Title: Modeling a Geothermal Development using methods in Data Science and Machine Learning

**Authors:** Munish Kumar1,2, Ivan Maulana1, Kanna Swaminathan1

1ERCE Singapore, 2Singapore University of Social Sciences

\*Corresponding author. Email: mkumar@erce.energy

**Abstract:** As an alternative energy source, geothermal energy is an appealing form of renewable energy and rapidly gaining mainstream acceptance, not only in Europe but also in parts of the Asia-Pacific as well. As in the case of oil and gas resources, its usefulness lies in its ability to provide heating, and for the generation of electricity. However, unlike oil and gas, its use is immediate. In most cases, no conversion or processing is required, and the by-product (cold water) can directly be reinjected into the ground. As the main energy output is heat, however, it requires careful and efficient planning when it comes to well placement and configuration, since borehole length, pipeline distribution and distance from source to market can directly impact the economics of the project. Particularly in the prospective stage of the project, understanding the well placement and distribution can also allow for modelling of reservoir performance and the impact of the cold front on the future deliverability of the heated fluids from the subsurface.

This paper will describe our workflow in modelling for the efficient well placement and pipeline distribution of a hypothetical geothermal field, assuming 2 demand locations, which required 80 MW and 100 MW of thermal output to be delivered. We will demonstrate how static properties were interpreted along with the use of a fairway mapping to identify sweet spots for optimal well placement. A geothermal simulator (DARTS, a product of TU Delft) was utilized as part of this work, to model the flow rates and doublet performance for the specific well pairs. We will also discuss the use of machine learning via a clustering algorithm to group wells pairs with respect to demand location and finally build an economic model that evaluates the levelized cost of heat. We finally conclude by proposing a generalised machine learning algorithm that allows for rapid evaluation of development scenarios as a first pass tool, which can be utilized to grade other prospect and lead locations as well.

**One-Sentence Summary:** This paper aims to demonstrate how a prospective geothermal development can be modelled using data analytics and machine learning, including aspects of economic modeling.

**Introduction:** The challenges of ensuring energy security and economic growth are fundamentally tied to issues dictated by geopolitical tension and the increasing awareness of climate related adverse weather patterns. This has put a new spotlight on the importance of diversifying energy resources, through the use of alternative fuels that have less of an anthropogenic impact. Alternatives such as wind, solar, biomass and hydrogen are some examples, but perhaps the one that is the closest in terms of technology and methods to conventional fossil fuels production is geothermal energy. Commonalities include phases within projects such as subsurface exploration and drilling, well completion, reservoir management, facilities and surface equipment and abandonment, the heavy capital-intensive nature of the projects and even how procurement and logistics are managed and function.

If viewed holistically, the output from any geothermal fields is hot water and/or heat. By its very nature, the outputs are natural low CO2 emitters and therefore would not be impacted by carbon taxation or other such emissions-based taxes. Typical projects have long lifespan (>25 years) and because surface facilities have a minimal footprint, are therefore easy to clean-up and abandon. Given that technical considerations for a geothermal resource are akin to conventional fossil fuel production (seismic and geological analysis, drilling a well, logging, coring and casing, perforation, flow rate, temperature and pressure determination, fluid injection), no significant technical barriers are present. Rather, the barriers tend to be commercial in nature, with lack of risk insurance schemes and longer payback periods.

**Mass and Energy Balance:** The geothermal potential of a given location is a delicate balancing act between the processes of conduction, convection and radiation, as heat is transported from the earths core, through the Earth’s mantle, to the oceanic lithosphere and continental crust.

For geothermal reservoirs, prediction of heat flow is based primarily on conduction of heat through rock and groundwater, as governed by Fourier’s Law

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 1 |
|  |  | Eq. 2 |

where Q is heat flow (W), A is the cross-sectional area (m2), l is thermal conductivity (Wm-1K-1), is the temperature gradient (○Cm-1), is Darcy velocity (ms-1), t is time (s) and is the temperature difference. Subscripts ‘r’ and ‘w’ are for rock and water respectively. As hot rocks essentially provide the energy to the ground water, therefore the total amount of heat (heat capacity) in the reservoir is constant, and if a low enthalpy geothermal system is assumed, then

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 3 |

where is the volumetric heat capacity of the system. For a geothermal extraction process, there must exist a well for production of hot water/heat and another for injection/disposal of the cold water. These “doublet” pairs must satisfy energy conservation such that the heat extracted from the rock, , must be equal to the heat available on the surface, . Therefore

|  |  |  |
| --- | --- | --- |
|  |  | Eq. 3 |
|  |  | Eq. 5 |

|  |
| --- |
|  |

**Background of the Doublet Geothermal System:** Geothermal “doublet” pairs are a type of geothermal system where heated fluid is produced from a production well, and concurrently injected into a nearby disposal well within the same formation. Such systems are particularly advantageous because (a) there is no issue with treating the produced water (particularly if there are fears of aquifer contamination), (b) pressure is maintained at the production well, and (c) surface subsidence is limited. While the operations surrounding sedimentary-basin type geothermal projects have been in study since the early 1970s, new, better computer models now allow for improvements in our ability to understand how such reservoirs are likely to behave. Such systems are commonly used in Canada [1], Netherlands [2], Belgium [2], France [4], Korea [4], Indonesia [5] and China [6] to name a few.

Determination of a suitable disposal well location is particularly challenging as it directly impacts the useful lifetime available in such two-well systems. Disposal wells need to be sufficiently far such that cool water breakthrough does not occur too early in the project lifetime. However, if well spacing is excessive, the project will incur additional costs brought about by the additional laying of pipeline, casing that must be modified to handle longer well life expectancies, right of access to land separating the 2 wells and increasing subsurface and geological uncertainty between the well pairs. This becomes even more challenging when there is a lack of data with respect to geology, geophysics, petrophysics, hydrogeology and thermal conditions.

**Machine Learning Methods: XXXX**

**Scope & Methods:** Our team performed this work as part of the SPE GEOHackathon organized by SPE Europe in 2021 [1], with a mandate given to participating teams to design the field development plan for a hypothetical geothermal project. Of the 60 teams from all over the world, our team was one of the few to make it to the final round.

The premise of the problem was that two demand locations (D1, D2) had to be supplied, with D1 requiring ~80 MW and D2 requiring ~100 MW of thermal output. Both locations were co-located within an area of interest (AOI) which measured 12.5 km by 12.5 km. In the same area, there existed drilled (but plugged and abandoned) wells spread randomly within the vicinity.

Officially, we were tasked with investigating the resource capacity and economic potential of this theoretical geothermal project. However, aside from the formal goals prescribed by the hackathon organisers, we quickly realized that an opportunity had present itself to build a “generalise” solution, where we could (a) build a workflow that considered both analytical and numerical outcomes, which we could than later port over to real-world applications in (for example) prospect and lead maturation, (b) optimize drill well location and pipeline arrangement prior to project investment using methods in machine learning, (c) evaluate levelized cost of heat (LCOH) over project lifetime and perform “what-if” scenario analysis, (d) build a generic machine learning model to predict rates at an arbitrary location and finally (e) evaluate greenhouse gas reduction over project lifetime.

***Data Preparation and Preliminary Exploration:***

Data was provided in the form of information relating to 12 pre-existing wells locations, of which there were 3 producer-injector pairs (P01-I01, P02-I02, P03-I03) and six exploration wells (E01 to E06). As shown in Figure XX, the exploration wells were on the boundary of the producer-injector pairs and were abandoned. Only some very basic sonic data, temperature and pressure data was collected in the exploration wells. Some of the wells had production tests conducted on them as well. Given the dearth of information, we made the following assumptions: (a) that while the wells were drilled respecting some topographic variation (on the “highs”), there was not a huge amount of relief and that the area was relatively flat, (b) that the existing twelve wells were vertical, (c) any well logs measured across these wells were across the reservoir section only and (d) that the reservoir from which the hot water was flowing was a “typical” sandstone reservoir, with shale breaks between the sands. Therefore while there could be porosities at ≥ 45 p.u., these were aberrations and that any sandstone reservoir would fall within the porosity ranges < 45 p.u.

Existing exploration wells and producer-injector pairs were isolated from the demand locations; we decided that a few more test boreholes were required closer to the

**Results:**

Table 1: Blah Blah

|  |  |
| --- | --- |
|  | |
| *Moo* | *Moo* |

Figure 7:Moo Moo

***Discussion and Implications:***

***AAA****:*

***BBB:***

***CCC:***

***DDD:***

***EEE:***

**Limitations of Study and Conclusions:**

# **References**

**Acknowledgments:** We would like to thank ERCE for peer review of this work.