PROCEEDINGS, 50th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 10-12, 2025
SGP-TR-229

Forecasting Induced Seismicity Using Temporal Fusion Transformer: A Case Study in the Geysers Geothermal Field

Nori Nakata[1], Zhengfa Bi[1]

[1]Earth and Environmental Sciences Area, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720

nnakata@lbl.gov

Keywords: induced seismicity, forecasting, machine learning, Geysers geothermal field

ABSTRACT

Induced seismicity remains a critical challenge in the sustainable development of geothermal energy, with fluid injection and reservoir stimulation often triggering seismic events. This study proposes a data-driven framework using a modified Temporal Fusion Transformer to forecast induced seismicity rates in geothermal fields. The model integrates heterogeneous data, including geological features, historical seismicity, and operational metadata, to predict seismicity across multiple temporal horizons. By leveraging attention mechanisms and probabilistic forecasting, our approach identifies key drivers of seismicity, such as injection rate variability, while providing uncertainty bounds for informed decision-making. The framework's interpretability enables scenario testing to evaluate the impacts of operational adjustments, such as changes in injection rates, on seismic hazards. Results from a case study at the Geysers geothermal field demonstrate the model's ability to accurately capture nonlinear relationships between fluid injection and seismicity, offering insights into subsurface processes and informing risk mitigation strategies. This work advances seismicity forecasting by bridging the gap between statistical and physics-based methods, contributing to safer and more efficient geothermal energy production.

1. INTRODUCTION

Induced seismicity presents a critical challenge in geothermal reservoir management in the context of sustainable development of geothermal energy, a renewable resource that plays a pivotal role in achieving global carbon neutrality goals (Gaucher et al., 2015). The occurrence of large seismic events poses public safety concerns and can provoke societal and economic repercussions, necessitating the development of robust mitigation plans by operators. These concerns have garnered significant scientific and public attention as geothermal systems grow in importance, particularly through the deployment of Enhanced Geothermal Systems (EGS; Tester et al., 2006; Ziagos et al., 2013) designed to extract energy from low-permeability reservoirs. Seismic events triggered by fluid injection, extraction, or circulation are not limited to geothermal operations but also extend to other fluid-driven perturbations, such as reservoir impoundment, oil and gas production, and underground disposal of waste fluids. These activities have led to the suspension or abandonment of projects, underscoring the urgent need for robust and reliable seismicity forecasting frameworks (Baisch et al., 2009; Bertani, 2012).

Many efforts have been made in the last decade to understand and forecast induced seismicity, using statistical, physics-based, or hybrid methods that rely on assumptions or prior geological knowledge of seismicity mechanisms. Statistical approaches, such as the epidemic-type aftershock sequence (ETAS) model (Ogata 1988 and 1998), and Traffic-light system (Bommer et al., 2006) have been widely used to model seismicity as a stochastic point process in time, space, and magnitude domains. These models effectively capture background seismicity driven by tectonic loading and aftershock triggering (Bachmann et al., 2011), using well-established seismological laws such as the Omori-Utsu formula and the Gutenberg-Richter distribution. However, applying these models to induced seismicity requires modifications to account for anthropogenic factors like fluid injection. Recent studies have addressed these challenges by integrating linear response theory into statistical seismicity models, treating fluid injection as an external forcing function. 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 and Avouac, 2023). Despite these advancements, such frameworks often focus primarily on temporal evolution, neglecting the stochastic nature and complex interactions of seismicity-driving factors, which limits their predictive power in seismicity forecasting.

On the other hand, physics-based models simulate the dynamic changes in stress and strain caused by geothermal operations, offering deeper insights into the underlying physical processes. While these models provide valuable understanding to induced seismicity, they are computationally intensive and rely on detailed subsurface property knowledge, which is often heterogeneous and poorly constrained. Moreover, these methods struggle to accommodate the inherently stochastic and nonlinear nature of earthquake triggering, particularly in highly active geothermal regions. Hybrid approaches that combine observational data with physical modeling have been proposed to improve predictive accuracy and reliability. However, they remain limited in identifying the relative importance of contributing factors, offering minimal insights into the mechanisms that trigger seismic events and their dynamic evolution.

In this study, we adopt a data-driven deep learning framework to forecast induced seismicity rates in geothermal fields. Building on a modified Temporal Fusion Transformer (TFT; Lim et al., 2021) architecture, this approach integrates all available geological and operational parameters, including historical seismicity and operational metadata, to predict future rates of induced seismicity. Although our approach can also be classified as a statistical method, it distinguishes itself through its capability for multi-horizon forecasting, addressing the challenges of predicting variables across multiple future time steps. Multi-horizon forecasting allows for a comprehensive understanding of the evolution of target variables, enabling users to make informed decisions at multiple stages in the future. The model's

built-in attention mechanism identifies key contributing factors, enabling interpretable insights into induced seismicity without requiring prior assumptions about the importance of input features.

The model processes static covariates (e.g., borehole sites), time-dependent past observations (e.g., historical seismicity and injection rates), and known future inputs (e.g., planned injection schedules). Outputs include probabilistic predictions across multiple quantiles, providing estimations of uncertainty under various operational scenarios. This interpretable framework not only demonstrates superior predictive performance but also enhances understanding of the mechanisms underlying induced seismicity. It serves as a robust tool for risk mitigation and improved reservoir management strategies in geothermal operations. By addressing the limitations of traditional methods, our approach contributes to the sustainable development of geothermal energy while ensuring public safety and minimizing economic disruptions.

2. METHODOLOGY

The proposed approach leverages a transformer network to integrate heterogeneous time series, dynamically select relevant features, and capture complex temporal dependencies, enabling accurate and interpretable forecasting of induced seismicity.

2.1 Temporal Transformer Framework

The Temporal Fusion Transformer (TFT) represents a state-of-the-art architecture for multi-horizon time series forecasting, designed to integrate heterogeneous data and capture complex temporal dynamics. In this study, we extend the TFT framework to forecast induced seismicity, aligning its design with the characteristics of seismicity and operational metadata to achieve high accuracy and interpretable insights. Our approach processes three input types: static covariates (e.g., geological features and well sites), time-dependent past observations (e.g., historical seismicity, earthquake magnitudes, injection rates), and known future inputs (e.g., planned injection schedules and operational trends). These inputs are unified through embeddings for categorical and contextual information and linear transformations for continuous variables. As shown in Figure 1, key components of TFT include static covariate encoder, variable Selection Networks, sequence-to-sequence LSTM, and self-attention decoder. These elements collectively enable robust and interpretable seismicity forecasting.

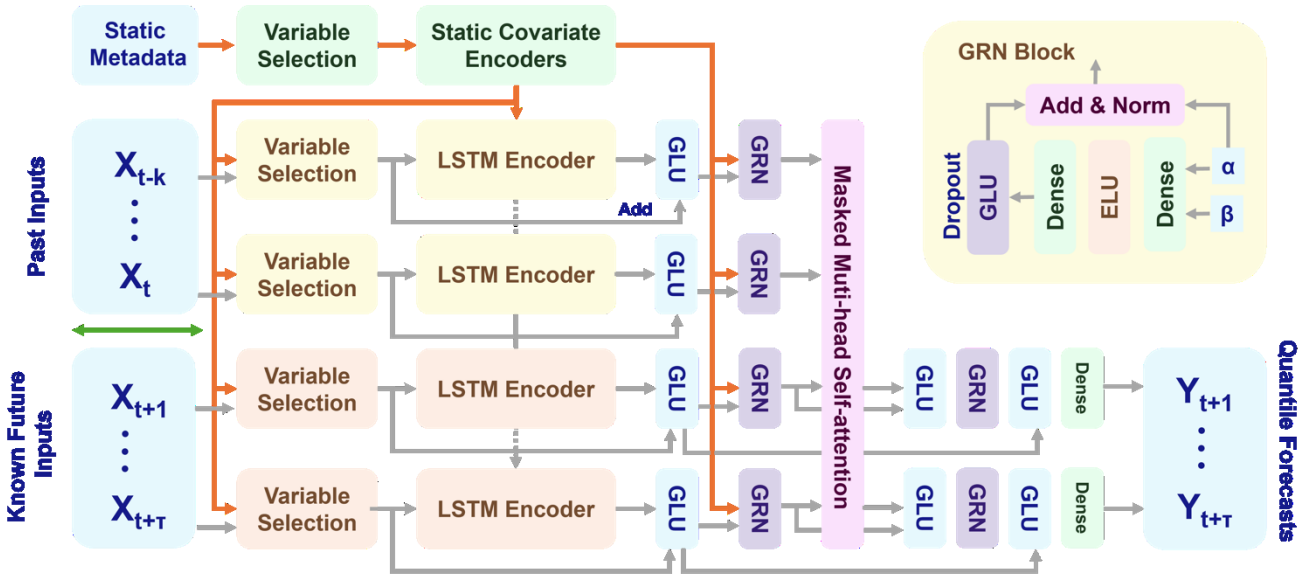

Figure 1: Overview of the Temporal Fusion Transformer (TFT) architecture for induced seismicity forecasting. The architecture effectively captures complex temporal dynamics, providing interpretable insights into the primary drivers of seismicity.

2.2 Static Covariance Encoder and Gating Mechanism

The model employs a Gated Residual Network (GRN) to encode static covariates that influence 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 geological knowledge and prior constraints. The GRN employs gating mechanisms to achieve dynamic feature selection, enabling the efficient processing of heterogeneous data with varying levels of complexity. The GRN is formulated as:

$$\mathrm{GRN}_{\omega}(a, c) = \mathrm{LayerNorm}(a + \mathrm{GLU}_{\omega}(\eta_1)), \tag{1}$$

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

where a and c are the primary input and the optional context vector, respectively. ω denotes weight sharing. The Exponential Linear Unit (ELU) activation behaves linearly for positive inputs and saturates for negative inputs, allowing the GRN to adapt its complexity dynamically. To control the non-linear contribution, we employ Gated Linear Units (GLUs) as gating mechanisms:

$$\mathrm{GLU}_{\omega}(\gamma) = \sigma(W_{4,\omega}\gamma + b_{4,\omega}) \odot (W_{5,\omega}\gamma + b_{5,\omega}), \tag{3}$$

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 applied before the gating layer to enhance generalization.

2.3 Variable Selection Network for Temporal Inputs

Time-varying past and known future inputs are processed through separate variable selection networks that dynamically assign weights to features at each time step, filtering out irrelevant or noisy inputs. Let $\xi_t^j \in R^{d_{model}}$ denote the transformed input of the j-th variable at time t, and $X_t = \left[(\xi_t^1)^T, \ldots, \left(\xi_t^{m_x}\right)^T\right]^T$ the concatenated vector of all past inputs. Variable selection weights are computed by passing X_t and an optional context vector c_s from static covariate encoder through GRN followed by a Softmax layer:

$$v_{\chi}^t = \mathrm{Softmax}(\mathrm{GRN}_{v_{\chi}}(\mathcal{X}_t, c_s)), \tag{4}$$

where $v_{\chi}^t \in R^{m_x}$ 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, in which weights of $GRN_{\tilde{\xi}}^j$ are shared across all time steps. The final weighted representation of past inputs is obtained by combining the refined features:

$$\tilde{\xi}_t = \sum_{j=1}^{m_{\chi}} v_{\chi}^{t,j}\tilde{\xi}_t^j = \sum_{j=1}^{m_{\chi}} v_{\chi}^{t,j}\mathrm{GRN}_{\tilde{\xi}}^j(\xi_t^j), \tag{5}$$

where $v_{\chi}^{t,j}$ is the j-th element of v_{χ}^t.

2.4 Sequence-to-Sequence LSTM

To capture the temporal dependencies for seismicity forecasting, TFT employs a hybrid approach combining sequence-to-sequence LSTM layers and interpretable multi-head attention mechanisms. The sequence-to-sequence LSTM layers process past observations and known future inputs separately, learning localized temporal patterns and initializing hidden states derived from static covariates. This design ensures that geological features and long-term operational trends are embedded into the temporal dynamics.

2.5 Interpretable Multi-Head Self-attention

The interpretable multi-head self-attention decoder further enhances the model's ability to identify long-term patterns and trends by highlighting significant time steps and feature interactions. This decoder builds upon multi-head attention in transformer-based architectures. In the standard attention mechanism, weighted representations of input values $\mathrm{V} \in \mathrm{R}^{\mathrm{N}\times\mathrm{d_v}}$ are computed as:

$$\mathrm{Attention}(Q, K, V) = A(Q, K)V = \mathrm{Softmax}\left(\frac{QK^{\top}}{\sqrt{d_{\mathrm{attn}}}}\right), \tag{6}$$

where $\mathrm{A(Q, K)}$is typically the scaled dot-product attention. Multi-head attention extends this by projecting queries, keys, and values into multiple subspaces, enabling the model to learn diverse temporal patterns. To improve interpretability, we modify multi-head attention to share values V across all heads and aggregate outputs additively:

$$\mathrm{MultiHead}(Q, K, V) = \tilde{A}(Q, K)VW_V, \tag{7}$$

where the unified attention weights are given by:

$$\tilde{A}(Q, K) = \frac{1}{m_H}\sum_{h=1}^{m_H} A(QW_Q^{(h)}, KW_K^{(h)}). \tag{8}$$

This approach allows each head to learn distinct temporal patterns while attending to a common set of input features V. The shared weights $\tilde{A}(Q, K)$$ provide a unified representation that enhances the analysis of significant temporal dependencies and feature interactions. This mechanism enables the identification of critical triggers of induced seismicity, such as abrupt increases in injection rates or changes in hydrothermal pressure.

2.6 Uncertainty Modeling and Quantile Foresting

To model the uncertainty in seismicity rate forecasting, the network outputs probabilistic predictions by generating multiple quantile forecasts for each time step. This enables robust decision-making by providing a range of plausible induced seismicity under varying operational scenarios. The model's interpretability is further enhanced through its ability to analyze variable importance. For example, the model highlights the dominant role of injection rate variability in driving seismicity, offering insights into how operational adjustments can mitigate seismic risks.

The Temporal Fusion Transformer (TFT) is trained by jointly minimizing the quantile loss, aggregated across all quantile outputs:

$$\mathcal{L}(\Omega, \theta) = \sum_{y_t \in \Omega} \sum_{q \in Q} \frac{1}{N\tau_{\max}} \sum_{\tau=1}^{\tau_{\max}} \mathcal{Q}(y_t, \hat{y}(q, t-\tau, \tau), q), \tag{9}$$

where

$$\mathcal{Q}(y, \hat{y}, q) = q(y - \hat{y})_+ + (1-q)(\hat{y} - y)_+, \tag{10}$$

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 $\tau_{\max}$. 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:

$$q_{\text{risk}} = 2\frac{\sum_{y_t \in \tilde{\Omega}} \sum_{\tau=1}^{\tau_{\max}} \mathcal{Q}(y_t, \hat{y}(q, t-\tau, \tau), q)}{\sum_{y_t \in \tilde{\Omega}} \sum_{\tau=1}^{\tau_{\max}} |y_t|}, \tag{11}$$

where $\tilde{\Omega}$ represents the test data domain.

3. APPLICATION

This study focuses on induced seismicity in the northwestern region of the Geysers geothermal field, leveraging high-resolution seismic and injection data from two key wells, Prati-9 and Prati-29 (Figure 2a). The Geysers, the world's largest geothermal field, has been operational since the 1960s and generates approximately 850 MW of electricity annually through steam production driven by extensive water injection. Over 70 injection wells, reaching depths of up to 5 km, are employed 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 350ºC below 2.75 km in the northwestern part of the field.

Fluid injection at the Geysers is gravity-driven, resulting in a negative gauge pressure at the wellhead, in contrast to the high-pressure methods typically used for reservoir stimulation. 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 subsurface forcing and seismic responses, with important implications for geothermal energy development and seismic hazard mitigation.

3.1 Earthquake Catalog and Injection Data

The earthquake catalog spans from 2009 to 2015 and includes 1254 seismic events near Prati-9 and Prati-29 (Figure 2b). These events are recorded by the Berkeley-Geysers (BG) seismic network, consisting of 31 three-component short-period geophones operated by Lawrence Berkeley National Laboratory. Events are relocated using the double-difference relocation technique, improving hypocenter precision to 50 m. Only events with magnitudes larger than 0 are considered to ensure consistency with the magnitude of completeness. To isolate induced seismicity, the catalog is declustered by removing aftershocks using spatial and temporal windows optimized for geothermal seismicity.

Continuous fluid injection at Prati-9 totals 7.21 Mm^3, while Prati-29 injects 3.29 Mm^3, with both wells exhibiting average injection rates of 8.79 x 10^4 m/month. Seasonal fluctuations in injection rates are observed, with peaks during winter months. 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 m, while only sparse activity (approximately 30 events) occurs near Prati-29 (Figure 2d), at an average distance of 800 m. Due to the limited seismicity near Prati-29, detailed analyses focuses on events near Prati-9, though the influence of total injection rates from both wells (Figure 2e) is considered.

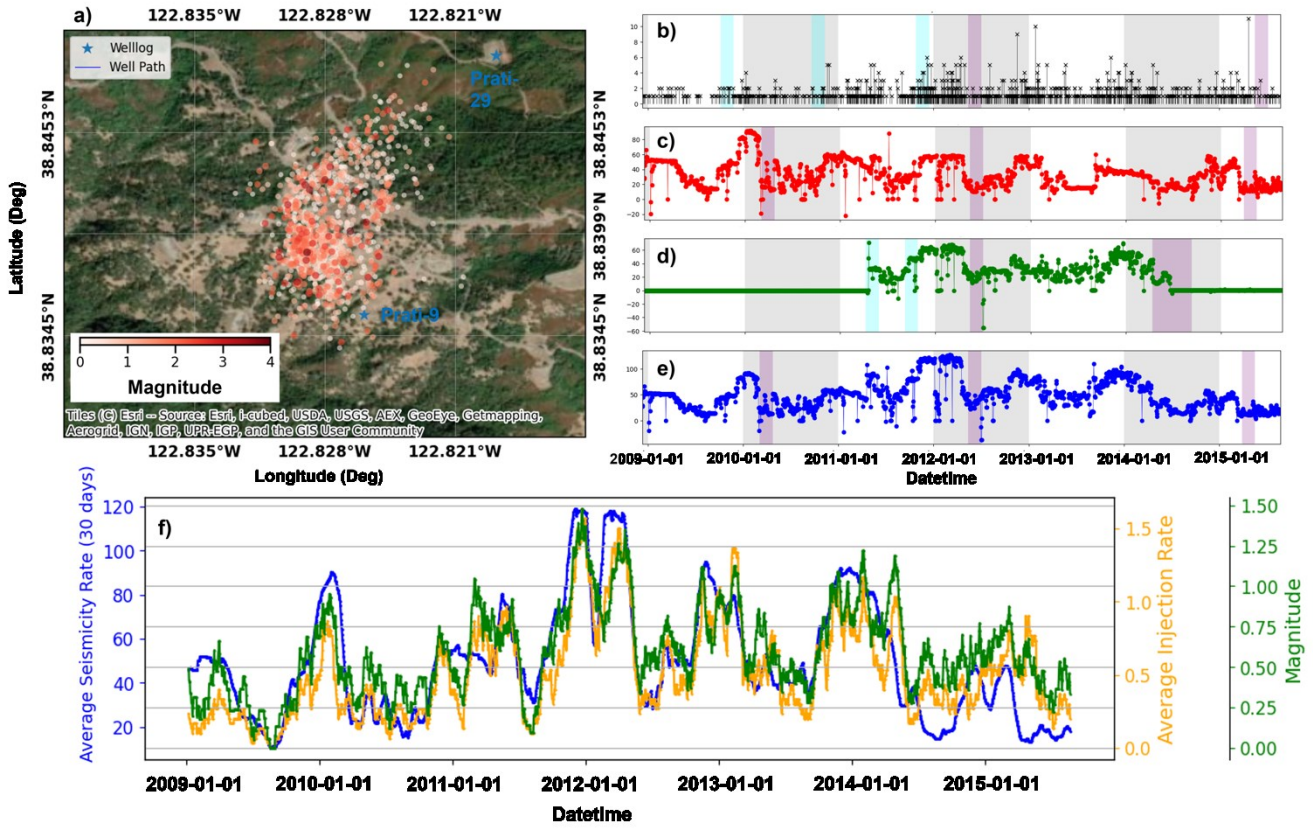

Figure 2: Study area and seismicity characteristics at the Geysers geothermal field. (a) Map of the locations of injection wells Prati-9 and Prati-29, as well as the spatial extent of the seismic events. (b) Spatial distribution of seismic events recorded from 2009 to 2015 near Prati-9 and Prati-29. Temporal variations in injection rates at Prati-9. Injection rates at Prati-29. Monthly injection rates for Prati-9 and Prati-29, showing seasonal fluctuations and the dominance of Prati-9 in overall injection volume. (f) Temporal trends of seismicity within 30-day intervals (blue) and earthquake magnitudes (green) compared to injection rates (orange).

3.2 Feature Engineering and Target

Monthly fluid injection data are preprocessed to calculate injection parameters, including normalized injection rates, cumulative injection volumes, and injection pressures. Anomalous spikes in injection volumes, defined as values exceeding ten times the average, are excluded to minimize errors. Geological features, including the basement depth and the vertical distance between injection depth and the basement, are incorporated to account for subsurface variations influencing induced seismicity.

The training dataset comprised seismic and injection data from 2009 to 2014, while the validation dataset consisted of data after 2014. Historical features, representing the differences between current and previous months, are included to capture potential temporal delays in seismic responses to fluid injection. Features with high redundancy or low correlation (smaller than 0.35) with the target variable are removed to ensure the robustness of the model. The target variable is the number of earthquakes occurring within 30, 60, 90, and 180 days, allowing forecasts across multiple temporal horizons.

3.3 Training and Model Validation

To validate the model, forecasts are performed using progressively extended training intervals. Earthquakes occurring between 2014 and 2015 (Figure 2f) were used, with a fixed forecasting window of 120 days. Predictions are evaluated against observed seismicity, with accuracy assessed using confidence intervals and statistical metrics. The results (Figure 3a) demonstrate that the our approach accurately captures seismicity trends and temporal dynamics, with predictions closely matching observed earthquake counts. The model's probabilistic forecasts including uncertainty bounds provide a comprehensive understanding of potential seismic outcomes and enhance its applicability for risk assessment. The findings emphasize model's ability to capture the nonlinear relationships between injection operations and seismicity, offering insights into the interactions between operational processes and seismic responses.

To further demonstrate the robustness of our approach, we analyze scenarios where injection rates are artificially enlarged or reduced after prediction times (marked by green lines in Figures 3b and 3c). These scenarios reveal trends that align well with physical expectations and provide meaningful interpretations of operational adjustments. The significance of this analysis lies in its potential to simulate and evaluate the effects of varying operational conditions, such as sudden changes in injection rates or shut-in operations. By enlarging injection rates, the model reflects potential scenarios where increased injection volumes or pressures amplify seismic activity, enabling proactive hazard assessments. Conversely, shrinking injection rates or stabilization of reservoir pressures corresponds to scenarios involving reduced seismicity rates, which align with mitigation strategies designed to control seismicity. These trends highlight the model's capacity to evaluate how operational interventions influence seismic hazards over time, offering critical insights for optimizing injection strategies.

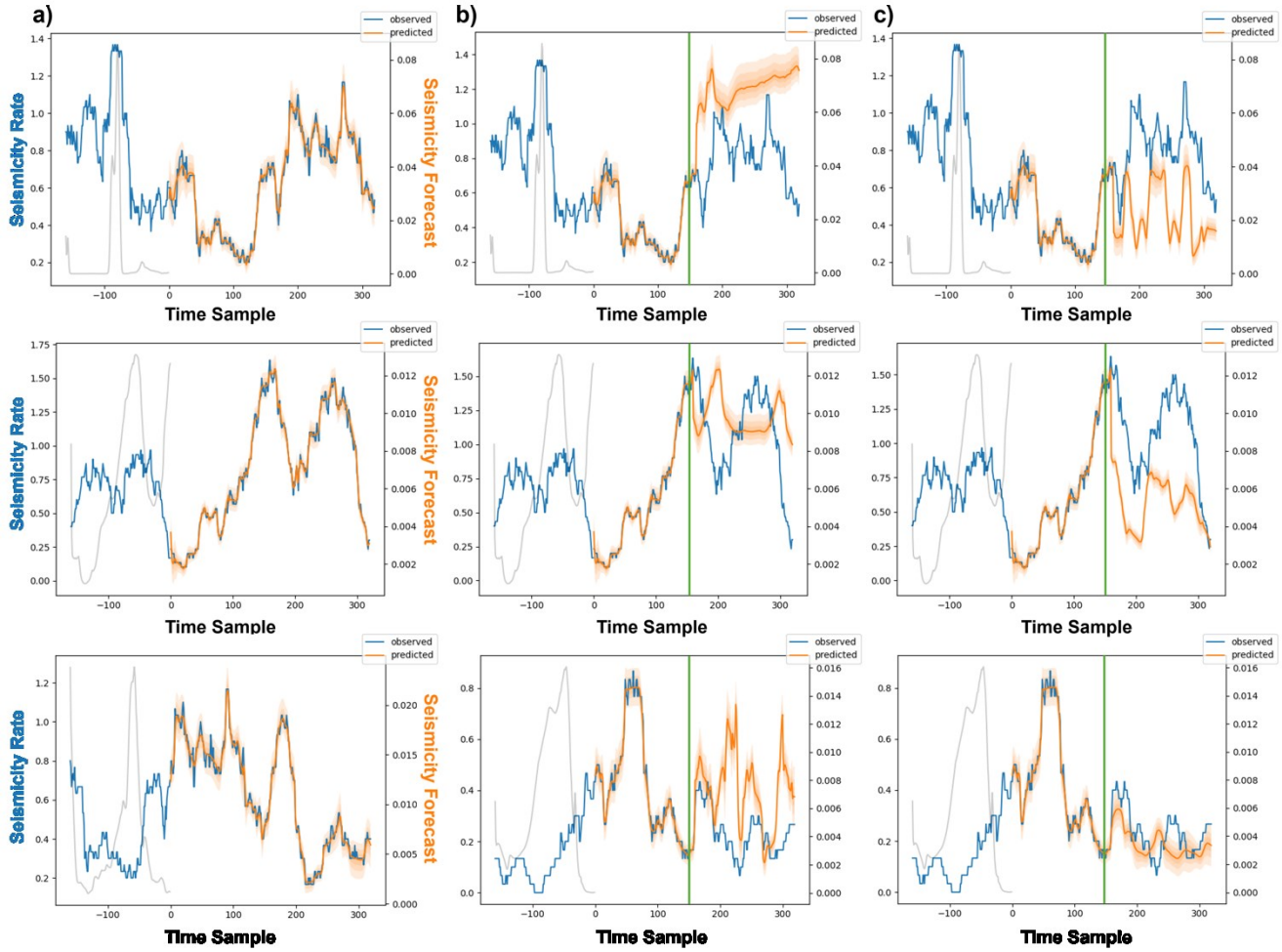

Figure 2: Validation and performance analysis of the proposed seismicity forecasting framework. (a) Predicted versus observed earthquake counts for 120-day forecasting intervals. The model accurately captures seismicity trends and temporal dynamics, with probabilistic forecasts shown as shaded uncertainty bounds (orange). (b) Scenario analysis with artificially enlarged fluid injection after a prediction time (green lines). (c) Scenario analysis with reduced injection rates after a prediction time. The model reflects amplified seismic activity consistent with expected physical responses to operational changes.

4. DISCUSSION

This study demonstrates the utility of deep neural network for forecasting induced seismicity and managing geothermal reservoir operations. Unlike traditional approaches that rely heavily on assumptions or static models, TFT dynamically incorporates heterogeneous data and captures the nonlinear relationships between fluid injection and seismic responses. This capability is critical for understanding the complex interplay of operational parameters and subsurface processes in active geothermal fields.

4.1 Interpretability and Decision Support

The interpretability of TFT is an advancement in seismicity forecasting, as it provides insights into the relative importance of different input features. By identifying dominant factors such as injection rates and geological attributes, TFT enables operators to prioritize key parameters for risk mitigation. The ability to simulate hypothetical scenarios, such as sudden increases or decreases in seismicity rates, further underscores its value in operational decision-making. These simulations allow stakeholders to evaluate the potential impacts of operational changes, including injection rate adjustments or reservoir shut-ins, providing a proactive mechanism for seismic hazard management.

For instance, the model's ability to reproduce trends where seismicity rates are amplified or diminished after specific time intervals demonstrates its robustness and alignment with physical expectations. These trends reflect real-world scenarios, such as the effects of increased injection rates on pore pressure diffusion and the corresponding seismic activity. Conversely, scenarios with reduced seismicity rates illustrate the effectiveness of mitigation strategies aimed at stabilizing reservoir pressures. This functionality positions TFT as a valuable tool for scenario testing, enhancing the safety and sustainability of geothermal operations.

4.2 Addressing Challenges in Seismic Hazard Mitigation

Induced seismicity remains a major challenge for geothermal energy development, particularly in regions with high societal and environmental sensitivities. The probabilistic forecasting capabilities of TFT, including quantile predictions, provide a comprehensive view of potential seismic outcomes under varying operational scenarios. By offering uncertainty bounds and probabilistic insights, the framework ensures that risk assessments are grounded in a range of plausible outcomes rather than deterministic predictions. This capability is essential for designing robust mitigation plans that account for the inherent variability and stochastic nature of induced seismicity.

Moreover, the integration of physics-informed trends, such as transient stress and pore pressure dynamics, into the forecasting process bridges the gap between statistical and physics-based approaches. While traditional physics-based models often require detailed subsurface data that may be unavailable or uncertain, TFT circumvents these limitations by leveraging historical data and real-time operational inputs. This enables reliable forecasting even in data-sparse environments, expanding the applicability of the model to diverse geothermal fields.

4.3 Limitations and Future Directions

While the proposed framework demonstrates advancements, several limitations warrant further investigation. First, the performance of TFT is dependent on the degree of coupling between well operations and seismicity, particularly in spatially heterogeneous regions where geological and operational conditions vary. Incorporating datasets with stronger correlations to induced seismicity, alongside comprehensive monitoring networks, could substantially improve the model's ability to generalize across diverse geological settings.

Second, although the model captures nonlinear relationships among operational parameters and seismicity, the underlying causal mechanisms require further exploration. Integrating domain-specific physics into the modeling framework represents a promising direction for bridging the gap between empirical observations and physical understanding. Coupling TFT with geomechanical models could provide new insights into the physical processes driving induced seismicity. For instance, integrating geological surveys, seismic catalogs, and 3-D fault models could explore how fluid injection interacts with pre-existing fractures to induce seismicity. This approach would provide insights into the stress redistribution, pore pressure migration, and fault reactivation mechanisms to identify critical thresholds and operational parameters for risk mitigation.

Third, our framework currently focuses on temporal trends in induced seismicity but does not explicitly address spatial distributions of seismic events. Extending the model to forecast the spatial propagation of induced seismicity would greatly enhance its utility for designing localized mitigation strategies. This could be achieved by coupling the TFT with spatially explicit machine learning approaches that account for subsurface heterogeneity. For example, we partition the study area in the Geysers geothermal field into triangulated grids based on historical seismicity. In regions with higher seismic density, the grid cells are finer, allowing for more localized representation, while injection and geological parameters are aggregated within each grid to construct the feature set. This approach could provide spatiotemporal characterization of induced seismicity, offering valuable insights for geothermal field operators to optimize injection strategies in high-risk zones.

Finally, while the framework achieves accurate multi-horizon forecasts, extending the temporal prediction window without compromising accuracy remains a challenge. This challenge is relevant for long-term planning in geothermal operations and seismic hazard mitigation. Future research could explore adaptive methods, such as dynamic model updating, to continuously recalibrate parameters based on new data. Incorporating online learning algorithms and real-time monitoring systems would allow the model to adapt to evolving operational contexts and provide actionable forecasts in near real time. By addressing these limitations, future studies could establish a more holistic framework for understanding induced seismicity, advancing the sustainability and safety of geothermal energy operations.

5. CONCLUSIONS

This study demonstrates the transformative potential of data-driven frameworks constructed from TFT in forecasting induced seismicity and managing geothermal reservoir operations. By integrating heterogeneous data, capturing nonlinear relationships, and providing interpretable insights, the proposed approach advances the state of the art in seismic hazard mitigation. Beyond forecasting, TFT serves as a decision-support tool, enabling operators to simulate and evaluate the impacts of operational adjustments on seismic risks. These

capabilities align with the global push for sustainable energy development, ensuring that geothermal energy continues to play a critical role in achieving carbon neutrality while prioritizing public safety.

6. ACKNOWLEDGEMENTS

We are grateful to Calpine and Northern California Earthquake Data Center to make the operation and seismicity data available. This work is supported by Utah FORGE, sponsored by the U.S. Department of Energy, under the award number 6-3656.

REFERENCES

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

Baisch, S., Carbon, D., Dannwolf, U., Delacou, B., Devaux, M., Dunand, F., ... others (2009). Deep heat mining basel: seismic risk analysis. SERIANEX Group, Departe- ment fu¨r Wirtschaft, Soziales und Umwelt des Kantons Basel-Stadt, Basel.

Bertani, R. (2012). Geothermal power generation in the world 2005–2010 update report. geothermics, 41, 1–29.

Bommer, J. J., Oates, S., Cepeda, J. M., Lindholm, C., Bird, J., Torres, R., ... Rivas, J. (2006). Control of hazard due to seismicity induced by a hot fractured rock geothermal project. Engineering geology, 83(4), 287–306.

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.

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

Lim,B.,Arık,S.O¨.,Loeff,N.,&Pfister,T.(2021).Temporalfusiontransformersforinter- pretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748–1764.

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.

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.

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).