Detection and Diagnosis of Abnormal Conditions in the Feed Pumps of a Geothermal Binary Power Plant Using Feature-based Time-series Analytics

Paul Michael B. Abrasaldo[1,2], Sadiq J. Zarrouk[1], Andreas W. Kempa-Liehr[1], Alan Mudie[2], Jonathan Cen[2], and Christine Siega[3]

[1] Department of Engineering Science and Biomedical Engineering, The University of Auckland, Private Bag 92019, Auckland, New Zealand

[2] Contact Energy Ltd., Wairakei Power Station, Taupo, New Zealand

[3] Mercury NZ Ltd, 283 Vaughan Rd, Rotorua 3010, New Zealand

pabr612@aucklanduni.ac.nz

Keywords: *Geothermal binary plant, artificial intelligence, machine learning, feature engineering, time-series analysis, fault detection, condition monitoring.*

ABSTRACT

Industries worldwide have seen innovative growth fueled by volumes of data generated by inexpensive sensors deployed across critical points of various systems. These large datasets have led to artificial intelligence and machine learning applications in designing and operating energy systems. Furthermore, analytics-driven decision-making in industrial operations is a field of research that has benefitted from the confluence of developments in data storage, capture, and processing technologies.

This study investigates the application of time-series analytics in detecting and diagnosing abnormal operating conditions in circulating feed pumps operating in a geothermal binary power plant. Pump power and speed data have shown that the pumps experience episodes of cavitation that have led to reduced performance and reliability in the past. Operators have documented five cavitation events within one year, but operators believe there may have been more events that were not recorded.

Systematic time-series feature engineering, supervised, and other machine learning methodologies were applied to detect and diagnose the drivers of the pump cavitation events. A fully-labelled dataset was used to develop predictive models to forecast the occurrence of the target events and diagnose the primary drivers that may have triggered such events in the past. The workflow deployed in this study can be used as a foundation for developing fit-for-purpose tools that can automatically detect abnormal operating conditions for different components and systems across other geothermal operations.

1. INTRODUCTION

The rise of affordable and dependable sensor technologies has triggered a rapid surge in global data accumulation, giving rise to the Internet of Things (IoT) concept. Similar to other industries, the geothermal sector has encountered the transformative impact of IoT, resulting in an unparalleled influx of data generated across every stage of the geothermal value chain. Lately, there has been a noticeable increase in scholarly articles focused on the examination and utilisation of these datasets to enhance the efficacy and dependability of geothermal systems (He et al., 2018; Huang et al., 2020; Jiang, Qin, Cladouhos, et al., 2022; Jiang, Qin, Faulder, et al., 2022; Keçebaş & Yabanova, 2013; Michael et al., 2023; Ruliandi, 2015; Zulkarnain et al., 2019). Research endeavours exploring data mining, machine learning, and other data-intensive methodologies have greatly benefited from recent technological strides. A significant body of research has been dedicated to employing machine learning to recognise and characterise subsurface geothermal reservoirs (Okoroafor et al., 2022). Furthermore, certain studies are exploring the applicability of similar techniques for the surface aspects of geothermal operations, with diverse models designed for tasks like refining designs (Arslan & Yetik, 2011; Senturk Acar, 2021; Xu et al., 2020; Zhi et al., 2019), condition monitoring (Ruliandi et al., 2021; Siratovich et al., 2022; Yan et al., 2021), boosting performance (He et al., 2018; Keçebaş & Yabanova, 2013; Langiu et al., 2022; Ling et al., 2022; Yilmaz & Koyuncu, 2021), and detecting faults (Liu et al., 2021; Michael et al., 2023; Zulkarnain et al., 2019). The current literature also encompasses various systems for converting geothermal energy, such as traditional geothermal power plants, binary cycle facilities, ground source heat pumps, and geothermal-fueled district heating systems.

The comprehensive performance of a geothermal power plant hinges not only on its output power and system efficiencies but also on the reliability and availability it offers. Consequently, the central objective of this investigation is to augment the overall reliability of the plant by discerning and identifying abnormal behaviours within the circulating feed pumps of a geothermal organic Rankine cycle (ORC) power facility. This research employs a methodological approach that combines time-series analytics with fundamental principles of thermodynamics, identifying potential early indicators for the target abnormal events and enabling the prediction of their occurrence. This prognostication is based on operational time-series data downloaded from sensors installed throughout the geothermal binary power plant. Information from the time series and time-series features that prove to be significant in predicting the target events feed into the domain analysis using thermodynamics principles.

2. DESCRIPTION OF THE PLANT AND THE ANOMALOUS EVENTS

The Te Huka power plant is situated within the Tauhara-North geothermal field in the Taupo Volcanic Zone, New Zealand (Castillo Ruiz et al., 2021). Commissioned in 2010, the Te Huka power station has been actively generating environmentally friendly and sustainable energy, exhibiting an average annual electricity output of 209 GWh. The facility relies on two production wells that supply geothermal fluid to two generating units, each with a nominal capacity of 24MWe. N-pentane serves as the primary working fluid in the binary plant's operations at Te Huka.

Proceedings 45th New Zealand Geothermal Workshop
15-17 November, 2023
Auckland, New Zealand
ISSN 2703-4275

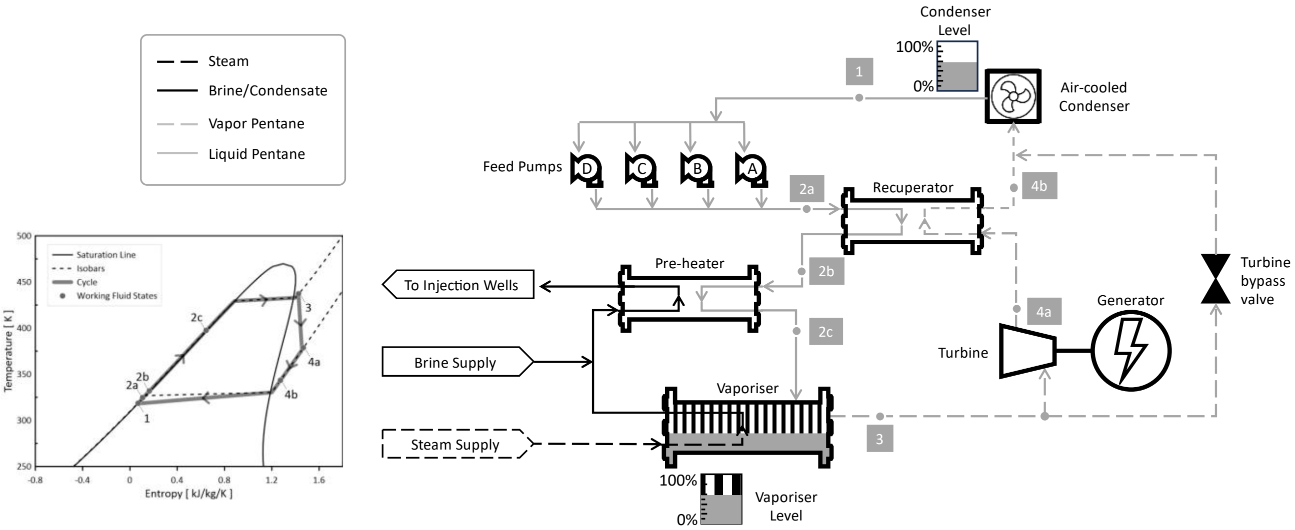

Figure 1. A simplified visual representation of the Te Huka binary plant situated in New Zealand, illustrating the key components of the surface infrastructure and the various stages of the working fluid within the cycle (numbered boxes). The graph in the lower left is a temperature-entropy (T-s) diagram depicting the states of the working fluid across the power generation cycle.

In this arrangement, steam, derived from the two-phase fluid originating from the production wells, serves as the primary heat source in the vaporisers, where the working fluid undergoes vaporisation and superheating. The separated brine from heat exchangers also contributes supplementary energy before entering the vaporiser. The superheated n-pentane propels turbines and subsequently undergoes cooling to return to a liquid state, facilitated by an air-cooled condenser unit. The liquid n-pentane is then directed through four centrifugal feed pumps, which push the working fluid back into the vaporisers, restarting the heating cycle. The condensed geothermal steam and the separated brine are redirected from the heat exchanges and reinjected into the subsurface reservoir. A schematic overview of the Te Huka power plant is depicted in Figure 1.

Feed pumps constitute a fundamental component within most ORC plants (Castillo Ruiz et al., 2021; Langiu et al., 2022). These pumps are responsible for maintaining adequate pressure and fluid flow throughout the broader system. Functionally, these pumps are regulated to maintain a setpoint of the working fluid level within the vaporiser (Figure 1). A reservoir accumulates liquid pentane cooled down from the condenser unit, ensuring a consistent influx of working fluid to the pumps, thereby mitigating issues like cavitation.

During typical operational scenarios, an automated control system governs the pump operation, striving to sustain the working fluid level within the vaporiser. Should the fluid level fall below the designated threshold in the vaporiser, a signal is dispatched to the variable frequency drives of the pumps, instigating an increase in pump speeds or initialisation of additional pumps, if any are available. Conversely, when the fluid level exceeds the desired level, such as when turbines are operated to produce energy at a reduced capacity for a defined duration, the control system modulates the pump speed to a lower setting. Examination of pump data reveals that under standard conditions, higher speeds correspond to increased power consumption by the pumps, whereas lower speeds correlate with reduced power consumption during operation.

On the other hand, operators have noted instances wherein the power of the pump gradually diminishes even without a change in speed (Figure 2). This phenomenon arises when the pump contains a lower quantity of working fluid than expected or when the fluid's density is influenced by partial vaporisation. In either scenario, the impellers within the pump experience reduced resistance, prompting the pump to consume less power while maintaining the set speed. The first situation is improbable due to an accumulating reservoir upstream of the pump system to ensure a sufficient supply of the working fluid to the pumps, as previously mentioned.

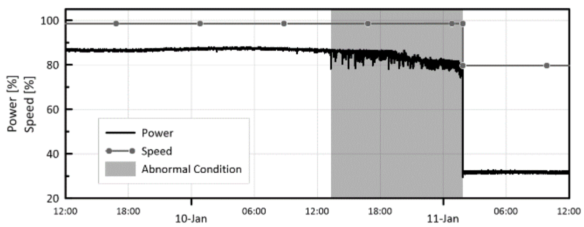

Figure 2. Time series data depict variations in pump power and speed, illustrating typical and anomalous operational states. The anomalies being studied are distinguished by an irregular and gradual decline evident in the pump power time series, despite the pump speed maintaining a consistent value.

The second scenario may transpire when localised vaporisation of the working fluid is driven by unexpected pressure drops often observed during instances of cavitation. Although the pump and other interconnected systems within the facility are designed and operated to avoid such incidents, the recurrent manifestation of these abnormal events has considerably disrupted plant operations, resulting in escalated maintenance expenses and operational interruptions.

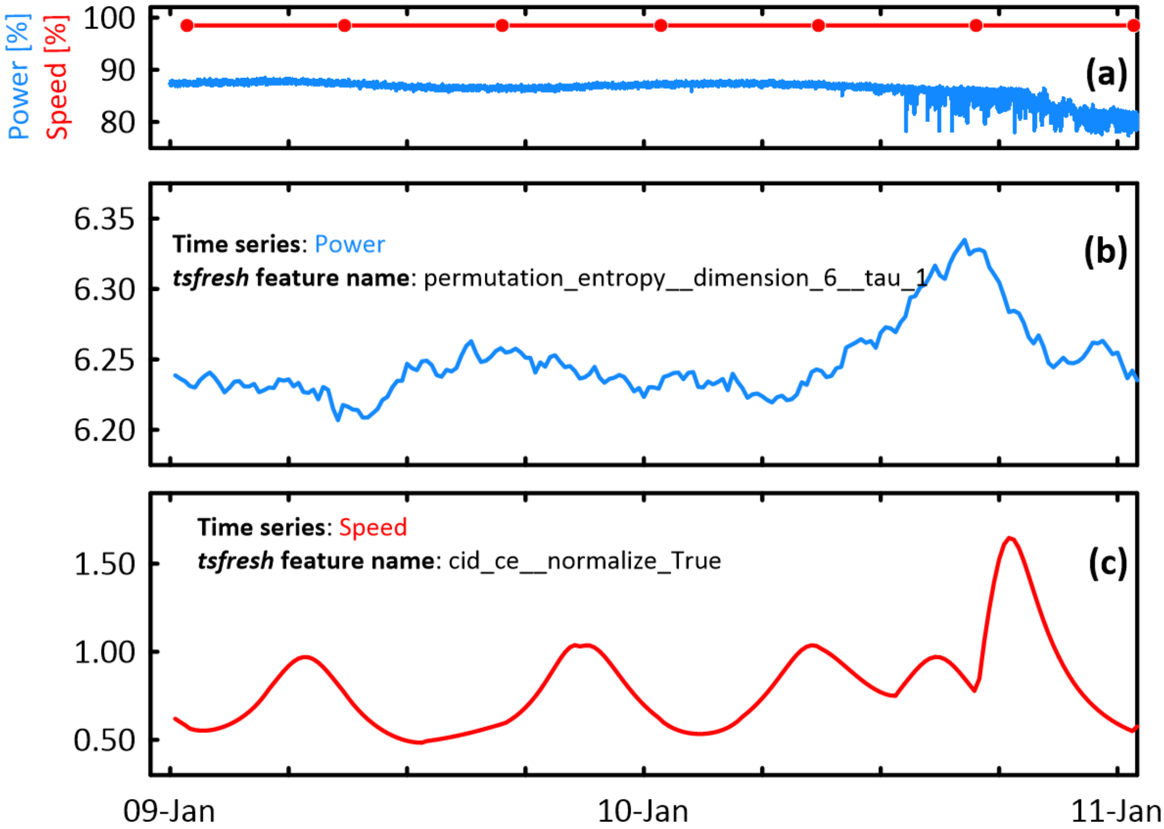

Figure 3. Visualisation of (a) pump power and pump speed as input time series and (b-c) several characteristic features extracted from the inputs using the tsfresh Python library.

3. SYSTEMATIC TIME-SERIES FEATURE ENGINEERING

Conventional time series analysis primarily involves using historical or lagged raw data values to train models with specialised functions, such as forecasting, classification, or regression (Box et al., 2016). Nonetheless, substantial research has highlighted instances where a feature-based representation of the time series can offer a deeper understanding of the phenomenon being studied and yield more robust models (Fulcher, 2017; Kennedy et al., 2021; Simmons et al., 2021). However, embracing a feature-based methodology comes with the trade-off of increased computational costs due to feature calculation and the supplementary task of selecting statistically relevant features for a given modelling objective.

The Python package *tsfresh* (v0.19.0) can leverage parallel computing technologies to automatically extract 794 features for each input time series (Christ et al., 2018) (Figure 3). Furthermore, a module to check the statistical relevance of each extracted feature using univariate hypothesis testing is already built into the package. The extracted features are then used as inputs to develop and train machine learning models that can perform the desired task.

4. FORECASTING AND IDENTIFICATION OF EVENT PRECURSORS

An examination of the control systems governing the power plant was undertaken to identify potential components that influence the operation of the feed pumps. Subsequently, time-series attributes were extracted from the operational data obtained from the pertinent elements of the power plant system. Predictive models were trained using the extracted time-series features and the fully labelled data with a forecasting horizon of 48 hours. Statistically relevant and important time-series features in predicting the target event were singled out and forwarded to the domain expert for validation.

Time-series features extracted from the vaporiser and condenser level time series were identified as the top explanatory features for predicting the target events (

Table 1). These relevant time-series features were then used to train, calibrate, and optimise a *lightGBM* classifier (Ke et al., 2017), aiming to predict the occurrence of the target events with a 48-hour lead time. Following hyperparameter optimisation via the *Optuna* package (Akiba et al., 2019) and probability calibration using the *scikit-learn* library (Pedregosa F. et al., 2011), the best-performing model attained an MCC score of 0.255 (Figure 4). While the MCC scores for the predictive models are relatively modest and indicate suboptimal forecasting performance, the importance rankings of features verify the pivotal role played by the condenser and vaporiser in projecting the target event.

ISSN2703-4275

Table 1. Selection of top explanatory time-series features extracted from various power plant time-series data identified after univariate hypothesis testing using the tsfresh library.

Time series	***tsfresh* feature name & description**	p-value
Condenser level	*absolute_maximum:* Largest value in the input time series	$\mathbf{6.43 \times 10^{-146}}$
Feed pump outlet pressure	*permutation_entropy__dimension_3__tau_1:* Calculates the permutation entropy, a time series complexity measure (Bandt & Pompe, 2002)	$\mathbf{2.54 \times 10^{-124}}$
Vaporiser pressure	*partial_autocorrelation__lag_6:* The value of the partial autocorrelation function at the given lag	$\mathbf{9..46 \times 10^{-94}}$

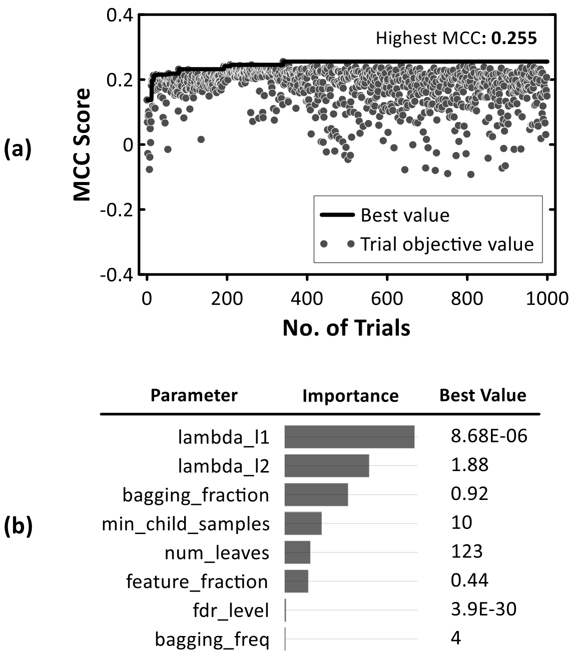

Figure 4. Hyperparameter optimisation results show (a) the best-performing model using data from various power plant components achieving an MCC score of 0.255 and (b) the parameter settings of the top model. Further information about the parameters can be accessed through the *tsfresh* and *lightGBM* documentation.

A second set of forecasting models were trained using only the power and speed data from the feed pumps to compare with the models trained on the data from the other power plant components. The forecasting models trained on the pump data performed significantly better, achieving the highest MCC score of 0.490 (Figure 5). The statistically relevant and important time-series features were mainly complexity measures of the pump power data and quantification of change in the pump speed. These features are expected since the pump speed data is relatively stable and is not likely to change during the target events. This model performance comparison indicates that the target events are driven mainly by the operating conditions immediate to the feed pump system.

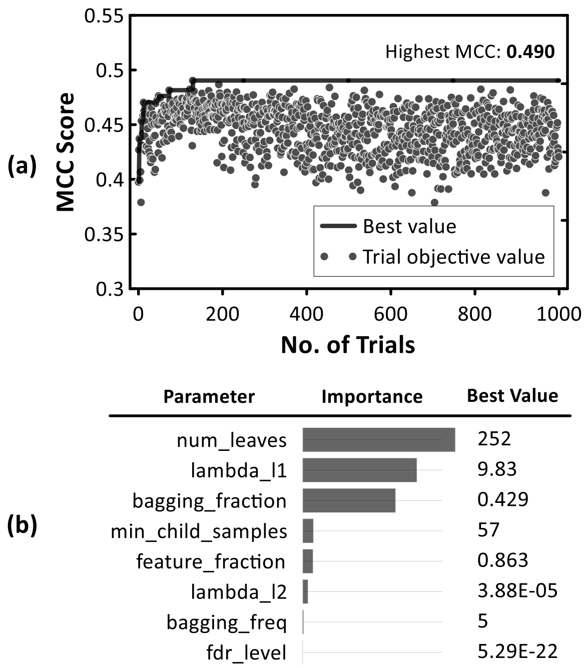

Figure 5. Hyperparameter optimisation results showing (a) the best-performing model trained on pump power and speed time series achieving an MCC score of 0.490 and (b) the parameter settings of the top model.

5. BAYESIAN A/B TESTING OF DIFFERENT PUMP OPERATING CONDITIONS

Upon closer examination of the operational history of the feed pump system, there were significant periods in which merely three out of the four available pumps were in operation (Figure 6). It became evident after looking at data aggregated daily that the binary power plant operated with three feed pumps (scenario A) for approximately 38% of the time, while periods with four feed pumps (scenario B) accounted for 44% of the observed intervals. Notably, the power plant remained non-operational during the remaining periods (18%) or instances when fewer than three pumps were running. Of the periods characterised by three operational pumps, roughly 8% were concurrent with the target event, whereas about 10% of the four-pump periods exhibited the target events.

Employing Bayesian A/B testing (Kamalbasha & Eugster, 2020) on the labelled dataset via the *PyMC* library (PyMC development team, 2017), an assessment was conducted to determine the likelihood of the target event taking place under the conditions of three versus four operational pumps.

The test outcomes reveal a 22% higher likelihood of target events transpiring when the feed pump system functioned with all four pumps, in comparison to scenarios involving the operation of only three pumps (Figure 7). These findings further indicate that the origins of these target events are closely associated with the local operating conditions of the feed pump system.

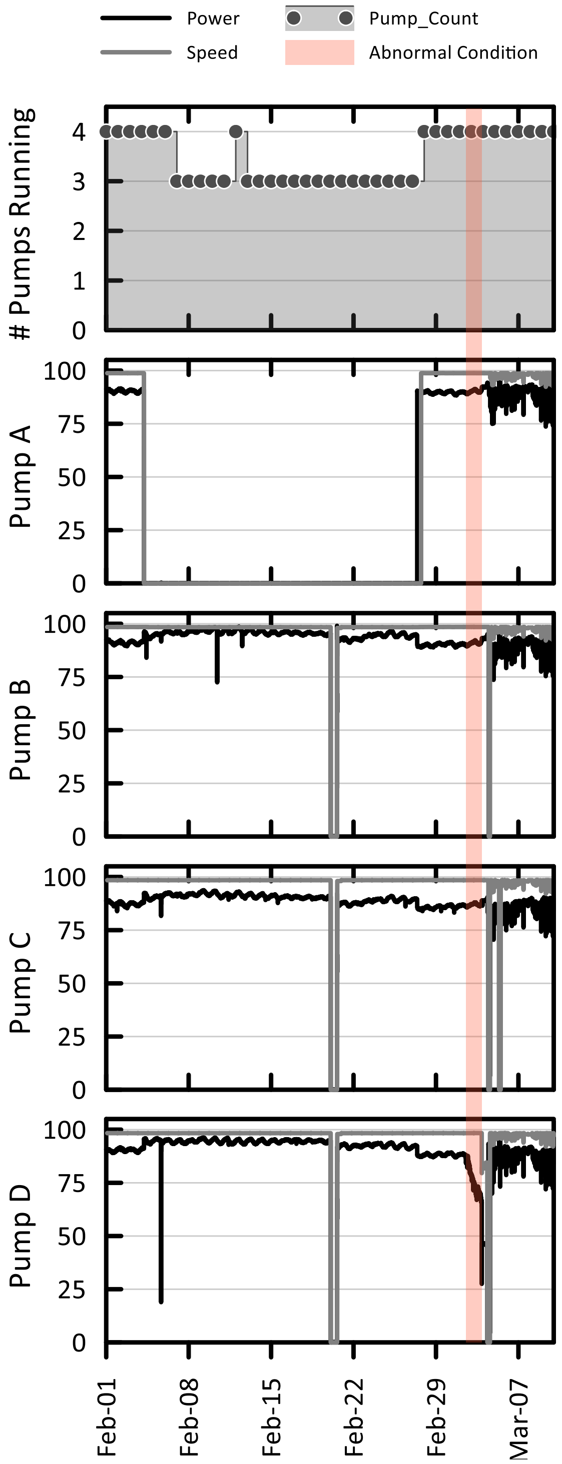

Figure 6. Exemplary plot showing periods with different numbers of feed pumps running. A cavitation event occurs (Pump D) after the fourth pump (Pump A) is brought back into operation after almost a month of running with only three pumps.

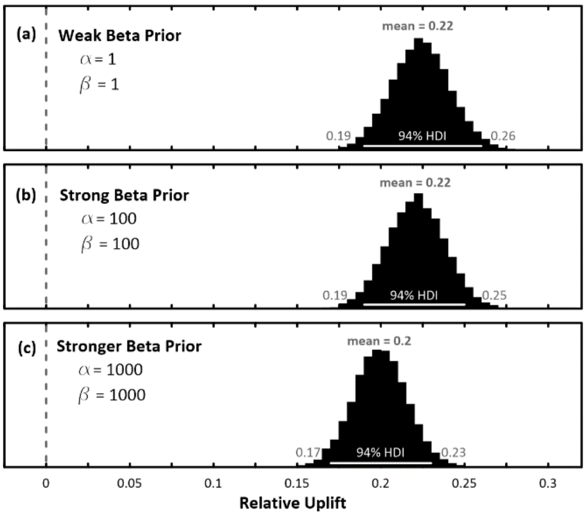

Figure 7. Bayesian A/B testing results show the posterior distribution for different beta priors plotting to the right of the zero-line, indicating a 22% higher likelihood of the target event occurring in scenario B (4-pump operation) versus scenario A (3-pump operation).

6. THERMODYNAMIC INDICATORS FOR POTENTIAL PUMP CAVITATION

Taking off from the precursor modelling and A/B testing results, a more in-depth investigation was undertaken to study the fluid conditions before and after the pumps. The measured outlet temperatures and pressures of the feed pump system were mapped onto a temperature-entropy (T-s) diagram, affirming that the fluid emerging from the pumps is in the expected compressed liquid state (Figure 8).

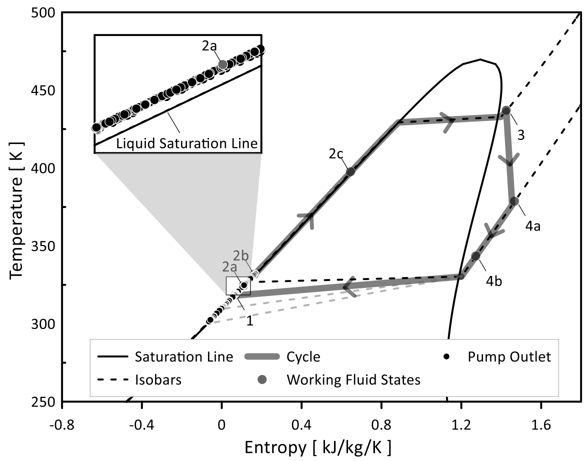

Figure 8. Temperature-Entropy diagram showing the state of the working fluid at the feed pump outlet. All the observations plot above and to the left of the liquid saturation line.

Ideally, a similar exercise should be done for fluid conditions measured at the pump inlet, but no temperature sensor is currently installed upstream of the pumps. Thus, a similar analysis is done upstream of the pump using the measured inlet pump pressure, while the pump inlet temperature is estimated assuming the pumping process is isentropic. A brief sensitivity analysis was done to see how different temperature difference values between the inlet and outlet fluid temperatures may affect the T-s diagram of fluid conditions at the pump inlet. From the resulting T-s diagrams, we can observe data points plotted in the vapour or superheated phase even with the conservative assumption that the pump operation is ideally isentropic (Figure 9).

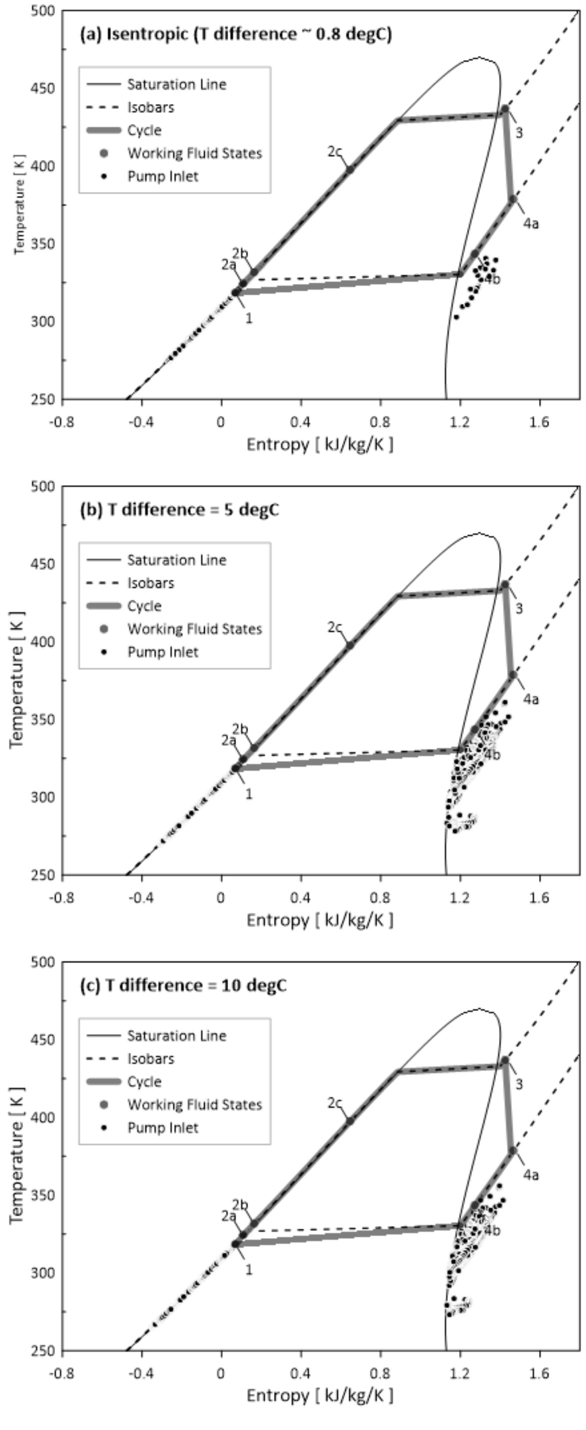

Figure 9. Temperature-Entropy diagram showing the state of the working fluid at the feed pump inlet at different inlet temperature estimates with (a) assuming the pumps are isentropic, (b) a temperature difference of 5°C, and (c) a temperature difference of 10°C between the pump inlet and outlet temperature.

At the pump intake, the working fluid is expected to be entirely in a compressed liquid state at the measured pump inlet pressure. The data points located in the vapour or superheated phase of the temperature-entropy (T-s) diagram signify instances when the pumps are operating outside the parameters for which they were designed. The partial vaporisation of the working fluid at the pump inlets can trigger cavitation, leading to degradation in pump performance due to the erosion of pump impellers.

The findings from this analysis require validation through the assessment of actual pressures and temperatures of the working fluid measured at the inlet of each feed pump. Furthermore, it is important to note that pump inlet conditions presented in Figures 8 and 9 can only be plotted on either the compressed liquid region or the superheated region of the T-s diagrams. These inlet conditions cannot move directly from the compressed or saturated liquid to superheat because they should first go through the two-phase region. However, additional information in the form of the working fluid dryness fraction is required to accurately plot the states of the working fluid during two-phase conditions. Measuring the dryness fraction is complicated and may not be possible or practical. Therefore, these points show up in the superheated region, which only requires the temperature and pressure of the working fluid. As previously noted, it will indicate partial vaporisation at the pump inlet, which can be the source of pump cavitation.

7. CONCLUSION

This research has showcased the use of an analysis workflow that combined engineering expertise with time-series analytics to characterise and predict abnormal states in the feed pumps of a geothermal binary power plant. The fully labelled dataset was used to develop new forecasting models. This modelling exercise revealed that the models trained using the feed pump data performed better than those trained on the time-series data of other power plant components (MCC score of 0.490 vs 0.255). The results of the forecasting models indicate that the feed pump data acted as better precursors to the target event over the other power plant data, which may point to the abnormal events being driven by operating conditions local to the feed pumps.

Upon deeper investigation into the operational history of the feed pump system, notable intervals were identified during which merely three out of the four available feed pumps were operating. Bayesian A/B testing was performed to ascertain whether the likelihood of target events varied under distinct operational conditions. The testing outcomes revealed that the occurrence of target events exhibited a 22% increase in probability when all four pumps were running, as opposed to when only three were functioning. This outcome underscores the initial findings that the events are intricately linked to the operational conditions of the feed pumps.

Subsequently, building upon the outcomes of the precursor modelling, a thermodynamic examination of the feed pump operation was performed, including an assessment of the temperature-entropy diagram representing the conditions of the working fluid across the pumps. Although the conditions observed at the pump outlet adhered to design parameters, the investigation unveiled intervals during which the working fluid might have encountered partial vaporisation as it entered the pump. This shift in the phase of the working fluid at the pump inlet is undesirable and could serve as the catalyst for potential cavitation events within the feed pump system.

Effectively applying a feature-based time-series analytics strategy to address operational challenges within a geothermal power plant underscores the considerable promise that comparable technologies hold for the industry, particularly considering the substantial data volumes produced by these power facilities. The hurdle resides in devising these solutions with the agility required to match the dynamic nature of complex industrial systems. A critical component for attaining robust resolutions to issues in the field of geothermal energy is the collaborative synergy between domain expertise and artificial intelligence.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to the New Zealand Ministry of Business, Innovation and Employment (MBIE) for their financial support through the Empowering Geothermal Energy funds and Contact Energy Ltd for the kind permission to use operational data from the Te Huka binary power plant and publish the results.

The authors wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities, consulting support and training services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. URL https://www.nesi.org.nz.

REFERENCES

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631. https://doi.org/10.1145/3292500.3330701

Arslan, O., & Yetik, O. (2011). ANN based optimisation of supercritical ORC-Binary geothermal power plant: Simav case study. *Applied Thermal Engineering*, *31*(17–18), 3922–3928. https://doi.org/10.1016/j.applthermaleng.2011.07.041

Bandt, C., & Pompe, B. (2002). Permutation Entropy: A Natural Complexity Measure for Time Series. *Physical Review Letters*, *88*(17), 174102. https://doi.org/10.1103/PhysRevLett.88.174102

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time Series Analysis : Forecasting and Control* (5th ed.). John Wiley & Sons, Inc.

Castillo Ruiz, N., Mclean, K., Richardson, I., Misa, T., Ferguson, A., Altar, D. E., & Kaya, E. (2021). Passive NCG Reinjection at Te Huka Geothermal Binary Power Plant. *Proceedings 43rd New Zealand Geothermal Workshop*. http://www.geothermal-energy.org/pdf/IGAstandard/NZGW/2021/80.pdf

Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W. (2018). Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing*, *307*, 72–77. https://doi.org/10.1016/j.neucom.2018.03.067

Fulcher, B. D. (2017). Feature-based time-series analysis. *Feature Engineering for Machine Learning and Data Analytics*, 87–116. https://doi.org/10.48550/arxiv.1709.08055

He, R., Wang, Z., & Fu, W. (2018). Application Of PID Control Based On BP Neural Network In The Expansion Machine Of Organic Rankine Cycle System. *GRC Transactions*, *42*. https://publications.mygeoenergynow.org/grc/1033998.pdf

Huang, R., Zhang, Z., Zhang, W., Mou, J., Zhou, P., & Wang, Y. (2020). Energy performance prediction of the centrifugal pumps by using a hybrid neural network. *Energy*, *213*, 119005. https://doi.org/10.1016/j.energy.2020.119005

Jiang, A., Qin, Z., Cladouhos, T. T., Faulder, D., & Jafarpour, B. (2022). A Multiscale Recurrent Neural Network Model for Long-Term Prediction of Geothermal Energy Production. *PROCEEDINGS, 47th Workshop on Geothermal Reservoir Engineering*. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2022/Jiang.pdf

Jiang, A., Qin, Z., Faulder, D., Cladouhos, T. T., & Jafarpour, B. (2022). Recurrent neural networks for short-term and long-term prediction of geothermal reservoirs. *Geothermics*, *104*, 102439. https://doi.org/10.1016/j.geothermics.2022.102439

Kamalbasha, S., & Eugster, M. J. A. (2020). Bayesian A/B Testing for Business Decisions. *Data Science – Analytics and Applications*, 50–57. https://doi.org/10.1007/978-3-658-32182-6_9

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 3149–3157. https://doi.org/10.5555/3294996

Keçebaş, A., & Yabanova, İ. (2013). Economic analysis of exergy efficiency based control strategy for geothermal district heating system. *Energy Conversion and Management*, *73*, 1–9. https://doi.org/10.1016/j.enconman.2013.03.036

Kennedy, A., Nash, G., Rattenbury, N. J., & Kempa-Liehr, A. W. (2021). Modelling the projected separation of microlensing events using systematic time-series feature engineering. *Astronomy and Computing*, *35*. https://doi.org/10.1016/j.ascom.2021.100460

Langiu, M., Dahmen, M., & Mitsos, A. (2022). Simultaneous optimisation of design and operation of an air-cooled geothermal ORC under consideration of multiple operating points. *Computers & Chemical Engineering*, *161*, 107745. https://doi.org/10.1016/j.compchemeng.2022.107745

Ling, W., Liu, Y., & Young, R. (2022). Deep Learning Models for Prediction and Optimisation of Air-Cooled Binary Cycle Geothermal Operation. *PROCEEDINGS, 47th Workshop on Geothermal Reservoir Engineering*. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2022/Ling.pdf

Liu, Y., Ling, W., Young, R., & Hsieh, M. (2021). Deep Learning for Prediction and Fault Detection in Geothermal Operations. *PROCEEDINGS, 46th*

Workshop on Geothermal Reservoir Engineering, *46*. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2021/Liu1.pdf

Michael, P., Abrasaldo, B., Zarrouk, S. J., & Kempa-Liehr, A. W. (2023). Characterising and predicting low discharge pressure events in less permeable geothermal production wells. *Geothermics*, *112*, 102756. https://doi.org/10.1016/j.geothermics.2023.102756

Okoroafor, E. R., Smith, C. M., Ochie, K. I., Nwosu, C. J., Gudmundsdottir, H., & (Jabs) Aljubran, M. (2022). Machine learning in subsurface geothermal energy: Two decades in review. *Geothermics*, *102*, 102401. https://doi.org/10.1016/j.geothermics.2022.102401

Pedregosa F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

PyMC development team. (2017). *PyMC*. https://www.pymc.io/welcome.html

Ruliandi, D. (2015). Geothermal power plant system performance prediction using artificial neural networks. *2015 IEEE Conference on Technologies for Sustainability (SusTech)*, 216–223. https://doi.org/10.1109/SusTech.2015.7314349

Ruliandi, D., Dwi Susanto, A., & Djanarto. (2021). Application of Artificial Neural Network to Exergy Performance Analysis of Geothermal Power Plant. *Proceedings World Geothermal Congress 2020+1*. http://www.geothermal-energy.org/pdf/IGAstandard/WGC/2020/26086.pdf

Senturk Acar, M. (2021). Multi-stage artificial neural network structure-based optimisation of geothermal energy powered Kalina cycle. *Journal of Thermal Analysis and Calorimetry*, *145*(3), 829–849. https://doi.org/10.1007/s10973-020-10125-y

Simmons, S., Jarvis, L., Dempsey, D., & Kempa-Liehr, A. W. (2021). Data Mining on Extremely Long Time-Series. *IEEE International Conference on Data Mining Workshops, ICDMW*, *2021-December*, 1057–1066. https://doi.org/10.1109/ICDMW53433.2021.00137

Siratovich, P., Buster, G., Taverna, N., Rossol, M., Weers, J., Blair, A., Huggins, J., Siega, C., Mannington, W., Urgel, A., Cen, J., Quinao, J., Watt, R., & Akerley, J. (2022). GOOML-Finding Optimization Opportunities for Geothermal Operations. *PROCEEDINGS, 47th Workshop on Geothermal Reservoir Engineering*. https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2022/Siratovich.pdf

Xu, X., Liu, J., Wang, Y., Xu, J., & Bao, J. (2020). Performance evaluation of ground source heat pump using linear and nonlinear regressions and artificial neural networks. *Applied Thermal Engineering*, *180*, 115914. https://doi.org/10.1016/j.applthermaleng.2020.115914

Yan, D., Yang, F., Yang, F., Zhang, H., Guo, Z., Li, J., & Wu, Y. (2021). Identifying the key system parameters of the organic Rankine cycle using the principal component analysis based on an experimental database. *Energy Conversion and Management*, *240*, 114252. https://doi.org/10.1016/j.enconman.2021.114252

Yilmaz, C., & Koyuncu, I. (2021). Thermoeconomic modeling and artificial neural network optimisation of Afyon geothermal power plant. *Renewable Energy*, *163*, 1166–1181. https://doi.org/10.1016/j.renene.2020.09.024

Zhi, L.-H., Hu, P., Chen, L.-X., & Zhao, G. (2019). Multiple parametric analysis, optimisation and efficiency prediction of transcritical organic Rankine cycle using trans-1,3,3,3-tetrafluoropropene (R1234ze(E)) for low grade waste heat recovery. *Energy Conversion and Management*, *180*, 44–59. https://doi.org/10.1016/j.enconman.2018.10.086

Zulkarnain, Surjandari, I., Bramasta, R. R., & Laoh, E. (2019). Fault Detection System Using Machine Learning on Geothermal Power Plant. *2019 16th International Conference on Service Systems and Service Management (ICSSSM)*, 1–5. https://doi.org/10.1109/ICSSSM.2019.8887710