PROCEEDINGS, 51st Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 9-11, 2026
SGP-TR-230

QuakeCastNet: An Interpretable Deep Learning Framework for Induced Seismicity Forecasting in Geothermal Fields

Zhengfa Bi[1], Nori Nakata[1]

[1]Lawrence Berkeley National Laboratory

zfbi@lbl.gov, nnakata@lbl.gov

Keywords: seismicity forecasting, induced seismicity, EGS, deep learning

ABSTRACT

Induced seismicity presents a critical challenge in geothermal reservoir management, as the occurrence of large seismic events raises public safety concerns and affects social acceptance of geothermal operations. Accurate forecasting of induced seismicity provides valuable information for operators and improves understanding of the underlying physical processes. However, physics-based modeling requires detailed knowledge of subsurface properties and intensive computation, while conventional statistical methods struggle to capture the nonlinear relationships between operational parameters and seismicity.

In this study, we introduce QuakeCastNet, an interpretable deep learning framework for forecasting induced seismicity, including seismicity rates, the spatiotemporal evolution, and magnitude distribution in geothermal fields, and demonstrate it using data recorded at Utah FORGE and The Geysers. QuakeCastNet combines a modified Temporal Fusion Transformer (TFT) for time-series forecasting with a Graph Neural Network (GNN) for spatial representation learning, enabling joint modeling of complex dependencies across both time and space. The framework integrates heterogeneous datasets—including geological information, historical seismicity catalogs, and injection and production metadata—to predict future seismicity rates and magnitude-frequency distributions within gridded regions of interest. The estimated magnitude distribution provides insights into the probability of larger seismic events in the future to support risk assessment.

By coupling explainable AI with data-driven learning, QuakeCastNet demonstrates improved predictive performance and interpretability, advancing our understanding of induced seismicity mechanisms and supporting safer, adaptive reservoir management strategies in geothermal operations..

1. INTRODUCTION

Energy and resource extraction reshape subsurface stress and pore pressure, potentially triggering seismic events (Ellsworth, 2013; Chen et al., 2017; Grigoli et al., 2017; Schultz et al., 2020). Incorporating these factors into quantitative models is essential for understanding earthquake processes and mitigating seismic hazards (McGarr, 2014; Chen & Nakata, 2017; Goebel & Brodsky, 2018; Lee et al., 2019; Hager et al., 2021; Ritz et al., 2022). As geothermal energy becomes a key pillar of global carbon neutrality efforts (Gaucher et al., 2015; Chang & Nakata, 2022), the challenge of managing induced seismicity and associated ground motion in geothermal fields grows increasingly critical. Large seismic events raise public safety concerns, drive economic repercussions, and necessitate stringent mitigation strategies. These concerns are magnified with the expansion of Enhanced Geothermal Systems (EGS) (Tester et al., 2006; Ziagos et al., 2013) designed to extract energy from low-permeability reservoirs. To create a geothermal reservoir at preferable depth and temperature, we stimulate fractures using, for example, hydraulic fracturing and shearing, potentially with proppant. This process alters the subsurface stress significantly. Large induced earthquakes can potentially pause and/or terminate the geothermal resource exploration, and hence the accurate forecasting of induced seismicity becomes an urgent task. Such perturbations even extend beyond geothermal operations to include reservoir impoundment, oil and gas production, and underground waste disposal (Spica, Nakata, et al., 2018; Spica, Perton, et al., 2018).

Many efforts have been made in the last two decades to understand and forecast induced seismicity using statistical or physics-based models that rely on assumptions or prior geological knowledge of seismicity mechanisms. Statistical approaches, such as the epidemic-type aftershock sequence (ETAS) model (Ogata, 1988, 1998), have been widely used to model seismicity as a stochastic point process in time, space, and magnitude domains (Vere-Jones, 2010). The ETAS model captures both background seismicity driven by tectonic loading and triggered aftershocks, and has been applied extensively to tectonic and induced seismicity (Bachmann et al., 2011; Brodsky & Lajoie, 2013; Llenos & Michael, 2016; Kothari et al., 2020; Aochi et al., 2021; Sedghizadeh et al., 2024). However, its applicability to induced seismicity requires modifications to incorporate anthropogenic influences, such as fluid injection and poroelastic stress changes (Weingarten et al., 2015; Alghannam & Juanes, 2020; Shcherbakov, 2024).

To address these limitations, models that directly link seismicity rate to fluid injection have been proposed, assuming a linear relationship between seismicity and injection (Hainzl & Ogata, 2005; Bachmann et al., 2011; Brodsky & Lajoie, 2013; Holtzman et al., 2018). These approaches have been applied to geothermal fields, hydraulic fracturing, and wastewater injection sites worldwide, demonstrating that controlling injection rates can mitigate seismic hazard (Improta et al., 2015; Garcia-Aristizabal, 2018; Langenbruch et al., 2018; Dempsey & Riffault, 2019). Recent advancements integrate linear response theory into statistical models, treating fluid injection as an external forcing function, where the seismicity rate is modeled as a convolution of the injection rate with a kernel function that encapsulates the transient response of the subsurface (Kim & Avouac, 2023). Building on this framework, induced seismicity is further formulated as the response of a nonequilibrium thermodynamic system to external perturbations, with the rheology-dependent kernel function capturing the viscoelastic properties of the subsurface (Shcherbakov, 2024; White & Nakata, 2025). Additionally, rate-and-state friction models have been employed to improve seismicity rate forecasts by incorporating stress redistribution effects (Segall & Lu, 2015; Deng et al., 2016; Norbeck & Rubinstein, 2018; Zhai et al., 2019). Despite these advancements, statistical frameworks often prioritize temporal evolution while neglecting the stochastic nature and complex interactions among seismicity-driving factors.

Physics-based models simulate the dynamic changes in stress and strain induced by geothermal operations, offering deeper insights into the underlying physical processes. Several mechanisms have been proposed to explain induced seismicity. One prominent mechanism is pore pressure diffusion, which reduces effective normal stress and promotes fault slip (Shapiro & Dinske, 2009; Bachmann et al., 2012). Another is poroelastic stress transfer, where fluid injection redistributes stresses beyond the fluid diffusion front and potentially triggers seismic events (Atkinson et al., 2016; Bao & Eaton, 2016). Additionally, aseismic slip propagation has been observed to induce earthquakes at distances far from injection sites, sometimes outpacing the diffusion of pore pressure (Guglielmi et al., 2015; Bhattacharya & Viesca, 2019; Yeo et al., 2020; Wynants-Morel et al., 2020). While these models provide valuable insights into the conditions leading to induced seismicity, they struggle to quantify the relative contributions of different triggering mechanisms or resolve the complex interactions driving seismic activity. Geomechanical modeling plays an important role in understanding the physical mechanisms of induced seismicity, but it is also limited by the randomness of earthquake occurrence and the heterogeneous properties and structure of the subsurface. Although numerical modeling of induced seismicity using thermal-hydrological-mechanical-chemical (THMC) modeling is another path to explore (Smith et al., 2024), this requires significant computational resources and detailed subsurface knowledge for accurate modeling.

Beyond traditional models, modern data-driven approaches have shown promise. For example, random forest models using physics-based features have been applied to forecast induced seismicity in Oklahoma, which identifies stress rate changes as dominant drivers of seismicity rate variations (Qin et al., 2022). A spatiotemporal ConvLSTM network has been proposed to capture both spatial correlations and temporal dependencies across global seismic maps, demonstrating improved predictive accuracy over classical methods (Zhang & Wang, 2023). Real-time forecasting efforts using big data streams, such as geo-acoustic and electromagnetic signals, have also shown promise and achieve up to 70% accuracy in weekly predictions over a seven-month campaign in China (Saad et al., 2023). Meanwhile, deep learning has been applied to synthetic megathrust simulations to identify precursory signals and predict complete ruptures (Blank & Morgan, 2021), and foundation models like transformers and graph neural networks have been evaluated for earthquake nowcasting over large spatiotemporal domains (Jafari et al., 2024). These developments exhibit the increasing potential of modern AI frameworks to model complex, multiscale interactions in seismic processes beyond the capabilities of traditional statistical tools.

In this study, we develop an interpretable deep learning framework to forecast the occurrence of induced seismicity in geothermal fields. Building on the Temporal Fusion Transformer (TFT) (Lim et al., 2021), the model integrates attention mechanism with multi-horizon quantile forecasting to capture complex interactions among contributing factors. Multi-horizon forecasting enables simultaneous prediction across multiple future time steps, which allows the model to characterize evolving temporal dynamics beyond a single-step forecast. Quantile forecasting estimates various percentiles of the predicted distribution to evaluate forecast uncertainty and range. The attention mechanism offers interpretable insights of key driving factors of induced seismicity without imposing prior assumptions. The model processes static parameters, time-dependent past observations, and known future inputs. Outputs include probabilistic predictions across multiple quantiles, with uncertainty estimates under varying operational conditions. While primarily designed for forecasting, the framework offers potential to enhance understanding of the physical processes governing seismicity by identifying dominant factors and their temporal influence through the attention weights. We apply the model to two geothermal systems characterized by induced seismicity, including the Geysers and the Utah FORGE fields. By demonstrating superior forecasting performance in both datasets, our approach contributes to the sustainable development of geothermal energy while enhancing risk assessment and public safety.

2. METHODOLOGY

The proposed model leverages a deep transformer to integrate multiple types of time series, dynamically select relevant features, and capture complex temporal dependencies, which enable accurate and interpretable forecasting of induced seismicity.

2.1. Temporal Transformer Framework

The Temporal Fusion Transformer (TFT) is a state-of-the-art architecture for multi-horizon time series forecasting that integrates heterogeneous data and captures complex temporal dynamics. Compared to conventional models such as Long Short-Term Memory (LSTM) or Sequence-to-Sequence architectures, TFT offers broader applicability for complex forecasting scenarios. It supports both static and time-varying covariates, handles known future inputs, and incorporates quantile forecasting for probabilistic output.

We extend this network to align with the characteristics of seismicity and injection data to achieve reliable and interpretable forecasting. Our model processes three input types: static metadata, time-dependent past observations, and planned future schedules. These inputs are processed through embeddings for categorical and contextual information and linear transformations for continuous variables. The model architecture comprises a static covariate encoder, variable selection networks, a sequence-to-sequence LSTM, and a self-attention decoder (Figure 1). These elements collectively enable robust and interpretable seismicity forecasting. The following sections provide a detailed exposition of these elements.

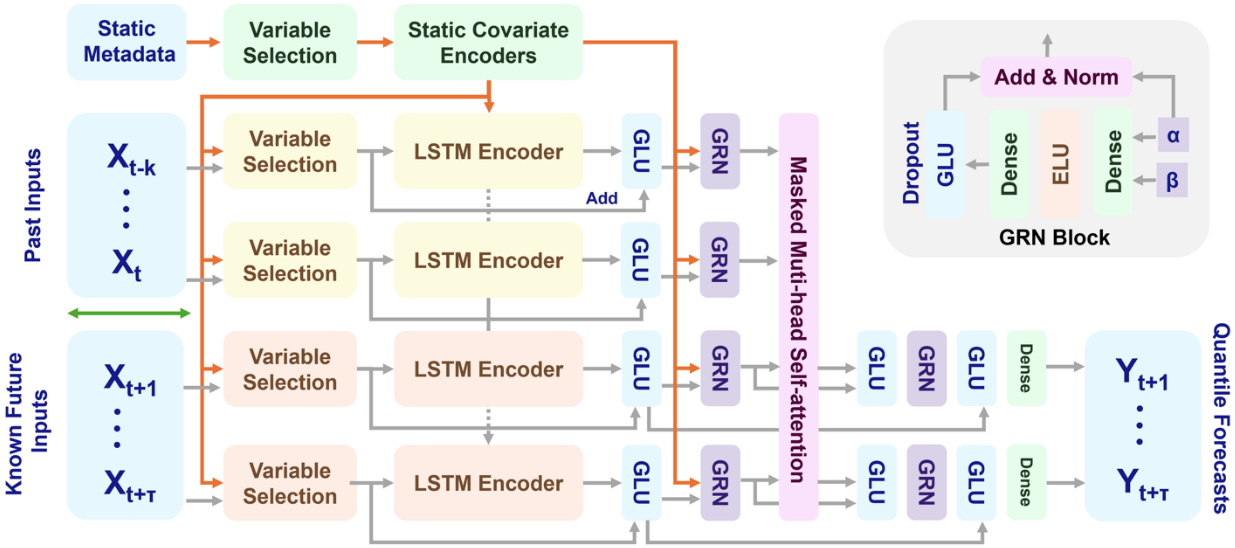

Figure 1: Overview of the Temporal Fusion Transformer (TFT) architecture for induced seismicity forecasting. The architecture effectively captures complex temporal dynamics, providing interpretable insights into the primary drivers of seismicity.

2.1.1. Static Covariate Encoder and Gating Mechanism

The model employs a Gated Residual Network (GRN) to encode static covariates that constrain temporal dynamics across the network, enabling downstream components to leverage contextual information. The static covariates are further reintroduced at multiple stages through a static enrichment layer, ensuring that predictions remain consistent with the constraints. The GRN employs gating mechanisms to achieve dynamic feature selection, enabling the efficient processing of heterogeneous data with varying complexity. The GRN is formulated as:

$$GRN_\omega(a, c) = \mathrm{LayerNorm}(a + GLU_\omega(\eta_1))$$

$$\eta_1 = W_{1,\omega}\,\eta_2 + b_{1,\omega}, \quad \eta_2 = \mathrm{ELU}(W_{2,\omega}\, a + W_{3,\omega}\, c + b_{2,\omega}),$$

where a and c are the primary input and the optional context vector, respectively, and ω denotes weight sharing. The Exponential Linear Unit (ELU) activation maintains sensitivity to informative positive signals while suppressing noisy negative inputs. This dynamic control over magnitudes allows the GRN to adjust its expressiveness based on the relevance of the inputs. To control the nonlinear contribution, we employ Gated Linear Units (GLUs) as gating mechanisms:

$$GLU_\omega(\gamma) = \sigma(W_{4,\omega}\, \gamma + b_{4,\omega}) \odot (W_{5,\omega}\, \gamma + b_{5,\omega}),$$

where $\sigma(\cdot)$ is the sigmoid activation, $W(\cdot)$ and $b(\cdot)$ are trainable weights and biases, and $\odot$ represents element-wise multiplication. GLUs suppress non-essential layers by producing outputs close to zero. Dropout is further applied before the gating layer to enhance generalization.

2.1.2. Variable Selection Network

Time-varying past and known future inputs are processed through separate variable selection networks, which dynamically assign feature weights at each time step to filter out irrelevant or noisy inputs. Let $\xi_t^j \in \mathbb{R}^{d\_model}$ denote the transformed input of the j-th variable at time t, and $\mathcal{X}_t = [\, (\xi_t^1)^T, \ldots, (\xi_t^{m_\chi})^T \,]^T$ represent the concatenated vector of all past inputs. Variable selection weights are computed by passing $\mathcal{X}_t$ and an optional context vector c_s from static covariate encoder through GRN followed by a Softmax layer:

$$v_\chi^t = \mathrm{Softmax}(GRN_{v_\chi}(\mathcal{X}_t, c_s))$$

where $v_\chi^t \in \mathbb{R}^{m_\chi}$ contains the selection weights for all m_χ variables at time t. For static variables, the context vector c_s is omitted. Each input ξ_t^j is further refined using the corresponding GRN, where weights of GRN_{ξ}^{j} are shared across all time steps and ξ denotes weighted representation of past inputs. The final weighted representation of past inputs is obtained as:

$$\xi_t = \Sigma_{j=1}^{m_\chi}\, v_\chi^{t,j} \cdot GRN_{\xi}^{j}(\xi_t^j)$$

2.1.3. Sequence-to-Sequence LSTM

To capture the temporal dependencies for seismicity forecasting, TFT integrates sequence-to-sequence LSTM layers with interpretable multi-head attention mechanisms. The sequence-to-sequence LSTM layers process past observations and known future inputs separately to learn localized temporal patterns while initializing hidden states derived from static covariates. This design ensures that static constraints and long-term operational trends are embedded into the temporal dynamics.

2.1.4. Interpretable Multi-Head Self-attention

The interpretable multi-head self-attention decoder enhances the model's ability to identify long-term dependencies by emphasizing significant time steps and feature interactions. It extends the transformer-based attention mechanism, where standard self-attention computes weighted representations of input values $V \in \mathbb{R}^{N \times d_V}$:

$$\text{Attention}(Q, K, V) = \text{Softmax}(QK^{T} / \sqrt{d_{attn}}) \cdot V$$

Multi-head attention extends this by projecting queries, keys, and values into multiple subspaces, enabling the learning of diverse temporal patterns. To improve interpretability, we constrain multi-head attention to share values V across all heads and aggregate outputs additively:

$$\text{MultiHead}(Q, K, V) = \bar{A}(Q, K) \cdot V \cdot W_V$$

where the unified attention weights are given by:

$$\bar{A}(Q, K) = (1 / m_H) \cdot \Sigma_{h=1}^{m_H} A(QW_Q^{(h)}, KW_K^{(h)})$$

This formulation allows each head to learn distinct temporal patterns while attending to a common feature space. The shared attention weights $\bar{A}(Q, K)$ provide a unified representation that enhances the analysis of significant temporal dependencies and seismicity drivers.

2.2. Uncertainty Modeling and Quantile Foresting

To quantify the uncertainty in seismicity forecasting, our model produces probabilistic predictions by generating multiple quantile forecasts for each time step. The TFT is trained by jointly minimizing the quantile loss, aggregated across all quantile outputs:

$$L(\Omega, \theta) = \Sigma_{y_t \in \Omega} \Sigma_{q \in Q} (1 / N\tau_{max}) \Sigma_{\tau=1}^{\tau_{max}} \mathcal{Q}(y_t, \hat{y}(q, t-\tau, \tau), q)$$

where

$$\mathcal{Q}(y, \hat{y}, q) = q \cdot (y - \hat{y})_{+} + (1 - q) \cdot (\hat{y} - y)_{+}, \quad \text{and } (x)_{+} = \max(0, x)$$

In this formulation, Ω represents the training data domain containing N samples, θ denotes the learnable weights of TFT, and Q is the set of output quantiles (we use $Q = \{0.1, 0.5, 0.9\}$ in our experiments). The loss is computed over the forecasting horizon τ_{max}.

For out-of-sample testing, we evaluate the normalized quantile losses across the entire forecasting horizon, focusing on P_{50} and P_{90} risk metrics:

$$q\_risk = 2 \cdot [\Sigma_{y_t \in \tilde{\Omega}} \Sigma_{\tau=1}^{\tau_{max}} \mathcal{Q}(y_t, \hat{y}(q, t-\tau, \tau), q)] / [\Sigma_{y_t \in \tilde{\Omega}} \Sigma_{\tau=1}^{\tau_{max}} |y_t|]$$

where $\tilde{\Omega}$ represents the test data domain. This probabilistic method enables informed risk assessment by providing a distribution of potential seismicity under varying operational conditions.

3. RESULT

In this section, two case studies from Geysers and Utah FORGE geothermal fields demonstrate the model's ability to capture relationships between fluid-injection and seismic response to evaluate earthquake hazards.

3.1. Application in Geysers Dataset

This section focuses on induced seismicity forecasting in the northwestern region of the Geysers geothermal field, using high-resolution seismic and injection data from two wells, Prati-9 and Prati-29 (Figure 2a). The Geysers is the world's largest geothermal field and has been operational since the 1960s. It generates approximately 850 MW of electricity annually through steam production driven by extensive water injection (Martınez-Garzon et al., 2014; Kwiatek et al., 2015). Over 70 injection wells reaching depths of up to 5 km are used to enhance steam production and mitigate reservoir pressure depletion. The vapor-dominated reservoir consists of fractured, hydrothermally altered greywacke sandstone with low porosity and temperatures exceeding 350°C below 2.75 km in the northwestern part of the field.

Fluid injection in this region is gravity-driven, resulting in a negative gauge pressure at the wellhead after permeability enhancement due to thermal contraction, in contrast to the high-pressure methods typically used for EGS stimulation (Garcia et al., 2016; Rutqvist et al., 2016). Induced seismicity in the region is primarily linked to water injection rather than steam production and is attributed to thermal stresses caused by subsurface rock expansion and contraction. The Geysers provides a unique environment for advancing methods to model the interplay between external forcing and seismic responses, with implications for geothermal energy development and seismic hazard mitigation.

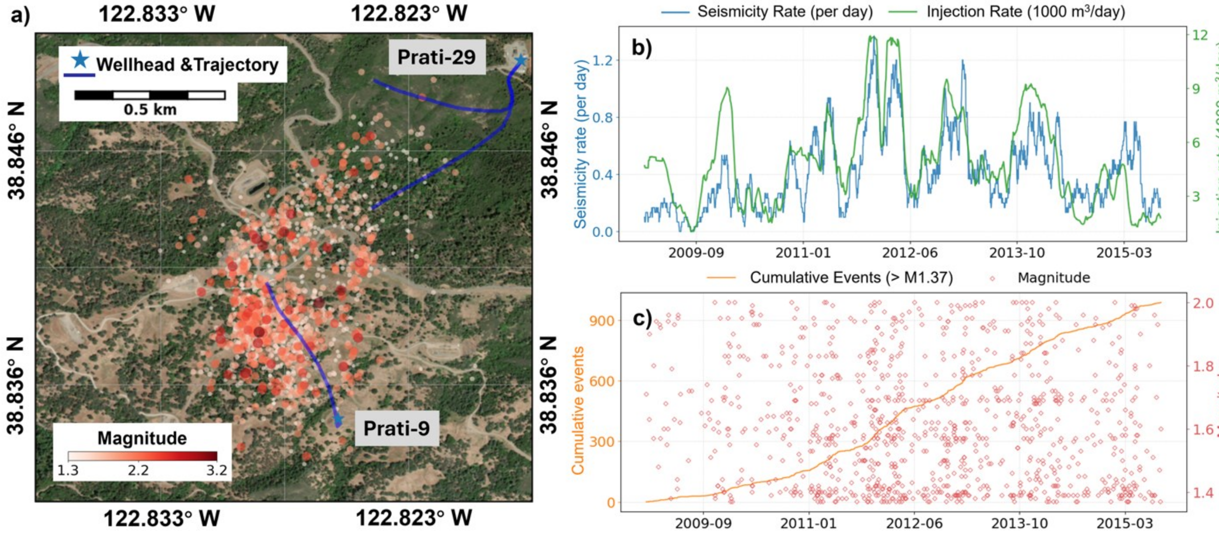

Figure 2: Seismicity and fluid injection data characteristics at the Geysers geothermal field. (a) Spatial distribution of seismic events recorded from January 2009 to August 2015, along with the locations of the Prati-9 and Prati-29 injection wells; (b) Temporal variations in seismicity rates and injection rates for Prati-9 and Prati-29 indicate their seasonal fluctuations due to the energy demands; (c) Seismic magnitude distribution and cumulative events over time. The minimum magnitude is 1.37.

3.1.1. Fluid Injection-Induced Seismicity Dynamics

The earthquake catalog (Martínez-Garzón et al., 2014; Leptokaropoulos et al., 2018) spans from January 2009 to August 2015 and includes 1,254 seismic events. These occurred up to a maximum distance of 600 meters from the open hole of the Prati-9 injection well (Figure 2b). The events were recorded by the Berkeley-Geysers (BG) seismic network, which consists of 31 three-component short-period geophones. Only events with magnitudes Mw ≥ 1.37 are considered to ensure consistency with the magnitude of completeness.

Injection at Prati-9 began in November 2007 and continued for approximately 18 years, while injection at Prati-29 started in April 2010 and lasted for 3 years. Continuous fluid injection at Prati-9 totals 7.21 million cubic meters (Mm^3), while Prati-29 injected 3.29 Mm^3. Both wells exhibited average injection rates of 8.79×10^4 m^3/month (Kwiatek et al., 2015).

The seismicity cloud surrounding Prati-9 (Figure 2c) shows a clear temporal correlation with injection activities, with seismicity pulsating in response to changes in injection rates. The average distance of seismic events from Prati-9 is 376 meters, while only sparse activity (approximately 30 events) occurs near Prati-29, at an average distance of 800 meters. Due to the limited seismicity near Prati-29, detailed analyses focus on events near Prati-9, although the influence of total injection rates from both wells is considered.

3.1.2. Data Conditioning for Time-Series Modeling

Seismicity rates are derived by counting earthquake events within a uniform daily time grid, while fluid injection rates are interpolated onto this grid to ensure temporal alignment. To reduce short-term fluctuations, a moving average with a one-month window is applied to smooth both time series. Given the large magnitude difference between seismicity and injection data, each variable is normalized separately using z-score normalization, while preserving the mean and standard deviation for subsequent de-normalization.

The training dataset spans from 2009 to 2014, while the validation dataset includes data from 2014 onward. Time series from both sets are divided into 480-day windows with a sliding step of two days, yielding 1,177 samples—approximately 80% of which are used for training and the remaining for validation.

Each segment comprises 320 days of historical seismicity rate and the entire injection rate with its gradient as inputs, and 160 days as the prediction target. This setup accounts for potential temporal lags in seismic responses to fluid injection. To improve model generalization and capture diverse temporal dependencies, variability is introduced into the input-output configurations. The input window length is randomly varied between 160 and 320 days, while the forecast horizon ranges from 80 to 160 days, enabling multi-horizon predictions.

Additionally, the mean and standard deviation of historical seismicity rates, along with the input window length, are incorporated as static covariates to provide contextual information for each time series segment.

3.1.3. Training and Forecast Validation

The model is trained to forecast seismicity rates by capturing complex temporal dependencies and uncertainty distributions. The optimization process employs an adaptive learning rate strategy that dynamically adjusts the learning rate based on model performance to enhance convergence. Using a reduction-on-plateau mechanism, the learning rate is initialized at 0.03 and decreases when the model fails to improve within four consecutive evaluations. Gradient clipping at 0.1 is applied to stabilize training by mitigating extreme gradient values. The model is trained for up to 100 epochs but incorporates early stopping with a patience of 10 epochs to avoid unnecessary computation once convergence is reached.

Hyperparameter optimization is conducted to improve model performance while maintaining computational efficiency. The search space includes hidden dimensions ranging from 8 to 128, attention head sizes from 1 to 4, and dropout rates between 0.1 and 0.3. The learning rate is tuned within the range of 0.01 to 0.1. The tuning process, conducted over 200 trials with a maximum of 60 training epochs per trial, evaluates model configurations to identify an optimal balance between learning speed and training stability. Gradient clipping values are also varied between 0.01 and 1.0 to determine the most effective stabilization level. The final selection of hyperparameters is guided by empirical performance metrics and overall stability, ensuring that the trained model generalizes well to unseen seismicity patterns.

To evaluate model performance, forecasts are generated using a validation set with a progressively expanding time window and a fixed 160-day forecasting window. Although this 160-day horizon is used for evaluation, the model is trained with a multi-horizon framework that allows flexible forecasting across varying lead times. The results (Figure 3) show that the model effectively captures earthquake trends and temporal dynamics, with predicted seismicity rates (orange) closely tracking observed rates (blue).

By incorporating uncertainty quantification, the forecasts provide a robust basis for seismic hazard assessment, enhancing the model's applicability in risk evaluation. The model successfully reconstructs the smoothed evolution of seismicity rates, capturing both long-term trends and smaller-scale fluctuations. Predictive uncertainty is expressed using the 5th and 95th percentiles of the forecast distribution, forming an interval (shown as an orange band) that contains most of the local peaks and troughs. This indicates that the model captures the variability of seismic response effectively.

However, some deviations occur during periods of rapid seismicity changes, highlighting the challenge of modeling abrupt fluctuations. Notably, the model predicts the elevated seismicity observed around April 2015, despite concurrently low injection rates. This behavior suggests the model captures delayed or memory effects, where seismic activity reflects both immediate stimulation and accumulated stress perturbations from prior injections. Nevertheless, forecast accuracy declines during sudden and isolated episodes that are not clearly associated with external drivers.

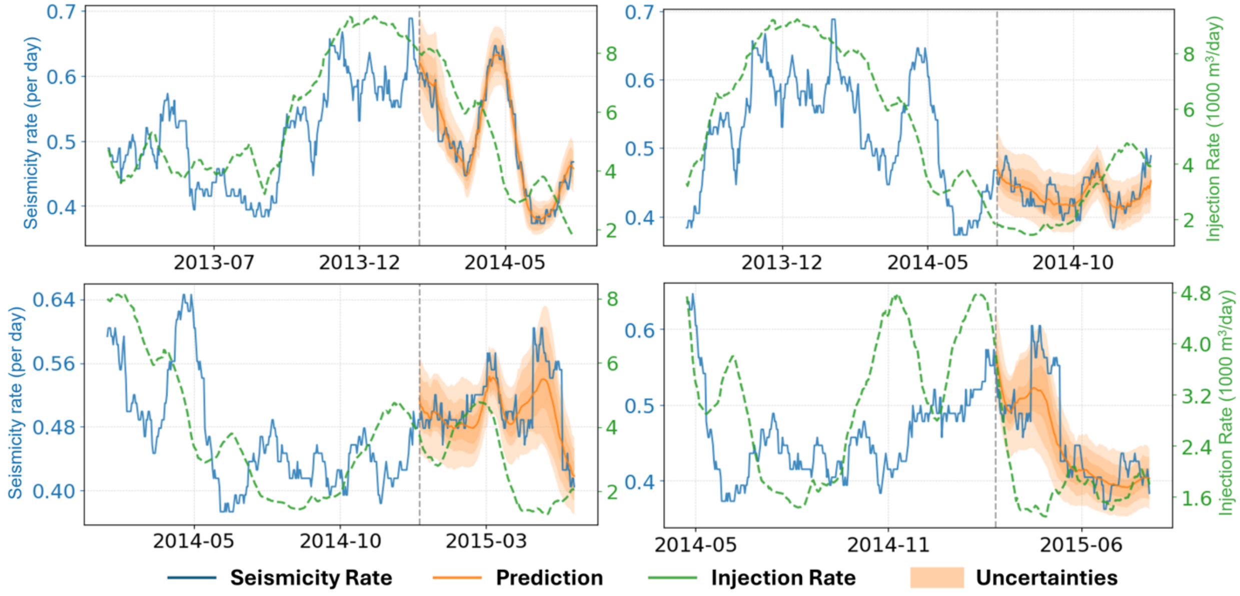

Figure 3: Comparison of observed seismicity rates (blue) and injection rates (green) with model predictions (orange) and uncertainty bounds.

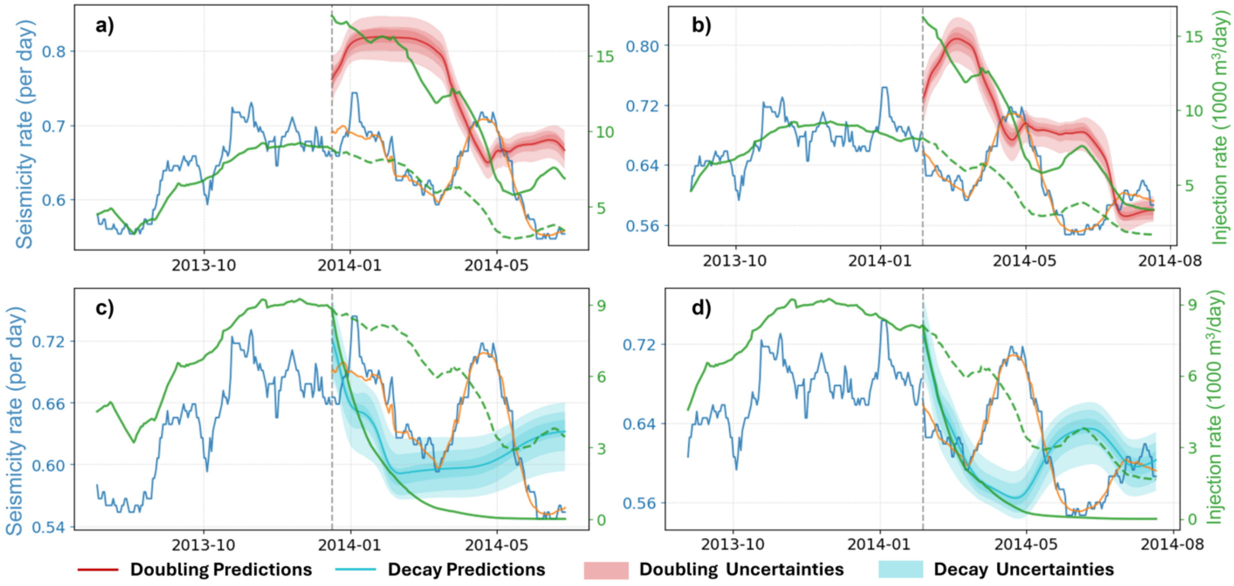

Figure 4: Modeled seismicity responses to varying injection rate scenarios in the Geysers geothermal field. (a) and (b) Doubling the injection rate increases seismic activity. (c) and (d) Reducing the injection rate to zero results in a smooth decline in seismicity. Solid blue lines represent observed seismicity rates, solid green lines denote injection rates, and dashed green lines indicate predicted trends. Shaded regions show predictive uncertainty.

3.1.4. Seismic Response to Fluid Injection

To further assess the robustness of our approach, we analyze scenarios with various operational conditions in the Geysers validation set, including doubling and halving injection rates, as well as implementing shut-in operations. Amplifying injection rates leads to elevated seismicity, reflecting the potential impact of increased fluid volumes and enabling proactive hazard assessment (Figure 4a and 4b). Conversely, a gradual decay of injection rates to zero (Figure 4c and 4d) reveals a smooth reduction in seismic response, consistent with reservoir stabilization processes.

These experiments reveal trends that closely align with physical expectations, offering meaningful interpretations of pore pressure diffusion and its impact on seismicity. The results demonstrate the model's capacity to quantify how operational interventions influence seismic hazards over time, and further provide insights for optimizing injection strategies.

By enabling the simulation of hypothetical scenarios, our framework serves as a valuable resource to help operators and stakeholders evaluate the impact of operational modifications and manage seismic risks effectively.

3.2. Application in Utah FORGE Dataset

This section focuses on induced seismicity forecasting at the Utah FORGE geothermal site, utilizing high-resolution seismic and injection data collected during the 2024 hydraulic stimulation of wells 16A(78)-32 and 16B(78)-32 (Figure 5a). Utah FORGE is a dedicated field laboratory for Enhanced Geothermal Systems (EGS) research, targeting a low-permeability granitic basement at approximately 2.5 km depth. The objective is to create a sustainable heat-exchange reservoir by enhancing permeability through hydraulic stimulation.

Unlike the Geysers geothermal field, where fluid injection is gravity-driven, Utah FORGE employs high-pressure hydraulic fracturing methods to control fracture growth and connectivity. The interplay between injection rate, pressure, and induced seismicity at this site provides a valuable benchmark for advancing forecasting approaches and improving our understanding of fluid-driven seismicity mechanisms.

3.2.1. Induced Seismicity and Injection Characteristics

The 2024 hydraulic stimulation at Utah FORGE spanned 10 stages, during which seismic events were continuously monitored to assess the reservoir's response to injection activities. The earthquake catalog, sourced from the Geothermal Data Repository (GDR) for the Utah FORGE project, includes over 2,900 induced seismic events, with magnitudes ranging predominantly from Mw –1.0 to 0.0, and a maximum recorded event of Mw 2.0.

Fluid injection was performed in stages at wells 16A(78)-32 and 16B(78)-32, involving the injection of approximately 118,000 barrels (4.96 million gallons) of water over a two-week period. The injection rate varied between 35 and 50 barrels per minute, with corresponding treating pressures reaching up to 50 MPa (7,000 psi) (Figure 5). The inclusion of treating pressure data provides additional insight into stress evolution within the reservoir and its impact on induced seismicity.

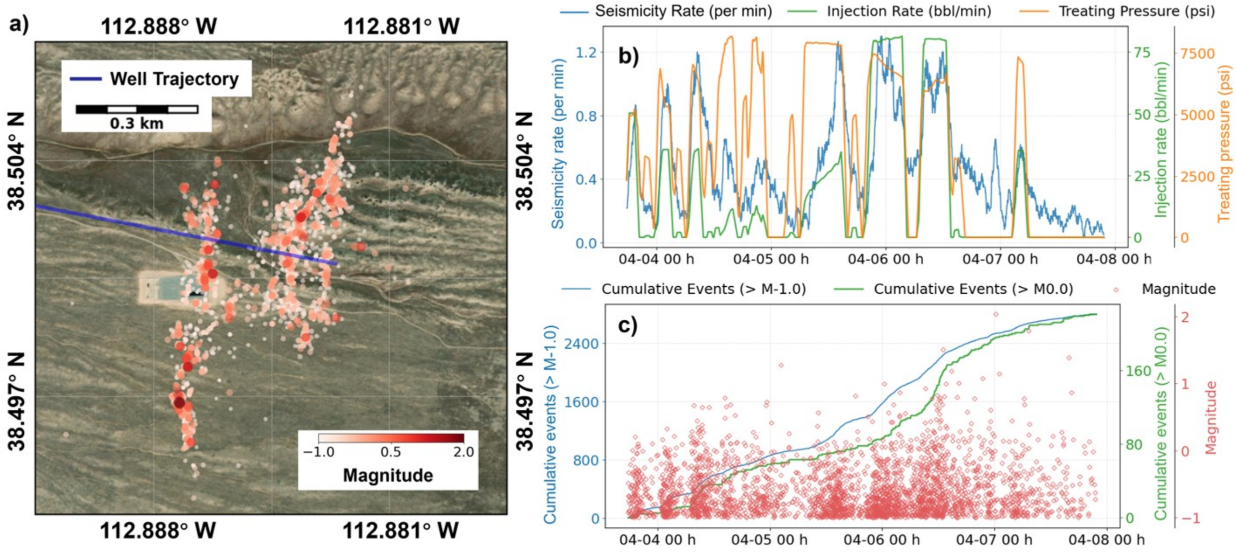

Figure 5: Seismicity and Fluid Injection Characteristics at the Utah FORGE Geothermal Site (a) Spatial distribution of induced seismic events recorded during the 2024 hydraulic stimulation phases, shown alongside the locations of the 16A(78)-32 and 16B(78)-32 injection wells. (b) Temporal variations in seismicity rates, injection rates, and treating pressures, illustrating their correlations and how changes in injection conditions align with seismic responses. (c) Seismic magnitude distribution and cumulative number of events over time. While the magnitudes in the Geysers dataset are mostly concentrated above 0, the magnitudes in the Utah FORGE dataset range between 0 and -1.

Seismicity was primarily concentrated around the injection wells, aligning with the extent of the stimulated fracture networks (Figure 5a). The spatial extent of seismicity varied across injection stages, with most events occurring within a few hundred meters of the wellbore. The seismicity cloud extended preferentially along the direction of the maximum horizontal stress (NNE–SSW), consistent with expected fracture orientations in the granitic host rock (Xing et al., 2020).

Temporal trends indicate a strong correlation between seismic activity and injection parameters, with increased seismicity observed during high-rate injection phases and periods of pressure buildup (Figure 5b). Notably, seismic events continued to occur after injection ceased, indicating a delayed response of the reservoir due to fluid diffusion and stress redistribution.

3.2.2. Data Preprocessing

Seismicity rates are calculated by counting events within a uniform time grid, while injection rates and treating pressures are interpolated onto this grid to ensure temporal alignment. To suppress short-term fluctuations in seismicity rates, injection rates, and treating pressures, a moving average with a 2-minute window is applied to smooth all time series. Each variable is normalized using z-score normalization, preserving the mean and standard deviation for subsequent de-normalization.

The dataset is divided into training and validation sets. Time series from both sets are segmented into 16-hour windows with a sliding step of two days, yielding 1,177 samples—approximately 70% of which are used for training, and the remainder for validation.

Each segment includes 12 hours of historical seismicity rate data, combined with full injection rate and treating pressure values, as well as their corresponding gradients, used as inputs. The subsequent 4 hours serve as the prediction target. To enhance model generalization and capture diverse temporal dependencies, input-output configurations are varied. The input window length is randomly varied between 6 and 12 hours, while the forecast horizon ranges from 2 to 4 hours.

Additionally, the mean and standard deviation of the historical seismicity rates, along with the input window length, are incorporated as static covariates to provide context for each segment.

3.2.3. Forecasting Seismic Activity

The optimization process employs an adaptive learning rate strategy, starting at 0.03 and reducing when model improvement stagnates over four consecutive evaluations. Gradient clipping at 0.1 is used to stabilize training by mitigating the effect of extreme gradient values. The model is trained for up to 100 epochs, with early stopping implemented to prevent redundant computations once convergence is achieved. As with the application to the Geysers dataset, hyperparameter tuning is used to enhance performance while maintaining computational efficiency.

Model performance is evaluated using the validation set with a progressively expanding time window and a fixed 4-hour forecasting window. The results (Figure 6) demonstrate that the model successfully captures earthquake trends and temporal dynamics, with predicted seismicity rates (orange) closely aligning with observed rates (blue). Incorporating uncertainty quantification further enhances the model's robustness for seismic hazard assessment, reinforcing its applicability for risk evaluation.

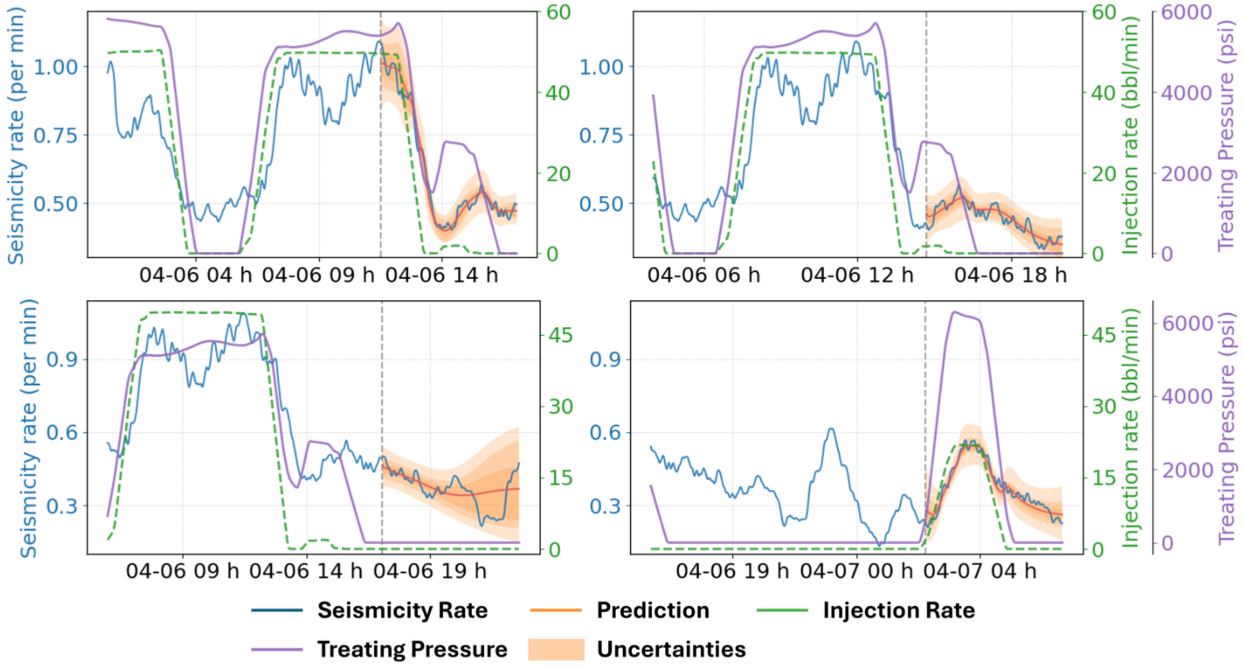

Figure 6: Comparison of observed seismicity rates (blue), injection rates (green), treating pressures (red), and model predictions (orange) with uncertainty bounds. The purple line indicates the treating pressure.

4. DISCUSSION

Having established the high performance of the model, we now discuss how it enables analysis of individual components of the input data to interpret the general relationships it has learned. This analysis is guided by measures of uncertainty and variable importance. We also demonstrate the model's ability to simulate seismicity rate responses under different fluid injection scenarios, allowing operators to test and optimize injection strategies for stimulation and/or production. Finally, we compare the results of our model with those of ETAS-based models.

4.1. Iterative Long-Horizon Forecasting

To extend the temporal coverage beyond a fixed forecasting horizon, we implement an iterative strategy in which the predicted seismicity rate is recursively used as input for subsequent predictions. The entire testing set from the Geysers dataset is divided into sequential, overlapping segments, each comprising 320 days of historical data and predicting the next 160 days. After each forecast, the output is concatenated with a portion of the previous data to form the new context for the next prediction step. This recursive process enables long-term forecasting across the entire dataset while preserving temporal consistency through smooth connections between overlapping time windows.

Our model captures not only short-term dependencies but also demonstrates robustness under evolving conditions (Figure 7). Even when trends in seismicity rate diverge from fluid injection parameters, the model maintains predictive accuracy with low uncertainty, as indicated by the orange shaded region. This suggests that the network has learned the triggering patterns and memory effects associated with injection-induced seismicity, rather than merely extrapolating from recent trends. The forecasts reflect sensitivity to earlier variations in injection history, even when those variations are temporally distant from the prediction window.

To evaluate forecast reliability over time, we compute cumulative uncertainty bounds (shown as the gray shaded region) using an exponentially decaying kernel applied to the averaged uncertainty (orange line). The cumulative upper and lower bounds are calculated as weighted sums of deviations from the mean prediction. In the early stage of forecasting, cumulative uncertainty grows slowly, reflecting stable predictions. However, beyond approximately eight months into the forecast horizon, the accumulation error increases more rapidly. This phase coincides with a period in which seismicity rates no longer co-vary with injection rates or their gradients, indicating a transition in temporal dynamics linked to evolving reservoir behavior in the geothermal system. This result highlights the model's sensitivity to underlying physical processes.

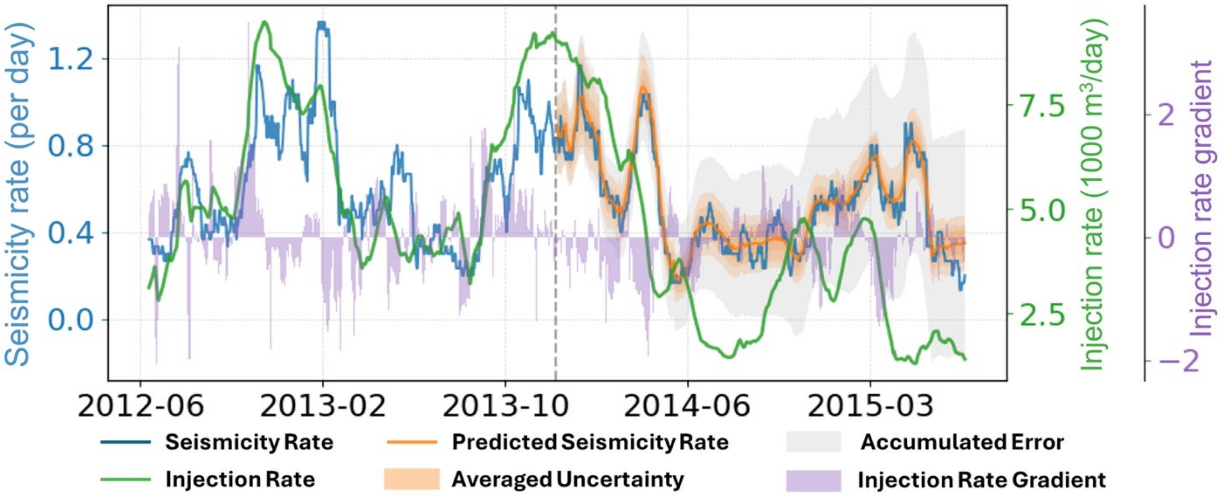

Figure 7: Iterative forecasting of seismicity rate in the Geysers geothermal field. The model is recursively applied across overlapping time windows, using previously estimated seismicity as input for subsequent forecasting steps.

4.2. Interpretability of Attention Weights

To explain the model's decision-making process, we quantify the relative contributions of input variables using percentiles of attention weights aggregated across the entire dataset. Percentiles are computed along the temporal axis for each input variable, and their average values are normalized to yield an importance score (Table 1).

In the Geysers dataset, which involves long-term forecasting, the model assigns dominant attention to the historical injection gradient (45.5%) and past seismicity rate (46.5%). This highlights the cumulative influence of historical injection dynamics and seismic feedback mechanisms. Notably, the scheduled injection gradient also ranks highly (85.0%), indicating the model's reliance on anticipated trends in injection variability. The high importance of the injection gradient suggests that the system is more sensitive to changes in injection rate than to the total volume injected. We speculate that this is due to the high permeability of the Geysers reservoir, which allows fluid to propagate rapidly via pressure diffusion, thereby stimulating faults under near-critical conditions.

In contrast, the Utah FORGE dataset reveals a different pattern, shaped by its shorter forecasting window and injection schedule. The future injection rate (46.7%) and treating pressure (32.9%) emerge as the most significant features, suggesting that induced seismicity is more directly influenced by ongoing operational conditions. The relatively low contribution from past seismicity rate (8.8%) indicates that short-term induced responses are less influenced by historical seismic activity.

Notable differences between the Utah FORGE and Geysers results—such as the high importance of injection gradient at the Geysers and injection rate at Utah FORGE—may reflect contrasting reservoir permeabilities. At the Geysers, pressure gradients may propagate more broadly and with higher amplitudes due to greater permeability, while at Utah FORGE, a certain injection volume may be needed to generate significant pressure changes at fault planes. In this scenario, pressure gradients resulting from changes in injection rate could dissipate more quickly at Utah FORGE. An alternative interpretation is that treating pressure and injection gradient have similar physical effects in the model, and their importance is divided across these correlated features. Further analysis accounting for feature covariance is needed to test this hypothesis. Overall, this analysis provides insight into how geological and operational regimes shape the key drivers of induced seismicity.

To understand how past events influence future variations, we analyze persistent temporal patterns that reveal the underlying mechanisms governing seismicity evolution. These patterns are identified through self-attention weights $\alpha(t, p, \tau)$, which quantify the influence of past features p on forecasts at horizon τ. Distributions of $\alpha(t, p, \tau)$ across time steps indicate the relevance of past observations.

For one-step-ahead forecasts, attention distributions across quantiles (10th, 50th, and 90th percentiles) highlight memory effects at each site. In the Geysers dataset (Figure 8a), past observations from approximately 160 days prior strongly influence seismicity forecasts, indicating a delayed response to injection variations. The rising attention weights near the present reflect the increasing importance of recent seismic and injection history. In contrast, the Utah FORGE dataset (Figure 8b) shows a more immediate response, with attention weights peaking close to the forecasting time and minimal long-term memory.

Figures 8c and 8d show how past data influence predictions across timescales for various forecast horizons ($\tau \in \{0, 30, 60, 90, 120\}$). In the Geysers dataset, consistently high attention around a 180-day lag across all horizons reinforces the presence of delayed seismic response. Shorter horizons rely primarily on recent injection changes, while longer horizons incorporate a broader range of seismic and injection history. This pattern suggests persistent memory effects driven by poroelastic and viscoelastic stress redistribution, where induced earthquakes follow delayed responses modulated by subsurface rheology.

In contrast, the Utah FORGE dataset exhibits a localized attention distribution, with weights peaking near the forecasting time for all horizons. This behavior suggests that seismicity is primarily dependent on immediate stress perturbations, with limited influence from earlier injection patterns beyond a few hours. This contrast highlights site-specific differences in the timescales of induced seismicity, driven by geological and hydromechanical conditions.

Table 1: Feature importance analysis for geothermal field data.

Category	Feature	Dataset	
		Geysers	FORGE
Past Inputs	Injection Rate	8.0	42.6
	Injection Gradient	45.5	11.9
	Treating Pressure	–	30.0
	Pressure Gradient	–	6.7
	Past Seismicity Rate	46.5	8.8
Future Inputs	Injection Rate	15.0	46.7
	Injection Gradient	85.0	13.0
	Treating Pressure	–	32.9
	Pressure Gradient	–	7.4

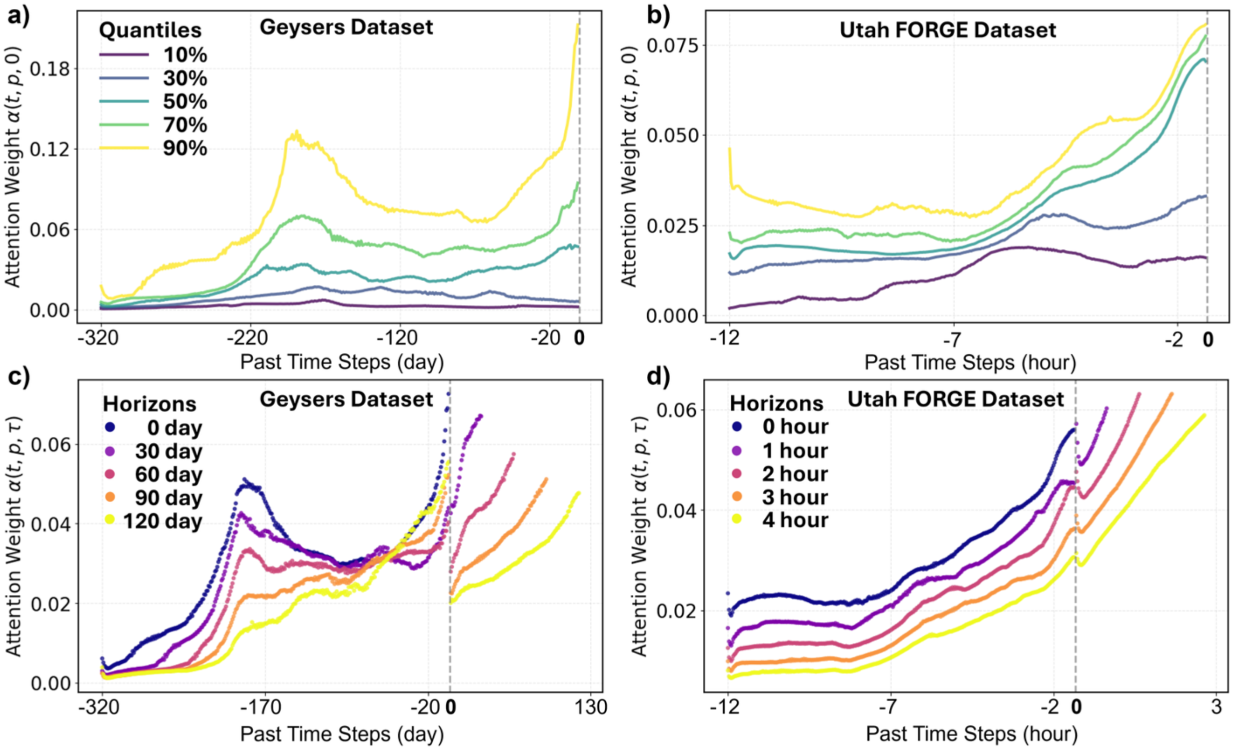

Figure 8: Feature importance and temporal attention patterns in seismicity forecasting. Percentiles of attention weights α(t, p, τ = 0) for one-step-ahead forecasting in the Geysers (a) and Utah FORGE (b) datasets, respectively. The attention distribution highlights varying levels of predictive significance across past time steps. Mean attention weights across different forecasting horizons in the Geysers (c) and Utah FORGE (d) datasets illustrate the evolving impact of historical information on predictions. The dashed vertical line denotes the current forecasting step.

4.3. Impact of Input Variables on Model Performance

We evaluate how input variables affect model performance on the Utah FORGE validation dataset (Table 2). Across all six metrics—mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), dynamic time warping (DTW), Pearson distance, and cosine distance—the model achieves its best performance when incorporating the full set of input variables. The inclusion of temporal gradients consistently reduces prediction errors and improves similarity measures.

While the injection rate represents the cumulative volume of fluid introduced into the reservoir, its temporal variation serves as a more sensitive indicator of seismic evolution. Abrupt pressure fluctuations can rapidly alter the stress state of faults, triggering seismic slip on critically stressed patches before cumulative fluid volumes reach levels sufficient to destabilize a wider region.

The importance of treating pressure further reinforces this interpretation, as it governs the effective stress acting on fractures. However, treating pressure alone yields inferior performance compared to injection rate, suggesting that pressure diffusion alone may not fully explain the spatiotemporal triggering patterns without accounting for dynamic injection behavior.

Thus, while each input variable contains partial information about the reservoir's evolving stress state, their integration enables the model to capture the complex nonlinear interactions that drive induced seismicity. This analysis also validates the model's ability to leverage multi-source temporal information and extract physically meaningful insights into the mechanisms of fluid-triggered seismicity.

Table 2: Quantitative evaluation metrics from our model applied to the Utah FORGE validation dataset using different combinations of input variables. The lowest value for each metric is highlighted in bold. Input abbreviations are as follows: I refers to injection rate, P to treating pressure, and G to the temporal gradients of injection rate and pressure.

Input Variables	MSE	RMSE	MAE	DTW Distance	Pearson Dist.	Cosine Dist.
Seismicity only	0.1239	0.3414	0.3015	3.8273	0.7237	0.4212
Seismicity + I	0.0242	0.1389	0.1038	1.3189	0.4049	0.0358
Seismicity + P	0.0287	0.1649	0.1442	1.0594	0.6566	0.0688
Seismicity + IP	0.0145	0.1168	0.0944	0.7098	0.5239	0.0516
Seismicity + IPG	**0.0018**	**0.0386**	**0.0306**	**0.3607**	**0.2157**	**0.0110**

4.4. Comparison with Physics-Informed Stochastic Models

We compare our model's forecasts with three physics-informed stochastic models in the Geysers validation set, including the classical ETAS model (brown), ETAS-Frac (purple), and ETAS-Conv (red), as described in Shcherbakov (2024) (Figure 9). All ETAS-based models are trained using seismicity and fluid injection data from the forecasting time back to the start of the dataset. The observed seismicity rate (blue) and injection rate (green) serve as benchmarks for evaluating forecast accuracy, while the ETAS-based predictions illustrate differences in model performance.

The classical ETAS model describes seismicity as a self-exciting point process driven solely by past earthquake events, without consideration of fluid injection. ETAS-Frac incorporates injection effects through a nonlinear log-linear term that links seismicity to the contemporaneous injection rate. ETAS-Conv introduces a convolutional kernel to model delayed and smoothed influences of injection, capturing reservoir memory effects. However, these models struggle to represent the nonlinear dynamics of earthquake triggering, resulting in significant deviations from observed seismicity rates in long-term forecasts beyond 40 days.

In contrast, our model (orange) effectively captures nonlinear dependencies between injection operations and seismicity evolution by learning directly from historical patterns without relying on explicit physical assumptions. Its uncertainty bounds approximately encompass the predictions of stochastic models near the forecasting time, suggesting robust generalization across varying injection regimes and stress redistribution conditions.

These findings underscore the advantages of deep learning in forecasting injection-induced seismicity and demonstrate its potential to outperform traditional physics-informed approaches.

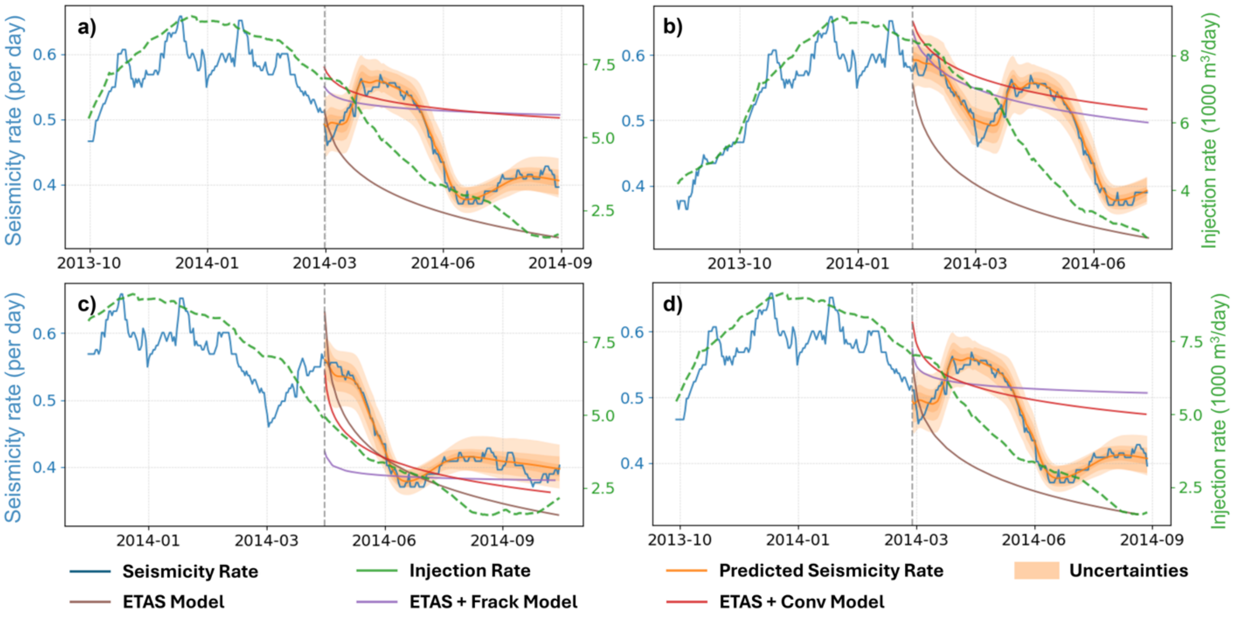

Figure 9: Comparison of model forecasts with physics-informed stochastic models using the Geysers dataset. Observed seismicity (blue) and injection rates (green) serve as benchmarks. Predictions from ETAS (brown), ETAS-Frac (purple), and ETAS-Conv (red) capture delayed seismic responses but deviate in long-term forecasts. In contrast, our model (orange) learns nonlinear dependencies without explicit physical constraints, demonstrating improved generalization.

4.5. Challenges and Future Directions

Despite advances in deep learning-based seismicity forecasting, several limitations remain. First, model performance is influenced by the coupling between fluid injection and seismicity rate, which varies across different geological settings. In this study, both the Geysers and Utah FORGE datasets exhibit relatively strong correlations between injection rates and seismicity, enabling the model to forecast seismic activity in a statistical sense. However, predicting individual high-magnitude events that occur after injection has ceased remains a challenge. Improving forecasts for such events will require more accurate modeling of delayed triggering mechanisms and post-injection stress evolution.

Additionally, prediction becomes increasingly difficult when seismicity exhibits irregular fluctuations, lacks clear periodic patterns, or is heavily influenced by external factors not included in the model inputs. Second, although the model captures nonlinear dependencies among multiple variables, the underlying causal mechanisms remain unclear. Integrating geomechanical models and three-dimensional fault structures could bridge empirical observations with physical principles, providing mechanistic insights into stress redistribution and fault reactivation.

Finally, the current framework focuses on temporal trends and lacks spatial resolution. Coupling the model with spatially aware architectures could enable localized risk assessments by partitioning seismic zones using geological priors. Addressing these challenges would further enhance the model’s applicability for managing injection-induced seismicity.

5. CONCLUSIONS

This study introduces a deep learning framework for forecasting induced seismicity in geothermal fields, leveraging the Temporal Fusion Transformer (TFT) to integrate heterogeneous time series data and extract key drivers of seismicity evolution. Applied to datasets from the Geysers and Utah FORGE, the model captures nonlinear dependencies between fluid injection and seismicity while providing probabilistic forecasts with quantified uncertainty.

Feature importance analysis and interpretation of attention weights reveal persistent temporal patterns, highlighting site-specific seismic responses to operational perturbations. Comparative analysis with physics-informed stochastic models, such as ETAS-Frac and ETAS-Conv, demonstrates superior generalization and predictive accuracy, underscoring the benefits of data-driven approaches in seismic hazard assessment.

Beyond forecasting, scenario-based simulations validate the model’s capacity to quantify seismic responses to operational changes, offering actionable insights for seismic risk mitigation. However, challenges remain in linking empirical predictions to underlying geophysical processes. Future research should incorporate geomechanical models and spatially adaptive frameworks to enhance physical interpretability and regional applicability.

By bridging data-driven forecasting with fundamental physics, this approach contributes to the sustainable development of geothermal energy while improving seismic hazard mitigation.

ACKNOWLEDGMENTS

Funding for this work was provided by the U.S. Department of Energy through Project DE-EE0007080: Enhanced Geothermal System Concept Testing and Development at the Milford City, Utah Frontier Observatory for Research in Geothermal Energy (Utah FORGE) Site. We are also grateful to Utah FORGE, Calpine, and GDR for maintaining the datasets used in this study. The authors additionally acknowledge the support of the Geothermal Data Repository, Core Service Anthropogenic Hazards, and the IS-EPOS Platform for providing access to seismicity-related data.

REFERENCES

Alghannam, M., & Juanes, R. (2020). Understanding rate effects in injection-induced earthquakes. Nature Communications, 11(1), 3053.

Aochi, H., Maury, J., & Le Guenan, T. (2021). How do statistical parameters of induced seismicity correlate with fluid injection? Case of Oklahoma. Seismological Society of America, 92(4), 2573–2590.

Atkinson, G. M., Eaton, D. W., Ghofrani, H., Walker, D., Cheadle, B., Schultz, R., ... others (2016). Hydraulic fracturing and seismicity in the Western Canada Sedimentary Basin. Seismological Research Letters, 87(3), 631–647.

Bachmann, C. E., Wiemer, S., Goertz-Allmann, B., & Woessner, J. (2012). Influence of pore-pressure on the event-size distribution of induced earthquakes. Geophysical Research Letters, 39(9).

Bachmann, C. E., Wiemer, S., Woessner, J., & Hainzl, S. (2011). Statistical analysis of the induced Basel 2006 earthquake sequence: Introducing a probability-based monitoring approach for enhanced geothermal systems. Geophysical Journal International, 186(2), 793–807.

Bao, X., & Eaton, D. W. (2016). Fault activation by hydraulic fracturing in Western Canada. Science, 354(6318), 1406–1409.

Bhattacharya, P., & Viesca, R. C. (2019). Fluid-induced aseismic fault slip outpaces pore-fluid migration. Science, 364(6439), 464–468.

Blank, D., & Morgan, J. (2021). Can deep learning predict complete ruptures in numerical megathrust faults? Geophysical Research Letters, 48(18), e2021GL092607.

Brodsky, E. E., & Lajoie, L. J. (2013). Anthropogenic seismicity rates and operational parameters at the Salton Sea geothermal field. Science, 341(6145), 543–546.

Chang, H., & Nakata, N. (2022). Investigation of time-lapse changes with DAS borehole data at the Brady geothermal field using deconvolution interferometry. Remote Sensing, 14(1), 185.

Chen, X., & Nakata, N. (2017). Preface to the focus section on the 3 September 2016 Pawnee, Oklahoma, earthquake. Seismological Research Letters, 88(4), 953–955.

Chen, X., Nakata, N., Pennington, C., Haffener, J., Chang, J. C., He, X., ... Walter, J. I. (2017). The Pawnee earthquake as a result of the interplay among injection, faults and foreshocks. Scientific Reports, 7(1), 4945.

Dempsey, D., & Riffault, J. (2019). Response of induced seismicity to injection rate reduction: Models of delay, decay, quiescence, recovery, and Oklahoma. Water Resources Research, 55(1), 656–681.

Deng, K., Liu, Y., & Harrington, R. M. (2016). Poroelastic stress triggering of the December 2013 Crooked Lake, Alberta, induced seismicity sequence. Geophysical Research Letters, 43(16), 8482–8491.

Ellsworth, W. L. (2013). Injection-induced earthquakes. Science, 341(6142), 1225942.

Garcia, J., Hartline, C., Walters, M., Wright, M., Rutqvist, J., Dobson, P. F., & Jeanne, P. (2016). The Northwest Geysers EGS demonstration project, California: Part 1: Characterization and reservoir response to injection. Geothermics, 63, 97–119.

Garcia-Aristizabal, A. (2018). Modelling fluid-induced seismicity rates associated with fluid injections: Examples related to fracture stimulations in geothermal areas. Geophysical Journal International, 215(1), 471–493.

Gaucher, E., Schoenball, M., Heidbach, O., Zang, A., Fokker, P. A., van Wees, J.-D., & Kohl, T. (2015). Induced seismicity in geothermal reservoirs: A review of forecasting approaches. Renewable and Sustainable Energy Reviews, 52, 1473–1490.

Goebel, T. H., & Brodsky, E. E. (2018). The spatial footprint of injection wells in a global compilation of induced earthquake sequences. Science, 361(6405), 899–904.

Grigoli, F., Cesca, S., Priolo, E., Rinaldi, A. P., Clinton, J. F., Stabile, T. A., ... Dahm, T. (2017). Current challenges in monitoring, discrimination, and management of induced seismicity related to underground industrial activities: A European perspective. Reviews of Geophysics, 55(2), 310–340.

Guglielmi, Y., Cappa, F., Avouac, J.-P., Henry, P., & Elsworth, D. (2015). Seismicity triggered by fluid injection–induced aseismic slip. Science, 348(6240), 1224–1226.

Hager, B. H., Dieterich, J., Frohlich, C., Juanes, R., Mantica, S., Shaw, J. H., ... others (2021). A process-based approach to understanding and managing triggered seismicity. Nature, 595(7869), 684–689.

Hainzl, S., & Ogata, Y. (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling. Journal of Geophysical Research: Solid Earth, 110(B5).

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. Science Advances, 4(5), eaao2929.

Improta, L., Valoroso, L., Piccinini, D., & Chiarabba, C. (2015). A detailed analysis of wastewater-induced seismicity in the Val d'Agri oil field (Italy). Geophysical Research Letters, 42(8), 2682–2690.

IS-EPOS Platform. (2022). Induced seismicity episode: The Geysers geothermal field, California. [Dataset]. IS-EPOS: Thematic Core Service Anthropogenic Hazards. https://doi.org/10.20366/ah-epos-2022.001

Jafari, A., Fox, G., Rundle, J. B., Donnellan, A., & Ludwig, L. G. (2024). Time series foundation models and deep learning architectures for earthquake temporal and spatial nowcasting. GeoHazards, 5(4), 1247–1274.

Kim, T., & Avouac, J.-P. (2023). Stress-based and convolutional forecasting of injection-induced seismicity: Application to the Otaniemi geothermal reservoir stimulation. Journal of Geophysical Research: Solid Earth, 128(4), e2022JB024960.

Kothari, S., Shcherbakov, R., & Atkinson, G. (2020). Statistical modeling and characterization of induced seismicity within the Western Canada Sedimentary Basin. Journal of Geophysical Research: Solid Earth, 125(12), e2020JB020606.

Kwiatek, G., Martínez-Garzón, P., Dresen, G., Bohnhoff, M., Sone, H., & Hartline, C. (2015). Effects of long-term fluid injection on induced seismicity parameters and maximum magnitude in northwestern part of the Geysers geothermal field. Journal of Geophysical Research: Solid Earth, 120(10), 7085–7101.

Langenbruch, C., Weingarten, M., & Zoback, M. D. (2018). Physics-based forecasting of man-made earthquake hazards in Oklahoma and Kansas. Nature Communications, 9(1), 3946.

Lee, K.-K., Ellsworth, W. L., Giardini, D., Townend, J., Ge, S., Shimamoto, T., ... others (2019). Managing injection-induced seismic risks. Science, 364(6442), 730–732.

Leptokaropoulos, K., Staszek, M., Lasocki, S., Martínez-Garzón, P., & Kwiatek, G. (2018). Evolution of seismicity in relation to fluid injection in the north-western part of the Geysers geothermal field. Geophysical Journal International, 212(2), 1157–1166.

Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748–1764.

Llenos, A. L., & Michael, A. J. (2016). Characterizing potentially induced earthquake rate changes in the Brawley seismic zone, Southern California. Bulletin of the Seismological Society of America, 106(5), 2045–2062.

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. Journal of Geophysical Research: Solid Earth, 119(11), 8378–8396.

McGarr, A. (2014). Maximum magnitude earthquakes induced by fluid injection. Journal of Geophysical Research: Solid Earth, 119(2), 1008–1019.

Norbeck, J., & Rubinstein, J. L. (2018). Hydromechanical earthquake nucleation model forecasts onset, peak, and falling rates of induced seismicity in Oklahoma and Kansas. Geophysical Research Letters, 45(7), 2963–2975.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association, 83(401), 9–27.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics, 50, 379–402.

Qin, Y., Chen, T., Ma, X., & Chen, X. (2022). Forecasting induced seismicity in Oklahoma using machine learning methods. Scientific Reports, 12(1), 9319.

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. Communications Earth & Environment, 3(1), 249.

Rutqvist, J., Jeanne, P., Dobson, P. F., Garcia, J., Hartline, C., Hutchings, L., ... Walters, M. (2016). The Northwest Geysers EGS demonstration project, California–Part 2: Modeling and interpretation. Geothermics, 63, 120–138.

Saad, O. M., Chen, Y., Savvaidis, A., Fomel, S., Jiang, X., Huang, D., ... others (2023). Earthquake forecasting using big data and artificial intelligence: A 30-week real-time case study in China. Bulletin of the Seismological Society of America, 113(6), 2461–2478.

Schultz, R., Skoumal, R. J., Brudzinski, M. R., Eaton, D., Baptie, B., & Ellsworth, W. (2020). Hydraulic fracturing-induced seismicity. Reviews of Geophysics, 58(3), e2019RG000695.

Sedghizadeh, M., van den Berghe, M., & Shcherbakov, R. (2024). Leveraging the ETAS model to forecast mining microseismicity. Geophysical Journal International, 238(3), 1491–1504.

Segall, P., & Lu, S. (2015). Injection-induced seismicity: Poroelastic and earthquake nucleation effects. Journal of Geophysical Research: Solid Earth, 120(7), 5082–5103.

Shapiro, S. A., & Dinske, C. (2009). Scaling of seismicity induced by nonlinear fluid-rock interaction. Journal of Geophysical Research: Solid Earth, 114(B9).

Shcherbakov, R. (2024). A stochastic model for induced seismicity at the geothermal systems: A case of the Geysers. Seismological Research Letters, 95(6), 3545–3556.

Smith, J. T., Sonnenthal, E. L., & Nakata, N. (2024). Thermal hydrological mechanical modelling of anticipated stimulation in Utah. Geothermal Reservoir Engineering, SGP-TR-227.

Spica, Z. J., Nakata, N., Liu, X., Campman, X., Tang, Z., & Beroza, G. C. (2018). The ambient seismic field at Groningen gas field: An overview from the surface to reservoir depth. Seismological Research Letters, 89(4), 1450–1466.

Spica, Z. J., Perton, M., Nakata, N., Liu, X., & Beroza, G. C. (2018). Site characterization at Groningen gas field area through joint surface-borehole H/V analysis. Geophysical Journal International, 212(1), 412–421.

Tester, J. W., Anderson, B., Batchelor, A., Blackwell, D., DiPippo, R., Drake, E., et al. (2006). The future of geothermal energy-impact of enhanced geothermal systems (EGS) on the United States in the 21st century: An assessment. Idaho Falls: Idaho National Laboratory, 1e8.

Utah FORGE. (2023). Utah FORGE: Geothermal field datasets. [Dataset]. U.S. Department of Energy Geothermal Data Repository. https://doi.org/10.15121/1452741

Vere-Jones, D. (2010). Foundations of statistical seismology. Pure and Applied Geophysics, 167, 645–653.

Weingarten, M., Ge, S., Godt, J. W., Bekins, B. A., & Rubinstein, J. L. (2015). High-rate injection is associated with the increase in US mid-continent seismicity. Science, 348(6241), 1336–1340.

White, M. C., & Nakata, N. (2025). Induced seismicity and geothermal energy production in the Salton Sea geothermal field, California. Scientific Reports, 15(1), 1–14.

Wynants-Morel, N., Cappa, F., De Barros, L., & Ampuero, J.-P. (2020). Stress perturbation from aseismic slip drives the seismic front during fluid injection in a permeable fault. Journal of Geophysical Research: Solid Earth, 125(7), e2019JB019179.

Xing, P., McLennan, J., & Moore, J. (2020). In-situ stress measurements at the Utah Frontier Observatory for Research in Geothermal Energy (FORGE) site. Energies, 13(21), 5842.

Yeo, I., Brown, M. R., Ge, S., & Lee, K. (2020). Causal mechanism of injection-induced earthquakes through the Mw 5.5 Pohang earthquake case study. Nature Communications, 11(1), 2614.

Zhai, G., Shirzaei, M., Manga, M., & Chen, X. (2019). Pore-pressure diffusion, enhanced by poroelastic stresses, controls induced seismicity in Oklahoma. Proceedings of the National Academy of Sciences, 116(33), 16228–16233.

Zhang, Z., & Wang, Y. (2023). A spatiotemporal model for global earthquake prediction based on convolutional LSTM. IEEE Transactions on Geoscience and Remote Sensing, 61, 1–12.

Ziagos, J., Phillips, B. R., Boyd, L., Jelacic, A., Stillman, G., & Hass, E. (2013). A technology roadmap for strategic development of enhanced geothermal systems (Tech. Rep.). Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States).