PROCEEDINGS, 49th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 12-14, 2024
SGP-TR-227

Machine-Learning Methods and Tools Designed for Community-Based Equitable and Inclusive Geothermal Development

Velimir ("Monty") Vesselinov, Hope Jasperson, Tracy Kliphuis

monty@envitrace.com

hope@envitrace.com

trais@envitrace.com

Keywords: Machine learning. Artificial intelligence, Robust, Trustworthy, Scientifically defensible, Science-Informed Machine Learning, Community based, Equitable, Inclusive, Geothermal, Development, Exploration

ABSTRACT

To achieve net-zero carbon emissions, it is essential to involve communities in the implementation of green energy technologies. This can be done through informed decision-making, community-centered research, and engagement of stakeholders at the local, state, and regional levels. Community-led research and implementation are fundamental to achieving success. These collaborations should include rule makers, environmental regulators, clean energy industries, and technology researchers and developers. Unfortunately, many green infrastructure initiatives still adhere to a top-down and expert-driven process of site selection and design without awareness and acknowledgment of public engagement needs. This can lead to costly delays, including lawsuits, and ultimately less than desired or lacking outcomes as well as missed opportunities. Geothermal energy is a promising green energy source, and it is important to ensure that its expansion is equitable and does not disproportionately impact disadvantaged and underprivileged communities. To address this, we have developed a novel web-based interactive software and user-friendly interface called GeoTGo (https://geotgo.com) that provides everything that is needed for communities to better understand and develop their geothermal resources. Our website provides access to our machine-learning method and tools. Our tool also includes a comprehensive and living Community Engagement Plan (CEP) that was collaboratively developed with members of underprivileged and underrepresented communities. It bridges the gap between technology advancements and community needs by facilitating the interactions between the geothermal industry, regulators, stakeholders, and end-users. GeoTGo merges data, software (including data analysis, text mining, artificial intelligence, and modeling tools), knowledge, expertise, and experience to provide fast processing and dissemination of the latest information about cutting-edge geothermal technologies to users and communities. Machine learning and artificial intelligence methods in GeoTGo are based on our existing open-source algorithms (SmartTensors, https://github.com/SmartTensors, SmartML, https://github.com/SmartTensors/SmartML.jl, MADS, https://github.com/madsjulia). GeoTGo is a valuable tool that can help communities to develop their geothermal resources in a way that is equitable and sustainable. We are currently working with several New Mexico Native Nations to pilot the tool and plan to expand it to other communities in the future. GeoTGo will help accelerate the development of geothermal energy and contribute to the achievement of net-zero carbon emissions.

1. INTRODUCTION

Geothermal energy is one of the most attractive renewable energy options; it is the only source capable of delivering consistent and reliable electricity 24 hours a day, 365 days a year, and it is right under our feet. Geothermal is critical for the socioeconomic development of underprivileged communities in the Southwest U.S., where substantial geothermal potential exists. The lack of geothermal development will negatively impact our society and environment through the continued utilization of fossil fuels.

Significant barriers to the wide use of geothermal energy include a lack of understanding, weak or non-existent supportive legislation/incentives, and high costs/risks associated with geothermal exploration and utilization (Prodi 2014). For effective and equitable outcomes towards achieving the national goal of net-zero carbon emissions, communities must be included and lead the implementation of innovative green-energy technologies. Collaborations with communities should happen through informed, community-centered research and local, state, and regional engagement. Unfortunately, many energy infrastructure initiatives still adhere to a top-down and expert-driven process of site selection and design without awareness and acknowledgment of public engagement needs. This can lead to costly delays, lawsuits, less-than-desired or lacking outcomes, and missed opportunities (Simons 2017).

Geothermal, like other new technologies whose impacts are not fully understood, may cause disproportionately high adverse effects on disadvantaged communities if the process does not fully account for their socioeconomic and cultural interests and concerns. That is why, under this project, we focus on the interplay between the technological and socioeconomic factors impacting geothermal utilization. A large portion of the potential geothermal resources in the U.S. is located in areas with predominantly Native American Nation (NAN) and underprivileged communities. That is why our work aims to engage with these communities and their workforce and businesses. Our work will support local workforce and business development in geothermal and AI/ML.

Our country has the largest known geothermal potential in the world (Bhatnagar et al. 2022). However, the resources are underutilized. The geothermal industry, decision-makers, and communities need tools that can help make decisions related to (1) where and how these resources can be tapped, (2) what the deployment options are, and (3) what the potential total energy output is. Geothermal utilization also depends on many other factors, including energy demand, socioeconomic needs, and existing infrastructure. Geothermal systems are expensive to construct and maintain. The advancements proposed under this project are critical to addressing these issues. Our primary goal is to remove the socioeconomic and technological barriers associated with geothermal exploration and utilization. Our work will support the development of local job opportunities and businesses. Our work will also address the methodological and technical risks associated with developing viable commercial software and services for geothermal exploration and utilization.

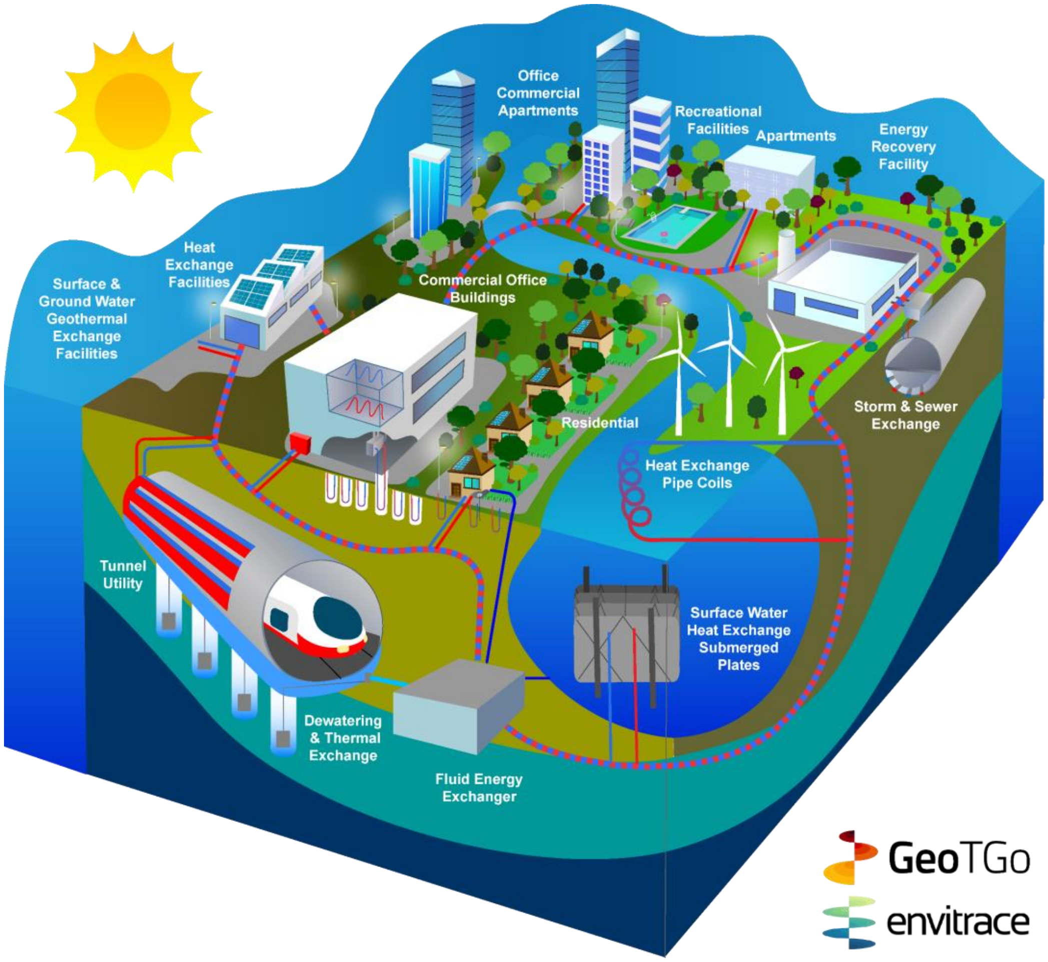


Figure 1: Geothermal uses in urban settings.

1. SCIENCE-INFORMED MACHINE LEARNING (SIML)

A critical aspect of our work is the development and demonstration of SIML (Science-Informed Machine Learning) technology for processing geological (geothermal, geochemical, geophysical, hydrogeological, etc.) data. These methods and tools are critical for the robust and efficient utilization of geothermal resources. They are also important for scientifically defensible assimilation of the available data and parameterization of the subsurface governing processes.

SIML methods are different from traditional ML techniques (**Figure 2**). In both cases, the ML models are trained to predict the spatial distribution of an output (e.g., pressure, temperature, heat flux) based on a series of inputs (e.g., permeability, porosity, etc.). The traditional ML (a) relies on deep and wide neural networks (NNs) based on simple algebraic mappings to represent complex processes. However, the conventional neurons (using standard activation functions) do not explicitly capture any physics. All constitutive relationships must be learned exclusively from the input data.

In contrast, the SIML neurons incorporate complex mappings (including constitutive relationships and physics/chemistry models) (**Figure 2**). This results in ML models that have a physical meaning and satisfy physics laws and constraints (e.g., Darcy's law, mass/energy conservation, stress/deformation relationships, etc.). SIML can also include automatically differentiable numerical models. As a result, the SIML NN as outlined in **Figure 2** can execute model selection. SIML can choose which of the provided alternative "flow" models is the most applicable to reproduce the data. SIML can also decide to use some combination of these models. For example, a series of alternative expressions can be provided relating fracture properties to the medium porosity and permeability. SIML can select or combine these relationships to represent the observed site data. SIML also accounts for the physics associated with different governing processes critical for the characterization and parameterization of the site conditions (e.g., flow, stress, deformation, etc.).

Alternatively, SIML analysis can be performed in cases where the loss function applied for optimization of the NN includes a differentiable (DP) model predicting a series of observable outputs (**Figure 3**). The discrepancies between model predictions and field measurements are applied to train an NN, parameterizing model inputs to represent the observed data.

As a result, SIML models are more robust, trustworthy, interpretable, and defensible than traditional ML models. SIML requires less data for training, testing, and validation and increases predictability and opportunity for knowledge transfer between different sites.

SIML methods can also be applied to develop robust, fast, surrogate (reduced-order) models predicting geothermal conditions and utilization based on site-specific data.

It is important to emphasize that the development of traditional ML is general, utilizing similar NN architectures. However, the SIML is problem specific. SIML for different science domains involves unique combinations of mathematical expressions, constitutive relationships, and numerical/analytical models tailored for the particular task.

SIML requires the development of a differential programming (DP) reservoir simulator. In our case, a DP simulator is integrated into our existing computational workflows for data analytics, model diagnostics, and SIML. The simulator is also coupled with existing products developed by our company (e.g., SmartTensors, SmartML, MADS). We have applied the DP simulator coupled with our SIML techniques to predict geothermal utilization prospectivity based on provided data. The obtained results are discussed below.

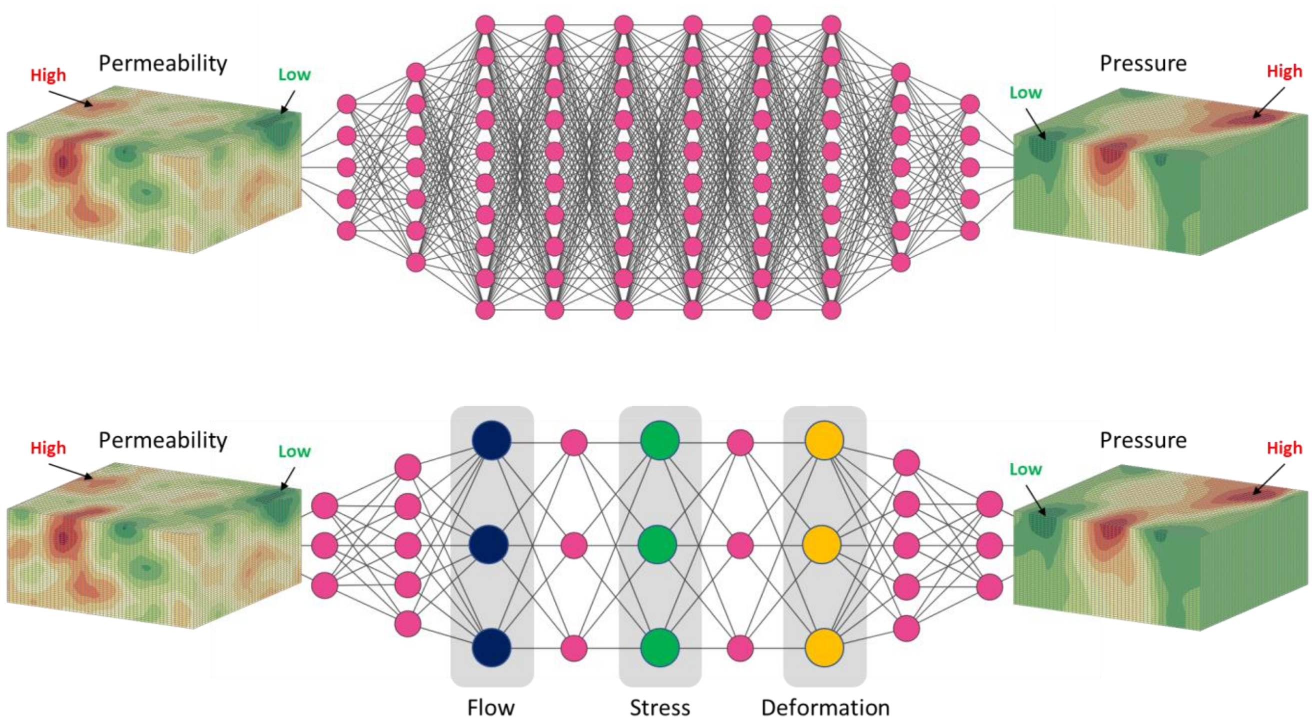


Figure 2: Traditional ML (top) vs SIML (bottom). In both cases, the ML models are trained to predict the spatial distribution of an output (e.g., pressure; right) based on a series of inputs (e.g., permeability, porosity, etc.; left). The traditional ML (a) relies on deep and wide neural networks (NNs) based on simple algebraic mappings to represent complex processes. However, the conventional neurons (using standard activation functions) do not explicitly capture any physics (a). In contrast, the SIML neurons incorporate complex mappings (including constitutive relationships and physics/chemistry models) (b). This results in ML models that have a physical meaning and satisfy physics laws and constraints (e.g., Darcy's law, mass/energy conservation, stress/deformation relationships, etc.). SIML can also include automatically differentiable numerical models. As a result, SIML can execute model selection: SIML can choose which "flow" model is the most applicable to reproduce the data. SIML also accounts for the physics associated with different governing processes (flow, stress, deformation, etc.).

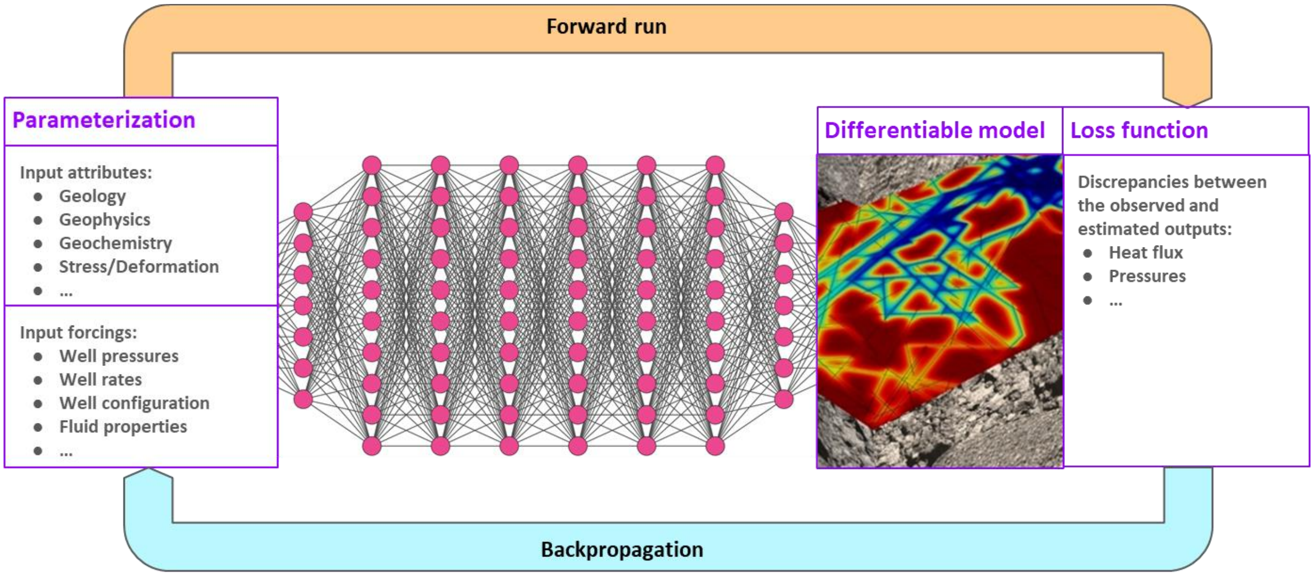


Figure 3: Example SIML analysis where the loss function includes a differentiable (DP) model predicting a series of observable outputs. The discrepancies between model predictions and field measurements are applied to train an NN, parameterizing model inputs to represent the observed data.

2. GeoTGo

We are developing a cloud-based interactive software called GeoTGo designed to provide everything needed to better understand and develop geothermal resources at local/regional scales, accounting for socioeconomic and geologic conditions. In summary, our software aims to:

- apply Science-Informed Machine Learning (SIML) to geothermal problems
- bridge the gap between geothermal technology advancements and community needs
- facilitate interactions between the industry, regulators, stakeholders, and users
- merge physics and engineering data, knowledge, expertise, experience, and models
- provide fast processing and analysis of available data
- suggest site-specific options for geothermal utilization depending on the subsurface conditions and energy needs
- propose alternative geothermal uses, including electricity/hydrogen generation, mineral and trace-element mining, greenhouse/district heating, recreation, and geoexchange heating and cooling.
- support equitable and inclusive community geothermal development

Key GeoTGo features that are designed to address our goals related to data processing and analysis are:

- cloud-based computing and data processing
- access to public datasets and management of proprietary data
- curing-edge data-analytics, model-diagnostics, and AI/ML methods and tools
- analyses incorporating socioeconomic, engineering, and geoscience data and knowledge
- transfer learning of information and experience between geothermal sites and regions
- data input-output capabilities supporting existing web databases (e.g., GDR, EDX, NGDS, and existing state-level and federal public data resources), their APIs, and commonly-used file formats

To support the needs of business customers, GeoTGo will provide a series of decision tools critical for geothermal exploration and utilization:

- cost and return of investment (RoI) estimates based on energy cost and needs
- prediction of geothermal energy demand and utilization
- optimization of data acquisition strategies for efficient geothermal exploration and utilization
- optimization of geothermal production
- design of geothermal systems that efficiently mitigate risks (e.g., environmental, economic, social) and are representative of community needs

To address the challenges associated with the community engagement:

- design based on community engagement inputs addressing community and local workforce needs
- access to training materials, reports, papers, informational videos, regulations, standards, and lessons learned
- dissolves barriers between technology capabilities and community needs
- intuitive interactive user interfaces for access from a wide range of devices, including mobile apps

GeoTGo is still in development. However, the software prototype and the existing dashboard are accessible through our GeoTGo (https://geotgo.com) website. Screenshots of the GeoTGo dashboard are showcased in **Figure 4**.

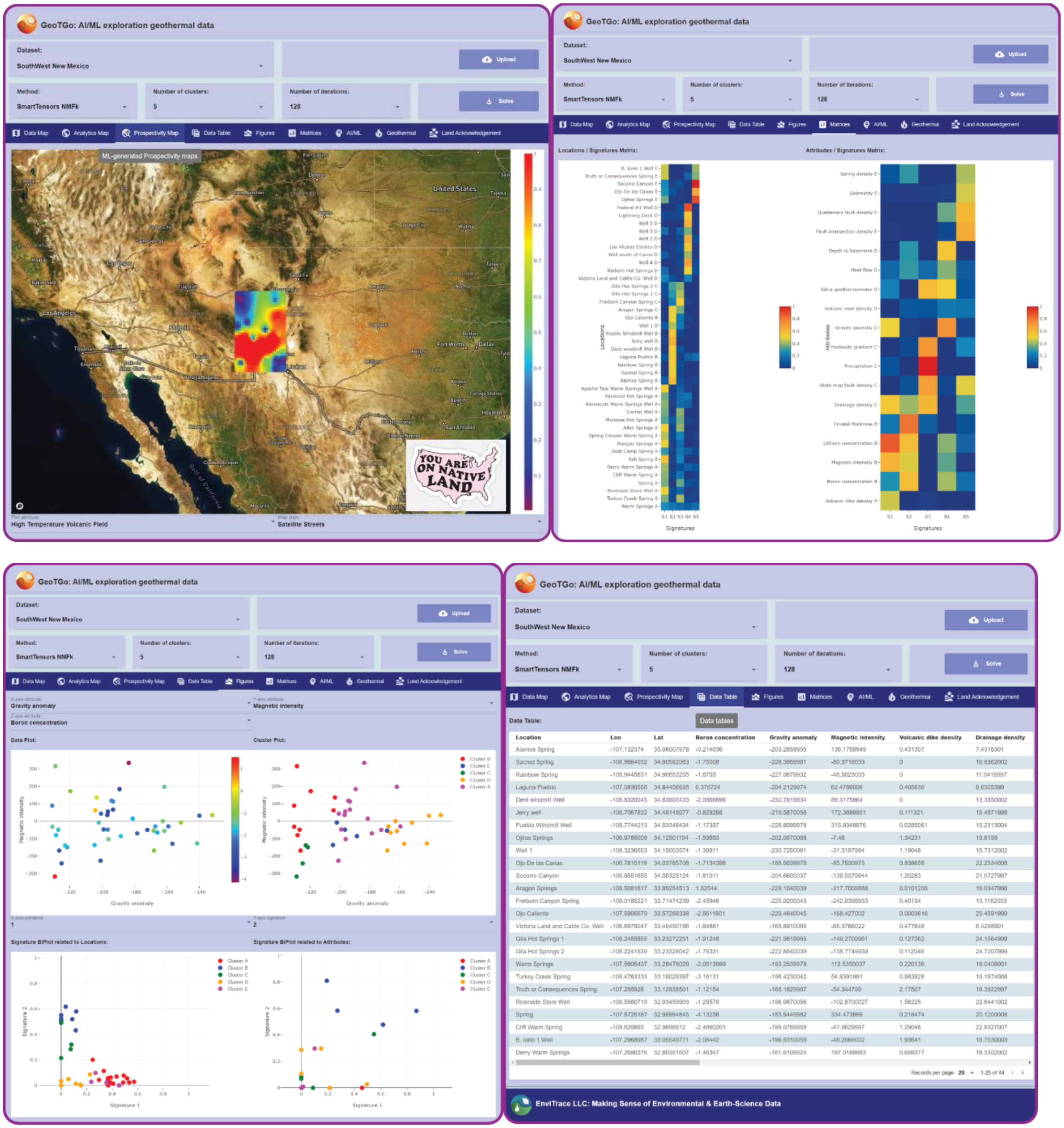


Figure 4: GeoTGo dashboard (http://geotgo.com) provides access to data and AI/ML methods and tools. Currently, only public data can be processed; proprietary data access is under development. A series of alternative datasets representing the geothermal conditions in New Mexico and the Great Basin can be selected. The dashboard provides cloud data management and computing. It also visualizes the data and results obtained. Results related to Great Based are showcased in Figure 7 below.

3. GEOTGO/SIML RESULTS

The setup of the developed DP simulator for integration in the SIML process is shown in **Figure 5**.

The obtained results are presented in **Figure 6**. Here, GeoTGo estimates for geologic-engineered systems consisting of injection and extraction wells. The 3 plots in **Figure 6** (**a**, **b**, and **c**) present solutions for different permeability fields. The plots show permeability fields (top left corner) and corresponding pressure fields predicted by GeoTGo. The SIML model also predicts the optimal rates constrained by a pressure of less than 1 m at a given distance from the system's center (here, 3000 m).

In addition, GeoTGo allows for the processing of a series of alternative datasets representing the geothermal conditions in New Mexico and the Great Basin. These datasets can be selected and processed through the GeoTGo dashboard (**Figure 4**).

For example, **Figure 7** demonstrates the results obtained for the Great Basin. The Great Basin is a vast region (covering areas of Nevada, Utah, California, Oregon, and Idaho) with significant geothermal potential. Many geothermal studies have been conducted over the last 50 years (Ayling 2022; Faulds 2015; 2021; GBCGE 2022; Wannamaker et al. 2020). Geothermal prospectivity so far has been evaluated through a laborious manual process requiring advanced subject-matter expertise. Now, GeoTGo can achieve similar results in minutes. GeoTGo has been applied to discover geothermal prospectively of the Great Basin from limited, sparse geochemical data (18 attributes; TDS, B, Mg, etc.) observed with gaps and uncertainties.

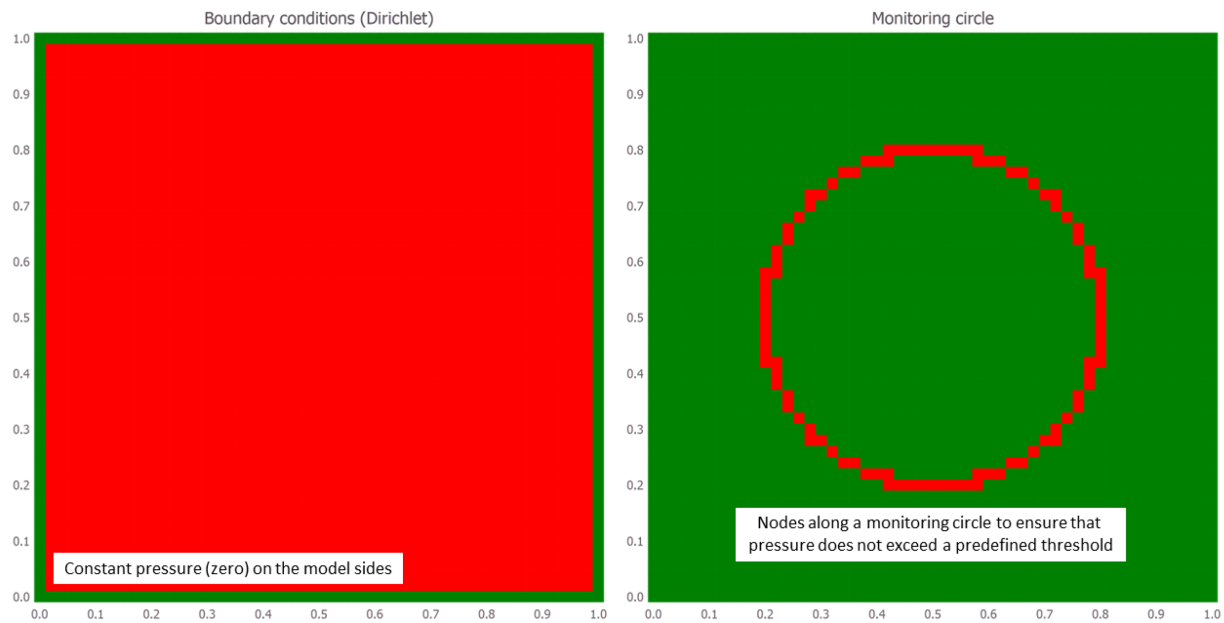
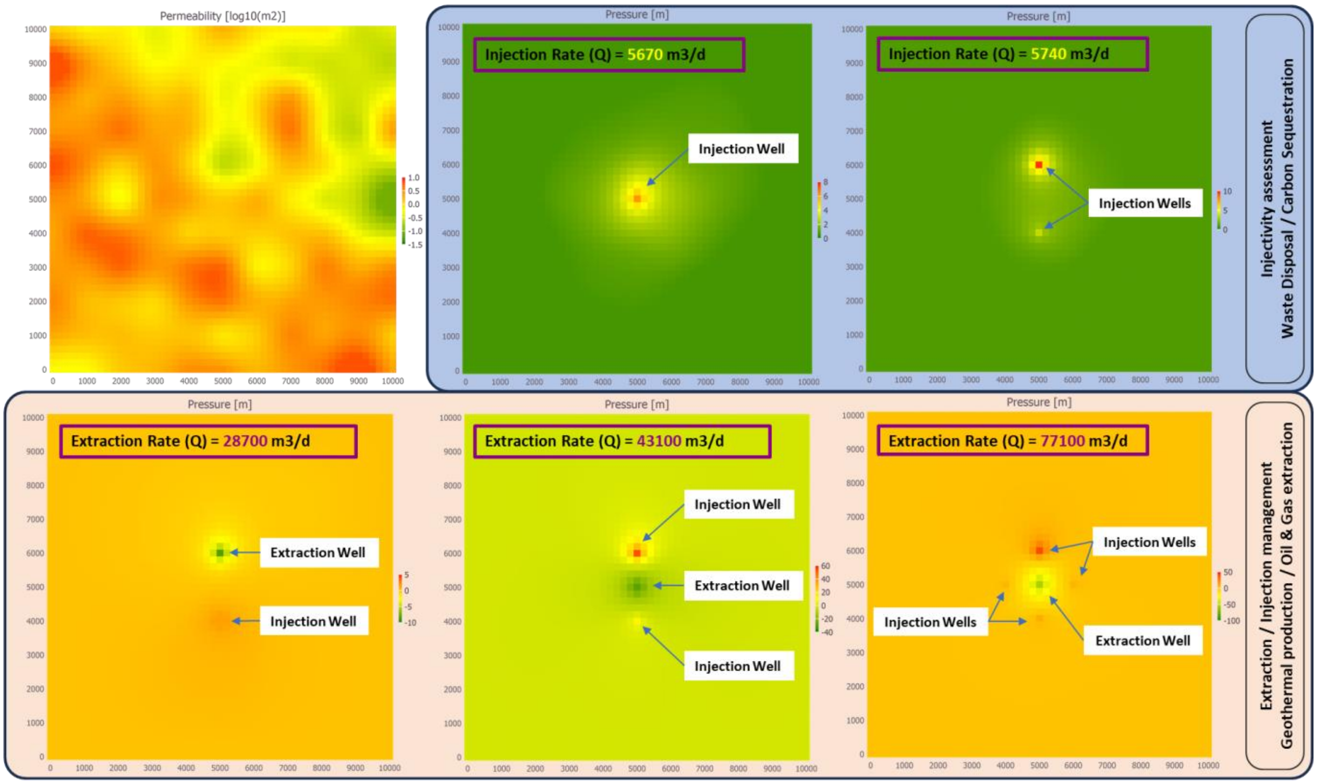
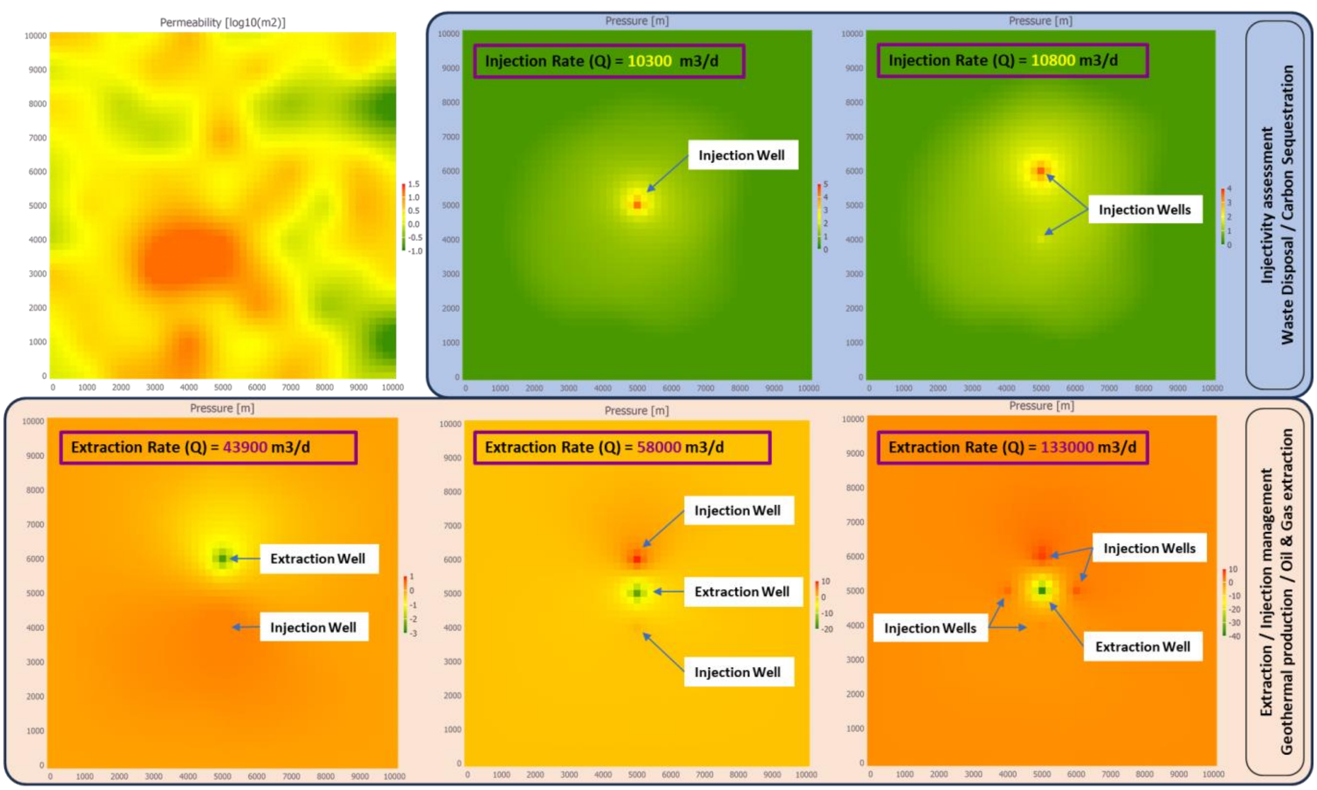


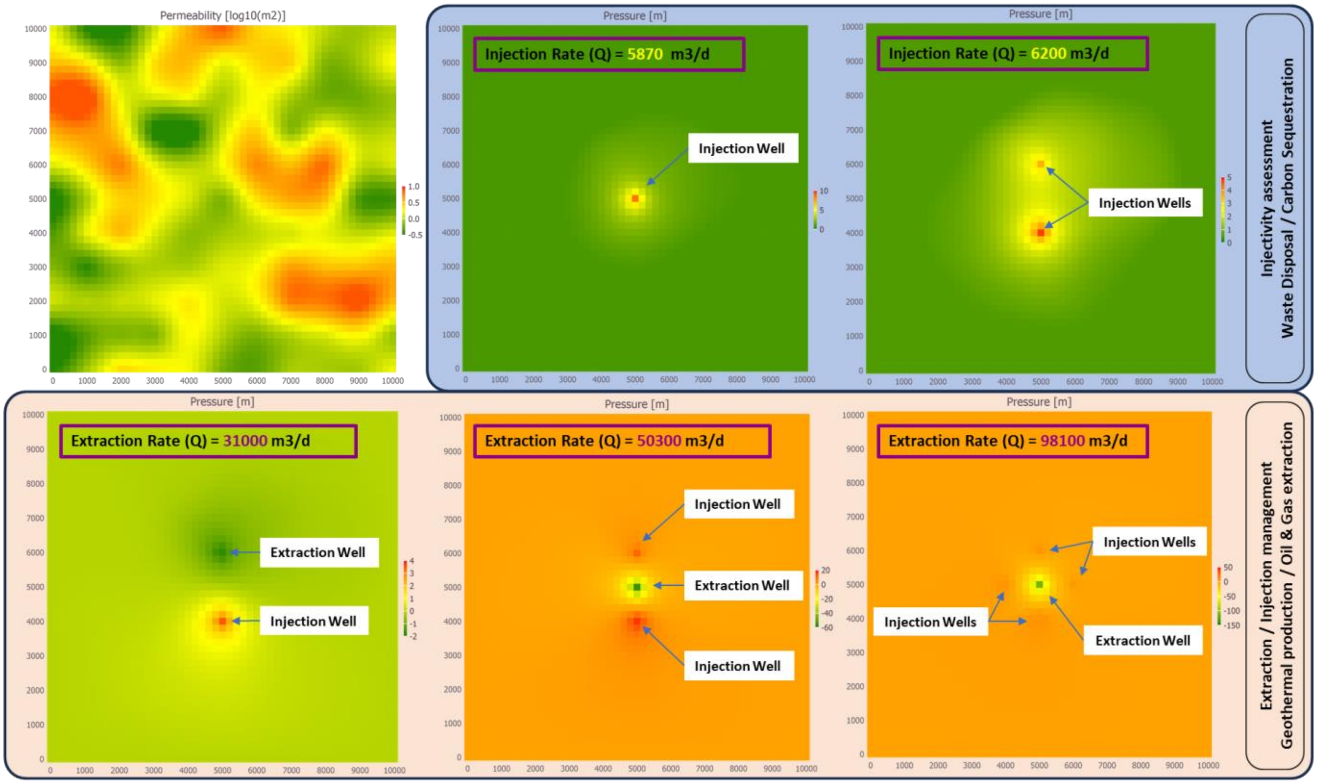
Figure 5: GeoTGo model setup. The model domain and other model parameters are considered to be dimensionless. This allows the developed SIML models to be applied to a range of reservoir conditions. Dimensionless model parameters include domain size, radius of monitoring circle, distance between wells, permeability, and pressure threshold.



(a) Random Permeability Case #1



(b) Random Permeability Case #2



(c) Random Permeability Case #3

Figure 6: Results demonstrating the applicability of SIML to predict injection/extraction rates. SIML provide predictions of geothermal reservoir utilization given different permeability fields (a, b, and c). The plots present SIML estimates about potential geothermal utilization assuming different configurations for well injection/extraction systems. The SIML model predicts the optimal rates by applying a constraint that the pressure buildup is less than a given threshold (here, 1 m) at a given distance from the center of the system (here, 3000 m). The SIML model also predicts the pressure field. The distance between the wells can also be varied and it also impacts the ML predictions. These analyses are also demonstrated on our GeoTGo software prototype website and dashboard (http://geotgo.com).

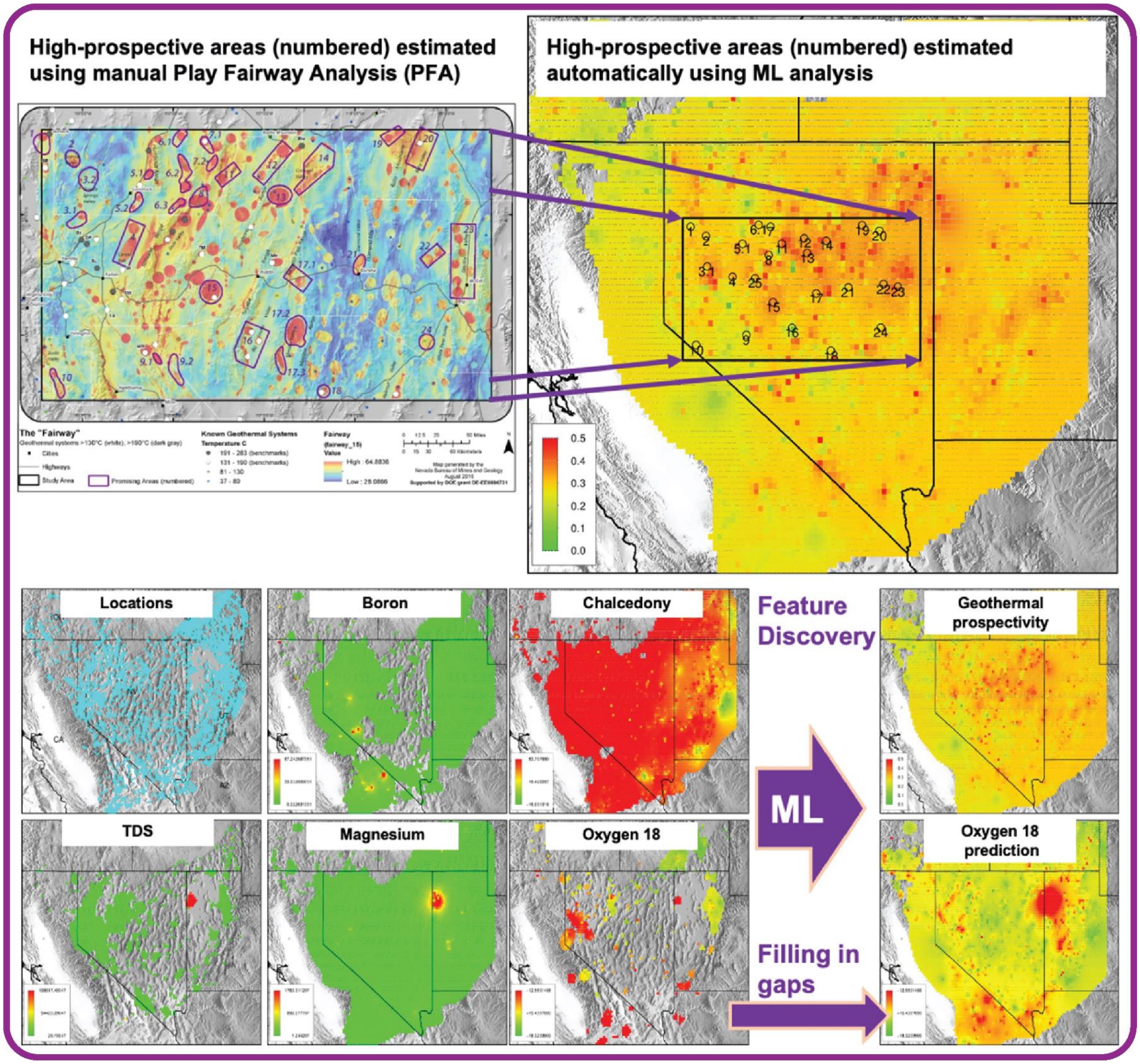


Figure 7: The Great Basin is a vast region (covering areas of Nevada, Utah, California, Oregon, and Idaho) with significant geothermal potential. Many geothermal studies have been conducted over the last 50 years. Geothermal prospectivity so far has been evaluated through a laborious manual process requiring advanced subject-matter expertise (top left). Now, GeoTGo can achieve similar results in minutes (top right). GeoTGo has been applied to discover geothermal prospectively of the Great Basin from limited, sparse geochemical data (18 attributes; TDS, B, Mg, etc.) observed with gaps and uncertainties (bottom). These analyses are also demonstrated on our software website and dashboard (http://geotgo.com).

3. CONCLUSIONS

We are actively developing novel Science-Informed Machine Learning (SIML) methods and tools designed to address geothermal needs associated with exploration and utilization. Our ML methodology is embedded in a cloud-based interactive software dashboard called GeoTGo (http://geotgo.com). It is designed to provide everything needed for communities to better understand and develop their geothermal resources. GeoTGo provides cloud data management and computing. Our work bridges the gap between technology advancements and community needs by facilitating the interactions between the geothermal industry, regulators, stakeholders, and end-users. GeoTGo merges data, software (including methods and tools for data analysis, text mining, artificial intelligence, numerical simulators, and visualization), knowledge, expertise, and experience to provide fast processing and dissemination of the latest information about cutting-edge geothermal technologies.

REFERENCES

Ayling, Bridget. 2022. “INGENIOUS - Great Basin Regional Dataset Compilation.” 1391. USDOE Geothermal Data Repository (United States); GBCGE, NBMG, UNR. https://doi.org/10.15121/1881483.

Bhatnagar, Dhruv, Hannah Rabinowitz, Daniel Boff, Barbara GR Pennell, Christine Holland, Mark Weimar, Bowen Huang, Di Wu, and Alain Bonneville. 2022. “Nontechnical Barriers to Geothermal Development.” PNNL-32760, 1879758. https://doi.org/10.2172/1879758.

Faulds, James. 2015. “Discovering Geothermal Systems in the Great Basin Region: An Integrated Geologic, Geochemical, and Geophysical Approach for Establishing Geothermal Play Fairways.” *PROCEEDINGS, 41st Workshop on Geothermal Reservoir Engineering*, September. https://www.osti.gov/biblio/1724098.

———. 2021. “Geochemistry and Paleo-Geothermal Features - Applications of Machine Learning Techniques to Geothermal Play Fairway Analysis in the Great Basin Region, Nevada.” 1352. USDOE Geothermal Data Repository (United States); Nevada Bureau of Mines and Geology. https://doi.org/10.15121/1832126.

GBCGE. 2022. “Nevada Play Fairway Phase 2.” Great Basin Center for Geothermal Energy. 2022. https://gbcge.org/recent-projects/nevada-play-fairway-phase-2/.

Prodi, Vitorio. 2014. “‘Significant Barriers’ to Geothermal Power Must Be Overcome,” 2014. https://www.theparliamentmagazine.eu/news/article/significant-barriers-to-geothermal-power-must-be-overcome.

Simons, Elise J. 2017. “Exploring Roles for Communities in Green Infrastructure Projects.” PhD Thesis, Tufts University.

Wannamaker, Philip E., Joseph N. Moore, Gregory D. Nash, Stuart F. Simmons, Virginia Maris, and Kevin Mendoza. 2020. “Play Fairway Analysis: Structurally Controlled Geothermal Systems in the Eastern Great Basin Extensional Regime, Utah.” Univ. of Utah, Salt Lake City, UT (United States). https://doi.org/10.2172/1695591.