PROCEEDINGS, 47th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 7-9, 2022
SGP-TR-223

Deep Learning Models for Prediction and Optimization of Air-Cooled Binary Cycle Geothermal Operation

Wei Ling (USC), Yingxiang Liu (USC), Robert Young (USC), Jalal Zia (Cyrq Energy Inc.), Michael Swyer (Cyrq Energy Inc.), Trenton T. Cladouhos (Cyrq Energy Inc.), Behnam Jafarpour (USC)

Mork Family Department of Chemical Engineering & Material Science, University of Southern California, Los Angeles, California, USA

wling@usc.edu

Keywords: Deep learning, neural networks, optimization, control, geothermal power plants

ABSTRACT

Real-time monitoring and optimization of geothermal operations require a reliable predictive model that incorporates the thermodynamic and thermoeconomic behavior in the power plant. The effect of ambient temperature on an air-cooled binary cycle power plant is complex and costly to model using a physics-based predictive model. Fit-for-purpose data-driven predictive models offer a cost-effective and efficient alternative to physics-based models for energy production prediction and online optimization. Neural networks offer powerful and flexible models that can be readily constructed, trained, and updated using feedback from the real process historical data. We present a deep learning-based approach for online prediction and optimization of geothermal operations while accounting for disturbances of ambient temperature and brine supply. The model predicts power output and thermal efficiency by propagating the measurement variables and by including the influence of control and disturbance variables in the form of a feed-forward neural network. Once trained, the deep learning model is used to predict and maximize the predicted power production by automatically adjusting the working fluid circulation rate via the pump speed. We present the workflow in detail and demonstrate its performance using a quasi-dynamic model in the form of thermodynamic flowsheet simulation as well as the field data from a binary cycle geothermal power plant.

1. INTRODUCTION

Geothermal energy conversion systems extract reservoir fluids from producing wells, utilize the heat to generate electricity in a power plant, and then reintroduce the fluids into the reservoir. ORC (Organic Rankine Cycle) systems are sustainable technologies that utilize low/medium-temperature heat sources in binary cycle power plants. Air-cooling is useful for the geothermal power plants that are operated in arid areas, where water is scarce. However, ambient air temperature significantly affects the power output of air-cooled thermal power plants (Varney & Bean, 2012). The automation of geothermal operations by considering the disturbance from the environment can result in significant performance and sustainability improvement. Predictive models are traditionally derived from the physics of the problem or a reduced version of the underlying governing equations. Recent developments in data science have motivated the development of data-driven prediction models for industrial applications, including real-time control.

Wang et al. (2020) present a machine learning-based predictive model for optimization of thermal efficiency with simulation data for off-design, which provides an efficient tool for choosing the layout, working fluid, and operation condition of the ORC operation. Hu et al. (2021) use a similar approach on a solar and geothermal hybrid system to predict and optimize the thermal efficiency of an off-design operation. For online control and automation of the geothermal system, Cupeiro Figueroa et al. (2020) present a physics-based model predictive control for hybrid geothermal systems. In general, detailed physics-based models are not straightforward to construct, calibrate and maintain. Ling et al. (2021) present an online economic MPC algorithm for optimization of ORC units with a dynamic deep neural network model of the ORC operation. In this study, we present an efficient workflow for prediction and optimization of air-cooled geothermal operations using a quasi-dynamic simulator that is used to predict the operation response of the field.

2. METHODOLOGY

2.1 Overview

We present an efficient prediction and optimization approach for binary geothermal power plants. Figure 1 depicts the overview of the workflow. The geothermal power plant provides measured historical data from the operation. Offline historical data are used to train a machine learning-based model, without requiring the detailed thermodynamic and other engineering information about the equipment. Once the data-driven model has been trained and validated, it can be used to predict the performance of the system under the reference control input that has been examined and optimized under physical constraints. To evaluate the proposed framework, we apply it to simulated data from a physics-based simulation model of a geothermal field that is operated by Cyrq Energy Inc.

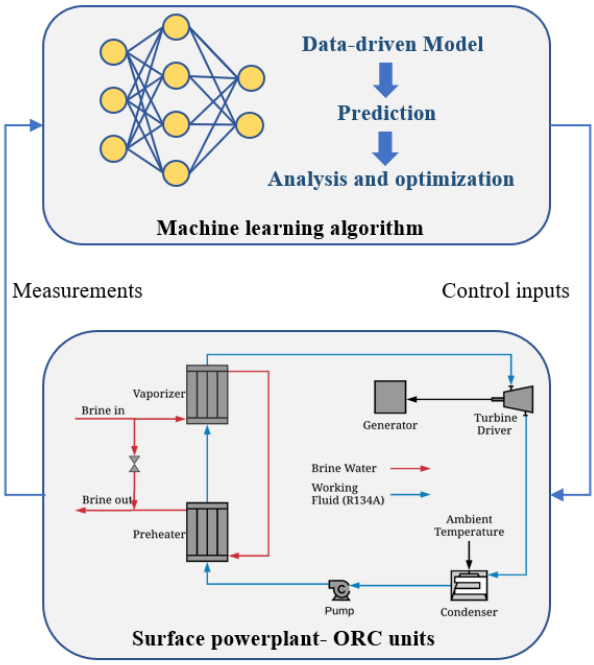

Figure 1: The overview of the developed approach.

2.2 Physics-based simulation

System description

The surface power plant part in Figure 1 shows the schematic description of a simplified ORC unit that is operated by Cyrq Energy Inc. In this process, the geothermal brine fluid is the heat source for preheating and vaporizing the organic working fluid. The higher-pressure working fluid vapor passes through a turbine and rotates the shaft connected to the generator to generate electricity. The turbine exhaust vapor then flows into an air-cooled condenser, which is condensed by ambient air. The organic liquid working fluid is then pumped back to the preheater to complete and start a new cycle. The physics-based model is developed using the DWSIM software (Daniel Medeiros, 2021), which is a chemical process open-source simulator. The challenges of building this simulation model are as follow:

- The design data has a significant discrepancy compared to historical data. While the temperature, pressure, and mass flow rate are not fully measured during the operation, the model parameters must be deduced from the available data.
- The process dynamics is only represented through the available process data, which has been post-processed to average hourly data.
- The system is affected by environmental disturbances, primarily the ambient temperature and brine supply.
- The air-cooled condenser that influences the turbine exhaust pressure (Kahraman et al., 2019) cannot be explained by the existing model.
- The working fluid R134a frequently operates above its critical pressure. The thermophysical property for some operation regions in the database is incomplete. Additionally, when the fluid is treated as the supercritical fluid the key variable of superheat temperature is unavailable.

In light of these challenges, we make the following assumptions:

- The ORC is modeled as a sequence of steady-state calculations with pump operation and disturbance as inputs. The inertial effect is neglected.
- The parameters in the thermodynamic model are estimated from the operation data.
- The fluid thermodynamics is based on the Peng-Robinson equation.
- The adiabatic efficiency of the pumps and the turbines are assumed to be 85% and 90%, respectively. The turbine is operated at a constant speed. The pump curve is fitted with a polynomial regression model.

- The turbine outlet pressure is determined by a polynomial regression model with ambient temperature input.
- The pressure drops across the piping and equipment as well as the heat losses are neglected.

Turbine:

As described in the above assumptions, the turbine is modeled under isentropic assumption and constant speed required by power grids. Figure 2 (a)&(b) depicts field data for ambient temperature, turbine exhaust temperature, and pressure for a month. The historical data indicate that ambient temperature has a strong correlation with turbine exhaust pressure. To model this relationship, using Eqs. (1) & (2), we fit a polynomial curve to predict the turbine outlet pressure when the ambient temperature $T_{amb} \geq 32$ (Figure 2); otherwise, the turbine outlet pressure p_{t_out} is a constant 0.42 MPa.

$$p_{t\_out} = 9.58e^{-5}T_{amb}^2 + 1.21e^{-3}T_{amb} + 2.84e^{-1} \quad (1)$$

$$32\ °\text{F} \leq T_{amb} \leq 100\ °\text{F} \quad (2)$$

Figure 2(c) shows the range of ambient temperature fluctuations in a year, which serves as input to the polynomial regression model in Figure 2(d).

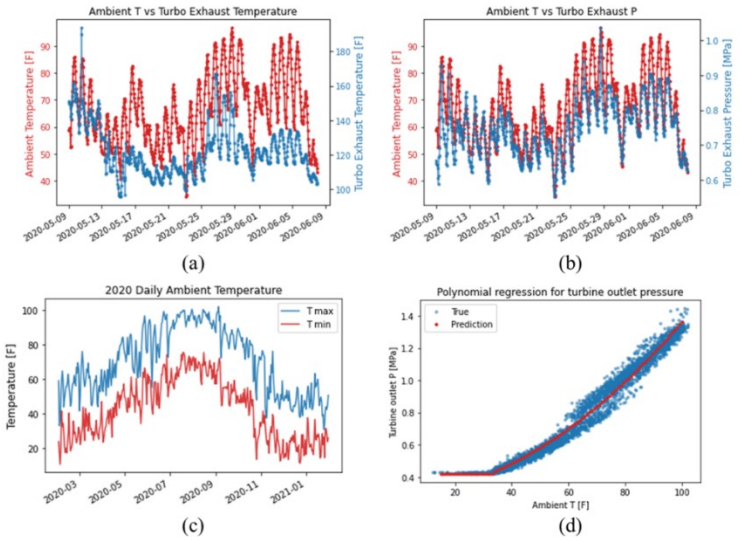

Figure 2: (a) Ambient temperature vs. turbine exhaust temperature (b) Ambient temperature vs. turbine exhaust pressure (c) Daily maximum temperature minimum temperature over a year (d) Polynomial regression model for turbine outlet pressure.

Multi-stage centrifugal pump

Due to the lack of the pump curve, we model the pump discharge pressure $p_{p_{out}}$ and volumetric flow $q_{p_{out}}$ using a 2nd order polynomial regression and linear regression with the pump speed n as input. Eqs. (3)&(4) show the exact expression for the pump curve, which is generated based on field data. Figure 3 depicts the pump curve results compared with the field data when it is available.

$$p_{p_{out}} = 1.16e^{-6}n^2 + 2.3e^{-3}n + 6.52e^{-1} \quad (3)$$

$$q_{p_{out}} = 5e^{-4}n \quad (4)$$

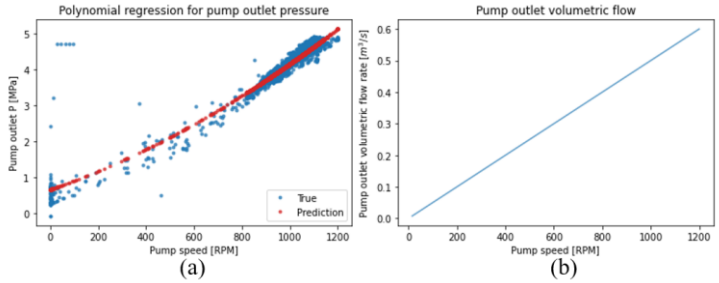

Figure 3: The simulated multi-stage pump curve of (a) discharge pressure and (b) the volumetric flow.

Heat exchangers:

There are three heat exchangers in the process: the preheater, the vaporizer, and the condenser. The preheater and vaporizer are modeled as the countercurrent heat exchangers with the potential phase described by the convection heat equation:

$$Q = UA\Delta T_{\mathrm{lm}} \tag{5}$$

where Q is the heat transferred, U is the heat transfer coefficient, A is the heat transfer area and ΔT_{lm} is the Logarithmic Mean Temperature Difference (LMTD) of the countercurrent flow of brine and R134a.

The air-cooled condenser condenses the R134a exhaust from the turbine at a pressure that is in equilibrium with the ambient temperature. The simulator utilizes two exchangers that separately condense and subcool the working fluid. The first one simply sets the outlet fluid phase to liquid, and the second one specifies the subcooled temperature drop correspondingly. Figure 4 displays the simulation model of the process in the surface power plant shown in Figure 1.

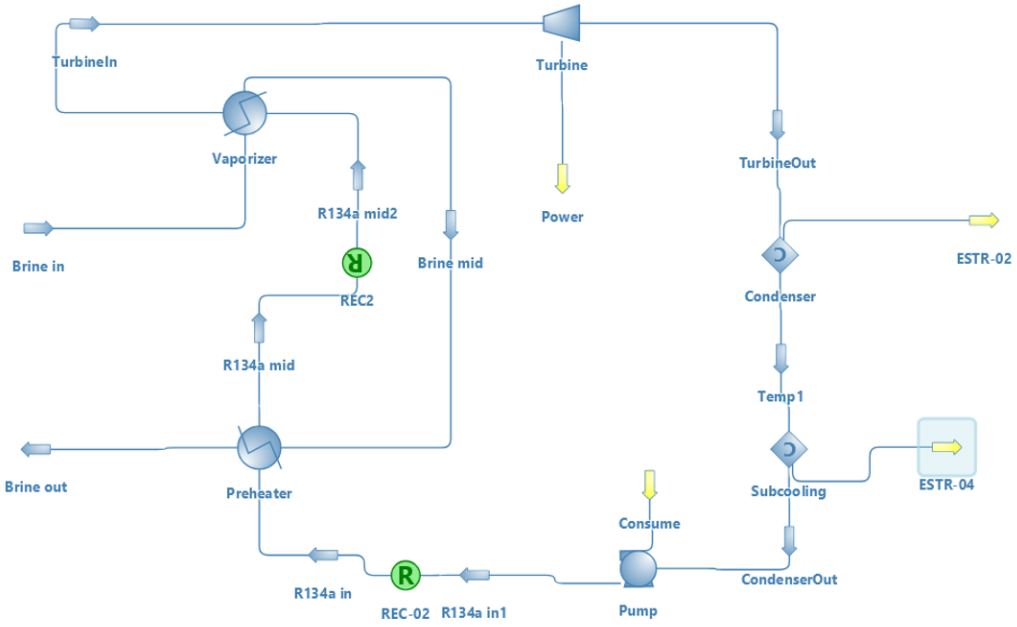

Figure 4: Diagram of the ORC model in DWSIM.

3.2 Data-driven model

Artificial Neural Network

An artificial neural network (ANN) model is utilized for learning the ORC process dataset. The multilayer structure of the model together with the use of nonlinear activation function enables a flexible structure to learn nonlinear patterns from the data. The model parameters are estimated by backpropagating the prediction error during training. Figure 5 depicts the ANN structure, which includes an input layer, several hidden layers, and an output layer. The inputs layer receives the ambient temperature T_{amb}, the brine inlet temperature T_{b_in}, pressure p_{b_in}, volumetric flow rate q_{b_out} and pump speed n. The output layer predicts the gross power generation P_g, the cost of pump operation, P_g, as well as other process variables. Once trained, the model can serve as an input-output model that can be used for optimization.

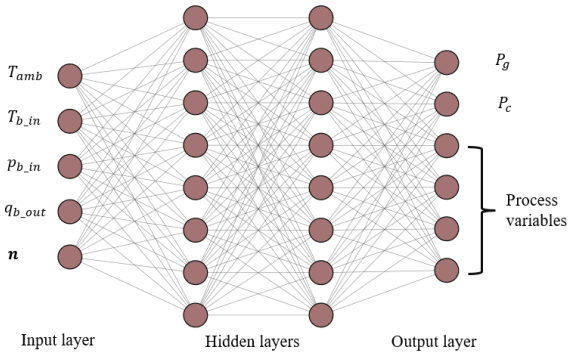

Figure 5: The schematic diagram of the ANN.

Optimization

In the presented examples, the net power generation is used as the objective function to maximize. This economic objective function is described in terms of the gross power generation and the pump operating cost as a function of the ANN model disturbances w_t and u_t, Eq.(6). Since the control input, the pump speed, is constrained to a specific range determined by the physical limits of the pump, the related lower and upper bounds are shown in Eq.(7).

$$\underset{u_t}{argmax}\ P_g(w_t, u_t) - P_c(w_t, u_t) \tag{6}$$

$$\text{subject to: } u_{min} < u_t < u_{max} \tag{7}$$

The ANN model is implemented in Pytorch (Paszke et al., 2019), and the optimization is performed using the 'COBYLA' algorithm from the Scipy library in Python 3.7 (Virtanen et al., 2020).

3. RESULT AND DISCUSSION

3.1 Evaluation of the physics-based simulator

The next step in the workflow is to evaluate the physics-based simulator by comparing the simulation results with the field data using the same historical disturbance and control input sequences. The four disturbance input variables are the ambient temperature, the brine inlet pressure, the brine inlet temperature, and the brine outlet volumetric flow rate. The control variable in this example is the pump speed, which is used to generate the output measurements. Figure 6 shows the disturbance and control inputs as well as the simulation process measurements.

The brine inputs in the first two subplots in Figure 6 (ordered from left to right and top to bottom) vary in a narrow range compared to the ambient temperature, which is considered as the main disturbance of interest during this period of operation. The net power generation (second row, first column) is calculated as the power generated by the turbine minus the power consumption by the pump, which shows a good agreement with the historical data from the field. The temperatures in the remaining subplots do not provide good agreements, especially when the ambient temperature is low. The discrepancy is likely caused by ignoring the heat losses in the system, which are larger when the ambient temperature is low.

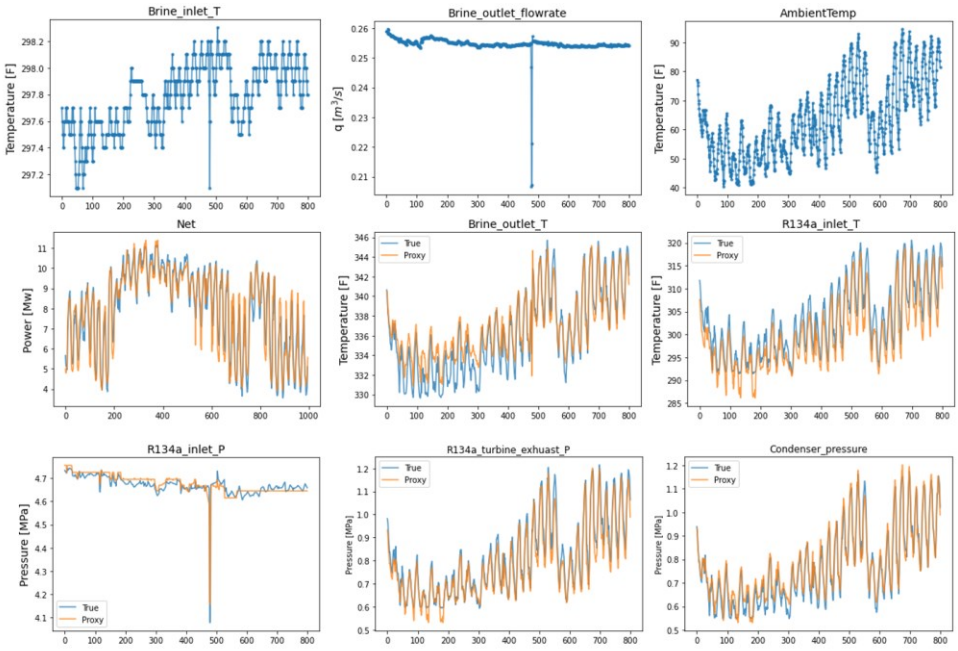

Figure 6: The simulation model result vs. the field data

Although the model is based on a steady-state assumption and there is mismatch between the simulated data and historical field data, the reproduction of the dominant variable in predicting the power generation ensures the simulator can be treated as a reasonable quasi-dynamic proxy model for generating the synthetic dataset. Given the close agreement between the simulation results and the historical data, the simulator is used to generate synthetic datasets for training the ANN model for prediction and optimization. A 10,000-hour dataset with disturbance and control inputs from the historical operation was generated to enable the training and testing of the ANN model.

3.2 Prediction

Synthetic dataset

The synthetic dataset is preprocessed by normalizing the data range between 0 and 1 and dividing it into training set and testing set with lengths 7000 and 3000, separately. The model is characterized by an input layer of 5 nodes, two hidden layers with 100 nodes, and rectified linear unit (ReLU) activation function. In this example, since there are no extra process variables monitored to address constraints in the optimization, the output layer has two nodes that represent the power production and pump operating cost. On each epoch during the training stage, the training dataset is randomly partitioned into 70% training and 30% validation. The prediction results are evaluated by Root Mean Square Error (RMSE) as shown in Eq. (8). The RMSEs for prediction of power production and pump cost are 0.03 and 0.01, respectively. The prediction results (red) are compared to the measured values (blue) in Figure 7, where the outputs from the model closely follow the general trend in the data.

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}} \sum_{t=t_0}^{N} (x_t - \hat{x}_t)^2} \tag{8}$$

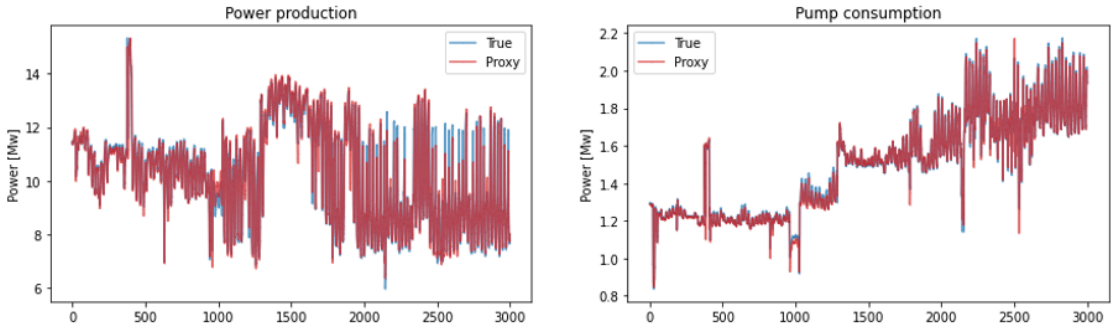

Figure 7: Prediction of power generation and pump operation consumption on the synthetic dataset.

Real field dataset

The same model is trained and tested on the Cryq field dataset to further evaluate the performance of the ANN model. The dataset is 8000 hours long, with 5000 hours for training and 3000 hours for testing. Figure 8 shows the 3000-hour prediction results for power generation and pump cost, given the disturbance and control inputs. The RMSEs for power production and pump cost are 0.069 and 0.044, respectively. The prediction performance is acceptable, although it is not as good as that of the performance in the synthetic example. Because the optimization performance cannot be evaluated without applying the controls in the field, the optimization is only demonstrated for the simulation example.

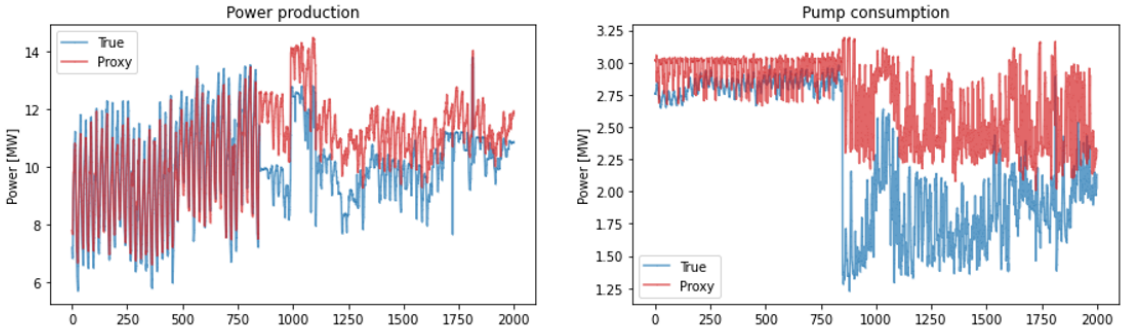

Figure 8: Prediction of power generation and pump operation consumption on actual field dataset.

3.3 Optimization

Once the ANN model is trained, the model is used to find the new control policy that maximizes the net power output under the pump physical constraints. Table 1 summarizes the average power generation, pump operating cost, and net power generation results for different methods. The optimized control inputs result in 180 kW average increase in a 10-megawatt ORC unit compared to the original control strategy. With the 0.22 $/kWh residential power price in October 2021 from U.S. Energy Information Administration website (n.d.), the 28,512$ worth of electricity is generated without extra equipment investment.

In Figure 9, the first three subplots (from left to right, top to bottom) show the simulation results after applying the new control inputs and compares that with the results from the original control. The fourth subplot compares the predicted and realized average output of the ANN model. The ANN model slightly overestimates the power generation and underestimates the pump cost, resulting in lower simulated net power generation than expected.

Table 1 Averaged output of the ANN model and simulation

Control inputs	Method	Power generation (MW)	Pump cost (MW)	Net power generation (MW)
Original control	Simulation	10.98	1.52	9.46
Optimized control	Simulation	11.64	2.00	9.64

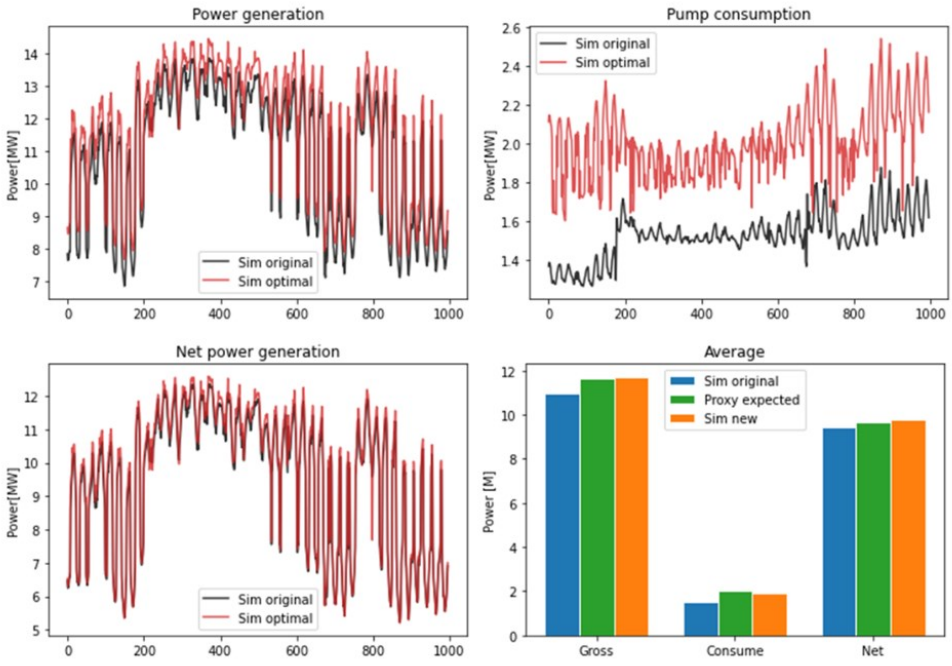

Figure 9: Optimization result

4. CONCLUSION

An efficient data-driven prediction and optimization framework is presented to improve the performance of air-cooled geothermal operations. The ANN predictive model is built purely from data and is used to predict the power generation and pump cost with a known disturbance and control inputs. The optimization results are evaluated using a physics-based simulation model that is parametrized based on an actual ORC power plant. Comparison between the data-driven predictive model and physics-based model simulator suggests that the physics-based model is challenging to implement when the process and/or engineering information is incomplete or missing, e.g., the operation curve of the pressure change equipment, the thermodynamic data on the supercritical fluid, and the system response to key disturbances. When the historical data is representative of the physics underlying the process, the data-driven model is a powerful tool for optimizing the process by only learning the patterns from the historical data. A promising area for future research is to deploy this approach online and gradually update the model with new samples to improve the performance of the predictive model and optimization result. Furthermore, the physics-based model is constructed under the steady-state assumption. An interesting avenue for future work is to integrate a dynamic simulator with the air-cooled condenser and use the resulting simulated data to train the dynamic neural network. However, the transition to the dynamic model requires detailed information about the operation and equation-based process simulator. The use of efficient data-driven prediction models in solving optimization problems can enable online automation for complex economic objective functions at a higher frequency.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Geothermal Technologies Office award number DE-EE0008765. Support from Energi Simulation Foundation Industry Chair Program is hereby acknowledged.

REFERENCES

Cupeiro Figueroa, I., Picard, D., & Helsen, L. (2020). Short-term modeling of hybrid geothermal systems for Model Predictive Control. *Energy and Buildings*, *215*, 109884. https://doi.org/10.1016/j.enbuild.2020.109884

Daniel Medeiros. (2021). *DWSIM* (6.5.3) [Computer software]. http://dwsim.inforside.com.br

Hu, S., Yang, Z., Li, J., & Duan, Y. (2021). Thermo-economic optimization of the hybrid geothermal-solar power system: A data-driven method based on lifetime off-design operation. *Energy Conversion and Management*, *229*, 113738. https://doi.org/10.1016/j.enconman.2020.113738

Kahraman, M., Olcay, A. B., & Sorgüven, E. (2019). Thermodynamic and thermoeconomic analysis of a 21 MW binary type air-cooled geothermal power plant and determination of the effect of ambient temperature variation on the plant performance. *Energy Conversion and Management*, *192*, 308–320. https://doi.org/10.1016/j.enconman.2019.04.036

Ling, W., Liu, Y., Young, R., Zia, J., Swyer, M., Cladouhos, T. T., & Jafarpour, B. (2021). Deep Learning-Based Predictive Control for Geothermal Operations. *Geothermal Rising Conference*.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 32* (pp. 8024–8035). Curran Associates, Inc. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Varney, J., & Bean, N. (2012). *Air-cooled binary rankine cycle performance with varying ambient temperature*. *36 2*, 1125–1131. Scopus.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang, W., Deng, S., Zhao, D., Zhao, L., Lin, S., & Chen, M. (2020). Application of machine learning into organic Rankine cycle for prediction and optimization of thermal and exergy efficiency. *Energy Conversion and Management*, *210*, 112700. https://doi.org/10.1016/j.enconman.2020.112700

(N.d.). In Electric Power Monthly—U.S. Energy Information Administration (EIA). https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=epmt_5_6_a