Development of Artificial Neural Networks for Estimating Static Formation Temperature

Shunsuke Kaneko, Shigemi Naganawa and Elvar K. Bjarkason

Graduate School of International Resource Sciences, Akita University, Japan

Keywords

Formation temperature, estimation, machine learning, neural networks, geothermal wells

ABSTRACT

This study looks at using artificial neural networks (ANNs) for estimating the static formation temperature (SFT) in geothermal wells. SFTs are conventionally estimated from temperature recovery data using some form of linear regression, such as the Horner method. However, there are a multitude of regression models to choose from, and the optimal one is not settled and may be situation specific. Moreover, commonly used linear regression methods, like the Horner method, require suitably long shut-in periods for their linearization assumptions to apply. As an alternative, machine learning methods have been considered previously to some extent for estimating SFTs. Nevertheless, the development of machine learning alternatives has been limited by available field data and especially accurate SFT data. In this study, we look at using ANNs for estimating the SFT in geothermal wells based on transient temperature recovery data. For training the ANNs, a large set of synthetic temperature recovery data was generated using a wellbore simulator called GEOTEMP2. The GEOTEMP2 simulations describe the temperature in a wellbore during drilling and the following recovery period after drilling. The developed ANN model was evaluated by comparing its estimation accuracy with the Horner method for synthetic validation data. On average, the ANN model provided SFT predictions of comparable quality to the standard Horner method.

1. Introduction

Accurate knowledge of static formation temperatures (SFTs) is important to ensure efficiency and safety when drilling wells in geothermal fields, as well as to determine whether development is possible (Sasaki, 1987). Furthermore, SFT estimates are used to understand the temperature distribution in geothermal fields and as reference data when calibrating geothermal reservoir models. SFTs can only be measured or estimated reliably by drilling wells and measuring temperatures down the wells. Logging equipment is lowered into the wellbore after drilling is completed or paused, and the formation temperature is determined based on temperature data

obtained from the logging. However, wellbore fluid circulation required during drilling shifts the wellbore temperature considerably below the formation temperature. Therefore, since it typically takes many days for the wellbore temperature to equilibrate with the surrounding formation temperature, multiple well logs are generally needed to understand how the wellbore temperature rebounds over time. Assuming negligible internal flow within the wellbore, the wellbore temperature can be expected to gradually rebound towards the SFT state that existed before drilling the well.

To estimate SFTs, analytical methods that are based on physical theory have been used in the past as well as simple empirical equations. The well-known Horner method is one of those methods that is commonly used in practice to estimate SFTs. The Horner method uses an approximate theoretical description of the thermal recovery of the temperature in the wellbore after the completion of drilling. The method requires an estimate of the time that fluid was circulated at the well depth of interest, and multiple bottom-hole temperatures measured at different times after shut in. Until now, the Horner method has been widely used for SFT estimation because of its simplicity, but this method requires sufficient shut-in time and an appropriate number of logging runs, which use up time and resources. A multitude of other SFT estimation methods have also been developed. However, it is not clear from the literature which methods should be preferred as the results appear to be generally inconclusive.

In response to these problems, this study proposes developing a new estimation method that applies machine learning as an alternative approach to SFT estimation. Previously, Bassam et al. (2010) developed an artificial neural network (ANN) model with one hidden layer to estimate SFTs. Their model used only three input values, including a bottom-hole temperature observation, the observation shut-in time, and an estimate of the derivative of the temperature with respect to time. A drawback of their neural network development approach was that the field data they used for training did not include reliable measurements for the SFTs. Because of that issue, they used multiple pre-existing SFT estimation methods to derive the reference SFT values used for training their neural network. Based on that, the training process could have been negatively impacted by inaccurate SFT training data.

As the Bassam et al. training procedure reflects, it is generally difficult to find reliable temperature data for training and validating SFT estimation methods. Therefore, the present study used a wellbore simulator to generate synthetic temperature recovery data to train neural networks. Differently to Bassam et al. (2010) we looked at developing deep neural networks containing multiple hidden layers to achieve better SFT estimates.

2. Methods

2.1 Horner Method

Figure 1 shows a schematic diagram of temperature recovery in a wellbore. The graph shows how the bottom-hole temperature is lowered when drilling fluid is circulated through the well. After fluid circulation ceases and the well is shut in, the well gradually warms up towards the SFT. As it can take a long time for the well temperature to equilibrate, the SFT is usually extrapolated from a series of well temperature logs recorded after shut in.

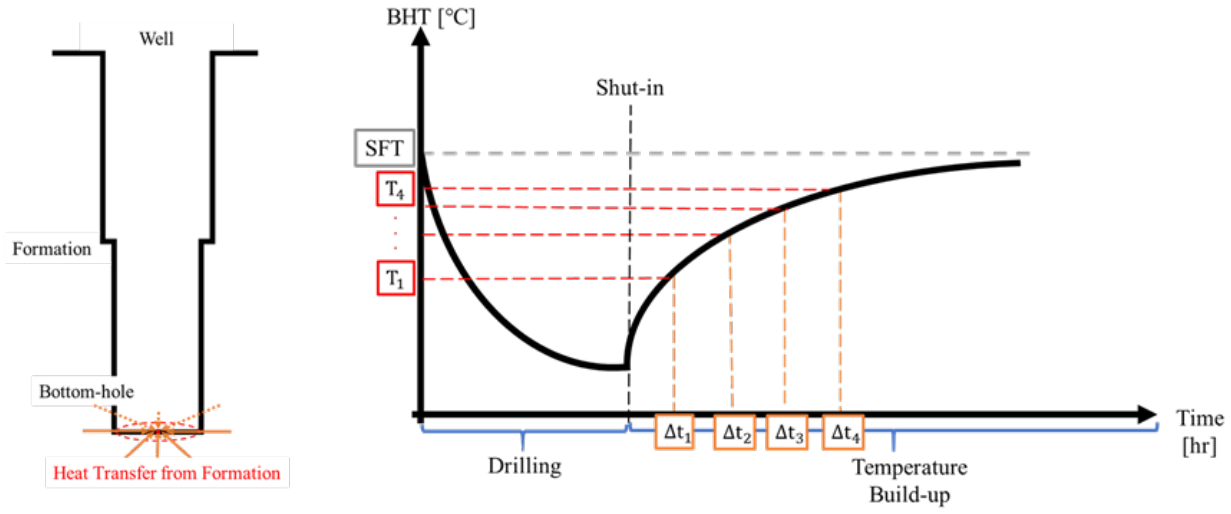

Figure 1: Recovery of bottom-hole temperature (BHT) after shutting in a well.

The Horner method has been widely used to estimate formation temperatures in geothermal wells. In the Horner approach, the SFT is estimated through linear regression by matching the following equation to the recorded temperature history.

$$T_{\mathrm{BHT}} = T_{\mathrm{SFT}} + m \times \log \frac{t_c + \Delta t}{\Delta t} \tag{1}$$

Here T_{BHT}, T_{SFT}, m, t_c, and Δt are the bottom-hole temperature, the SFT, slope, the fluid circulation time at bottom hole, and the elapsed time after shut in, respectively. The assumption is that the downhole temperature changes follow (1) when a sufficiently long time has elapsed.

The Horner plot equation (1) can be derived by describing the well as a line heat-sink drawing in a constant heat flux while the drilling fluid is circulating (see, e.g., (Kutasov and Eppelbaum, 2018)). However, that is a simplification of reality since the temperature conditions and heat flux across the well wall vary during operation. Similar simplifications have been used to derive various other analytical equations for SFT extrapolation. As a result, the applicability of those methods can be limited. Instead of using some of the standard analytical equations, we consider using ANNs to derive an empirical relationship between wellbore temperature recovery data and the SFT.

2.2 Neural Network Approach to SFT Estimation

2.2.1 Neural Network Architecture

This research aimed to develop an SFT estimation model using an ANN algorithm. The neural network model was developed using TensorFlow in Python. The ANN model in this study is a supervised learning model, in which the parameters of the model were adjusted so that it can estimate SFT values by learning hidden relationships between the ANN input data and the SFT values to be estimated.

Figure 2 shows a diagram of the multi-layered feed-forward neural network model developed in this study. The figure shows that the neural network outputs an SFT prediction based on nine model inputs: the circulation time, four observation times, and four corresponding bottom-hole

temperature observations. Figure 1 provides an example illustration of the four pairs of observation times (Δt_i) and temperatures (T_i) that are used as model inputs. In practice, available logging data can be limited and SFT estimation is sometimes carried out based on as few as three or even (with difficulty) two temperature observations. For reliability, however, we suspect that at least four observations might be needed to reveal useful information about the nonlinearity of the temperature transient. The number of temperature inputs were chosen based on that expectation and for the purpose of generating a simple neural network algorithm that would be applicable to cases with observations from only a few well logging runs.

A linear activation function was applied after the output layer of the model and a ReLU (rectified linear unit) activation function was used between all the layers. The number of nodes and hidden layers are not specified in Figure 2 since we tested out different combinations to optimize the ANN design. However, the tested neural networks were constructed with an equal number of nodes in each hidden layer. The following section describes the optimization of the network design.

2.2.2 Optimization of the Neural Network

To optimize the ANN model structure, we trained multiple neural networks using different combinations of layer and node numbers. Those hyperparameters were optimized using simple grid search with the number of hidden layers varied from one to five, and the number of nodes (neurons) in a layer varied from ten to a hundred at intervals of five nodes. The RMSProp (Root Mean Square Propagation) optimization algorithm was used to fit each ANN model to the training data. We chose to minimize the mean squared error (MSE) of the match to the SFT training data while monitoring the MSE of the SFT test data for early stopping. Each optimization run was stopped when the test MSE did not improve for 500 consecutive epochs. The maximum number of epochs was set to 5,000. The combination of network weights, biases, and hyperparameters (number of nodes and layers) that resulted in the lowest mean absolute error for the test data was adopted for the final ANN model. The final model obtained through this optimization process was a network with 3 hidden layers and 30 neurons in each layer. This model was subsequently tested on validation data to check its prediction performance.

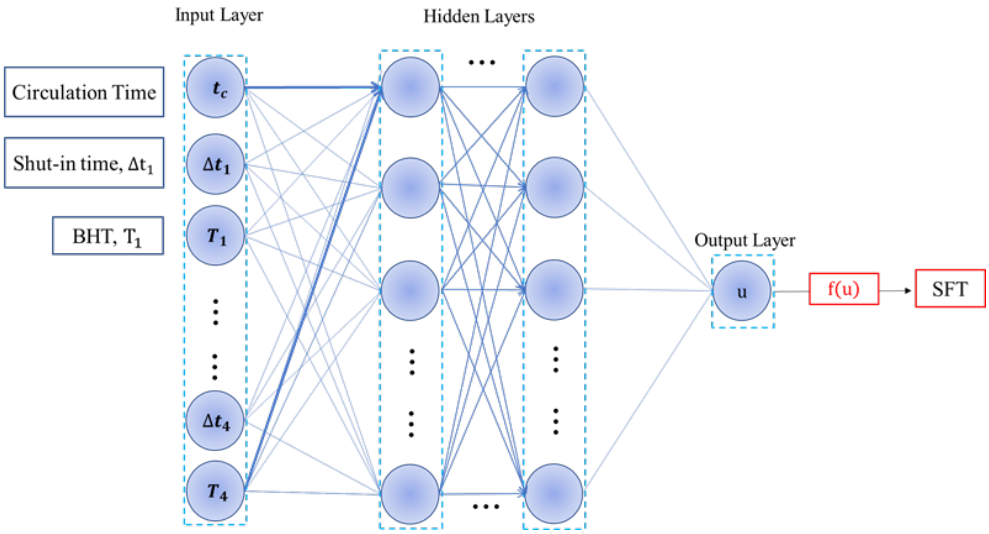

Figure 2: Conceptual diagram of the neural network used in this study.

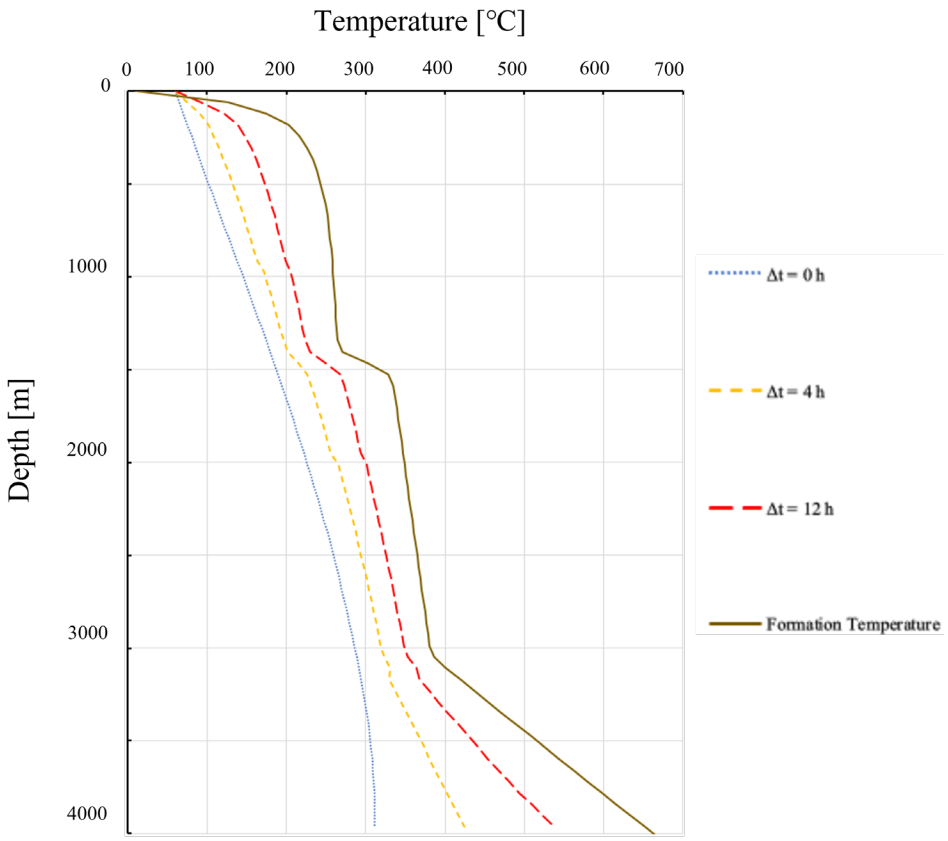

Figure 3: Example GEOTEMP2 simulation results for temperature recovery in a wellbore. Shut in is at $\Delta t = 0$. The static formation temperature shown is based on well WD-1a in the Kakkonda geothermal field, Japan.

2.3 Training Data

2.3.1 Creation of Training Data Using the GEOTEMP2 Well Simulator

The training data used to train the SFT estimation model was generated using a wellbore temperature simulation program called GEOTEMP2 developed at Sandia National Laboratories in the U.S. (Mondy et al., 1984). We used the simulator to model thermal transport within the wellbore for both the drilling period where the well is cooled by drilling fluid circulation and the warmup period when the well is shut in. GEOTEMP2 numerically solves the temperature distribution in a wellbore using a finite volume method for the heat transfer between the downhole fluids and rock formations surrounding the wellbore.

We note that the simulator describes thermal transport caused by both convection and conduction within the wellbore. However, the heat transport within the formation is limited to conduction. As a result, neural networks developed based on the GEOTEMP2 simulations are expected to be mainly applicable for conduction dominated formations.

Figure 3 shows examples of GEOTEMP2 simulated temperature profiles for a well warming up after halting fluid circulation. In this study, GEOTEMP2 was used to assemble a dataset containing multiple series of bottom-hole temperature recovery data. In total, we generated 20,000 instances of thermal recovery data from wellbore simulations. That dataset was used to develop the presented

neural network with 70% of the dataset used as training data and the remaining 30% of the data used as test data. Distinct data series were generated by randomly allocating the formation temperature and well operation conditions.

2.3.2 Range of Input Values for Well Simulations

The performance of a machine learning model depends largely on the quantity and quality of the training data used for the model development. In particular, the ANN generalization ability is expected to be improved by using data obtained from a variety of cases. Therefore, we diversified the training data by randomly allocating input values for the GEOTEMP2 simulations. Variable drilling parameters include the target drilling depth, the fluid circulation time (t_c) after reaching the target depth, and the fluid circulation rate. The target drilling depth varied between 100 and 4,000 m, while the drilling rate was set to a constant 50 m/day for all simulations. Similar drilling rates can be expected for typical geothermal wells. In fact, according to Ong'au (2012), the average drilling rate for wells in Kenya and Iceland is around 50 m/day. The circulation time was allowed to range from 2 to 48 hours, and the flow rate was randomly set to a value between 1,000 and 4,000 L/min (see, e.g., (Ajima et al., 2023; Ishikawa and Naganawa, 2022)).

The drilling fluid was assumed to be injected down the drill string from the surface at a constant temperature of 30°C. In addition, the formation temperature conditions were randomly assigned for each simulation. For simplicity, the temperature with depth profile was assumed to be linear. Thus, we only varied the temperature gradient and the ground surface temperature. This was achieved by randomly varying the surface temperature between 5 and 30°C, and the temperature at a 4 km depth was allowed to range between 100 and 800°C (Figure 4). Note that this SFT-depth profile assumption does not cover all field conditions since actual underground temperature profiles are often nonlinear—influenced by geological structures and groundwater flow. The thermal conductivity of the reservoir formation used in all cases the default value of 1.73 W/m·K in GEOTEMP2. However, we plan to vary the formation thermal conductivity in future work.

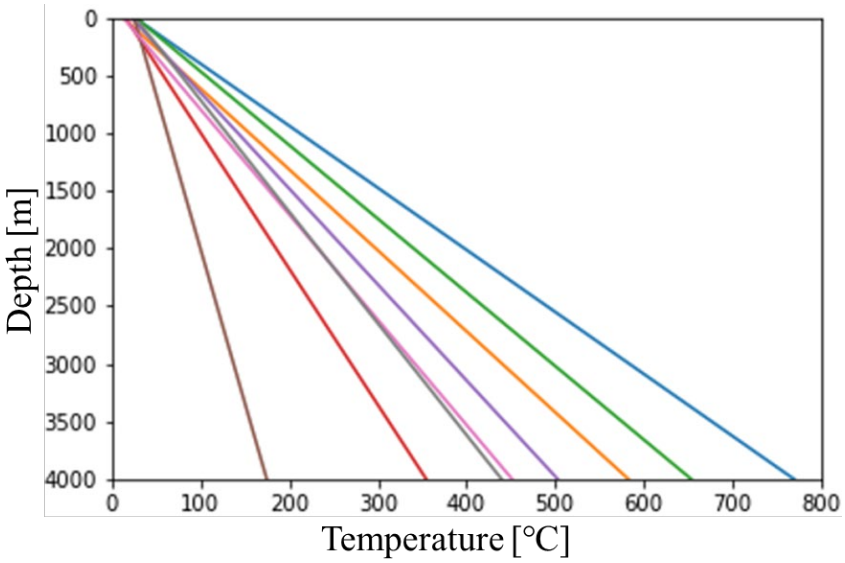

Figure 4: Example formation temperature gradients used for generating training data.

For the temperature observations, the constructed ANN model assumes that data from four logging runs are available. When generating the training dataset, the fourth observation time, Δt_4, was randomly varied from 10 to 250 hours, and the values for Δt_1 to Δt_3 were randomly selected within the range from one to Δt_4 hours. The temperature data obtained at those observation times was used as input into data for training the ANN model. The range of observation times was set to encompass the validation data used in this study.

2.4 Validation Data

The accuracy of the developed ANN static formation temperature estimation model was evaluated using data that is independent of the training data. In this study, the validation data includes synthetic data based on well WD-1a, a super-hot geothermal well drilled in the Kakkonda area of Iwate prefecture, Japan (Ikeuchi et al., 1998). The Kakkonda validation data was generated using the GEOTEMP2 simulator. However, unlike the formation temperature profiles used for training, the shallow Kakkonda temperature profile is characteristic of a convective, high-temperature geothermal field (Figure 3).

Table 1: Synthetic Kakkonda validation data for different bottom-hole depths.

2,000 m Depth		3,000 m Depth		3,700 m Depth		4,000 m Depth	
T_{SFT} = 349.3°C		T_{SFT} = 380.6°C		T_{SFT} = 580°C		T_{SFT} = 670°C	
Δt [h]	T_{BHT} [°C]	Δt [h]	T_{BHT} [°C]	Δt [h]	T_{BHT} [°C]	Δt [h]	T_{BHT} [°C]
46	304.7	5.2	295.8	73.1	561	70.5	647.9
64.3	310.9	24	353.7	75.7	561.5	78.8	649.7
74.7	313.77	113.4	368.9	145.1	568.1	104.9	653.2
86.3	316.33	163.4	371.4	180.6	569.8	132.1	655.7

Table 2: Validation data from the literature (Bassam et al., 2010).

Shen and Beck (1986)		Cao et al. (1988)		Iglesias et al. (1995)		Steingrimsson and Gudmundsson (2006)	
T_{SFT} = 80°C t_c = 5 h		T_{SFT} = 120°C t_c = 5 h		T_{SFT} = 169°C t_c = 5 h		T_{SFT} = 240°C t_c = 2.5 h	
Δt [h]	T_{BHT} [°C]	Δt [h]	T_{BHT} [°C]	Δt [h]	T_{BHT} [°C]	Δt [h]	T_{BHT} [°C]
15	69.6	24	115.8	12	94	30	125
20	71.7	30	117.1	26.57	122	52	158
30	74.1	40	118.4	47.1	139	96	184
40	75.5	50	119.1	95.27	152	242	218

Table 1 summarizes the temperature recovery data generated using GEOTEMP2 based on the temperature profile of well WD-1a in Kakkonda. The fluid circulation time was commonly set to 48 hours for the cooling operation after the target drilling depth was reached, and the simulations were run for four different bottom-hole depth cases. The flow rate was set to 2,000 L/min for all four cases.

Moreover, we also considered validation data used previously by Bassam et al. (2010). Table 2 summarizes this additional validation data. It includes synthetic warmup data reported by Shen and Beck (1986) and Cao et al. (1988). Those datasets have known SFT values that are useful for validation. For real geothermal field examples, we considered well tests reported by Iglesias et al. (1995) for well CH-A in El Salvador and Steingrimsson and Gudmundsson (2006) for well KJ-21 in Iceland. The values reported in Table 2 for those two wells are taken directly from (Bassam et al., 2010). We note, however, that we have not been able to confirm the reliability of those reference values and we consider that the reference SFT values reflect plausible values.

For the datasets taken from the literature, we selected the last four reported temperature observations. An exception is the data from well KJ-21 reported by Steingrimsson and Gudmundsson (2006). For well KJ-21, we ignored data recorded at 556 hours of shut in.

3. Results

The newly developed ANN model was evaluated using the validation data presented in Tables 1 and 2. Figure 5 shows the results of SFT estimation using the newly developed ANN model and the Horner method. The SFT estimates are compared against the reference SFT values for the validation data. The straight line drawn in the figure has a slope of 1, indicating that the closer the data points are to the straight line, the smaller the estimation error. The figure shows for both the ANN model and the Horner method that the estimation errors are within or close to ±5% for all validation data.

Table 3 summarizes the results of Figure 5 in more detail. From left to right, there are the reference SFT values for the test data, the estimated values by the ANN model, the estimated values by the Horner method, the deviation (error) between the ANN model and the reference (true) values, and the deviation (error) between the Horner method and the reference (true) values. The maximum estimation errors of the ANN model and Horner method for all validation data are respectively 8.0% and 6.8%. The average error is 3.2% for the ANN model and 2.7% for the Horner method. The ANN approach, therefore, performed nearly as well as the standard Horner method for the presented test cases.

The maximum estimation error of the ANN model is 0.6% for the Kakkonda synthetic data, which assumes a super-hot geothermal well. This indicates that the developed model can estimate formation temperatures above 400°C, which is considered to be in the high-temperature range, with good accuracy. On the other hand, the ANN approach showed larger errors for the low-temperature data, with a particularly high error for the synthetic dataset from Shen and Beck (1986). The error for the Cao et al. (1988) synthetic dataset is surprisingly high considering that the last temperature observation is within 1% of the true SFT. The Horner method gave considerably better estimates for those two datasets. However, the ANN model markedly outperformed the Horner method for the Kakkonda data. These results suggest that there might be

some features in the Shen and Beck, and Cao et al. datasets that are not reflected in the training dataset. This aspect needs to be explored further in future work.

The ANN estimates for the two geothermal wells (CH-A and KJ-21) appear to be plausible, which is an encouraging result considering that the field measurements are affected by measurement errors. Furthermore, the temperatures in those wells are affected by fluid flow through the surrounding formations, unlike the temperature simulations that generated the training data. Therefore, the ANN approach could potentially be robust to measurement errors and some conceptual differences between the heat transport governing the training data and field data.

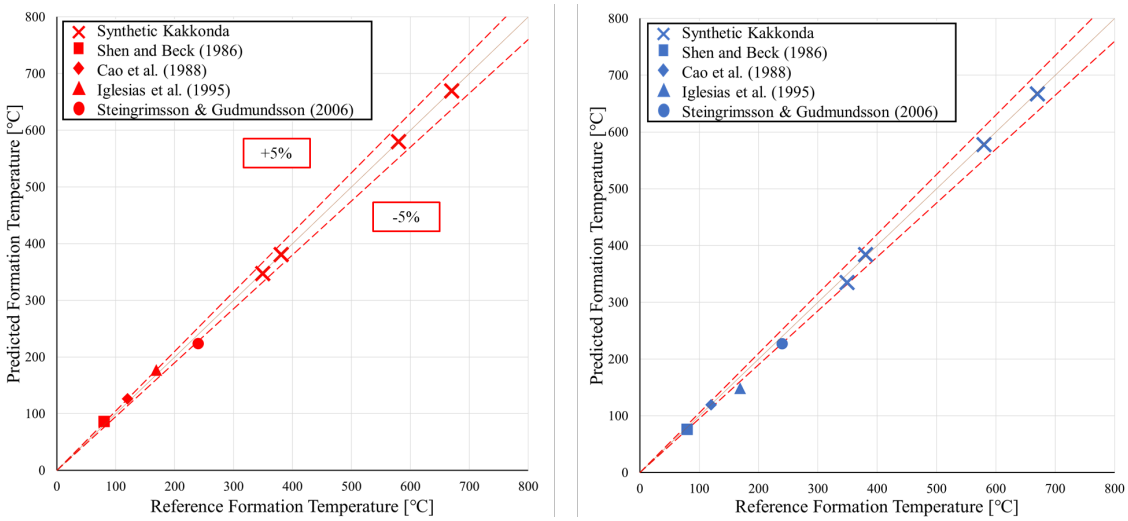

Figure 5: Comparison between reference static formation temperature data and estimated values using the developed ANN model (Left) and the Horner method (Right). The solid diagonal line indicates perfect agreement, while the red dashed lines signify errors of 5%.

Table 3: Results for estimated static formation temperature (SFT).

		Estimated SFT [°C]		Estimation error [%]	
	Reference SFT [°C]	ANN model	Horner method	ANN model	Horner method
Kakkonda 4 km	670	669.7	667.1	0.0	0.4
Kakkonda 3.7 km	580	580.0	577.5	0.0	0.4
Kakkonda 3 km	380.6	380.6	384.1	0.0	0.9
Kakkonda 2 km	349.3	347.3	334.6	0.6	4.2
Steingrimsson and Gudmundsson	240	223.5	223.59	6.9	6.8
Iglesias et al.	169	177.0	158.93	4.7	6.0
Cao et al.	120	126.3	122.5	5.3	2.1
Shen and Beck	80	86.4	79.5	8.0	0.6

4. Conclusions

In this study, we aimed to develop a new method for estimating the SFT in geothermal wells using an ANN algorithm. The neural network method estimates the SFT from transient temperature recovery data recorded while the well warms up after halting fluid circulation. The neural network was trained using a synthetic temperature recovery dataset generated using a wellbore simulator that describes temperature changes caused by fluid circulation during drilling and the following warmup period after the well is shut in.

The ANN estimation errors were found to be within 1% for synthetic validation data generated using the same wellbore simulator as was used to create the training data. However, the ANN errors were around 5–8% for synthetic validation data reported elsewhere in the literature. Moreover, when applied to field data from two geothermal wells influenced by natural convection, the ANN predictions deviated by about 5–7% from the expected reference values. On average, the performance of the ANN approach was similar to the standard Horner plot method. The Horner method resulted in an average percentage error of 2.7% while the average was 3.2% for the ANN approach. A more detailed comparison shows that the ANN method outperformed the Horner method for synthetic datasets generated using the GEOTEMP2 wellbore simulator used in this study. For the synthetic GEOTEMP2 validation data, the average error was 0.2% and 1.5% for the ANN approach and the Horner method, respectively. The ANN errors were noticeably worse for synthetic validation data generated by other means in the literature. This indicates that the data used to train the ANN model may lack relevant features and there is room to improve the training dataset.

Although additional testing is needed, these results indicate that this machine learning approach holds promise for estimating SFTs in geothermal wells. However, improvements are needed to make the presented ANN approach widely applicable. Future work should consider extending the training dataset and other ways of improving the optimization of the ANN model parameters and possible enhancements of the model architecture. The neural network presented here uses a fixed number of temperature inputs. In future work, it may therefore be worth considering extending this approach to allow more flexibility in the number of inputs. Furthermore, we are interested in exploring the potential advantages of this approach for short observation histories for which traditional methods, such as the Horner method, usually fail to deliver good predictions. We also think it is necessary to examine more field data for validation purposes.

REFERENCES

Ajima, K., Naganawa, S., and Bjarkason, E. K. "Application of insulated drill pipe to supercritical/super-hot geothermal well drilling." *Proceedings, ASME 2023 42nd International Conference on Ocean, Offshore and Arctic Engineering*, (2023).

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), 1191–1199, (2010).

Cao, S., Lerche, I., and Hermanrud, C. "Formation temperature estimation by inversion of borehole measurements." *Geophysics*, 53(7), 979–988, (1988).

Iglesias, E. R., Campos-Romero, A., and Torres, R. J. "A reservoir engineering assessment of the Chipilapa, El Salvador, geothermal field." *Proceedings, World Geothermal Congress*, Florence, Italy, 1531–1536, (1995).

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), 591–607, (1998).

Ishikawa, S., and Naganawa, S. "Development of downhole cooling charts to prevent drilling problems in geothermal wells through wellbore temperature simulation." *Proceedings, 44th New Zealand Geothermal Workshop*, (2022).

Kutasov, I. M., and Eppelbaum, L. V. "Utilization of the Horner plot for determining the temperature of frozen formations—A novel approach". *Geothermics*, 71, 259–263, (2018).

Mondy, L. A., and Duda, L. E. "Advanced Wellbore Thermal Simulator GEOTEMP2 User Manual." Sandia Report SAND-84-0857, Sandia National Laboratories, Albuquerque, NM, (1984).

Ong'au, M. T. "Controlled directional drilling in Kenya and Iceland (time analysis)." *GRC Transactions,* 36, (2012).

Sasaki, S. " A reliability of Horner-plot method for estimating static formation temperature from well log data." *Journal of the Japanese Association for Petroleum Technology*, 52(3), 243–252, (1987).

Shen, P. Y., and Beck, A. E. "Stabilization of bottom hole temperature with finite circulation time and fluid flow." *Geophysical Journal of the Royal Astronomical Society*, 86, 63–90, (1986).

Steingrimsson, B., and Gudmundsson, Á. "Geothermal borehole investigations during and after drilling." *Presented at Workshop for Decision Makers on Geothermal Projects in Central America*, El Salvador, 10 pages, (2006).