GOOML: Geothermal Operational Optimization with Machine Learning

Paul A. Siratovich[1], Andrea Blair[1], Jon Weers[2]

[1]Upflow, P.O. Box 61, Taupo, New Zealand 3330

[2]National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, United States

Keywords

big data, machine learning, algorithms, field optimization, GOOML

ABSTRACT

Geothermal Operational Optimization with Machine Learning (GOOML) is a project focused on maximizing increased availability and capacity from existing industrial-scale geothermal generation assets. The GOOML project will develop a suite of machine learning-based algorithms that analyze historical production datasets and provide predictive setpoints for geothermal field operations. Historical datasets from New Zealand and the US will provide the input to develop digital geothermal system twins which allow prediction of market conditions, maintenance operations and steamfield optimization. The algorithms will identify key parameters within fields and suggest setpoints for components of the system to maintain optimal generation. Set-points can be instructed to follow mass flow restrictions, generation maximization and optimal field/reservoir balance and give field operators a guide by which generation can be optimized. The datasets that will be used to develop GOOML are sourced from operating geothermal fields in New Zealand and the United States with varying degrees of complexity. This will ensure that most geothermal systems can utilize the GOOML tool to assist in optimizing operations. GOOML aims to achieve a step-change in geothermal operations by developing state-of-the-art machine learning algorithms, comprehensive data analytics, and a first-of-its-kind automated, intelligent geothermal system model.

1. Introduction

The Geothermal Operational Optimization with Machine Learning (GOOML) project is led by Upflow Ltd (Upflow), a geothermal service provider working in geothermal power projects in New Zealand and around the world. The GOOML project will utilize datasets from industry-leading geothermal operators, global specialists in big-data analysis and machine learning experts to develop state-of-the-art machine-learning (ML) algorithms, comprehensive data analytics, and a first-of-its-kind automated, intelligent geothermal system models. To achieve this, Upflow has partnered with the National Renewable Energy Laboratory's (NREL), data-scientists and high-performance computing capabilities. Datasets for this project will be sourced from US based Ormat Technologies operations (Ormat), and New Zealand's Contact Energy (Contact) and Ngati Tuwharetoa Geothermal Assets Limited (NTGA). This project team will come together to develop a series of \tools aimed at optimizing system processes and increasing efficiency in existing infrastructure.

This project realizes the aspiration that both countries have to cooperate in the development of advanced geothermal energy technologies; in particular the joint development and improvement of modeling tools, as per the agreement signed on the 22nd June 2018 between the US Department of Energy and New Zealand Ministry of Business, Innovation and Employment (MBIE). "This new research partnership with New Zealand will connect experts from both countries to collaborate on a mutually beneficial basis to advance and accelerate the development of geothermal technologies." Timothy Unruh, EERE's Deputy Assistant Secretary for Renewable Power.

Data from New Zealand's Wairakei geothermal field will serve as the baseline and training datasets from which analytical models, tools and machine learning experiments will be developed and verified. This includes data from more than 300 wells, immense bore-field infrastructure (>4,000 supervisory control and data acquisition SCADA sources), and multiple steam, flash, and binary power generation units. The complexity of Wairakei is an ideal location for the development and verification of geothermal ML algorithms aimed at full system optimization.

Lessons learned from Wairakei, (because of its complexity), will be broadly applicable to global geothermal operations. The ML algorithms will then be tested on Ormat's, NTGA's and Contact's other operations. Ormat's domestic US geothermal fleet will provide proving grounds for the developed technology on electricity generation operations; whilst NTGA's operations at Kawerau will allow testing of algorithms on a field that uses geothermal fluids to generate both electricity and drive industrial heat processes. This on-the-ground testing of the ML algorithms in complex forums will verify that we have developed a universally deployable, automated, intelligent digital geothermal system for geothermal operational optimization. The resulting models will help to increase operational efficiency, reduce downtime, increase capacity factors, and reduce cost of electricity.

2. Machine Learning and Energy

Machine learning in geothermal energy is a relatively new are of research that has seen some deployment in exploring for geothermal resources, (Arslan and Yetik, 2014; Ishitsuka et al, 2019) improving the efficiency of modeling geothermal reservoirs (Li et al, 2017; Tian and Horne, 2019) and providing insight to formation and bottom-hole temperatures in newly drilled wells (Gul et al, 2019). However, the application of machine learning has seen limited use in geothermal powerplant operations.

Geothermal operations can potentially be significantly improved through the use of ML experimentation and algorithm development; application of ML techniques has resulted in significant improvements in coal-fired and gas-turbine power plants through reduction of start times, increasing availability (~1% per annum) and total generation output; up to 15MWe instant availability in gas-turbine peaking plants (GE,2016; Collins, 2018; Rahat etl al, 2018). Machine learning models of wind conditions and turbine output have been used to anticipate power generation (Clifton et al, 2013) and the subsequent effects on other market factors e.g. conventional generation, storage, consumer behavior (Treiber et al, 2016). Such analysis applied to geothermal operations would allow greater demand forecasting and help operators to extract more information from large datasets to improve decision-making and, potentially increase operational efficiencies and forecasting of maintenance demands. Similar technologies (e.g., wind, coal, nuclear) have increased capacity factors, lowered O&M costs, and achieved faster start-up to full generation periods through ML applications (GE,2016; Islikaye and Cetin, 2018).

2.1 Innovating Geothermal Operations

The GOOML project intends to develop a suite of well-tested ML algorithms capable of analyzing and recommending optimal system parameters for a geothermal field as a digital system "twin." This digital twin will use real-time data to offer analytical predictions for power plant operation using intelligent analytics providing guidance for on-the-ground operations. The system twin will be a predictive tool that can be run alongside real-world operations and predict, diagnose, and optimize the implementation of operational improvements, thus raising operational performance. This will be a world-first geothermal development, allowing rigorous analytics to be provided from a digital tool developed on some of the world's most complex geothermal fields.

Current geothermal operations are managed by individuals and teams at differing performance levels and tend to be field-specific, where each respective field is operated based on a set of parameters catered to the configuration of production/injection wells and energy conversion technology. Through the field data and technologies provided by our data partners, GOOML will enable the development of a ML-based system twin that can be universally deployed to geothermal operations supporting better decision making and optimizing system outputs.

The outcome of the project if successful would serve to increase capacity factors by ~1%–5% across the US geothermal fleet, which translates to 315 to 1,575 additional annual GWh of geothermal electricity to the national grid (a value of $18–$94M at $0.06/kWh). Current capacity factors for geothermal electricity in the US are on average 74.4% (USEIA, 2018). Through increasing the availability factors of geothermal operations, the long-run marginal cost of electricity (LCOE) can be reduced and made more cost-efficient. In the New Zealand market, geothermal operators are governed by mass-flow takes from reservoirs, such that finding greater efficiencies in surface infrastructure optimization can result in overall lower cost of generation and increased station availability factors.

3. Project Team

The GOOML team is made of a multidisciplinary international team of engineers, data scientists, operational experts and project managers from New Zealand and the United States. It will leverage some of the most capable minds, innovations, and world-leading operational insights in the geothermal industry. The GOOML team is as follows:

New Zealand:

- Upflow (Upflow): Project manager, geothermal consultants and thought leaders in geothermal operations, technology, and business support functions. Upflow is the principal investigator (PI) for the GOOML project and will coordinate all project activities.
- Contact Energy (Contact): Commercial partner and provider of data, expertise, and guidance for project development. Contact Energy operates over 440 MWe of installed geothermal capacity including the Wairakei Geothermal Field which has been generating electricity from dual-phase fluids for well over 60 years.
- Ngati Tuwharetoa Geothermal Assets (NTGA): Commercial partner, provider of expertise, data, and project guidance. NTGA operates a portion of the Kawerau Geothermal Field and provides clean steam to direct use operations at the field (timber, paper, and dairy processing) as well as operation of a 25MWe binary station.
- Flow-State Solutions (FSS): Commercial partner and modeling expertise. FSS will provide model guidance and development of the ML algorithms to ensure that models are accurate, reflecting true geothermal operational conditions. FSS brings numerical modeling expertise and geothermal domain knowledge to the project.

United States:

- National Renewable Energy Laboratory (NREL): Co-principal investigator for the GOOML project. NREL will provide data science support, machine learning algorithm development and data curation expertise. NREL will also supply the computing power behind the GOOML project through the Eagle supercomputer located at NREL's campus.
- Ormat Technologies; Worldwide geothermal operator responsible for over 800 MWe of installed capacity around the globe. Ormat will provision data, expertise and guidance to machine learning algorithm development to ensure that the GOOML outputs are relevant and globally applicable.
- United States Department of Energy Geothermal Technologies Office (GTO): Project sponsor for the GOOML project and project governance. The GTO will provide direction and guidance throughout the entirety of the project.

To coordinate a resource intensive project such as GOOML requires a rigorous understanding of project roles and responsibilities.

Figure 1 shows the high-level project hierarchy and responsibilities of each respective institution. Each of the four main tasks will be overseen by the steering committee of the GTO and the commercial partners as well as the Principal Investigator (Upflow) and co-Principal Investigator (NREL).

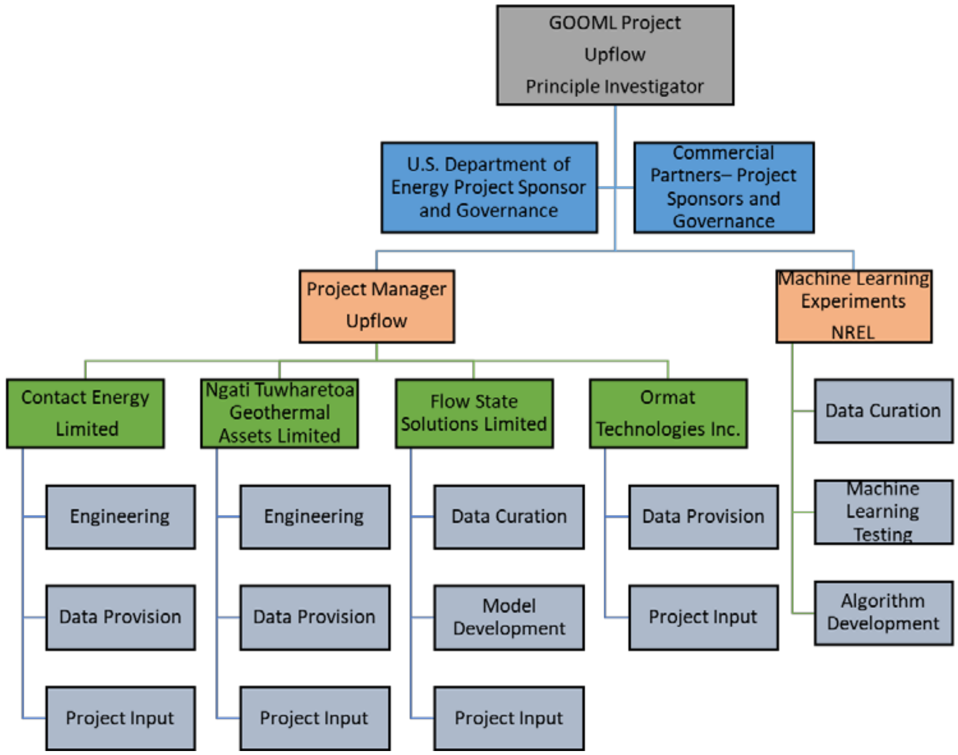

Figure 1: GOOML Project Organization Chart showing project sponsors, commercial partners, and roles/responsibilities of individual entities.

4. Project Workflow

GOOML will require significant volumes of data to be successful. In order to achieve this, data will be provisioned from Contact's, NTGA's and Ormat's commercial operations. These datasets will serve as the historical data against which algorithms will be trained, developed, and refined. Figure 2 shows the high-level workflow for the project showing the data curation, machine learning testing and ultimate utilization of the GOOML tools on active geothermal fields. The

GOOML data will come from the following sources:

- Production wells: massflow rates, enthalpies, pressures, temperature profiles, chemistry, and fluid phases
- Injection wells: massflow rates, pressures, temperature profiles, fluid chemistries
- Pipeline data: massflows, pressures, steam/water ratios, pressure drops and thermal losses
- Separator data: massflow, pressure drop, steam fraction, brine fraction
- Turbine/ORC data: massflow, enthalpies, vibration, maintenance schedules
- Generator data: gross output, net output, annual productions, generation efficiency
- System data: regulatory limits, generation targets, station outages, capacity/availability factors and other system histories

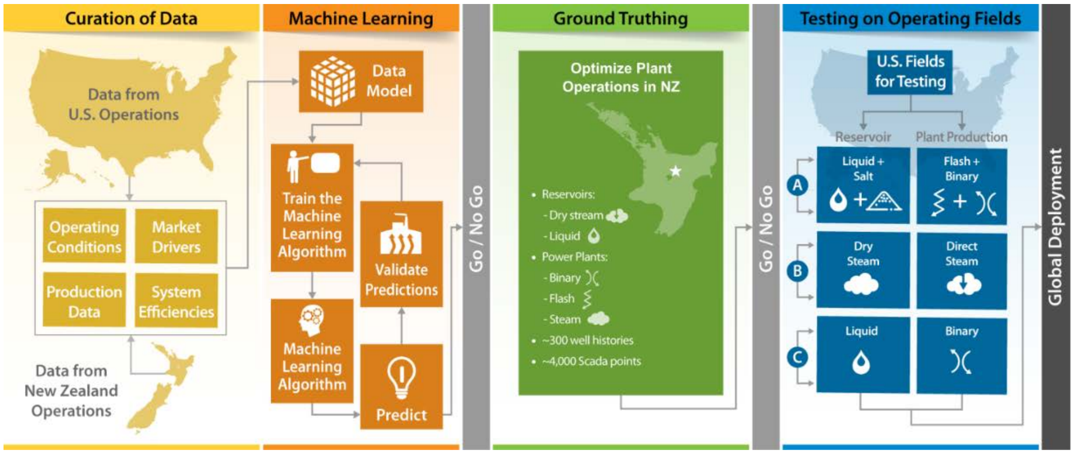

Figure 2: General Overview of GOOML Project Flow.

The datasets obtained for the project will be immense, so data analytics, which only ML techniques can provide, will result in tangible outcomes. The analytics of trained algorithms will be used to develop geothermal field/plant "twins" similarly used in modern combustion stations. Through this project, we envisage that such development is a natural progression for these already-successful applications of ML analytics.

4.1 Digital Twin Development

A principal component of the work effort for GOOML will be the development of the "digital twin"; an intelligent, agent-based model of a working geothermal system utilizing a nodal architecture to model the interconnectivity of individual plant/field components. Each node will represent a component of the steam field (borehole, valve, pipe, SCADA/PI node, etc.) and will possess the appropriate parameters, physics-based limitations, and awareness of its relationship to other nodes in the system. This will allow us to quickly adjust parameters within the system and model the outcome on the full system. Predictive analytics will be used to identify anticipated failure points based on historical data, which accounts for human-based decisions in plant operations and incorporated into the individual parameters for each node. This modular, nodal design will allow for easy extensibility and application of the "digital twin" model to other fields.

Geothermal operations are typically remotely controlled by centralized people (operators) who have a "feel" for the needs of the plant as well as guidance by field models, market conditions and operational protocols. Through these drivers, output parameters are adjusted accordingly to provide fuel to stations; adjusting output (and input to injection wells) from the bore fields to meet station needs. The use of ML-based "twins" to develop and predict optimal plant conditions would aid in more consistent station operation, reducing both variations in performance between individual operators and the risk of human error. Applying ML algorithms to optimize generation will be more efficient than that of a human operator (GE, 2016); which is not to replace operators but provide a tool to improve and test decisions and which are subsequently used on real-world operations.

4.2 Machine Learning Techniques

Substantial advances in the application of ML technologies have produced a variety of standardized methods, many of which have been successfully applied in the oil and gas industry and are now available as free, extensible software packages. We will begin by curating operational data at scale using Random Cut Forest, sorting with k-means clustering, and developing rule-bases using neural networks and regression techniques such as linear and logistic. Platforms to be used include scikit-learn, TensorFlow, and Amazon SageMaker, which is a fully-managed service with access to the computational power and storage scale of Amazon Web Services (AWS), enabling us to build, train, and deploy multiple ML algorithms quickly, in parallel, from a library of tried and tested, high-performance ML algorithms. The GOOML team will explore further plant optimization opportunities by applying ML algorithms such as DeepAR for forecasting infrastructure failure rates (e.g. turbines), Factorization Machines for estimating the interactions between nodes (e.g. wells, pipes and separators), and Random Forrest to test different system configurations and identify optimal settings across the system to provide optimal electricity generation.

5. Project Tasks

The following project objectives describe the workstreams for the GOOML program:

Objective 1: Curation of a quality-assured and standardized electronic dataset suitable for machine learning analytics and algorithm development. The datasets are planned to be sourced from two liquid-dominated geothermal fields and from a prominent US based operation.

Objective 2: Utilizing the curated and standardized datasets to create models and develop machine learning algorithms trained on real-world geothermal operational data. The algorithms will then be optimized and tested on the historical data and highlight where efficiencies may have been gained and give direction for future operations.

Objective 3: Verification of the developed algorithms and stress testing of their capabilities on the full history of existing geothermal fields. This will involve testing the algorithms alongside ongoing operations, harvesting, and utilizing new data and further refining the algorithms.

Objective 4: Deployment of the developed digital twin models for real-world operational testing and verification of algorithm capabilities on our partner's fields. Fields will be tested based on a variety of generation technologies and reservoir conditions. This will be followed by publications, reporting and workshopping of how the developed tools are used.

These objectives will be achieved through the tasks outlined in Figure 3 which shows the high-level work-breakdown structure (WBS) for the GOOML project. We have divided the plan into four specific tasks, with a fifth task covering the life of the project. Each task will culminate in the deployment of the GOOML tool to national and international fields.

Figure 3: Work Break-Down Structure for the GOOML project showing each significant task and milestone needed to achieve the outcomes and transferability to the wider geothermal industry.

6. Technology Transfer

Technology transfer between the US and NZ is a key element of the project delivery. The US team will apply machine learning expertise to New Zealand geothermal fields as well as providing additional access to Ormat's operated field which will help the GOOML technology. The New Zealand team will provide operational insights, that have led to some of the highest availability factors (e.g. capacity factors) in geothermal electricity generation. Whilst the US has the largest installed geothermal capacity in the world, capacity factors are c.a. 76% which is behind the average capacity factor for New Zealand which sits around 92-95% capacity in a given year. This international technology transfer will enable our findings to go beyond the shores of the partner countries and will make a meaningful contribution to keeping geothermal technology relevant and cost-effective for geothermal operators worldwide.

7. Conclusion

Through a combined approach of machine learning techniques and data science research, and partnership with commercial operations, we intend to understand the critical elements of optimal geothermal energy production and develop a tool that mirrors geothermal system activities and provides guidance on increasing operational efficiencies. These algorithms will be used to predict and suggest maintenance schedules, increase operational efficiency, and reduce station downtime. Benefits will result in increased capacity factors and ultimately drive down the levelized cost of energy ensuring that geothermal electricity generation remains competitive in the wider energy marketplace.

Acknowledgement

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 Technology Office Award Number DE-EE0008766. We would also like to acknowledge our commercial partners of Contact Energy Limited, Ngati Tuwharetoa Geothermal Assets Limited and Ormat Technologies Inc. for their support of this project, without access to their data, experts and institutional knowledge, this project would not be possible.

REFERENCES

Arslan, O. & Yetik, O. ANN Modeling of an ORC-Binary Geothermal Power Plant: Simav Case Study, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 36:4, 418-428, DOI: 10.1080/15567036.2010.542437. 2014.

Collins, B. July, GE Uses AI to Increase Responsiveness of Thermal Power: Q&A. Bloomberg Finance. 2018.

Clifton et al. Using machine learning to predict wind turbine power output. Environ. Res. Lett. 8 (2013) 024009 (8pp). 2013.

EERE, Energy Department to Collaborate with New Zealand on Geothermal Energy Advancement https://www.energy.gov/eere/articles/energy-department-collaborate-new-zealand-geothermal-energy-advancement. June 2018.

General Electric. 2016. GE Digital Twin: Analytic Engine for the Digital Powerplant. GE Power Digital Solutions.

Gul, S., Aslanoglu, V., Kaan Tuzen, M., Senturk, E.: Estimation of Bottom Hole and Formation Temperature by Drilling Fluid Data: A Machine Learning Approach. Proceedings, 44th Workshop on Geothermal Reservoir Engineering, Stanford University (2019).

Ishitsuka et al. Resistivity-Based Temperature Estimation of the Kakkonda Geothermal Field, Japan, Using a Neural Network and Neural Kriging. IEEE Geoscience and Remote Sensing Letters. Vol. 15, No. 8. 2018.

Islikaye and Cetin, Performance of ML Methods in Estimating Net Energy Produced in a Combined Cycle Power Plant. 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG). p217-220. 2018.

Li, Y., Júlíusson, E., Pálssond, H., Stefánsson, H., Valfells, A.:Machine learning for creation of generalized lumped parameter tank models. Geothermics 70 (2017). 62-84.

Rahat et al. 2018. Data-driven multi-objective optimization of coal-fired boiler combustion systems. Applied Energy. Vol 229 (2018) 446-458. Tan et al., 2016. Modeling and reduction of NOx emissions for a 700 M coal-fired boiler with the advanced machine learning method. Energy 94 (2016) 672-679.

Tian, C., Horne, N.R.: Applying machine-learning techniques to interpret flow rate, pressure, and temperature data from permanent downhole gauges. SPE Reservoir Evaluation & Engineering (2019).

Treiber et al. Wind Power Prediction with Machine Learning. Computational Sustainability. s645. Springer. 2016.

U.S. Energy Information Administration. Electric Power Monthly with Data for June 2018. USDOE. 2018

Word Energy Council. World Energy Resources: Geothermal 2016. WEC. 2016.