GOOML- Real World Applications of Machine Learning in Geothermal Operations

Paul Siratovich[1], Andrea Blair[1], Andrew Marsh[1], Grant Buster[2], Nicole Taverna[2], Jon Weers[2], Christine Siega[3], Alex Urgel[3], Warren Mannington[3], Jonathan Cen[3], Jaime Quinao[4], Robbie Watt[4], John Akerley[5]

[1]Upflow Limited, New Zealand

[2]National Renewable Energy Laboratory (NREL), Golden, Colorado

[3]Contact Energy Limited, New Zealand

[4]Ngati Tuwharetoa Geothermal Assets Limited, New Zealand

[5]Ormat Technologies Inc, Reno, Nevada

Keywords

Machine Learning, GOOML, forecasting, shut planning, optimization, operations

ABSTRACT

GOOML (Geothermal Operational Optimization with Machine Learning) is a machine-learning based framework that enables geothermal power plant operators to explore optimization opportunities for their assets in an efficient and robust digital environment. Backed by real-world data sources, thermodynamic constraints and steamfield intelligence, the GOOML environment provides new tools to explore how to best operate steamfields as well as test new scenarios and configurations prior to implementation in the field.

To prove the effectiveness of GOOML, we have undertaken optimization experiments using reinforcement learning (RL) to generate operational suggestions using a balance of mass-take targets, sustainability considerations and net generation. Our experiments use the GOOML construct to explore different field parameters and perform multiple reinforcement learning experiments. Like a comprehensive laboratory workbench, we can change out components of a steamfield to perform testing under a variety of conditions (restrict mass, increase pressure, re-route steam, etc.). This flexibility allows us to explore conditions that would require significant infrastructure changes in a real-world setting at a fraction of the cost and time in a digital environment. The results highlight the benefits of using digital twins and advanced data analytics for the geothermal industry.

1. Introduction

GOOML is a digital environment that allows users to interact with comprehensive geothermal datasets and construct artificial digital twins of their operations. These digital twins can then be used to interrogate historical data records and perform scenario analysis of various steamfield configurations. We have developed a platform that goes beyond data interrogation and have applied technology that provides hypothetical scenarios to intelligent agents (reinforcement learning) to seek optimal solutions for power generation operations.

The GOOML environment can be likened to a laboratory chemist's workbench where the user can perform a myriad of experiments using the tools provided to solve operational problems (Figure 1). The digital 'workbench' incorporates unique characteristics of several key components required to operate a geothermal steam field. These include but are not limited to: Wells (single-phase liquid, steam and two-phase), pipelines, wellhead and centralized separators, pipeline junctions, split junctions, weir boxes, steam turbines, binary generators, pumped production wells, and pumped injection wells. Additionally, to develop GOOML components we have utilized datasets from historical operations which include but are not limited to well output (single and two-phase), steam/liquid mass flows and pressures, power output, reinjection flows, tracer-flow testing (TFT) output equations, decline curves, etc. These components and architecture are discussed in full detail in Buster et al., (2021).

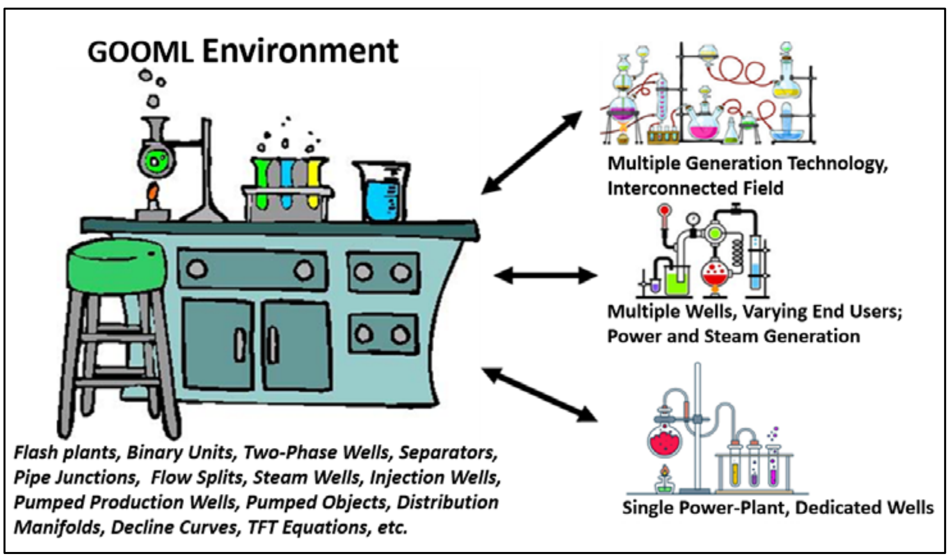

Figure 1: Cartoon representation of the GOOML environment and the capabilities that can be represented in the digital benchtop. Multiple inter-connected objects can be put together to make digital twins of geothermal fields of varying size and complexity.

A user must compile and curate their operational datasets prior to use in GOOML. Best practices for these processes are elaborated in Taverna et al., (2022) which breaks data curation for machine learning into the following steps: (1) data acquisition, (2) data digestion, (3) data transformation, (4) quality assurance (QA) and quality control (QC), (5) use in machine learning models, (6) reiteration through the previous steps until model performance is deemed adequate, and (7) dissemination of data and results. Once the data has made it through basic QA and QC, the user can then configure a steamfield and model field components. This model, based on historical data, has some built-in more rigorous QA and QC features, and is ultimately used to train forecast models which incorporate regression and prediction models, providing an intelligent framework to solve geothermal production scenarios. Having established field models, a user can then explore many configurations and scenarios and use genetic algorithms and reinforcement learning to find potential opportunities for optimization in their steamfield based on discovered idealized configurations.

2. GOOML Model Framework - Gateway to Optimization Investigations

In order to utilize the GOOML environment to its full potential, users must first ensure that their data are as complete as and of the highest quality possible (see Taverna et al., 2022). Once confident the datasets are in adequate order, a user can begin to integrate their historical records into the environment and train forecast components. This integration begins upstream at the production wells (Figure 2a, b). From there, integration of other steamfield features such as flash-plants (Figure 2c), cumulative mass-flows (Figure 2d), power generators (Figure 2e) and injection (Figure 2f) allow users to make hindcast predictions for validation purposes and then subsequently to extrapolate into future forecast predictions (see Buster et al., 2021 and Siratovich et al., 2022 for more on model architecture and system setup).

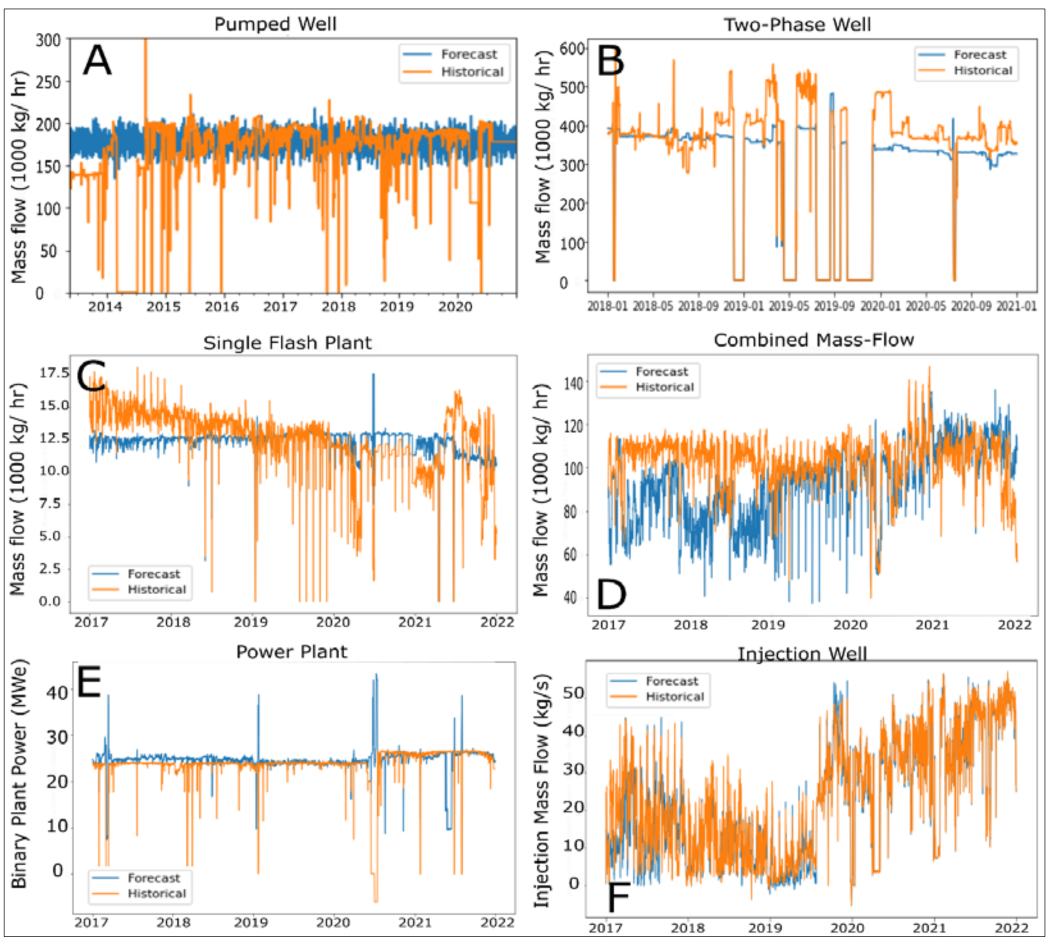

Figure 2: Components used in the GOOML environment to establish system performance and allow user interrogation. A&B show well historical and forecast mass flows, C is separation flow from a single flash-plant, D is total system separation flow, E is total system power output and F is matching of shallow and deep reinjection flows.

3. Plant Configuration - Genetic Algorithms

When the historical and hindcast models are aligned, the modeling framework can be used to make predictions (forecast models) and investigate other areas of interest. One area of research that has been considered is genetic optimization whereby components of a steamfield are chosen for their peak performance and contribution to system performance, much in the same way that natural selection selects for favorable evolutionary advantages (see Cheng, 2018). This approach to steamfield design identifies solutions that a traditional engineering approach may not. Whilst genetic algorithms should not be used to entirely replace traditional engineering techniques, insights from genetic optimization may give designers and operators other avenues to consider in steamfield design.

We used genetic optimization to interrogate a fictional plant within the GOOML framework called Big Kahuna (see the Open EI Geothermal Data Repository, Siratovich et al., 2021). Big Kahuna was built using a simple configuration of five production wells, two flash-plants and a single generator, each consisting of fictional but realistic dimensions, operating parameters, and time series data. These components were founded in the constructs of the GOOML models built on real-world historical plant data (Figure 3).

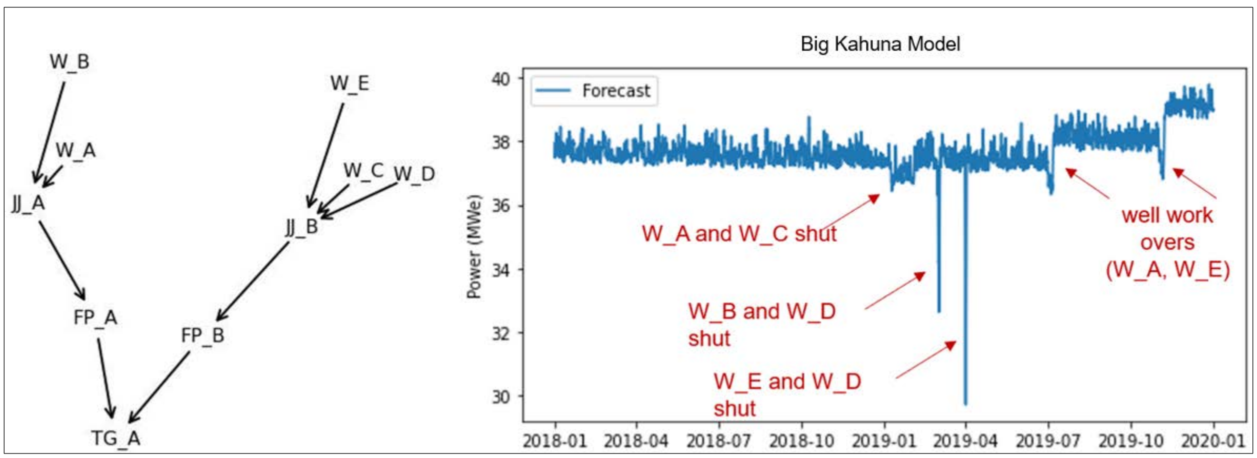

Figure 3: Big Kahuna architecture showing hypothetical performance based on well output, shuts and maintenance of wells in the GOOML modeling framework.

We can use genetic optimization to evaluate if there are configurations better suited to existing well outputs and power generation technology for a given steamfield. Figure 4 shows how performance of the plant can evolve based on the separator design parameters (e.g., number of flash tanks, pipe diameters, sizing, etc.) and ultimately provides an improved performance in power output. This modeling technique utilized genetic algorithms (GA) to achieve a best fit solution.

GA generally operate through the following phases: (1) generate, where a population is generated with individuals possessing a potential solution to the problem of interest, (2) computation of fitness, where each population member is given a fitness score representing its probability for reproduction, (3) selection, where the fittest population members pass on their genes to the next generation, (4) crossover, where genes are passed over to offspring, (5) mutation, where new offspring possess genetically mutated parameters and (6) termination, at which point the population has converged with genetically similar offspring providing a suitable solution to the problem proposed.

Using this approach, iterations of populations provide better solutions towards fitting objective functions. In the case investigated, we targeted power generation fitness from the population of possible separator designs. Users can investigate the optimization of any design parameter “gene”. This type of investigation produces configurations that otherwise would likely not have emerged through traditional engineering approaches to steamfield design. Through the random nature and selection process, better solutions emerge and population fitness increases. The solutions given using GA in our experiments are thought-provoking, but we suggest that the outputs of optimization experiments should still be fully verified through traditional safety and thermodynamic rigor to provide optimal steamfield design solutions.

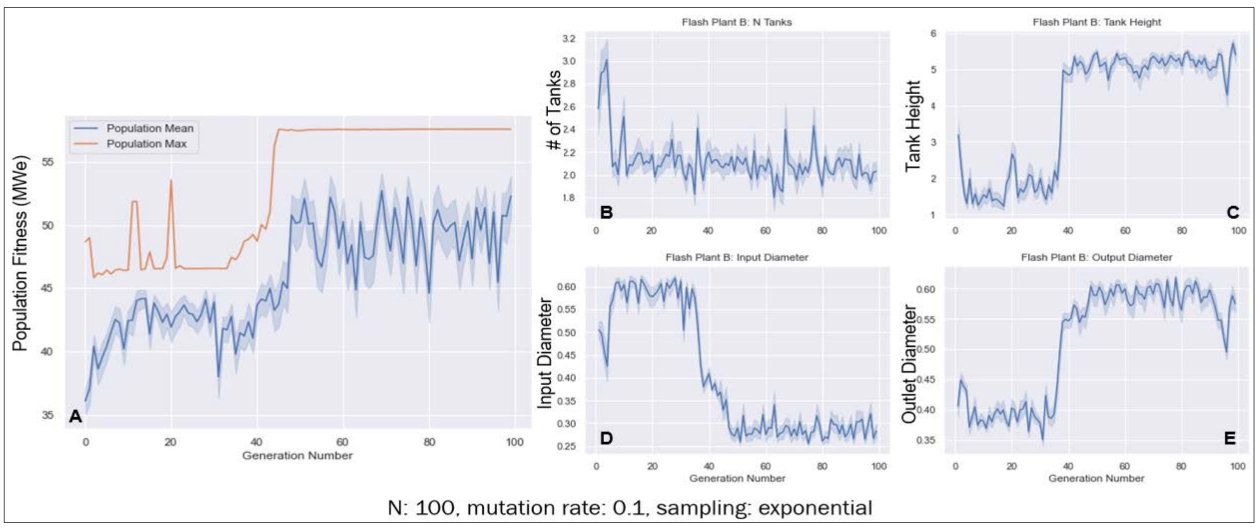

Figure 4: Big Kahuna during genetic optimization. A-showing improvement through overall system output through an optimization of B-number of flash-plant tanks, C- Tank height, D-Flash-plant input diameter and E-Flash-plant output diameter.

4. Operational Optimization - Reinforcement Learning

Like the genetic algorithms, reinforcement learning (RL) gives insight to how an operating plant might be optimized through allocation of plant flows and pressure distributions. Figure 5 shows the optimization of a power system hindcast from the baseline performance and hypothetical suggestions of an RL agent. The agent can provide an uplift in total system power output at a cost of higher mass-take for a large steamfield. Outputs from the GOOML environment can potentially guide an operator towards an uplift in power output by allowing the agent to change conditions that yield optimal mass flow and enthalpy within a generating plant or series of plants.

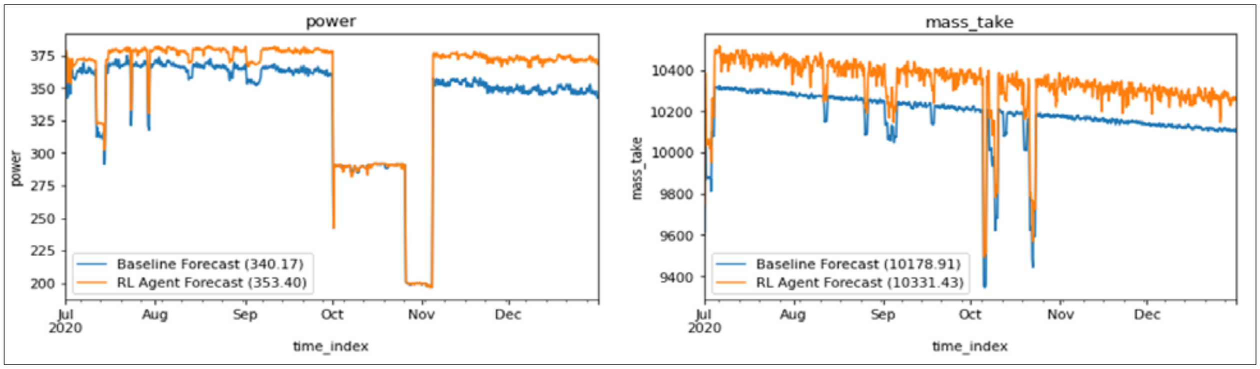

Figure 5: Output from optimization using an RL agent to optimize a complex steamfield. The baseline forecast (blue) is how the GOOML system would provide power and mass-take from a baseline model. The RL agent forecast (orange) is the optimized solution which provides a higher overall power output, but which is penalized by greater mass-take from the system.

To develop the optimization output shown in Figure 5, we allowed the reinforcement learning agent to change pressure on control wells within the field, bypass wellhead separators to favor larger centralized separators, and manage the direction of flow in pipelines that service different generators on the field of interest. We provided a reward for its activity that was founded in the baseline forecast model; any improvements above this forecast resulted in a reward to the agent. Further, initial experiments showed that the agent was taking liberties with flow direction switching (swinging pipe junctions) such that it attained rewards that were unrealistic in real world operations. To address this, the agent was penalized for each action that resulted in a flow direction change.

The activity of the agent shows that there could be possible further optimization of the system as the model showed an uplift of 13 MWe over the period investigated for the field of interest. This did however come with a mass-take extraction penalty of 150 tonnes/hour from the reservoir. These results should be considered carefully as the digital twin is only a representation of the system and may not account for all of the safety factors in place in surface plant (though we have endeavored to implement as many limitations and constraints as possible). It should be noted that additional mass-take may not be available for all reservoirs by all operators and a balance of resource sustainability and mass extraction should be considered against these results.

5. Future works and the future of GOOML

The use of machine learning in geothermal operations is still a nascent technology that has not seen wide-reaching application in power plant optimization. This may be attributed to the fact that each steamfield has unique characteristics and configurations that make off-the-shelf software solutions less suitable as seen in technologies like gas and wind turbine applications. However, we believe that the GOOML environment and the component-based approach could yield significant tools for geothermal operators to greater understand their operations and provide insight to areas for further optimization strategies.

The GOOML digital environment allows a programmatic and machine learning-driven approach to geothermal operational decision making. We have built data curation and historical modeling frameworks that clean and integrate datasets into graph data structures for use in training algorithms. We have also built a forecast modeling framework that can use trained models and simple seed data to predict how future decisions will impact operations. Thus, we have built a series of tools that are useful for interrogating past, present, and future geothermal operational decisions.

Our RL experiments demonstrate that the developed GOOML digital space is a powerful platform to investigate scenarios of operational conditions and system optimization. Future work will be focused on the deployment of GOOML tools for the geothermal industry. This will be done collaboratively with our commercial partners, ensuring they are relevant and useful to operations.

Acknowledgements

The authors of this paper would like to thank to our commercial partners: Contact Energy Limited, Ngati Tuwharetoa Geothermal Assets Limited and Ormat Technologies, Inc. for their gracious support of this project. Access to datasets, staffing, knowledge and institutional expertise has been critical to assure the research presented has solid foundations in geothermal operations.

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. This work was authored by Upflow, Ltd. and the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308 with funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy (EERE) Geothermal Technologies Office (GTO). The views expressed in the article do not necessarily represent the views of the DOE or the United States Government. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid- up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for United States Government purposes.

REFERENCES

Buster, G., Siratovich, P., Taverna, N., Rossol, M., Weers, J., Blair, A., Huggins, J., Siega, C., Mannington, W., Urgel, A., Cen, J., Quinao, J., Watt, R., Akerley, J. "A New Modeling Framework for Geothermal Operational Optimization with Machine Learning (GOOML)" *Energies*, (2021), 14, 6852. https://doi.org/10.3390/en14206852

Cheng, J. "Evolutionary Optimization: A Review and Implementation of Several Algorithms" Technical Discussion, *Strong.io*, October, (2018). https://www.strong.io/blog/evolutionary-optimization.

Siratovich et al., "GOOML - Finding Optimization Opportunities for Geothermal Operations", *Proceedings, Stanford Geothermal Workshop on Geothermal Reservoir Engineering.*, Stanford University, Stanford, CA, (2022), February 7-9.

Taverna et al., "*Data Curation for Machine Learning Applied to Geothermal Power Plant Operational Data for GOOML: Geothermal Operational Optimization with Machine Learning*" Stanford University, Stanford, CA, (2022), February 7-9.