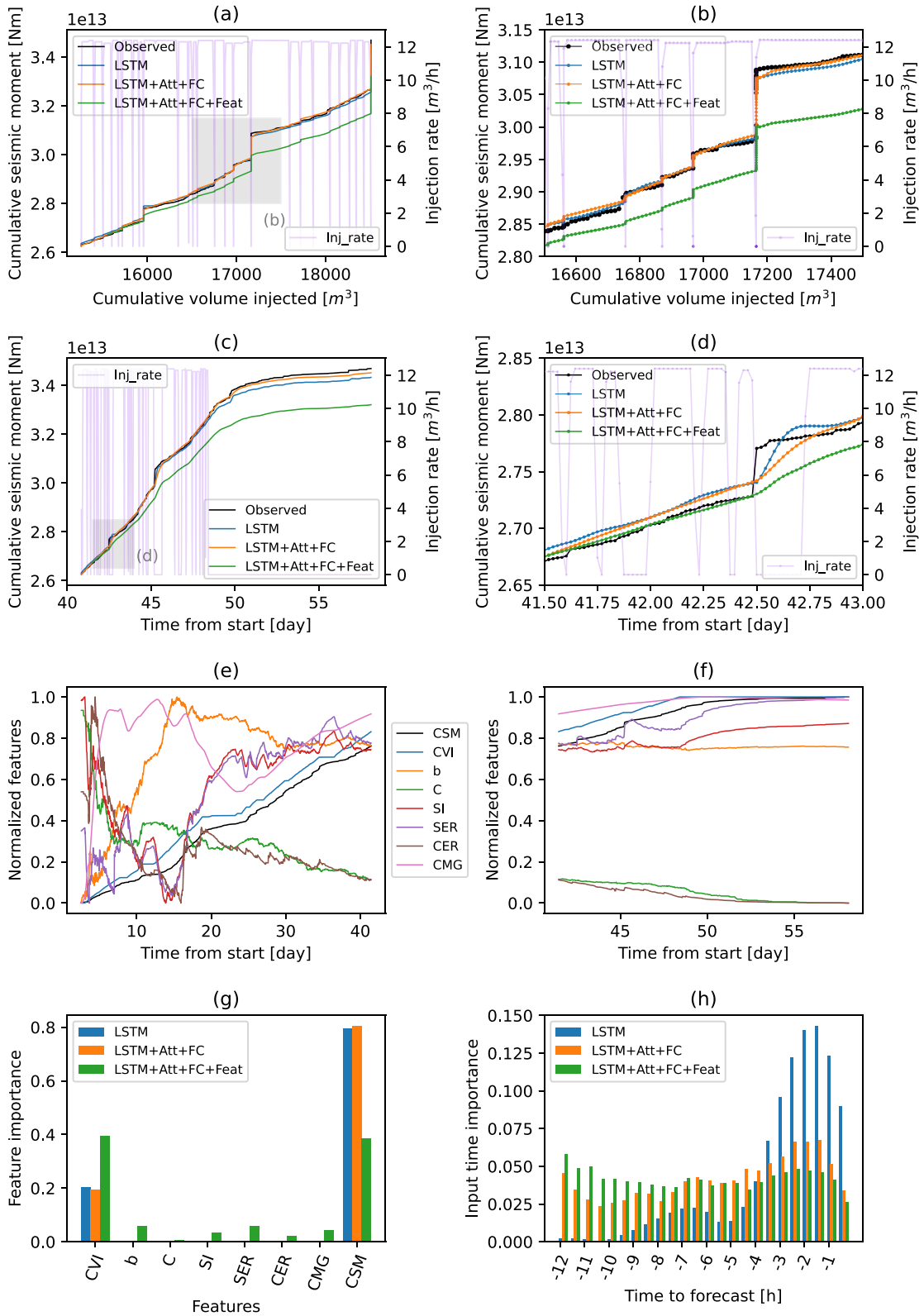

Figure 7. Similar to Fig. 6, but for the St1 Helsinki.

Unlike for Cooper Basin and the lab data, CVI plays an important role in the St1 Helsinki case. Figs 7(b) and (d) show that while fluid is being injected, the model prediction benefits from the high correlation between CSM and CVI (see Fig. 4(b) and also the slope of the forecasting plot in Figs 7(a) and (b)). When injection stops, the model uses the last CSMs in the input time-series.

We computed the importance of each feature and also of each time-step in the trained models (Figs 6, 7, 8g, h) using the

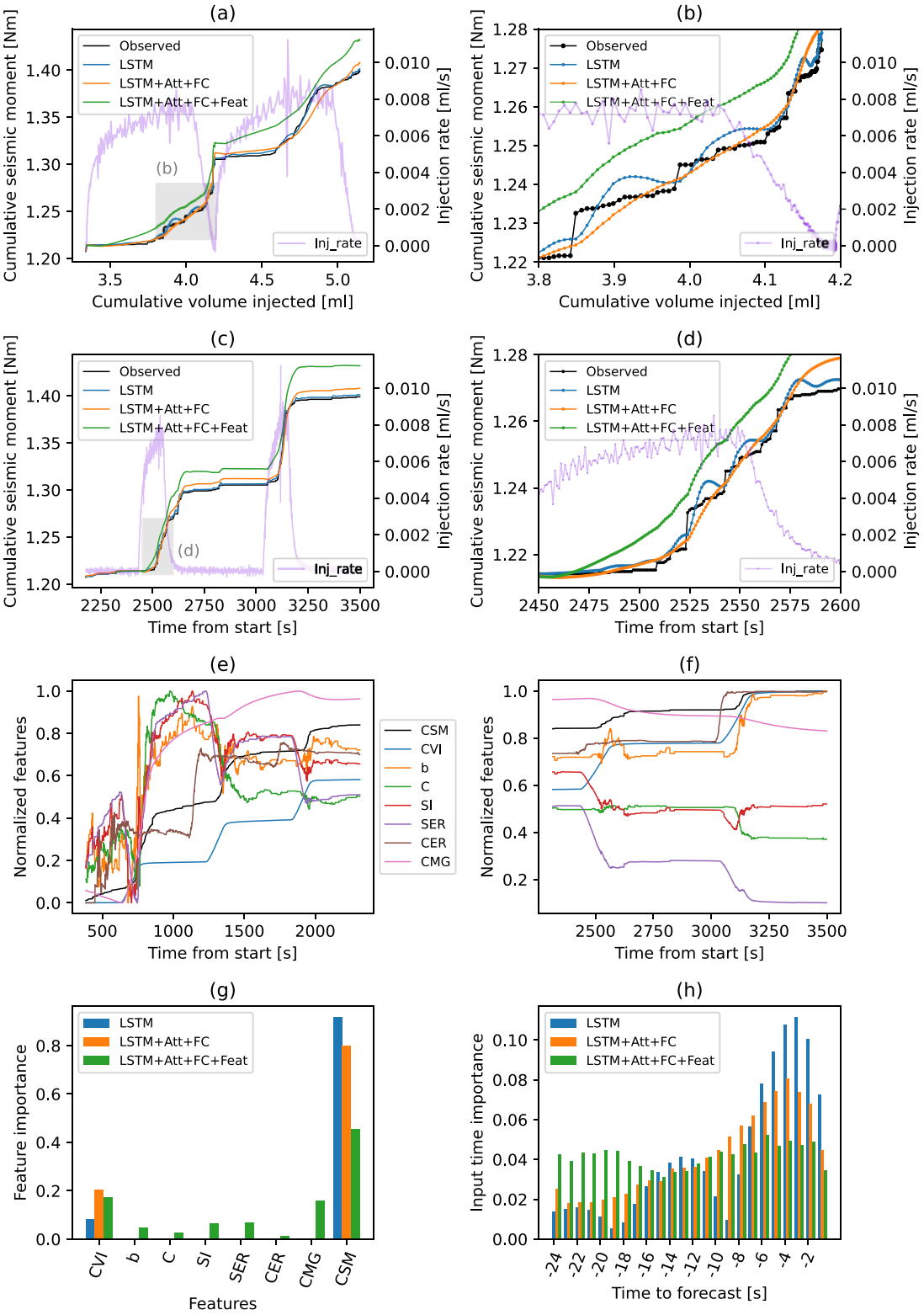

Figure 8. Similar to Fig. 6, but for the experimental data.

integrated gradients method (Sundararajan *et al.* 2017). A comparison of feature importance among different models reveals that almost all models prefer to use past CSM data rather than past CVI data to forecast the CSM in the next time window. All models (but especially LSTM) simply learn using past values of CSM to predict the next upcoming CSM, largely ignoring the CVI evolution. Adding an attention layer and additional features progressively decreases the CSM importance and increases the importance of remaining parameters. In addition, we found that the LSTM mostly uses the most recent time windows to forecast CSM. Adding the

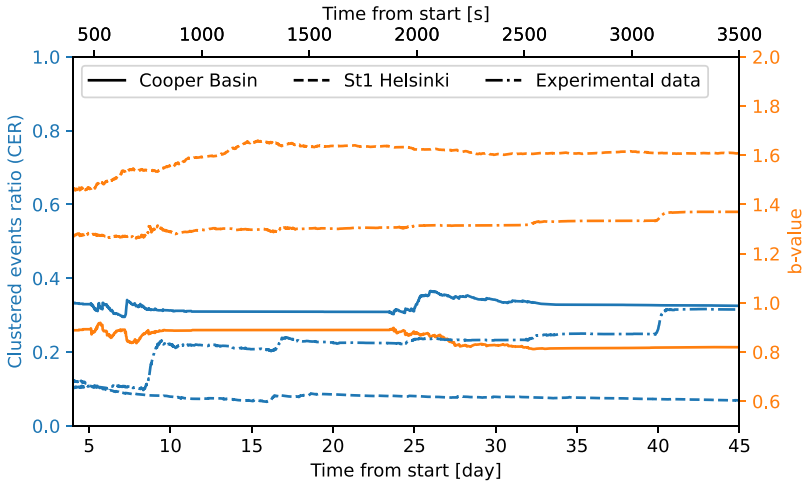

Figure 9. Comparison of clustered events ratio (CER) and *b*-value evolution in all three cases. The time axis for Cooper Basin and St1 Helsinki is in days (lower axis) and for experimental data is in seconds (upper axis).

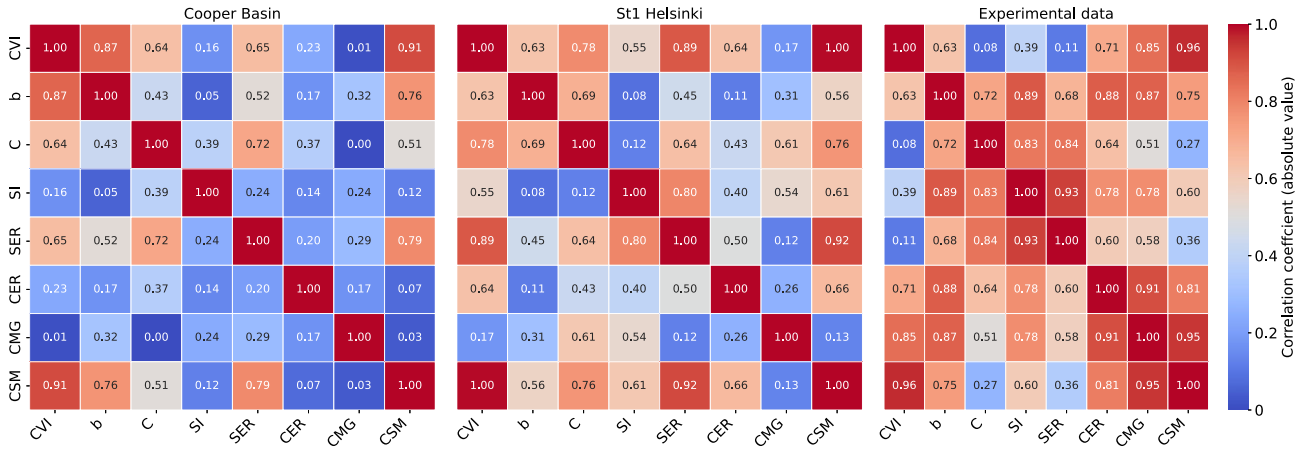

Figure 10. Correlation coefficient (absolute value) among all inputs for every case study.

attention layer enables other models to use more information from a longer time period of the past. This effect is even more emphasized in the third model, in which additional features are used.

5 DISCUSSION

In this study, we forecast cumulative seismic moments and evaluate the stability of the geothermal reservoir in the next injection time step. In addition to the two main inputs (CSM and CVI), we apply catalogue/operational features to introduce more information to the ML model relative to reservoir characteristics, such as potential changes in the structural inventory of the reservoir, or changing stress conditions (Martínez-Garzón *et al.* 2020).

Results show that in the Cooper Basin project and in the laboratory experiments, the addition of seismic features leads to a forecast improvement. In the case of the St1 Helsinki project, however, adding features to the input data does not improve the forecasting model, regardless of the technique applied. In this case, it is conceivable that the additional features calculated do not hold any significant additional information that helps improve model performance. This could be due to several reasons. First, Kwiatek *et al.* (2024) found that the St1 seismicity does not display significant clustering and earthquake interaction, suggesting a passive response to injection operation. In contrast, Fig. 9 shows that for Cooper Basin and similarly experimental data, a high amount of clustered seismicity with low *b*-value may indicate increased mechanical instability induced by the injection within the fault system. Secondly, the St1 seismicity activated a broad network of distributed fractures, without evidence for a prominent fault (Kwiatek *et al.* 2019; Leonhardt *et al.* 2021). Differently, in both Cooper basin and the laboratory experiment, the hypocentre clouds are distributed along a planar fault zone (Király *et al.* 2014; Wang *et al.* 2020a, c). Such a structural difference can strongly affect the (stability of) seismic response of the reservoir as fluid is injected. This is also reflected in the different evolution of seismological features in the studied cases.

Fig. 10 shows the correlation coefficient of input features in each case. The CSM shows a perfect correlation with CVI for St1 Helsinki, but only a relatively low correlation for the other two case studies. Figs 3, 4, 5(b) show that although linear correlation of CSM and CVI is relatively high for the Cooper Basin and experimental data, the jumps in CSM (due to relatively large events) could not be explained only by the CVI. This is why the ML models in these cases benefit from other features to fill these gaps. This is also supported by calculating feature importance (Figs 6–8g). Only in the case of St1 Helsinki, even by adding features, the model increases the importance of CVI, although it tries to use other features. In two

other cases, the reduction of CSM importance is compensated by the other features, and the CVI importance remains similar to the other model (i.e. LSTM + Att + FC).

Feature importance plots for Cooper Basin and experimental cases (Figs 6g, 8g) show that SER, CMG and SI are among the most important features. As the best feature in the Cooper Basin case, SER shows the highest correlation with CSM (Fig. 10) and decreases before large events (Figs 6e and f). CMG is the most important feature in the experimental data, showing a high correlation with CSM in this case (Fig. 10). It helps the ML model to increase forecasting CSM by CVI even if the observed CSM does not increase. SI does not show high correlation with CSM in any of the cases, although ML models still employ its information to forecast the output.

Some limitations exist with regard to applying ML models for large earthquake forecasting in EGS. The main limitation is training data. Since we are limited to the data from one operational well (catalogue and injection rate), ML models could not be trained unless enough data is obtained covering all possible situations. For example, in this study we use about 30–40 d of data to train the model. Although one may start training the model at the very first injection steps (although they are noisy) and update (fine-tune or retrain) it with new data, lacking enough intervals with deviation from the stable phase (e.g. by emerging large events) in the training data is another limiting point. It means that there should be some deviations in the training data to help the model learn the logic of such forecasting. A potential solution to overcome this limitation could be data augmentation by numerical simulation (Gaucher *et al.* 2015). By assuming different subsurface possibilities for each ESG, numerical modelling could be applied to generate numerous injection scenarios and to prepare subsequent catalogues of seismicity. Some specific characteristics such as delayed fault activation, like in Pohang EGS (Woo *et al.* 2019), may also be implemented by numerical modelling (e.g. a digital twin). Such a huge training data set including various subsurface/fault scenarios could then be used to pre-train a more robust ML model. Following the potentials shown in this study, preparing such numerical training data would be a future step of this study.

The other limitation would be the range of data. Due to the cumulative nature of the output value (CSM), the ML model needs to extrapolate the next CSM, which is highly error prone for most ML algorithms. However, with the problem in this study (Fig. 1), we can have a rough estimation of the maximum (end) values in the future. For example, based on the scenario of an injection operation, the maximum cumulative fluid volume injected (CVI) is known. Using this value, we could forecast the final CSM with respect to the theoretical models (e.g. McGarr 2014). These final values then define the full range of variables and allow for data normalization. Such a strategy enables us to predict new outputs in the test steps (regarding the values in training data), considering the increasing slope of CSM with CVI.

If the above limitations are properly considered, implementation of the method used in this study is somewhat straightforward in real operational wells. At each time step (i.e. every 30 min in this study), the catalogue and injection rate are obtained from the operational unit, input features are computed, training data are reproduced and the ML model is either fine-tuned or retrained. The training time of a new model is on the order of a few minutes in this study (CPU: Core-i9, GPU: RTX-A2000, RAM: 32 GB), which is relatively small compared to the time step (of 30 min). In addition to all operational data, the estimated CSM for the next step provides valuable additional information for operational decision-makers on site to potentially adjust the injection schedule.

6 CONCLUSIONS

In this study we have employed ML to forecast the evolution of cumulative seismic moment, and hence the seismic stability of a reservoir with respect to fluid injection activities. The seismic and injection parameters from two geothermal experiments and one laboratory experiment involving fault reactivation due to fluid injection have been analysed and compared. Our results demonstrate the feasibility of predicting CSM for subsequent injection phases. They also indicate that feature-rich models outperform simpler frameworks in cases with complex seismic responses, as observed for Cooper Basin and laboratory data. The most important features in these cases are SER, CMG and SI. However, in environments such as St1 Helsinki, characterized by distributed seismicity with weak clustering, additional features offered limited predictive benefit. A comparison of the tested data sets points towards a strong role of earthquake clustering and localization causing differences in seismic responses (i.e. stable versus unstable), here related to the existence of prominent structural features (faults) in the case of Cooper Basin and experimental data.

Key limitations of the presented forecasting approach include the scarcity of training data from operational wells, the extrapolation demands of cumulative outputs and the underestimation of rapid CSM jumps linked to large induced earthquakes. Addressing these limitations through synthetic data augmentation by numerical simulation could improve model robustness and predictive accuracy.

The findings highlight the necessity of incorporating real-time monitoring, enhanced data features and adaptable ML frameworks to ensure the safe and efficient operation of EGS projects. As geothermal energy expands its role in sustainable energy systems, such predictive methodologies will be vital in mitigating seismic risks, while optimizing energy extraction.

ACKNOWLEDGMENTS

This work was performed in the frame of EU HORIZON DT-GEO project (HORIZON-INFRA-2021-TECH-01, project number 101058129). This work was supported by Helmholtz Association's Initiative and Networking Fund through Helmholtz AI (grant number: ZT-I-PF-5-01).

AUTHORS' CONTRIBUTION

Sadegh Karimpouli (Methodology, Validation, Visualization, Writing—original draft), Grzegorz Kwiatek (Conceptualization, Supervision, Validation, Writing—review and editing), Patricia Martínez-Garzón (Conceptualization, Validation, Writing—review and editing), Danu Caus (Methodology, Validation, Writing—review and editing), Lei Wang (Data curation, Validation, Writing—review and editing) ,Georg Dresen (Conceptualization, Validation, Writing—review and editing), and Marco Bohnhoff (Conceptualization, Validation, Writing—review and editing)

DATA AVAILABILITY

Seismic and hydraulic data for the Cooper Basin case are accessible on the geodata repository of the EPISODES platform at https://doi.org/10.25171/InstGeoph_PAS_ISEPOS-2020-001. The St1 Deep Head data are available from Leonhardt, Kwiatek, Martínez-Garzón, & Heikkinen (2021) at https://doi.org/10.5880/GFZ.4.2.2021.001. Experimental data are available from Wang *et al.* (2020b) at https://doi.org/10.17632/cbhrs6d8rr.1

REFERENCES

Amitrano, D., 2003. Brittle-ductile transition and associated seismicity: experimental and numerical studies and relationship with the b value, *J. geophys. Res. Solid Earth,* **108,** 2044. doi:10.1029/2001JB000680

Bahdanau, D., 2014. Neural machine translation by jointly learning to align and translate, *arXiv* (arXiv:1409.0473)

Baiesi, M. & Paczuski, M., 2005. Complex networks of earthquakes and aftershocks, *Nonlinear Process Geophys.,* **12,** 1–11.

Baisch, S., Vörös, R., Weidler, R. & Wyborn, D., 2009. Investigation of fault mechanisms during geothermal reservoir stimulation experiments in the Cooper Basin, Australia, *Bull. seism. Soc. Am.,* **99,** 148–158.

Baisch, S., Weidler, R., Vörös, R., Wyborn, D. & Graaf, L.d., 2006. Induced seismicity during the stimulation of a geothermal HFR reservoir in the Cooper Basin, Australia, *Bull. seism. Soc. Am.,* **96,** 2242–2256.

Bentz, S., Kwiatek, G., Martínez-Garzón, P., Bohnhoff, M. & Dresen, G., 2020. Seismic moment evolution during hydraulic stimulations, *Geophys. Res. Lett.,* **47,** e2019GL086185. doi:10.1029/2019GL086185

Deichmann, N. & Giardini, D., 2009. Earthquakes induced by the stimulation of an enhanced geothermal system below Basel (Switzerland), *Seismol. Res. Lett.,* **80,** 784–798.

Ellsworth, W.L., 2013. Injection-induced earthquakes, *Science,* **341,** doi:10.1126/SCIENCE.1225942/SUPPL_FILE/1225942.MP3

Elman, J.L., 1990. Finding structure in time, *Cogn. Sci.,* **14,** 179–211.

Elst, N.J.v.d., Page, M.T., Weiser, D.A., Goebel, T.H.W. & Hosseini, S.M., 2016. Induced earthquake magnitudes are as large as (statistically) expected, *J. geophys. Res. Solid Earth,* **121,** 4575–4590.

Galis, M., Ampuero, J.P., Mai, P.M. & Cappa, F., 2017. Induced seismicity provides insight into why earthquake ruptures stop, *Sci. Adv.,* **3**. doi:10.1126/SCIADV.AAP7528/SUPPL_FILE/AAP7528_SM.PDF

Gaucher, E., Schoenball, M., Heidbach, O., Zang, A., Fokker, P.A., Wees, J.D.V. & Kohl, T., 2015. Induced seismicity in geothermal reservoirs: a review of forecasting approaches, *Renew. Sust. Energy Rev.,* **52,** 1473–1490.

Giardini, D., 2009. Geothermal quake risks must be faced, *Nature,* **462,** 848–849.

Hallo, M., Eisner, L. & Ali, M.Y., 2012. Expected level of seismic activity caused by volumetric changes, *First Break,* **30,** 97–100.

Hallo, M., Oprsal, I., Eisner, L. & Ali, M.Y., 2014. Prediction of magnitude of the largest potentially induced seismic event, *J. Seismol.,* **18,** 421–431.

Hincks, T., Aspinall, W., Cooke, R. & Gernon, T., 2018. Oklahoma's induced seismicity strongly linked to wastewater injection depth, *Science (1979),* **359,** 1251–1255.

Hochreiter, S. & Schmidhuber, J., 1997. Long short-term memory, *Neural Comput.,* **9,** 1735–1780

Holtzman, B.K., Paté, A., Paisley, J., Waldhauser, F. & Repetto, D., 2018. Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field, *Sci. Adv.,* **4,** doi:10.1126/SCIADV.AAO2929/SUPPL_FILE/AAO2929_SM.PDF

Hui, G. *et al.*, 2023. A machine learning-based study of multifactor susceptibility and risk control of induced seismicity in unconventional reservoirs, *Pet. Sci.,* **20,** 2232–2243.

Karimpouli, S. *et al.*, 2023. Explainable machine learning for labquake prediction using catalog-driven features, *Earth planet. Sci. Lett.,* **622,** 118383. doi:10.1016/J.EPSL.2023.118383

Karimpouli, S., Kwiatek, G., Ben-Zion, Y., Martínez-Garzón, P., Dresen, G. & Bohnhoff, M., 2024. Empowering machine learning forecasting of labquake using event-based features and clustering characteristics, *J. geophys. Res.: Machine Learning and Computation,* **1,** e2024JH000160. doi:10.1029/2024JH000160

Karimpouli, S., Kwiatek, G., Martínez-Garzón, P., Dresen, G. & Bohnhoff, M., 2024. Unsupervised clustering of catalogue-driven features for characterizing temporal evolution of labquake stress, *Geophys. J. Int.,* **237,** 755–771

Kingma, D.P. & Ba, J., 2014. *Adam: a Method for Stochastic Optimization,* arXiv(arXiv:1412.6980) Retrieved from http://arxiv.org/abs/1412.6980

Király, E., Gischig, V., Karvounis, D. & Weimer, S., 2014. Validating models to forecasting induced seismicity related to deep geothermal energy projects, in *Proceeding of 39th Workshop on Geothermal Reservoir Engineering,* Stanford, CA, pp. 1–9

Kwiatek, G., Grigoratos, I. & Wiemer, S., 2024. Variability of seismicity rates and maximum magnitude for adjacent hydraulic stimulations, *Seismol. Res. Lett.,* **96,** 920–932

Kwiatek, G., Martínez-Garzón, P., Davidsen, J., Malin, P., Karjalainen, A., Bohnhoff, M. & Dresen, G., 2022. Limited earthquake interaction during a geothermal hydraulic stimulation in Helsinki, Finland, *J. geophys. Res. Solid Earth,* **127,** e2022JB024354. doi:10.1029/2022JB024354

Kwiatek, G. *et al.*, 2019. Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland, *Sci. Adv.,* **5,** doi:10.1126/SCIADV.AAV7224/SUPPL_FILE/AAV7224_SM.PDF

Leong, Z.X. & Zhu, T., 2024. Machine learning-assisted microearthquake location workflow for monitoring the Newberry Enhanced geothermal system, *J. geophys. Res.: Machine Learning and Computation,* **1,** e2024JH000159 doi:10.1029/2024JH000159

Leonhardt, M., Kwiatek, G., Martínez-Garzón, P., Bohnhoff, M., Saarno, T., Heikkinen, P. & Dresen, G., 2021. Seismicity during and after stimulation of a 6.1 km deep enhanced geothermal system in Helsinki, Finland, *Solid Earth,* **12,** 581–594.

Leonhardt, M., Kwiatek, G., Martínez-Garzón, P. & Heikkinen, P., 2021. Earthquake catalog of induced seismicity recorded during and after stimulation of Enhanced Geothermal System in Helsinki, Finland, GFZ Data Services doi:10.5880/GFZ.4.2.2021.001

Li, D., Huang, L., Li, Y., Zheng, Y. & Moore, J., 2024. Seismic monitoring of EGS fracture stimulations at Utah FORGE (Part 1): time-lapse variations of b-values and shear-wave splitting rates of induced microearthquakes, *Geothermics,* **120,** 103005. doi:10.1016/J.GEOTHERMICS.2024.103005

Li, Z., Eaton, D. & Davidsen, J., 2022. Short-term forecasting of Mmax during hydraulic fracturing, *Sci. Rep.,* **12,** 1–12.

Limbeck, J. *et al.*, 2021. Using machine learning for model benchmarking and forecasting of depletion-induced seismicity in the Groningen gas field, *Comput. Geosci.,* **25,** 529–551.

Majer, E., Nelson, J., Robertson-Tait, A., Savy, J. & Wong, I., 2012. Protocol for addressing induced seismicity associated with enhanced geothermal systems, Washington, DC (United States), doi:10.2172/1219482

Martínez-Garzón, P., Kwiatek, G., Bentz, S., Bohnhoff, M. & Dresen, G., 2020. Induced earthquake potential in geothermal reservoirs: insights from The Geysers, California, *The Leading Edge,* **39,** 873–882.

Martínez-Garzón, P., Kwiatek, G., Sone, H., Bohnhoff, M., Dresen, G. & Hartline, C., 2014. Spatiotemporal changes, faulting regimes, and source parameters of induced seismicity: a case study from the Geysers geothermal field, *J. geophys. Res. Solid Earth,* **119,** 8378–8396.

Martínez-Garzón, P., Zaliapin, I., Ben-Zion, Y., Kwiatek, G. & Bohnhoff, M., 2018. Comparative study of earthquake clustering in relation to hydraulic activities at geothermal fields in California, *J. geophys. Res. Solid Earth,* **123,** 4041–4062.

McGarr, A., 2014. Maximum magnitude earthquakes induced by fluid injection, *J. geophys. Res. Solid Earth,* **119,** 1008–1019.

Mehrabifard, A. & Eberhardt, E., 2021. Investigation of the dependence of induced seismicity magnitudes on differential stress and pore pressure using supervised machine learning, northeastern British Columbia (NTS 093, 094A, B, G, H) and globally, *Geosci. BC Summary of Activities,* **2020,** 2021–2022.

Mukuhira, Y. *et al.*, 2023. Scaling microseismic cloud shape during hydraulic stimulation using *in situ* stress and permeability, *J. geophys. Res. Solid Earth,* **128,** e2023JB026839. doi:10.1029/2023JB026839

Norbeck, J.H. & Horne, R.N., 2018. Maximum magnitude of injection-induced earthquakes: a criterion to assess the influence of pressure migration along faults, *Tectonophysics,* **733,** 108–118.

Okamoto, K., Mukuhira, Y., Darisma, D., Asanuma, H. & Moriya, H., 2024. Machine learning automatic picker for geothermal microseismicity analysis for practical procedure to reveal fine reservoir structures, *Geothermics,* **116,** 102832 doi:10.1016/J.GEOTHERMICS.2023.102832

Park, S. *et al.*, 2020. Observations and analyses of the first two hydraulic stimulations in the Pohang geothermal development site, South Korea, *Geothermics,* **88,** 101905. doi:10.1016/J.GEOTHERMICS.2020.101905

Picozzi, M., Iaccarino, A.G. & Spallarossa, D., 2023. The preparatory process of the 2023 M_w 7.8 Türkiye earthquake, *Sci. Rep.,* **13,** 1–10

Qin, Y., Chen, T., Ma, X. & Chen, X., 2022. Forecasting induced seismicity in Oklahoma using machine learning methods, *Sci. Rep.,* **12,** 1–8.

Ripperger, J., Ampuero, J.P., Mai, P.M. & Giardini, D., 2007. Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress, *J. geophys. Res. Solid Earth,* **112,** 4311. doi:10.1029/2006JB004515

Ritz, V.A., Rinaldi, A.P. & Wiemer, S., 2022. Transient evolution of the relative size distribution of earthquakes as a risk indicator for induced seismicity, *Commun. Earth Environ.,* **3,** 1–11.

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J. & Johnson, P.A., 2017. Machine learning predicts laboratory earthquakes, *Geophys. Res. Lett.,* **44,** 9276–9282.

Rumelhart, D.E., Hinton, G.E. & Williams, R.J., 1986. Learning representations by back-propagating errors, *Nature,* **323,** 533–536.

Saad, O.M. *et al.*, 2023. Earthquake forecasting using big data and artificial intelligence: a 30-week real-time case study in China, *Bull. seism. Soc. Am.,* **113,** 2461–2478.

Scholz, C.H., 2015. On the stress dependence of the earthquake b value, *Geophys. Res. Lett.,* **42,** 1399–1402.

Shapiro, S.A., Dinske, C., Langenbruch, C. & Wenzel, F., 2010. Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations, *The Leading Edge,* **29,** 304–309.

Sundararajan, M., Taly, A. & Yan, Q., 2017. Axiomatic attribution for deep networks, in *the 34th International Conference on Machine Learning,* pp. 3319–3328, PMLR. Retrieved from https://proceedings.mlr.press/v70/sundararajan17a.html

Vaswani, A. *et al.*, 2017. Attention is all you need. Advances in neural information processing systems, *Adv. Neur. Inf. Process Syst.,* **30,** 1–11

Wang, L., Kwiatek, G., Rybacki, E., Bohnhoff, M. & Dresen, G., 2020a. Injection-induced seismic moment release and laboratory fault slip: implications for fluid-induced seismicity, *Geophys. Res. Lett.,* **47,** e2020GL089576. doi:10.1029/2020GL089576

Wang, L., Kwiatek, G., Rybacki, E., Bohnhoff, M. & Dresen, G., 2020b. *Injection-induced Seismic Moment Release of Acoustic Emission during Laboratory Fault Slip,* **1**, Mendeley Data. doi:10.17632/CBHRS6D8RR.1

Wang, L., Kwiatek, G., Rybacki, E., Bonnelye, A., Bohnhoff, M. & Dresen, G., 2020c. Laboratory study on fluid-induced fault slip behavior: the role of fluid pressurization rate, *Geophys. Res. Lett.,* **47,** e2019GL086627. doi:10.1029/2019GL086627

Woo, J.U. *et al.*, 2019. An In-depth seismological analysis revealing a causal link between the 2017 MW 5.5 Pohang earthquake and EGS Project, *J. geophys. Res. Solid Earth,* **124,** 13 060–13 078.

Wozniakowska, P. & Eaton, D.W., 2020. Machine learning-based analysis of geological susceptibility to induced seismicity in the Montney Formation, Canada, *Geophys. Res. Lett.,* **47,** e2020GL089651. doi:10.1029/2020GL089651

Zaliapin, I. & Ben-Zion, Y., 2013. Earthquake clusters in southern California I: identification and stability, *J. geophys. Res. Solid Earth,* **118,** 2847–2864.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V. & Wong, H., 2008. Clustering analysis of seismicity and aftershock identification, *Phys. Rev. Lett.,* **101,** 018501. Available at: https://doi.org/10.1103/PHYSREVLETT.101.018501/FIGURES/5/MEDIUM