Proceedings World Geothermal Congress 2026
Calgary, Canada, June 8 – 11, 2026

Development of a Recurrent Neural Network for Formation Temperature Estimation During Drilling

Yuta Takahashi[1], Shigemi Naganawa[1] and Elvar K. Bjarkason[1]*
Graduate School of International Resource Sciences, Akita University, Japan[1]
elvar@mine.akita-u.ac.jp

Keywords: Formation temperature, estimation, machine learning, recurrent neural networks, geothermal wells

ABSTRACT

Estimation of the static formation temperature (SFT) is essential for evaluating geothermal potential, understanding reservoir characteristics, and ensuring safe drilling operations within the thermal tolerance of downhole instruments. However, real-time estimation of SFT during drilling remains technically challenging due to thermal disturbances caused by circulating drilling fluids. This study aims to develop an artificial intelligence-based method for predicting the SFT in real time using surface and downhole logging data acquired during drilling operations. Real-time estimation of the SFT during drilling could facilitate flexible decision-making, including adjustment of drilling or completion depth, and early evaluation of reservoir characteristics.

Synthetic datasets were generated using the GEOTEMP2 wellbore thermal simulator under various geothermal gradients and drilling conditions to reflect diverse geothermal environments. Recurrent neural network (RNN) and Long Short-Term Memory (LSTM) models were trained on these datasets to learn the relationship between the wellbore thermal behavior and the SFT. The input features consisted of bottom-hole temperature (BHT), and mud outlet temperature, while the target variable was the SFT. The proposed models are designed to estimate SFTs using only temperature data recorded during drilling, without requiring wellbore temperature recovery observations during post-drilling shut-in periods.

Model performance was evaluated using synthetic datasets generated based on reference SFT profiles from five geothermal wells. The simpler RNN model achieved higher prediction accuracy than the LSTM model. These results demonstrate that the relatively simple RNN effectively captured nonlinear temporal variations in temperature and achieved reliable SFT estimation. This study integrates artificial intelligence into geothermal wellbore analysis and provides a foundation for real-time thermal diagnostics and operational decision support in high-temperature geothermal drilling environments.

1 INTRODUCTION

In the development of high-temperature geothermal wells, accurate determination of the static formation temperature (SFT) is crucial for ensuring both safe drilling operations and reliable assessment of geothermal resource potential. The SFT serves as a fundamental input for the thermal design of drilling fluids, cement slurries, and bottom-hole assemblies (BHA), which must maintain performance and structural integrity under high-temperature downhole conditions (Gul et al., 2019). In addition to thermal system design, SFT estimation plays a critical role in geothermal resource assessment. Reliable estimation of formation temperature enables early evaluation of geothermal gradients and subsurface heat flow, which are essential for understanding reservoir behavior, estimating energy potential, and supporting the design of surface facilities such as power generation systems (Bassam et al., 2010; Espinosa-Paredes and Espinosa-Martínez, 2009). However, during drilling, the continuous circulation of drilling mud cools the borehole, causing the recorded bottom-hole temperature (BHT) to be lower than the undisturbed SFT. Therefore, SFT cannot be directly measured during drilling. Real-time SFT prediction during drilling would enable operators to assess whether the thermal target has been reached, thus facilitating prompt decisions on final drilling depth and completion design (Tekin, 2010; Gul et al., 2019; Khaled, et al., 2024)

Traditionally, SFT has been estimated through temperature recovery tests, in which drilling operations are halted and the well is shut in to record the temperature recovery curve. After shut-in, the wellbore temperature gradually increases over time due to heat transfer from the surrounding formation and asymptotically approaches the original formation temperature (SFT), as illustrated in Figure 1. Typically, the SFT is estimated through some form of linear regression, whereby wellbore temperature recovery observations are extrapolated to infinite time (Deming, 1989; Espinosa-Paredes and Garcia-Gutierrez, 2003; Kutasov and Eppelbaum, 2005).

However, this method may require suspending drilling operations for several days or even weeks, and experienced temperature conditions may be high enough to exceed the operational limits of downhole sensors, thereby hindering accurate data acquisition (Tekin, 2010).

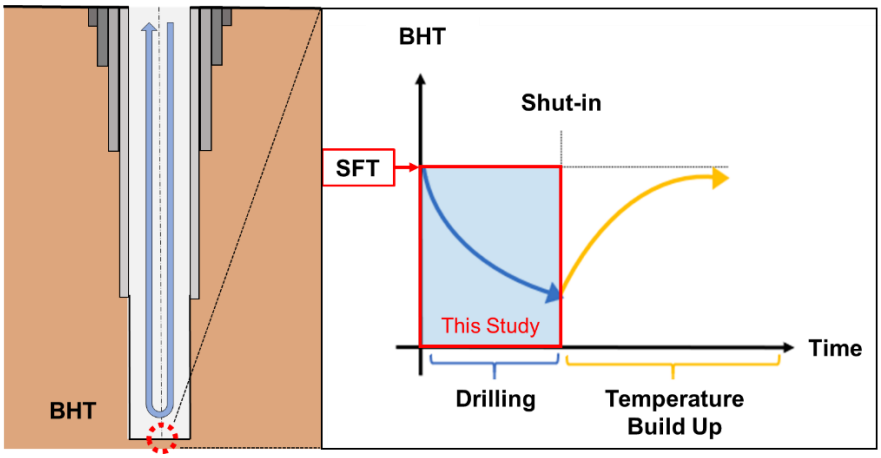


Figure 1. Variation of bottom hole temperature (BHT) during drilling and after shut-in.

In recent years, artificial intelligence (AI) has been progressively integrated into geothermal energy research across multiple domains, including reservoir characterization, seismic risk assessment, and drilling optimization. Publication trends indicate a clear acceleration in AI adoption since 2020, highlighting its expanding role in addressing technical and operational challenges in geothermal systems (Dashtgoli et al., 2026).

A few studies have investigated the estimation of bottom-hole circulating temperature (BHCT) and SFT using both numerical simulations and data-driven approaches. Tekin and Akin (2011) performed numerical analyses of wellbore temperature distributions using the GTEMP wellbore simulator and reported good agreement between simulated and measured SFT values. Espinosa-Paredes and Espinosa-Martínez (2009) introduced a feedback-based inverse heat transfer method, which improved the stability and accuracy of SFT estimation compared to conventional least-squares-based techniques. However, these numerical methods require a large number of input parameters and are computationally demanding, making real-time application challenging.

Recently, machine learning methods have emerged as a promising alternative for estimating wellbore temperatures. Gul et al. (2019) developed and compared several machine learning models—including Random Forest, XGBoost, Support Vector Machines (SVMs), and artificial neural networks (ANNs) using mud-logging data such as surface temperature, flow rate, and fluid properties to predict the BHCT and SFT. Khaled et al. (2024) compared multiple machine learning algorithms using simulated drilling data from the Utah FORGE project and found that the Long Short-Term Memory (LSTM) model achieved the lowest mean absolute error (MAE). These studies indicate the capability of machine learning to provide real-time, high-accuracy temperature predictions compared with conventional physics-based models.

The objective of this study is to develop a machine learning–based model capable of estimating the SFT in real-time during drilling operations. The proposed approach employs recurrent neural network (RNN) and LSTM architectures trained on synthetic temperature datasets generated using the GEOTEMP2 wellbore thermal simulator (Mondy and Duda, 1984). The input features include mud-logging parameters and temperature data monitored at the surface, which can be continuously acquired in real time during drilling. By establishing a data-driven SFT prediction framework, this study aims to overcome the temporal and operational limitations of conventional temperature recovery and numerical simulation methods. Real-time estimation of SFT would facilitate dynamic decision-making for casing design, target depth adjustment, and well completion planning, ultimately enhancing the safety and operational efficiency of geothermal drilling. This research contributes to the integration of artificial intelligence into geothermal wellbore thermal analysis and demonstrates its potential for thermal behavior diagnostics and decision support in high-temperature geothermal environments.

2 METHODS

2.1 Dataset Generation Using Wellbore Thermal Simulator

The prediction accuracy of machine learning models strongly depends on the quality and quantity of training data. However, because the availability of field data from actual geothermal wells is limited, this study generated synthetic temperature datasets for model training and validation using a wellbore thermal simulator. The

simulator employed was GEOTEMP2 (Mondy and Duda, 1984), a wellbore thermal-fluid analysis program developed at Sandia National Laboratories (USA). GEOTEMP2 numerically solves the temperature distribution in and around the wellbore using the finite volume method based on the principle of energy conservation. It accounts for both convective and conductive heat transfer within the wellbore, as well as thermal conduction in the surrounding formation, thereby enabling temporal and spatial modeling of temperature profiles during drilling operations. By utilizing GEOTEMP2, it is possible to compute BHTs and mud outlet temperatures under various drilling conditions, allowing the generation of a comprehensive set of synthetic datasets even in the absence of measured field data. These datasets serve as the foundation for training and evaluating the machine learning models developed in this study.

2.2 Description of Training and Testing Datasets

In this study, wellbore temperatures during drilling were simulated using GEOTEMP2, and the generated synthetic temperature data were used to train and validate RNN and LSTM models. The input parameters for GEOTEMP2, including mud inlet temperature and SFT profiles were defined with reference to field data and randomly varied within realistic ranges to generate multiple simulation cases. The corresponding input values are summarized in Table 1. For each case, BHT and mud outlet temperature calculated at 200 ft intervals were recorded to construct the dataset. The explanatory variables used in the machine learning models were mud outlet temperature and BHT, while the target variable was the SFT. An example of the dataset is shown in Figure 2. Each dataset corresponds to a well with a total depth of 3 km, containing temperature data recorded at 200 ft intervals during drilling. Finally, 100 well cases were generated, and the dataset was divided into 80% for training and 20% for validation to train the models.

Table 1. Input parameters and their ranges used in GEOTEMP2 for dataset generation.

GEOTEMP2 Input Parameter	Minimum	Maximum
Geothermal Gradient [℃/m]	0.025	0.12 (Maximum 374 ℃ at 3000 m)
Mud Inlet Temperature [℃]	10	50

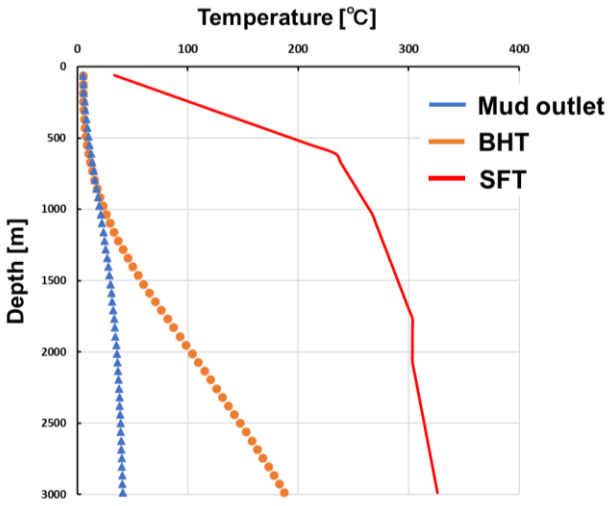


Figure 2. An example of the temperature dataset generated using GEOTEMP2 and used for model training.

For model demonstration, five geothermal wells were used: Kakkonda WD-1a in Japan (Ikeuchi et al., 1998) and four Icelandic wells: RGJ-34, K-11, HE-55, and well NJ-11 at Nesjavellir (Steingrimsson, 2018). Reported estimates of SFT values in these wells served as reference data, while other parameters were configured in GEOTEMP2 within the same parameter ranges as those applied during training data generation. The resulting synthetic datasets were subsequently used to test the predictive performance of the optimized models.

2.3 RNN/LSTM-Based Models for SFT Estimation

The RNN employed in this study is a type of neural network that belongs to the broader category of deep learning models, which itself is a subfield of machine learning. RNNs possess recurrent connections within their

architecture, allowing them to retain information from previous inputs as hidden states, thereby capturing temporal dependencies inherent in time-series data. This property makes RNNs particularly effective for analyzing nonlinear and history-dependent continuous data, such as downhole temperature variations during drilling. However, conventional RNNs suffer from the vanishing gradient problem, which limits their ability to learn long-term dependencies. To overcome this limitation, the LSTM model was developed. Although the LSTM model is an RNN variant, it features a structure that dynamically regulates the retention and forgetting of information, enabling stable learning of long-term dependencies. As a result, LSTM models usually demonstrate higher representational capability and learning stability than simpler RNNs when applied to long time-scale drilling data. In this study, a simple RNN was examined and compared with a more advanced LSTM. Figure 3 shows the basic RNN architecture used in this study. For the implemented models, all temperature data was normalized by dividing temperatures by 500°C, which gave normalized temperature values ranging between zero and one.

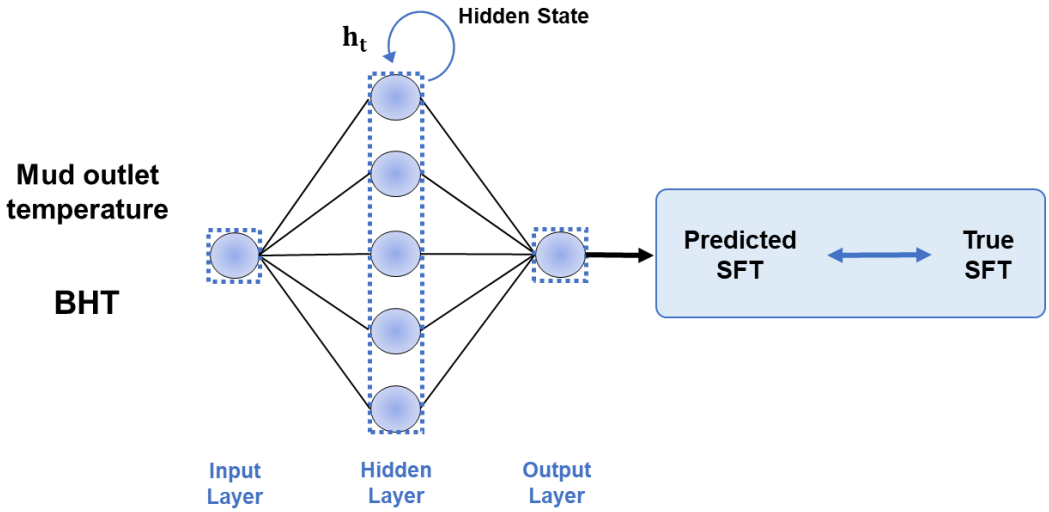


Figure 3. Conceptual diagram of the recurrent neural network structure used in this study.

2.4 Training and Optimization of RNN/LSTM Models for SFT Prediction

In this study, deep learning models based on the RNN and LSTM architectures were developed to estimate the SFT during drilling. Both models were implemented using the TensorFlow/Keras library in Python. A linear activation function was applied to the output layer, while the Rectified Linear Unit (ReLU) activation function was used in all hidden layers.

Optimization of the network architecture and training process was performed by tuning key hyperparameters, including the number of nodes, batch size, learning rate, number of hidden layers, and number of neurons per layer. These parameters were systematically explored using a grid search approach, and multiple combinations within predefined ranges were evaluated (Table 2).

The models were trained using the Adam optimization algorithm (Adaptive Moment Estimation) with the mean squared error (MSE) as the loss function. Model performance on the validation data was continuously monitored, and early stopping was applied when no improvement was observed in MSE for the validation data over 2,000 epochs. The maximum number of epochs was limited to 15,000.

Table 2. Ranges of hyperparameters used for model optimization in grid search.

Hyperparameter	Min	Max
Number of Hidden Layers	1	5
Number of Nodes	5	50
Learning Rate	0.0001	0.01
Batch Size	1	10

After training, the model configuration that yielded the smallest MAE on the validation data was adopted as the optimized model. To further evaluate the generalization capability of the optimized RNN and LSTM models, their predictive performance was assessed using independent test datasets. The evaluation metrics included the

coefficient of determination (R^2), MAE, and mean absolute percentage error (MAPE), all computed using the scikit-learn library in Python.

3 RESULTS AND DISCUSSION

3.1 Model Training and Optimization Results

Table 3 summarizes the 20 models that gave the lowest validation loss across all hyperparameter combinations. Eighteen of these models were LSTM architectures, while only two were RNN models. Models 1 and 15 were selected as the optimal LSTM and RNN models, respectively.

Table 3. Grid search results for model optimization and loss performance.

No.	Model	Number of Nodes	Learning Rate	Batch Size	Number of Epochs	Training Loss [$\times 10^{-5}$]	Validation Loss [$\times 10^{-5}$]
1	LSTM	30	0.001	1	10653	2.566	2.692
2	LSTM	40	0.001	3	9667	2.627	2.746
3	LSTM	40	0.001	5	10902	2.811	2.868
4	LSTM	40	0.001	1	6584	2.979	2.924
5	LSTM	50	0.001	5	7386	2.893	2.937
6	LSTM	30	0.0001	3	15000	3.117	3.01
7	LSTM	50	0.01	7	7195	3.089	3.019
8	LSTM	50	0.01	3	7669	3.44	3.08
9	LSTM	50	0.001	7	10945	2.819	3.08
10	LSTM	50	0.001	1	5003	2.974	3.081
11	LSTM	10	0.001	1	13353	2.943	3.115
12	LSTM	30	0.001	3	12189	2.992	3.12
13	LSTM	40	0.001	7	12178	2.781	3.158
14	LSTM	30	0.001	7	10711	3.255	3.194
15	RNN	50	0.001	5	6031	3.70	3.237
16	RNN	50	0.0001	3	15000	3.086	3.271
17	LSTM	20	0.001	5	14698	3.279	3.272
18	LSTM	20	0.001	1	6536	3.566	3.284
19	LSTM	30	0.01	3	5455	3.618	3.318
20	LSTM	10	0.01	1	7952	4.137	3.345

Table 4. Optimized RNN and LSTM model architectures and hyperparameters.

Model	RNN	LSTM
Number of hidden layers	1	1
Number of nodes	50	30
Learning rate	0.001	0.001
Batch size	5	1
Trained epochs	6031	10653
Activation function (hidden layers)	ReLU	ReLU
Activation function (output layer)	Linear	Linear
Optimization algorithm	Adam	Adam

Table 4 presents the model architecture and training parameters of the optimal RNN and LSTM models. The validation loss values were 2.692×10^{-5} for the optimal LSTM model and 3.237×10^{-5} for the optimal RNN model, corresponding to average temperature errors of 2.60°C and 2.84°C, respectively. The small differences observed between the training and validation losses indicate that both models achieved stable learning behavior without significant overfitting. Subsequently, these two optimized models were further evaluated using additional synthetic test datasets to assess their final SFT prediction performance. The results are presented in Section 3.2.

Despite the hyperparameter optimization, the top 20 models exhibited diverse parameter combinations. This suggests that model performance depends more strongly on the quality and quantity of the dataset rather than on specific hyperparameter configurations.

3.2 Optimized Model Evaluation Results

Table 5 presents the evaluation results of the models on the synthetic test dataset. For the five geothermal wells, the average performance metrics were R^2 = 0.984, MAE = 3.800, and MAPE = 2.654 for the RNN model, and R^2 = 0.979, MAE = 4.701, and MAPE = 3.172 for the LSTM model. These results indicate that the RNN model achieved higher predictive accuracy than the LSTM model for the test data.

In contrast, as described in Section 3.1, the LSTM model exhibited smaller losses for the training and validation datasets, revealing a different trend between the datasets. Figures 4–6 present the prediction results for the test data. Both models demonstrated high accuracy in estimating the SFT, although their prediction accuracy decreased at the shallowest depths near the surface. This reduced accuracy at shallow depths is likely attributable to the fact that the data were recorded at 200 ft intervals, which limits the model's ability to capture rapid temperature variations near the surface.

In addition, the LSTM model tended to show larger prediction errors than the RNN model at depths of approximately 1,500 m in the Kakkonda WD-1a and Krafla K-11 wells, where the geothermal gradient changes sharply. These findings suggest that the LSTM model may have partially overfitted the training data, resulting in reduced generalization performance under previously unseen geological conditions.

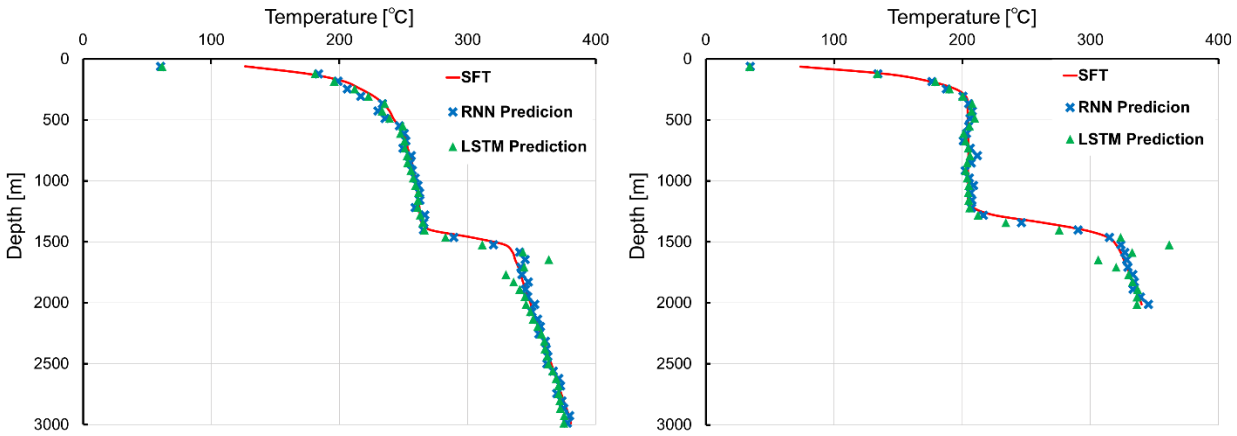


Figure 4. Comparison of SFT values predicted by the RNN and LSTM models with the reference SFT profiles for Kakkonda WD-1a (left) and Krafla K-11 (right).

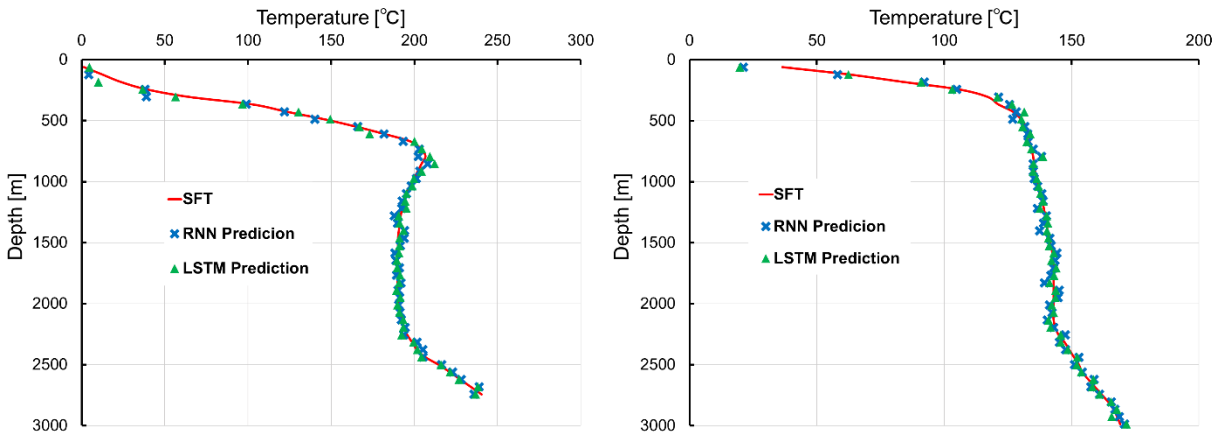


Figure 5. Comparison of SFT values predicted by the RNN and LSTM models with the reference SFT profiles for Hellisheidi HE-55 (left) and Reykjavik RGJ-34 (right).

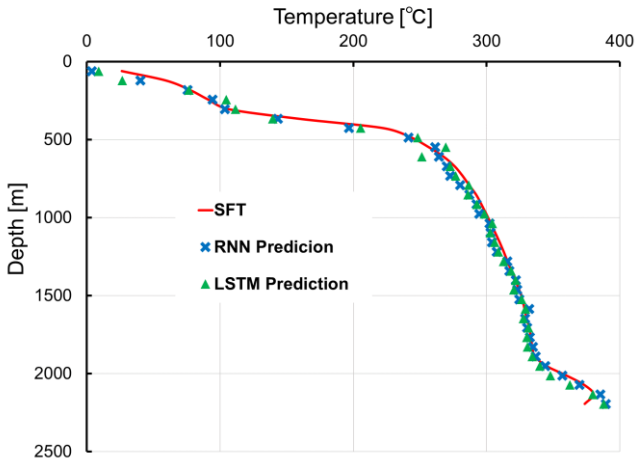


Figure 6. Comparison of SFT values predicted by RNN and LSTM models with the reference SFT profiles for the Nesjavellir NJ-11 well.

Table 5. SFT prediction results of the optimized RNN and LSTM models using the test datasets.

Model	RNN			LSTM		
Geothermal Well	R^2	MAE [°C]	MAPE [%]	R^2	MAE [°C]	MAPE [%]
Kakkonda WD-1a	0.973	4.341	2.164	0.968	5.287	2.379
Hellisheidi HE-55	0.990	3.059	0.942	0.994	2.749	1.897
Reykjavik RGJ-34	0.982	1.905	2.165	0.983	1.659	1.970
Krafla K-11	0.985	4.351	2.994	0.963	7.366	3.986
Nesjavellir NJ-11	0.992	5.345	5.005	0.989	6.446	5.625
Average	0.984	3.800	2.654	0.979	4.701	3.172

4 CONCLUSION

The RNN and LSTM models developed in this study demonstrated high accuracy in predicting the SFT during drilling operations. However, both models exhibited reduced accuracy near the surface, where available temperature data were limited. The LSTM model showed signs of overfitting under certain conditions, leading to lower predictive accuracy for the test data, whereas the RNN model more consistently captured the overall temperature trends and achieved better generalization performance.

The dataset used in this study was a simplified representation of data typically obtained from mud-logging and measurement-while-drilling (MWD) operations and did not incorporate complex geological structures or reservoir properties. Therefore, the models may not fully represent the complex thermal behavior and subsurface conditions of actual geothermal wells. Future work should involve retraining and validating the models using field temperature data obtained from real geothermal wells to assess their predictive accuracy and generalization capability under realistic geological conditions.

The findings of this study demonstrate the potential effectiveness of real-time SFT estimation using drilling data, compared to conventional post-drilling SFT estimation methods such as temperature recovery tests. This approach shows strong potential for supporting real-time thermal analysis and decision-making in geothermal drilling operations, thereby contributing to safer and more efficient well design and temperature management.

REFERENCES

Bassam, A., Santoyo, E., Andaverde, J., Hernández, J.A., and Espinoza-Ojeda, O.M.: Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach, *Computers & Geosciences*, 36(9), (2010), 1191–1199.

Dashtgoli, D.S., Giustiniani, M., Busetti, M., Cherubini, C., Alessandrini, G., and Narsilio, G.A.: Artificial Intelligence in Deep Geothermal Energy: Trends, Insights, and Future Perspectives, *Renewable Energy Focus*, 56, (2026), 100781.

Deming, D.: Application of bottom-hole temperature corrections in geothermal studies, *Geothermics*, 18(5/6), (1989), 775–786.

Espinosa-Paredes, G., and Espinosa-Martínez, E.G.: A feedback-based inverse heat transfer method to estimate unperturbed temperatures in wellbores, *Energy Conversion and Management*, 50(1), (2009), 140–148.

Espinosa-Paredes, G., and Garcia-Gutierrez, A.: Estimation of static formation temperatures in geothermal wells, *Energy Conversion and Management*, 44(8), (2003), 1343–1355.

Gul, S., Aslanoglu, V., Tuzen, M.K., and Senturk, E.: Estimation of bottom hole and formation temperature by drilling fluid data: A machine learning approach, *Proceedings, 44th Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, USA, (2019).

Ikeuchi, K., Doi, N., Sakagawa, Y., Kamenosono, H., and Uchida, T.: High-temperature measurements in well WD-1a and the thermal structure of the Kakkonda geothermal system, Japan, *Geothermics*, 27(5/6), (1998), 591–607.

Khaled, M.S., Wang, N., Ashok, P., van Oort, E., and Wisian, K.: Real-time prediction of bottom-hole circulating temperature in geothermal wells using machine learning models, *Geoenergy Science and Engineering*, 238, (2024), 212891.

Kutasov, I.M., and Eppelbaum, L.V.: Determination of formation temperature from bottom-hole temperature logs—a generalized Horner method, *Journal of Geophysics and Engineering*, 2(2), (2005), 90–96.

Mondy, L.A., and Duda, L.E.: Advanced wellbore thermal simulator GEOTEMP2 user manual, Sandia National Laboratories, Albuquerque, NM, USA, SAND-84-0857, (1984).

Steingrimsson, B.: Geothermal Well Logging: Temperature and Pressure Logs, *Presented at SDG Short Course III on Geothermal Reservoir Characterization: Well Logging, Well Testing and Chemical Analysis*, organized by UNU-GTP and LaGeo, Santa Tecla, El Salvador, (2018).

Tekin, S.: Estimation of the formation temperature from the inlet and outlet mud temperatures while drilling geothermal formations, master's thesis, Middle East Technical University, Turkey, (2010).

Tekin, S., and Akin, S.: Estimation of the formation temperature from the inlet and outlet mud temperatures while drilling geothermal formations, *Proceedings, 36th Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, USA, (2011).