Artificial Intelligence Approaches for Sustainable Geothermal Energy Systems: With A Case Study

Füsun S. Tut Haklıdr1, Feyza Özen2, Rusen Halepmollarî3, Zeynep Korkmazi4, Mehmet Haklıdr2

1Zorlu Enerji, İstanbul, Türkiye, 2TÜBITAK BILGEM b3LAB, Kocaeli, Türkiye

fusuntut@gmail.com, feyza.ozen@tubitak.gov.tr, rusen.halepmollaris@tubitak.gov.tr, zeynep.korkmaz@zorlu.com, mehmet.haklidir@tubitak.gov.tr

Keywords: Geothermal energy, artificial intelligence, predictive maintenance, digital twin, geothermal power plants

ABSTRACT
The energy production potential of renewable energy sources is expanding rapidly, and increased CO2 emissions in the atmosphere indicate the importance of using environmentally friendly energy production options. Although these clean energy options offer substantial solutions for a better world, the energy efficiency of most renewable energy systems is still not competitive with other energy sources, and the systems need to be improved. Artificial intelligence (AI) technologies have begun to be applied in various areas and can be a good solution to increase the existing system efficiency of power production from renewables. Geothermal energy is one of the major players among renewables; however, it is costly and a risky investment, and it requires comprehensive resources and plant management. AI technologies can be integrated to the geothermal energy systems and these new technologies to increase the system efficiency and data management effectiveness and sustainability by applying correct methods for optimization and control systems with different AI algorithms both geothermal reservoir and the power plant parts especially in complex geothermal power systems such as flash, multi-flash and advanced geothermal power cycles.

The present study proposed that to apply digital twin in geothermal power systems to increase the plant efficiency and to extend the lifetime of power plant equipment such as reinjection pump, which is quite critical for reinjection application in geothermal power plants. This paper mainly focuses on the predictive maintenance part of a collaborative effort to build a digital twin of reinjection pumps in a geothermal power plant.

1. INTRODUCTION
Although using geothermal energy and other renewables for energy production is a good option for the future, there are some technical challenges to the expansion of the renewable energy industry. These include capacity factors of the systems, critical equipment failures, balancing of the flow of energy, the requirement of effective energy storage technologies and controlling of hybrid energy systems. Thus, effective management and optimization of an energy system and using energy sources more efficiently look like better options than to install a new energy system in the future. Beside all these issues, geothermal energy systems require additional attention for optimization of reservoir conditions and efficient geothermal fluid production to generate long-term energy. At this point, artificial intelligence-based control and optimization systems are expected to increase plant efficiency.

As of 2022, geothermal power capacity is recorded around 17 GWe and the top countries for geothermal power production are the United States (3,794 GWe), Indonesia (2,356 GWe), the Philippines (1,935 GWe), Turkey (1,682 GWe), New Zealand (1,037 GWe) (thinkgeoenergy.com).

Different geothermal power cycles can be used for power production based on geothermal reservoir types and reservoir temperatures. Basically, dry steam power plants are suitable for power generation for steam-dominated reservoirs, while flash, binary and advanced (flash + binary cycles) plants have been used for power production in water-dominated reservoirs. Worldwide, total installed geothermal power capacity is mainly provided from flash and multi-flash power plants (around 63%), and the rest of the capacity is provided from dry steam (21%), binary (15%) and back pressure systems (1%) (Huttrter, 2020). Hot dry rock (HDR) is different from other systems in that there is no flow from the reservoir.

Technological developments started to provide new opportunities for geothermal energy production by hybrid energy systems such as geothermal-solar (McTigue et al., 2018) like Alaşehir (Manisa-Türkiye) geothermal-solar hybrid power plant, geothermal-hydrogen production (Karakiçik et al., 2019), zero-emission power plants (Bonalumi et al., 2017) and downhole geothermal power production from oil and gas wells (Wang et al., 2018). One other issue is to use AI technologies to increase system efficiency with system optimization, determination equipment failure and fault detections by predictive maintenance approach and controlling of different processes or well management at geothermal systems.

AI based applications have been started using in fault detection systems for geothermal heat exchangers (Casteleiro-Roca et al., 2016), optimizing well emplacement in geothermal reservoirs (Akin et al., 2010), prediction rate of penetration of drilling (Díaz et al., 2019), controlling calcium carbonate and silica scaling (Tut Haklıdr and Haklıdr, 2017), prediction of geothermal reservoir temperatures
Tut Haklıdr et al.

(Porkial et al., 2015; Pérez-Zárate et al., 2019; Tut Haklıdr and Haklıdr, 2019), evaluation of the reservoir tests (Tian and Horne, 2019), using machine learning for geological CO\textsubscript{2} sequestration monitoring design (Chen et al., 2018) and geothermal well optimization (Tut Haklıdr 2020) Previous studies have generally used the ANN-based approaches for prediction, and new AI approaches, such as machine learning and deep learning methodologies, are still open new research areas for now.

2. DIGITAL TWIN TECHNOLOGY

The Internet of Things (IoT), which has become very widespread with Industry 4.0, facilitates the digitization of data and enables the use of many advanced technologies such as artificial intelligence, cloud computing, and digital twin. A combination of artificial intelligence, software systems, and human insight is increasingly utilized to deliver unequalled performance in many industries including energy, manufacturing, infrastructure, energy, power and chemical where digital transformation is rapidly embraced. In this context, IoT and digital transformation offer a new technology, digital twin (DT) technology, by enabling data collection and communication. According to the Gartner report, DT was one of the top 10 strategic technology trends for 2017-2019 (Ross, 2018). DT, coined by Grieves in 2002 as a concept underlying product lifecycle management (Grieves, 2002), is a virtual representation of assets, processes, or services in the physical World (Grieves, 2005). It can separate information about a physical system, and then mirror of twin that system (Grieves, 2019). DT enables the communication between the virtual and physical worlds and allows data analysis and system monitoring to able to predictive maintenance, avoid downtime, develop new opportunities and plan for the future. DT studies are a growing field in both academia and industry in various domains such as manufacturing, aeronautics, energy, 5G/6G networks, smart city, and smart grid systems (Fuller et al., 2020). Also, manufacturers from various industries have adopted the DT concept and its applications to improve real-time monitoring of assets and save maintenance and operational costs.

2.1 Digital Twin in Energy

The utilization of DT is important for both researchers and practitioners focusing on energy efficiency/flexibility/quality, demand response, smart grid infrastructures, etc. Meanwhile, there are many research studies that aim to achieve real-time operation/system monitoring and control (Zhou, 2019), optimization (Howard, 2020), energy efficiency (Seo, 2022), predictive maintenance (Hosamo, 2022) and reduction of failure risks (Sleiti, 2022) (Figure 1). There are also industrial studies using DT to monitor the behavior of the equipment in the power transmission lines (Siemens, 2018), to increase the energy production of wind turbines (GE, 2015), and to investigate minor leaks and malfunctions to prevent serious incidents (Shell, 2016). However, in the last decade, while the body of research on DT in the energy sector has grown rapidly, there is a relatively small number of studies on geothermal energy. Osinde et al. provided a case study that designed and implemented a DT for a drilling tool head, which is a part of a mechatronic assembly system, to real-time monitor the equipment and to develop its predictive maintenance during the geothermal drilling.

Figure 1: DT Application Areas in Energy https://www.challenge.org/insights/composite-digital-twin/

2.1.1 Possible Benefits of Applying Digital Twin Technology in Geothermal Energy

In systems where energy is obtained from geothermal fluid, thermodynamic changes, equipment-fluid contact, control of chemical injection systems, management of production and reinjection wells, steam turbines, heat exchanger systems must be controlled simultaneously. It is often not possible for the human operator to continuously evaluate the data coming from different processes of the plants and to recognize the problem at the first onset.

Digital twin applications, especially in geothermal power plants, can enable the digital twinning of a certain process or a critical equipment in geothermal power plants, allowing the dynamic conditions in the power plant to be monitored one-to-one in a digital environment, allowing equipment failures, an anomaly that will occur in a process to be caught at the first moment. By allowing instant analysis of multiple data received from the system, it allows intervention before the problem starts, preventing the loss of equipment, stopping electricity generation during equipment replacement and support the sustainability of a geothermal power plants. Simultaneously, it may allow simulations of different applications in existing geothermal systems in a digital environment, allowing possible results to be revealed in the power plant.
3. THE CASE STUDY: PREDICTION MAINTANANCE APPROACH OF ALAŞEHİR-1 GEOTHERMAL POWER PLANT (MANISA-TURKIYE)

Alaşehir Geothermal Power Plant is located in the Alaşehir district of Manisa and has been operated by Zorlu Energy since 2015, with an installed capacity of 48.7 MWe. In Alaşehir-1 geothermal power plant, electricity is produced by using a double flash plus binary combined cycle.

In December 2022, a solar hybrid renewable energy power plant was commissioned to Alaşehir-1 geothermal power plant. Solar Power Plant, whose required certifications were completed in April 2021, obtained a license in December 2021. With solar energy, 3.588 MWM of power is supplied from the solar power plant installed next to the facility, and 0.1663 MWM from the solar power plant installed on the roof of the facility. The solar energy installed on the roof was commissioned in January 2023 and solar energy contributes to the internal consumption of Alaşehir-1 GPP and it is the first geothermal-solar hybrid power plant in Türkiye (Figure 2).

3.1 Working Principle of the Power Plant

The geothermal fluid is coming from 13 production wells converges in a single line within the power plant and enters the high pressure (HP) separator at approximately 160.1 °C. Liquid phase geothermal brine and steam are separated in the separator. The steam part is directed to the HP demister and the liquid phase geothermal brine part to low pressure (LP) separator by piping line. The saturated steam coming out of the HP separator enters the HP demister and the saturated steam is separated and sent to the HP turbine. The liquid phase geothermal brine coming out of the HP separator enters the LP separator and is separated again as liquid phase geothermal brine and steam. The steam part is sent to the LP demister and the liquid phase is sent to the brine injection pumps. Saturated steam from HP demister enters the HP turbine and saturated steam from the LP demister enters the LP turbine. A special double flash+binary combined cycle technology is utilized in this plant. The steam turbine generates electricity by utilizing HP and LP steam lines, while ORC or binary turbine generates electricity by using a working fluid that evaporates at lower temperatures. ORC unit has two different heat exchangers being fed by different sources as can be seen below. Steam vaporizer utilizes the energy available at the HP turbine exhaust while brine vaporizers use liquid-phase geothermal brine after the LP separator before reinjection. The mixed brine and steam condensate are pumped back into 9 re-injection wells using three re-injection pumps (Figure 3).

Figure 2: Alaşehir-1 geothermal-solar hybrid power plant in Türkiye (Photo credit: Zorlu Energy Company)

Figure 3: Process flow diagram of Alaşehir-1 GPP
3.2 The Methodology of the Study

3.2.1 Site Selection

Within the scope of this study, it was decided to apply digital twin technology in the reinjection pumps of Alaşehir-1 geothermal power plant and predictive maintenance was started (Figure 4).

Reinjection pumps are critical equipment for the power plant where the fluid temperature drops, the fluid is used to transfer the fluid to the reinjection wells smoothly, energy consumption is high, and minerals such as silica tend to precipitate due to the decrease in fluid temperature.

![Re-injection pumps in Alaşehir-1 GPP](image)

**Figure 4: Re-injection pumps in Alaşehir-1 GPP (Photo credit: Zorlu Energy Company)**

3.2.2 Accepting Benefits of Predictive Maintenance Approach

Predictive maintenance (PdM) is a concept that uses data-driven techniques to effectively manage asset maintenance plans by predicting failures. Data is collected over time in these scenarios to monitor the state of the equipment.

PdM uses AI (Machine Learning) techniques to learn from data collected over a certain period of time and uses live data to identify certain patterns of system failure, as opposed to conventional maintenance procedures that rely on the life cycle of machine parts.

The ML-based predictive approach analyzes live data and attempts to determine the correlation between certain parameters to predict system failure or schedule equipment maintenance.

ML technology helps identify fault lines by predicting failures at the right time and thus maximizing resource utilization. This ensures a balance between maintenance needs and resource utilization (Figure 5).

![Predictive Maintenance Process](image)

**Figure 5: Predictive Maintenance Process**

3.2.3. Data Collection

IoT in the geothermal industry can drive predictive maintenance, making equipment more efficient and increasing its availability. For instance, using sensors for pressure, temperature, vibration, corrosion, and flow can help track the performance of pumps, motors, and valves to generators and cooling towers. In this study, data obtained from 3 reinjection pumps labeled A, B and C of Alaşehir-1 geothermal power plant, collected between March 13, 2021 and August 27, 2023, were used. Data flow diagram in project is given Figure 6.
3.2.4. Preprocessing

Pump C was selected to be used in the detailed study on the analysis and predictive maintenance of A, B and C re-injection pumps. The operating times of the pumps were determined by analyzing the discharge flow data. It was found that only two pumps were running at any given time and one remained idle. Based on this analysis, periods when Pump C was not operational were identified and subsequently removed from the dataset.

The feature set of Pump C consists of vibration, temperature, pressure and flow rate measurements. In addition to these measurements, the dates of periodic maintenance operations such as lubrication and oil change are also known. Based on the examination of the data, it was seen that the pump was shut off and started again during these maintenance dates. Therefore, these maintenance dates were used as labels in the dataset and also became features.

During the time the data was collected, the pump malfunctioned only once. The entire time period between two weeks before the failure date and 5 days after the failure was labeled as an anomaly to test the model. Since the model used is an unsupervised learning model, anomaly labels were used only in the testing phase of the model. The data is divided into training and testing sets. At this point it was split down the middle to ensure that the malfunction date remained in the test data. The train set consists of normal values only. This division process is important for the Autoencoder model.

The statistical distribution of each variable was taken into account when scaling the data. The data were scaled by the method chosen according to the statistical distribution analysis of each variable. Quantile-Quantile Plot (QQ Plot) method was applied to analyze its distribution (Figure 7). As a result of this analysis, it was seen that the variables did not have a normal distribution, therefore the data was normalized and scaled by applying Min-Max Scaling.
Scaled time series data were converted into sequences so that they could be used as input to the LSTM autoencoder model. At this stage, time series data are divided into 240-hour sequences. This length was decided experimentally, taking into account computational limitations. Anomaly and normal labels corresponding to each sequence were added to the sequences to be used in the testing phase to evaluate the performance of the model. As a result of the separation into sequences, a 3-dimensional array was formed, each showing the 10-day behavior of the pump (5888, 240, 13).

After the sequences were created, the training sequences were shuffled, thus preventing the model from learning in a fixed order and preventing overfitting problems that may occur. Lastly, a validation set was created by splitting the training sequences into two parts at ratios of 0.8 and 0.2.

3.2.5. LSTM-Autoencoder Model

To solve the predictive maintenance problem, the LSTM-Auto-Encoder model was used, which is a combination of LSTM (Long Short Term Memory), which is a neural network architecture suitable for processing time series data, and Auto-Encoder, which is used for anomaly detection in cases where anomalies are rare (Elsayed et al., 2020). Traditional machine learning models are not good at learning the sequential structures and time dependencies of time series data. LSTM neural networks solve this problem and make the model used suitable for the problem. However, since it is not sufficient on its own, it has been combined with the Autoencoder model. Autoencoders are trained with normal data only. In this way, they reconstruct the normal input values they encode as normal data. When test data is given as input to the Autoencoder model, it will reconstruct it according to the normal data it has learned, and points with high reconstruction errors can be detected by the model as an anomaly, that is, a malfunction, in this problem.

The LSTM-Autoencoder architecture consists of two main components: encoder and decoder and both consist of LSTM layers. The encoder part processes the input time series data and reduces it to a condensed representation. Thus, key features and temporal patterns in the data are captured and the input sequence is encoded in a compressed format. The first layer of the model is the 100-unit LSTM layer. This layer, which forms the first part of the encoder, starts compressing the input sequences and the “Relu” activation function is used in this layer (Ide and Kurita, 2017). A dropout layer was added immediately after the LSTM layer. This layer prevents overfitting by randomly skipping 20% of features during training and the encoding process is completed at this point.

The decoder part starts reconstructing the original time series data by decoding the representation compressed by the encoder. The reconstructed data must match the original data. Anomalies, i.e. malfunction, in this problem are defined based on the reconstruction error. Larger reconstruction errors indicate that there are significant deviations from learned normal patterns, meaning that these situations are anomalies. The first layer of the decoder is the RepeatVector layer. It is a connection layer between the encoder and decoder. The second LSTM layer is a reflection of the first. Similar to the encoder, the LSTM layer in the decoder is followed by a dropout layer. The last layer of the model consists of the TimeDistributed wrapper applied to the Dense layer.

During the training phase, the model was fine-tuned with the Adam optimizer (Lugsi et al., 2021) and mean square error (MSE) was used to calculate the accuracy. In this way, the difference between the values predicted by the model and the actual values was tried to be reduced. If the reconstruction error calculated with MSE exceeds the 99.5% of the MSE distribution, the processed and predicted sequence is labeled as an anomaly. This threshold value was chosen as the balance value to reduce false positives and ensure early anomaly detection.
3.2.5. Model Evaluation

The model was run for 10 epochs, and as a result, it was observed that anomaly detections were made in three separate regions throughout the time series (Figure 8). One of the three detected regions is the region where the actual malfunction is located, and the model successfully completed the early malfunction detection task by detecting anomaly during the 9 days and 2 hours remaining until the failure. As a result of the evaluation process, confusion matrix was obtained and according to this matrix, 0.88 precision, 0.70 recall and 0.78 f1-score were obtained (Figure 9). Recall value is the ratio of the situations that the model detects as anomalies and are actually anomalies (true positive) to the sum of the true positives and the situations that the model predicts as "normal" but are anomalies (false negative). Precision is the ratio of true positive values to the sum of true positives and false positives that the model predicts as anomalies but are not actually anomalies. F1-score is the harmonic average of precision and recall values. Since the true value in our problem, that is, the cases labeled as anomalies, are quite rare, the recall value is even more important. High recall indicates that anomalies were not missed, which is the main issue in the predictive maintenance problem.

![Discharge Flow with Anomalies](image1)

**Figure 8: Identified Anomaly Indexes on Time Series Data**

![Confusion Matrix](image2)

**Figure 9: Confusion Matrix of LSTM-Autoencoder Model**
3. CONCLUSION

To ensure the sustainability of geothermal systems and power plants, it is important to optimize the field so that the reservoir can be properly managed. In addition, especially in geothermal power plants, new technologies should be utilized to extend the life cycles of the equipment, to prevent power outages in the system and to increase the efficiency of the system. At this stage, the digital twin in application, which is one of the artificial intelligence-supported technologies, is considered as a new and effective perspective in the operation of geothermal power plants. In this context, it is expected that the continuous monitoring of the system by making a digital twin of the reinjection pump in a geothermal power plant in Türkiye will provide both an effective reinjection and an opportunity to examine the working conditions of the reinjection pump and a good case study.

In the case study, a predictive maintenance study was conducted using time series data, where anomalies were quite a minority, which is a challenging scenario. This situation was successfully overcome by using the LSTM Autoencoder model and early malfunction detection was successfully achieved.

REFERENCES


