PROCEEDINGS, 50th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 10-12, 2025
SGP-TR-229

GeoCLUSTER v2.0: A Closed-Loop, Techno-Economic Simulator Supporting New Case Studies

Anastasia Bernat[1], Alexander Buchko[1], Koenraad Beckers[2], and Aaron Moreno[1]

[1]Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99354

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

anastasia.bernat@pnnl.gov, alexander.buchko@pnnl.gov, Koenraad.Beckers@nrel.gov, aaron.moreno@pnnl.gov

Keywords: closed-loop, techno-economic, simulator, Slender-Body Theory (SBT)

ABSTRACT

Closed-loop geothermal systems (CLGSs) use a closed-loop heat exchanger such as U-loop or co-axial systems for subsurface heat extraction. These systems have recently received significant attention and investment, with several companies developing and commercializing this technology. Additionally, access to consolidated, independent, high-quality simulations for early scoping and/or project management purposes has become increasingly useful. However, open-source software for rapid data exploration and decision-making for closed-loop projects is limited. Most existing packages or tools serve as general-purpose subsurface simulators, and they may not provide above-ground economic modules. Likewise, their software may operate on legacy code, require local setup, and rely on the user scripting default conditions or researching hundreds of parameters to roughly mimic a closed-loop system. Furthermore, proprietary applications that comprehensively serve both below-ground and above-ground techno-economics currently need to be purchased. To this end, we present GeoCLUSTER v2.0: an open-source, cloud-native, techno-economic web simulator that enables start-up developers and venture capitalists to rapidly explore the economic viability of closed-loop geothermal systems, such as capital and levelized costs. Users can explore scenarios through several methods: 1) toggling between the heat-exchanger designs, working fluids, and end-use, 2) optimizing power output and economic competitiveness by clicking on the scenario buttons and moving easy-to-use sliders, and 3) visualizing simultaneous graphics and summaries. In GeoCLUSTER's v2.0 release, we integrate a Slender-Body Theory (SBT) model that enables users to rapidly and semi-analytically simulate any type of U-loop and co-axial system which were not originally included in its first release, such as depths deeper than 5 km and geothermal gradients larger than 70°C/km. This speed and parameter space expansion is due to the SBT's 1D discretization along the wellbore that, unlike 3D finite element (FEM) or finite volume methods (FVM) that must solve the entire reservoir volume extending outwards from the well, can cut down on orders of magnitude and mesh size. In turn, with data and compute increasingly integrating into the tool, we profile the application's memory usage and significantly reduce its memory footprint (97% drop). Thus, GeoCLUSTER acts as an agile software for quickly adapting to emergent geothermal feasibility research, and in this paper, we present GeoCLUSTER's most recent updates as it supports continued systematic evaluation of the technical performance and cost-competitiveness of closed-loop geothermal systems.

1. INTRODUCTION

Risk averse compared to enhanced geothermal systems (EGSs) (Budiono et al., 2022) but capital-intensive (Liu et al., 2024; White et al., 2023), closed-loop geothermal systems (CLGSs) hold the promise of clean electricity on the upwards of 300+ GW by 2050 across the United States (Blankenship et al., 2023; INL, 2006) but also face the challenges of high drilling costs (White et al, 2023; Fervo, 2023) and efficient heat production over extended operational periods. These systems are a next-generation geothermal technology whose heat-exchanger design circulates a fluid through closed wellbores drilled in the subsurface, and their commercial viability has largely depended on drilling wells at depths far deeper than those of current geothermal wells in order to reach commercial levels of power production (Fervo, 2023). Their next-gen counterpart, the EGS, conducts far greater heat extraction than CLGSs by requiring hydraulic fracturing to create pathways for the fluid to directly penetrate the reservoir, but closed-loop projects are afforded the benefit of avoiding risks from hydraulic fracturing, such as induced seismicity, reservoir sustainability, soil contamination, and water drawback (White et al., 2023; Alimonti et al., 2018). To then bridge the "heat gap", technical efforts have been underway to further enhance the heat extraction and reduce the costs of CLGSs by, for example, repurposing abandoned oil wells and adding insulation like polyurethane foam (Liu et al., 2024). In turn, identifying how to close these gaps for optimized heat extraction and economic feasibility of CLGSs can add further value for both current and future closed-loop projects that aim to commercialize CLGSs.

Efforts to increase investments towards commercializing closed-loop projects can then significantly benefit from decision support tools that easily enable early scoping on the performance of different system designs. This is becoming more relevant as geothermal tools continue to integrate sensitivity studies and model increasingly complex and wider parameter spaces that compute vast simulations in support of high-quality analyses (Budiono et al., 2022). In its most reporting, the Department of Energy's (DOE) Geothermal Data Repository (GDR) hosts 1,285 datasets amassing 287 TB of data for the geothermal community (Weers et al., 2025). And for closed-loop projects that have "just begun to come online" (Blankenship et al., 2023), they may anticipate data volumes similar to pioneers like DOE's FORGE that currently hosts over 98 TB of wellbore, rock property, seismic, and facture mapping data (FORGE, 2025). In turn, retrieving information and analytics across massive data pools and their various modalities, while balancing data storage and compute, becomes increasingly valuable – as recently architected by GeoMap™ (Project InnerSpace, 2025) –, with some of those retrieval and predictive tasks are already being uptaken by artificial intelligence (AI) and machine learning (ML) capabilities, as demonstrated by GeoTGo (Kliphuis et al., 2025). Simulators have then become increasingly useful for capturing and evaluating the techno-economic feasibility of next-generation geothermal technologies – in particular, easy-to-access online tools with high quality data and a low memory footprint for seamless scoping.

However, for closed-loop projects, such tools are still rare. Existing tools extensively consider geothermal system design or serve as general-purpose geothermal simulators, such as STOMP-GT (PNNL, 2023), T2Well (Pan et al., 2024), PLOTRAN (2024), FEHM (2017), OpenGeoSys (2025), and GEOPHIRES (NREL, 2018). However, these packages or modules often operate on legacy software, require setup in a local environment, or rely on the user knowing and researching how to script and tune their default wellbore and reservoir boundary conditions in order to mimic a "sealed" or closed-loop system. For easier setup, other independent tools, like GEOPHIRES or GES-CAL (Bláquez et al., 2020), provide interactive simulation, either through the web or a local user interface, but require user training to navigate, at times, hundreds of editable, techno-economic parameters. These interfaces are often suitable for subject matter experts, but for start-up developers, venture capitalists, or power purchasers looking to rapidly explore the economic viability of closed-loop geothermal systems, they may not be readily adopted. Likewise, state-of-the-art, commercial software often extensively consider intuitive user flows; however, their capabilities may focus largely on subsurface modules, simulating the thermal performance of ground source heat pump systems, while overlooking techno-economic modules for various closed-loop configurations (e.g., ResFrac (Fowler et al., 2021), CMG GEM™ (2017)). Even in the most comprehensive demonstrations of "out-of-the box", closed-loop modules, like Eavor-Loop™ that offers cloud-native, advanced builder capabilities for users to simulate their techno-economic case studies, still need to be purchased (Eavor, 2024).

To then equip the geothermal community with an agile tool that adapts quickly to state-of-the-art geothermal feasibility research, we introduce GeoCLUSTER v2.0 (https://geocluster.labworks.org/; Figure 1-5): a cloud-native, techno-economic web simulator that enables start-up developers and venture capitalists to explore the economic viability of closed-loop geothermal systems. In its latest release, GeoCLUSTER has been updated to support a parameter space of 31 user editable parameters, has significantly reduced its memory footprint (97% drop), and offers fast, new, on-the-fly computed case studies that can reach in the quadrillions due to the integration of the Slender-Body Theory (SBT) model (Beckers et al., 2015). Over the past few years, multiple studies have been undertaken by our closed-loop geothermal working group to create GeoCLUSTER, including a 1) general feasibility study by White et al. (2023), 2) a database of pre-calculated reservoir simulations of multiple closed-loop designs (Beckers et al., 2023), 3) a heat transfer performance study of closed-loop geothermal systems with thermally conductive enhancements (Beckers et al., 2024), and 4) an impact assessment of convection on performance of closed-loop systems (Hakes et al., 2024). In this paper, we will share GeoCLUSTER's latest updates in greater detail and conclude with anticipated future directions.

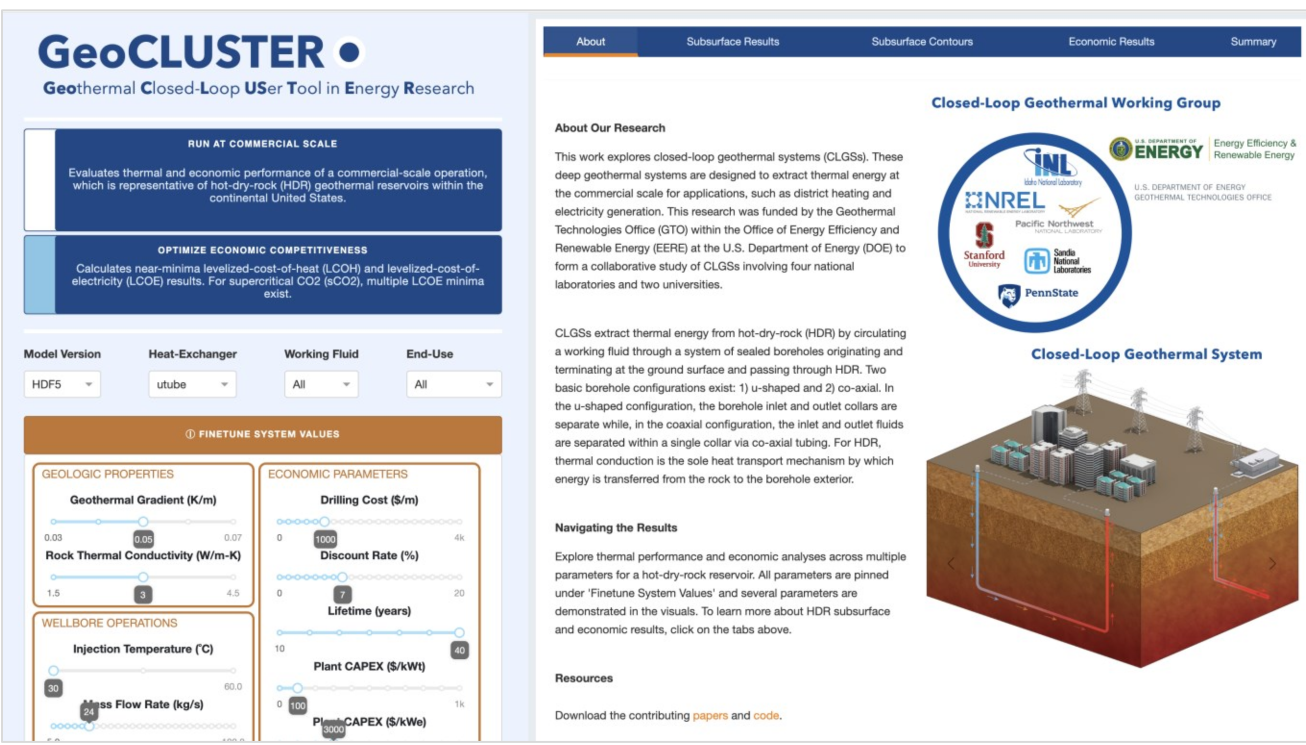

Figure 1: GeoCLUSTER v2.0 on its landing page.

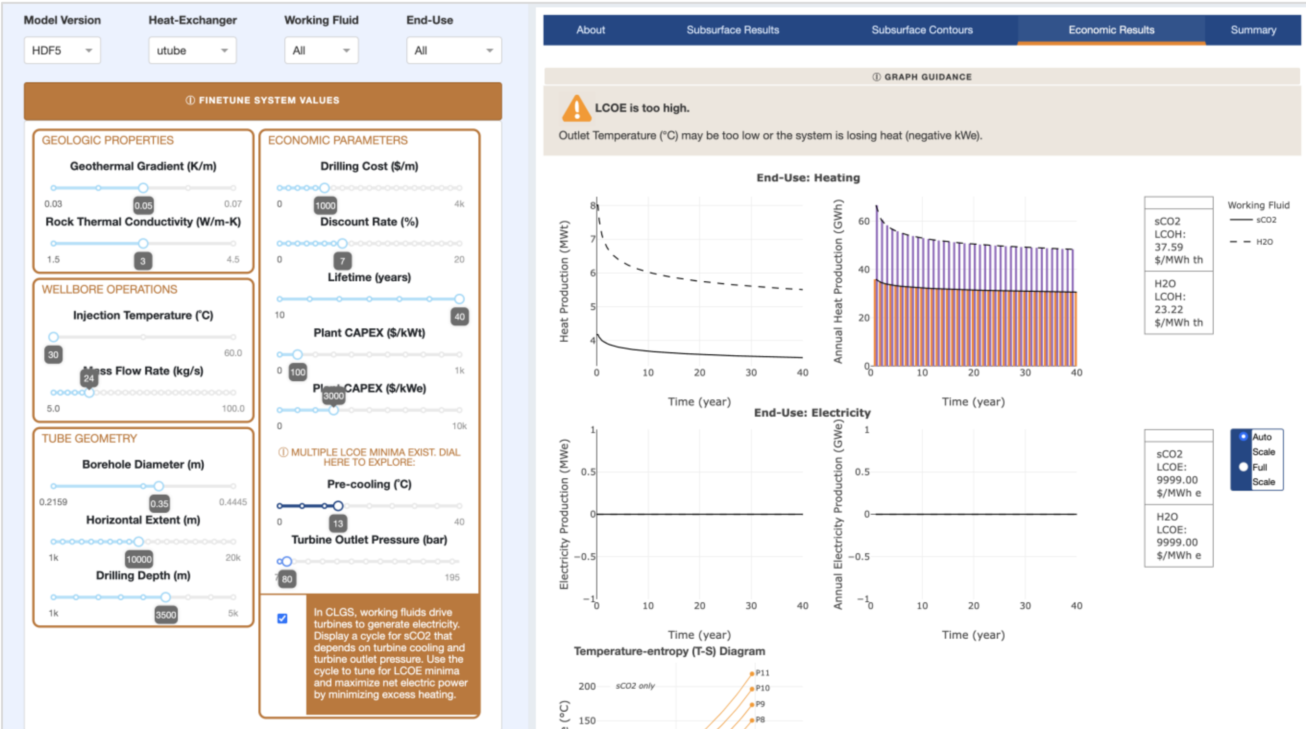

Figure 2: GeoCLUSTER v2.0 displaying the full, pre-calculated, parameter space (i.e., HDF5 model) and its economic results.

Figure 3: GeoCLUSTER v2.0 displaying HDF5 subsurface results.

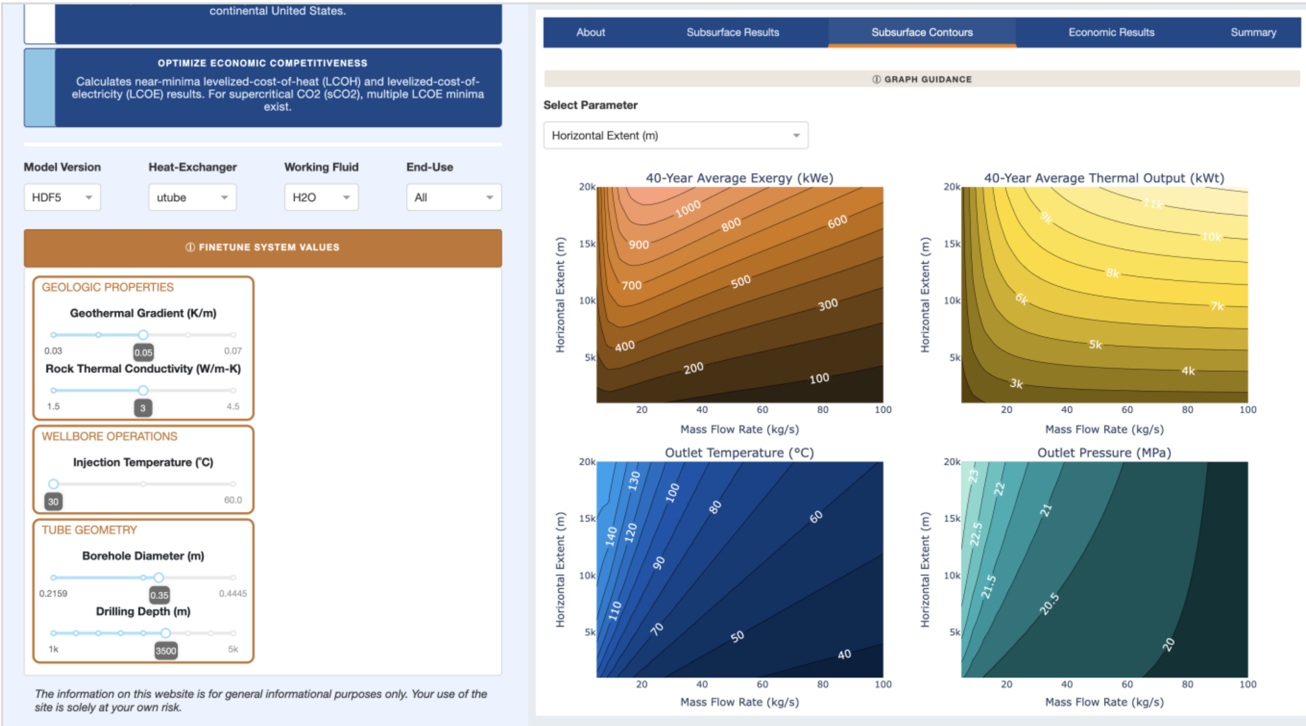

Figure 4: GeoCLUSTER v2.0 displaying HDF5 subsurface contours.

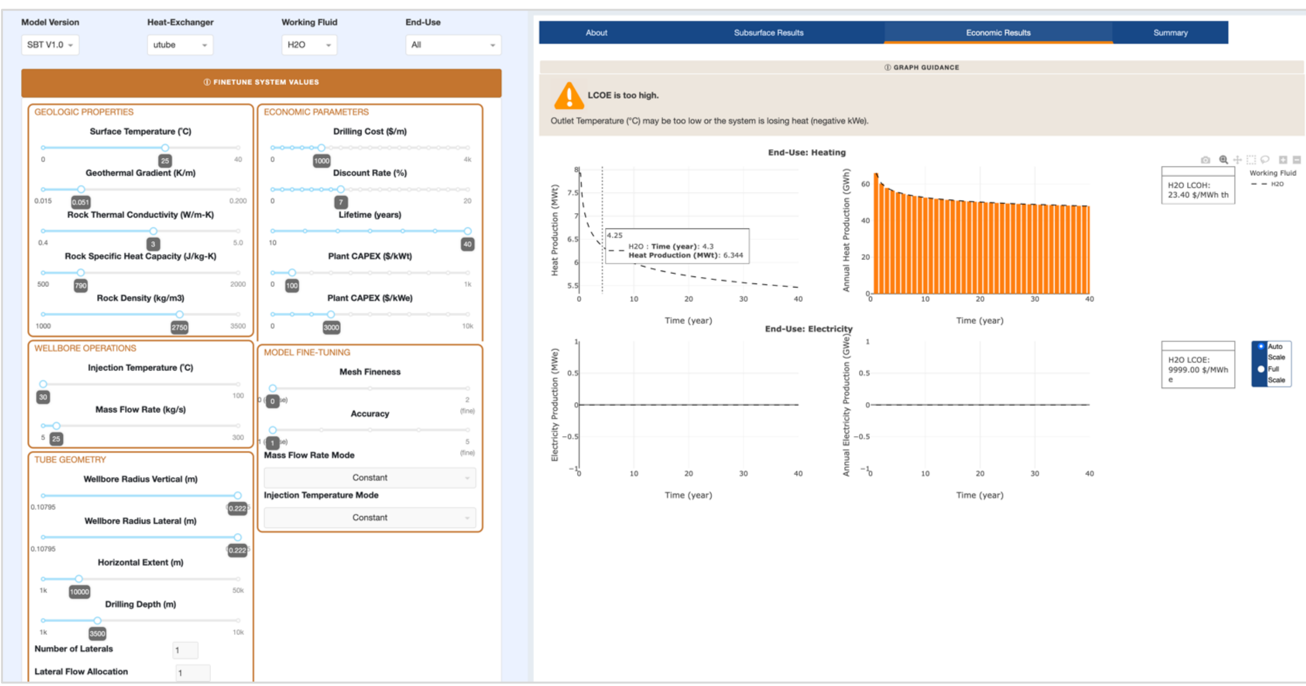

Figure 5: GeoCLUSTER v2.0 displaying the latest, on-the-fly, parameter space (i.e., SBT v1.0 model) and its economic results.

2. UPGRADES

GeoCLUSTER is a user-friendly, closed-loop, techno-economic simulator hosted on Amazon Web Services (AWS) and publicly accessible via the web (URL: https://geocluster.labworks.org/) or on GitHub (codebase: https://github.com/pnnl/GeoCLUSTER). The tool is intended to enable users to rapidly explore numerous techno-economic simulations for closed-loop geothermal systems. GeoCLUSTER v1.0 was originally developed by a team of multiple national laboratories (i.e., Closed-Loop Geothermal Working Group as mentioned previously); likewise, GeoCLUSTER has now been enhanced in part by the Pacific Northwest National Laboratory to deploy a more lightweight and cost-effective 2.0 version. GeoCLUSTER v2.0 also integrates a Slender-Body Theory (SBT) model developed by the National Renewable Energy Laboratory (NREL) and Sandia National Laboratory (Sandia). Between GeoCLUSTER v1.0 and v2.0, major differences and similarities are documented in Table 1 and further described below.

Table 1: Parameter metrics and their ranges between versions of GeoCLUSTER. Rows shaded grey showed no change between v1.0 and v2.0 of GeoCLUSTER. In total, 31 parameters are editable in GeoCLUSTER v2.0.

Parameter	**GeoCLUSTER v1.0**	**GeoCLUSTER v2.0 (SBT Integrated)**
Heat-Exchanger Design	U-Loop, Co-axial	U-Loop, Co-axial
Working Fluid	H_2O, sCO_2	H_2O, sCO_2
End-Use	Heating, Electricity	Heating, Electricity
Mass Flow Rate	5 kg/s to 100 kg/s	5 kg/s to 300 kg/s
Horizontal Extent	1,000 m to 20,000 m	1,000 m to 50,000 m
Drilling Depth	1,000 m to 5,000 m	1,000 m to 10,000 m
Geothermal Gradient	0.03 K/m to 0.07 K/m	0.015 K/m to 0.20 K/m
Borehole Diameter	0.2159 m to 0.4445 m	0. 2159 m to 0.4445 m
Annulus Radius (coaxial)	--	0.1080 m to 0.2223 m
Center Pipe Radius (coaxial)	--	0.06350 m to 0.1740 m
Center Pipe Thickness (coaxial)	--	0.0050 m to 0.025 m
Insulation Thermal Conductivity	--	0.025 W/m-K to 0.50 W/m-K
Wellbore Radius Vertical (utube)	--	0.1080 m to 0.2223 m
Wellbore Radius Lateral (utube)	--	0.1080 m to 0.2223 m
Number of Laterals	1	0 to 30
Mesh Fineness	--	0 (coarse) to 2 (fine)
Accuracy	--	1 (coarse) to 5 (fine)
Mass Flow Rate Mode	--	Constant, Variable
Injection Temperature Mode	--	Constant, Variable
Injection Temperature	30°C to 60°C	30°C to 100°C
Injection Pressure	200 bar (**fixed parameter**)	200 bar (**fixed parameter**)
Rock Thermal Conductivity	1.5 W/m-K to 4.5 W/m-K	0.4 W/m-K to 5 W/m-K
Rock Specific Heat Capacity	790 J/kg-K (**fixed parameter**)	500 J/kg-K to 2,000 J/kg-K
Rock Density	2,750 kg/m^3 (**fixed parameter**)	1000 kg/m^3 to 3,500 kg/m^3
Surface Temperature	25°C (**fixed parameter**)	0°C to 40°C
System lifetime	10 years to 40 years	10 years to 40 years
Drilling Cost	0 $/m to 4,000 $/m	0 $/m to 4,000 $/m
O&M Cost Plant as Percentage of Capital Cost	1.5% (**fixed parameter**)	1.5% (**fixed parameter**)
Discount Rate	0 % to 20 %	0 % to 20 %
Direct Use Heat Plant CAPEX	0 $/kWt to 1,000 $/kWt	0 $/kWt to 1,000 $/kWt
Power Plant CAPEX (for electricity generation)	0 $/kWe to 10,000 $/kWe	0 $/kWe to 10,000 $/kWe
Pre-cooling	0 °C to 40°C	0 °C to 40°C
Turbine Outlet Pressure	75 bar to 200 bar	75 bar to 200 bar
Dead-State Temperature	20°C (**fixed parameter**)	20°C (**fixed parameter**)
Dead-State Pressure	1 bar (**fixed parameter**)	1 bar (**fixed parameter**)
Turbine Isentropic Efficiency (for sCO_2 electricity)	90% (**fixed parameter**)	90% (**fixed parameter**)
Generator Efficiency (for sCO_2 electricity)	98% (**fixed parameter**)	98% (**fixed parameter**)
Compressor Isentropic Efficiency (for sCO_2 electricity)	90% (**fixed parameter**)	90% (**fixed parameter**)
Turbine Outlet Pressure (for sCO_2 electricity)	79 bar (**fixed parameter**)	79 bar (**fixed parameter**)
Pre-cooling Temperature Decline (for sCO_2 electricity)	5°C (**fixed parameter**)	5°C (**fixed parameter**)

Although not comprehensive, it should be noted that compared to published sensitivity parameter studies of closed-loop geothermal systems, GeoCLUSTER leads considerably, even when only comparing a fraction of its parameter space to recent reports (Table 2). Users have wide flexibility to input data based on their specific closed-loop project by tuning various geologic properties, wellbore operations, tube geometry, economic parameters, and model specifications (a total of 31 parameters). This broad coverage ensures results remain relevant for practical feasibility studies in diverse geologic sites, but now GeoCLUSTER v2.0 enables users to assess the performance of closed-loop projects operating in more extreme cases.

Table 2: Sensitivity parameters of closed-loop geothermal systems (CLGSs) by publication. Table was recreated from publications with more than five sensitivity parameters in Table 6 of the literature view conducted by Budiono et al. 2022. Modifications to the table include adding GeoCLUSTER v1.0 and v2.0 as rows and adding new heat transfer and economic parameters added as columns. There are a total of 23 sensitivity parameters considered and compared.

	Heat Transfer Performance Parameters and Sensitivity Variables																	**Economic Parameters**					
Reference	**T**	**Qr**	**P**	**Pth**	**H**	**q**	**t**	**Tgrd**	**Kins**	**Pin**	**Tin**	**d**	**F**	**Lhor**	**Kres**	**Ws**	**HE**	**EU**	**D**	**DR**	**L**	**PCt**	**PCe**
Song et al.	x			x	x	x	x				x												
Sun et al.	x	x	x			x				x	x												
Zhang et al.	x			x		x	x		x				x										
Sun et al.	x		x			x	x		x					x									
Yu et al.	x	x	x		x	x				x													
Sun et al.	x	x	x		x	x				x													
Yuan et al.	x			x			x				x				x	x							
Zhang et al.	x	x				x	x							x	x								
Wang et al.	x	x				x	x					x	x	x									
Sun et al.	x	x	x		x	x		x	x	x			x	x									
GeoCLUSTER v1.0					x	x	x	x	x		x	x	x	x	x		x	x	x	x	x	x	x
GeoCLUSTER v2.0					x	x	x	x	x		x	x	x	x	x		x	x	x	x	x	x	x

Note: T—Temperature distribution; Qr—Heat extraction rate; P—Pressure/Pressure drop; Pth—Thermal power; H—Wellbore depth; q—Flowrate; t—Production time; Tgrd—Reservoir gradient temperature; Kins—Conductivity coefficient of the insulation; Pin—Inlet pressure; Tin—Inlet temperature; d—Wellbore diameter; F—Fluid type; Lhor—Horizontal length; Kres—Conductivity coefficient of the reservoir; Ws—Wellbore spacing (multi-lateral type CLGSs only); HE—Heat-exchanger; EU—End-use; DC—Drilling cost; DR—Discount rate; L—Lifetime; PCt—Plant CAPEX ($/kWt); PCe—Plant CAPEX ($/kWe).

2.1 Backend Optimizations

To ensure that GeoCLUSTER is served to users via web in ways that are lightweight and cost-effective, we upgraded our servers, deployment strategies, and the application's memory performance as summarized in Table 3 and outlined in further detail below.

Table 3: Setup, performance, and parameter metrics between versions of GeoCLUSTER.

Setup	**GeoCLUSTER v1.0**	**GeoCLUSTER v2.0 (SBT Integrated)**
Web Framework	Dash-Plotly	Dash-Plotly
Python Version	3.8	3.11
Memory Performance (Data File)	6500 MB (HDF5 ~90 MB compressed)	200 MB
EC2 Instance (Memory; Cost)	r6i.large (16 GB; $90; Memory Optimized; 2 vCPUs)	t3.small (2 GB; $15; Burstable Performance; 4 CPUs)
Number of Simulations	631,800 simulations per each of the four combinations of heat-exchanger / fluid type, totaling to over 2.5 million simulation runs (if all results were precomputed, it would require approximately 500 petabytes).	Additional simulations in the quadrillions can be generated on the fly. Most notably, users can run case studies with depths deeper than 5 km and geothermal gradients larger than 70°C/km

2.1.1 Cloud Architecture

Between GeoCLUSTER v1.0 and v2.0, there was no change in cloud architecture. Figure 6 shows a minimal setup on Amazon Web Services where the Apache web server, Dash-Plotly web framework, and data files are all on the same Elastic Container Compute (EC2) instance, with traffic distributed by an Elastic Load Balancer (ELB). In this architecture, the web server does not just serve GeoCLUSTER, but it is also involved in processing requests, making it responsible for both the view (the display of the data) and the controller (the processing behind requests). The HDF5 file serves as the model (storage of data).

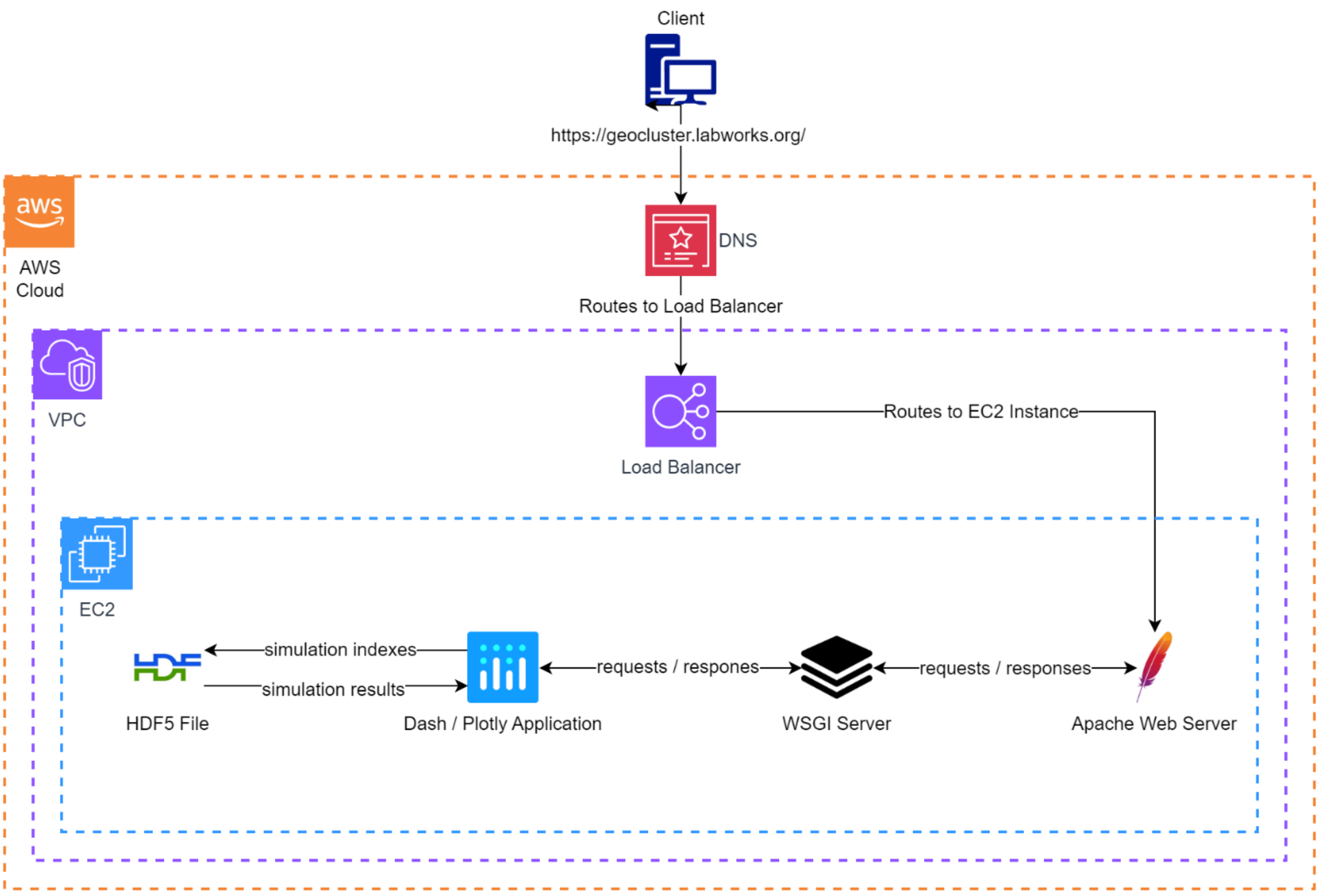

Figure 6: Simple AWS architecture diagram of GeoCLUSTER deployed on AWS Cloud.

2.1.2 Memory Optimizations

The previous version of GeoCLUSTER was light on CPU load but required 6.5 GB of RAM. As such, backend optimizations to reduce the memory needed were required to make GeoCLUSTER more energy- and cost-efficient. Due to project constraints, the application was limited to being hosted on an EC2 instance (described above) and optionally a MySQL database. So, optimization efforts were focused on reducing the number of computational resources the hosting instance needed for the application instead of moving towards a traditional split model-view-controller paradigm.

Some explanation of the working of GeoCLUSTER is a prerequisite to explaining the memory optimizations. GeoCLUSTER contains a subsurface model (i.e., clgs.py) that calculates the output of a CLGS, and an economic model (i.e., clg_tea_module.py) that uses the subsurface results to calculate the economic performance of that CLGS. The subsurface model leverages Beckers et al. (2023) dataset, which contains the output heat, and pressure over time of a CLGS at 631,800 different configurations for each of the 4 combinations of working fluid (H2O or sCO_2) and tube shape (coaxial or utube). This dataset was generated by simulating a CLGS over the Cartesian production of 7 different parameters describing a CLGS. For any combination of these parameters, the dataset contains the output pressure and temperature of the CLGS at that configuration over 161 points in time, representing the output of the CLGS over 40 years. This data is stored as an 8D matrix where the first 7 dimensions are the CLGS configuration parameters, and the 8th dimension is the output over time for either the heat or the pressure. The dataset contains 8 of these 8D matrices: one for the output pressure and one for the output temperature for each of the 4 possible combinations of working fluid and tube shape.

While these 8D matrices take only 90MB of disk space when compressed into an HDF5 file as a result of singular value decomposition compression, they take much more when decompressed and read into memory. The 8D matrices were stored in float64s at runtime, so each 8D matrix required $(631{,}800 * 161)\text{datapoints} * 8 \frac{bytes}{datapoint} \approx 0.81\text{GB}$. Since the application uses 8 of these matrices, GeoCLUSTER requires 8 times this amount, about 6.5GB.

These 8D matrices are used in two ways. The first is to look up the output over time for a given set of input parameters, as shown in Figure 7. This is equivalent to slicing the matrix by holding the first 7 dimensions constant as the input parameters and fetching the output over time.

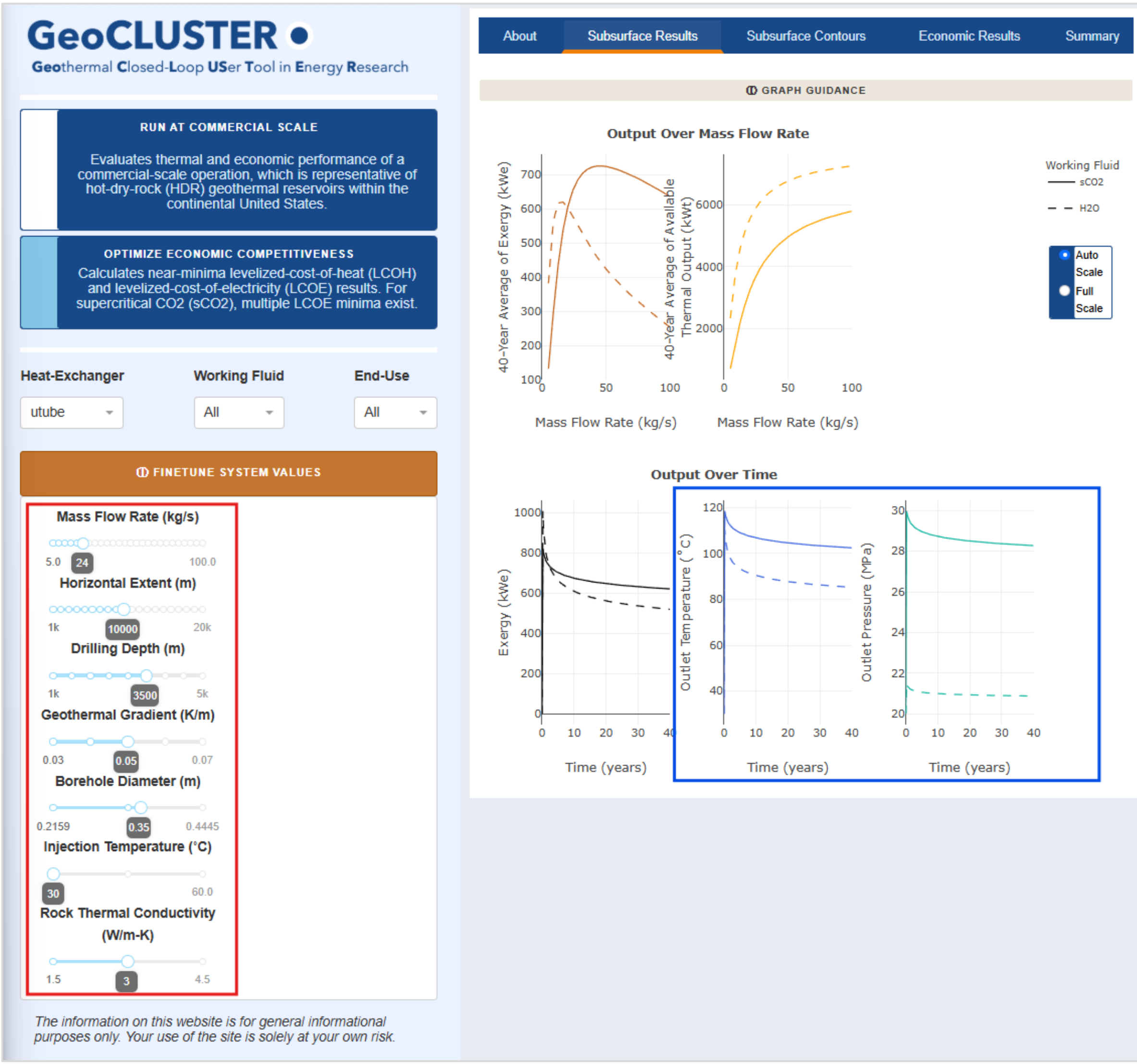

Figure 7. Input parameters in red and output temperature and pressure over time in blue

Secondly, we generate the subsurface contours, which show how the outlet state changes with respect to mass flow rate and a second parameter (any one of the 6 other input parameters), at a CLGS configuration defined by the remaining 5 parameters. This is essentially slicing the matrix by mass flow rate and the second parameter, holding the 5 other input parameters and time constant. In both cases, only a small subset of the data from the entire matrix is needed at a given moment.

Figure 8. Subsurface contours with the second parameter in green, variation of outlet state with respect to mass flow rate and the second parameter in red, and parameters held constant in blue.

However, the access patterns are complicated by the use of interpolation to approximate the outlet states of the CLGS when the input parameters are not equal to their precomputed values. This is because the precomputed dataset computed the results at many different discrete parameter combinations, but the user can input any value on a nearly continuous[1] range for each parameter. For example, in Figure 8, the borehole diameter is 0.35m, which is not one of the precomputed values for borehole diameter. In this case, GeoCLUSTER cannot simply slice into the matrix where the borehole diameter is equal to 0.35m, as that data point does not exist. Instead, the application uses multidimensional linear interpolation to approximate the output value at that point. GeoCLUSTER treats the 8D matrix like an 8D grid, where the location of the point is the parameters of the CLGS configuration, and the value at that point is the outlet state for that configuration. When given a CLGS configuration that isn't precomputed, it falls "between" the points on the grid, and the application uses linear interpolation and the precomputed outputs at points near the given point to approximate the output at the given point.

To remove this large memory burden, a way was needed to offload the storage of these matrices from directly into memory to a database. One considered option was to simply avoid the challenge of interpolation entirely and pre-interpolate the outlet states at every possible value in GeoCLUSTER, save the output to Amazon's Simple Storage Service (S3), and fetch the simulation output at the given parameters. However, due to the curse of dimensionality, this would take an astronomical amount of space. In order to have the same degree of granularity in a pre-interpolated matrix as the granularity offered with interpolation on GeoCLUSTER, the shape would need to change from (26, 20, 9, 5, 3, 3, 3, 161) to (190, 190, 80, 79, 114, 150, 150, 161). This would require $190 * 190 * 80 * 79 * 114 * 150 * 150 * 161 \approx 9.42 * 10^{16}$ datapoints, which would require approximately 377 petabytes if they were stored as float32s. Before even constructing an access pattern for this data, the raw amount of storage needed alone disqualified this approach.

Instead, focus was turned to offloading the data from memory and then only fetching the points the application needed to approximate the output of a given CLGS configuration. In the same way only the 2 sample points around a given point are needed to linearly approximate in 1D, and only the corners of the square around a point are needed to bilinearly approximate in 2D, only the corners of the surrounding 8-cube in the 8D matrix are needed to interpolate a point on an 8D grid. Code was written to compute the indices of the corners of the 8-cube around the point to interpolate for. Then, the outlet states at these points could be fetched from a database and fed into scipy's interpolation function (i.e., internpn), which would perform the linear interpolation.

To ease integration with the pre-existing codebase, the database chosen was an 8D HDF5 dataset in the same format as the 8D NumPy matrix in memory, but the key difference is that this new dataset is chunked. Instead of the data being stored in one contiguous file on disk that's then all read into memory, it's stored as many different individual subsets of the data, chunks, which can be read into memory independent of each other. When the application slices into this chunked matrix, it loads only the chunks containing the slice into memory and reads the points requested in the slice from those chunks. Given our 2 access patterns, a chunk size of (26, 1, 1, 1, 1, 1, 1, 161) was chosen as a balance between reducing the amount of unnecessary data per chunk and reducing the number of chunks needed per lookup. Another difference between the 2 versions is that the new dataset is not stored using singular value decomposition, but instead the 8D

[1] The input for each parameter is not actually continuous but is instead hundreds of very small steps.

matrix is stored as-is to disk, increasing the disk size needed to 3.2GB. This was done to minimize latency in the application, as GeoCLUSTER would need to decompress the data every time it read a chunk from disk if it was stored in a compressed format.

This chunked HDF5 dataset on the disk of the EC2 instance acts as the database for GeoCLUSTER and meets all the demands of the application (Table 4). It supports concurrent reads (meaning that two threads can read from the dataset at the same time when the website is serving multiple users simultaneously) and partial I/O, which allows us to only read a portion of the dataset into memory at a time, instead of the entire dataset. Since the application only reads the dataset and never writes to it, a traditional database is not needed. A prototype version of the application using MySQL as a database instead was tested, but the memory overhead and performance were comparable between the two versions, and the MySQL version unnecessarily increased the complexity of the application with no added benefit.

In all, the output of a CLGS configuration at a point in time is treated as a point on an 8D grid. When that configuration is not precomputed, the point is essentially "between" the grid lines. The application finds the corners of the 8-cube that contains the point, fetches the output at those corners from a database, and then linearly interpolates on them to approximate the output at the given point. With this strategy of offloading the 8D matrices from memory onto disk, the RAM needed for the application was reduced by roughly 97%, from 6.5GB to 200MB.

Table 4: Available output datasets, [hx] = "utube" or coaxial", [fluid] = "H2O" or "sCO2". Note the left, right singular vectors correspond to a rank k approximation. This table has been expanded from Table 3 of Beckers et al. 2023.

HDF5 v1.0 dataset path	Description	HDF5 v2.0 dataset path	Description
/[hx][fluid]/output/Tout/U	Left singular vectors for outlet temperature state	/[hx][fluid]/output/chunked_tout	Chunked temperature outlet state 8D matrix
/[hx][fluid]/output/Tout/sigma	Singular values for outlet temperature state	/[hx][fluid]/output/chunked_tout	
/[hx][fluid]/output/Tout/Vt	Right singular vectors for outlet temperature state	/[hx][fluid]/output/chunked_tout	
/[hx][fluid]/output/Pout/U	Left singular vectors for outlet pressure state	/[hx][fluid]/output/chunked_pout	Chunked pressure outlet state 8D matrix
/[hx][fluid]/output/Pout/sigma	Singular values for outlet pressure state	/[hx][fluid]/output/chunked_pout	
/[hx][fluid]/output/Pout/Vt	Right singular vectors for outlet pressure state	/[hx][fluid]/output/chunked_pout	
/[hx][fluid]/output/We	Available work over forty years, units [GWhr]	/[hx][fluid]/output/We	Available work over forty years, units [GWhr]
/[hx][fluid]/output/Wt	Heat output over forty years, units [GWhr]	/[hx][fluid]/output/Wt	Heat output over forty years, units [GWhr]

2.2 Frontend Profiling

A detailed memory and object data analysis of GeoCLUSTER was made using Python's memory_profiler and guppy3 utility libraries. The analysis focused on two primary areas: NumPy objects used for model calculations, and Plotly graph objects associated with the Dash components rendered on the front-end web page. The analysis revealed that the data class gets instantiated 4 times on app startup accumulating 10MB of memory in the heap each time. Each time new parameters were passed and outputs required interpolation, it involved manipulating NumPy array and matrix objects that caused additional memory spikes ranging from 1-12MB. This was then output through the generation and rendering of new, or existing graphs. Consequently, Dash components were revealed to cause significant memory spikes ranging from 1-15MB when rendering these graphs.

Further heap analysis using the guppy3 utility revealed that objects of the numpy.ndarray type, while few in number (159 instances), accounted for the largest portion of memory. Of the entire heap storage while GeoCLUSTER was running, 40MB was from storing NumPy objects in memory. This was primarily caused by GeoCLUSTER's data class, which is responsible for parsing the data from the HDF5 file (Table 4). String objects and tuples accounted for 26MB and 23MB in the heap memory, respectively.

In exploring solutions to optimize memory usage, Pyodide presented interesting possibilities. Pyodide enables Python code to be run in the web browser by compiling CPython to WebAssembly, effectively allowing computational tasks to be offloaded from the GeoCLUSTER server to the end user's internet browser. This decentralization was thought to bring significant reductions to the server's memory load by utilizing the user's local resources for storing data and running memory-intensive computations but proved to not be necessary.

2.3 Slender-Body Theory (SBT) Model

We integrated the Slender-Body Theory (SBT) model into GeoCLUSTER to allow simulating closed-loop geothermal scenarios and configurations that were not originally included in the pre-calculated database Beckers et al. (2023). For example, with the SBT model, designs can be simulated for depths deeper than 5 km, geothermal gradients larger than 70°C/km and with number of laterals greater than 1, which were originally upper limits considered for the respective parameters when generating the database. The SBT model combines Green's functions with a one-dimensional discretization along a closed-loop heat exchanger for computationally fast simulations of heat transfer with closed-loop systems in conduction-only reservoirs. The model was originally developed by Beckers et al. (2015) for simulating shallow geothermal heat exchanger and was updated by Beckers et al. (2022) for simulating deep co-axial and U-loop closed loop systems. The model requires constant rock properties but can dynamically update the heat transfer fluid properties over time and along the heat exchanger as a function of temperature and pressure. Thermal interference between laterals is captured and variable flow rate and injection temperature can be simulated. Recently, Aljubran et al. (2024) simulated dispatchable geothermal power production with closed-loop systems using the SBT model by varying injection temperature and flow rate on an hourly time step over a 20-year lifetime. For integration into GeoCLUSTER, we converted the SBT model code from MATLAB to Python.

3. CONCLUSION

In approximately six months, GeoCLUSTER v2.0 has integrated new capabilities to provide an agile and cutting-edge closed-loop geothermal simulator for both subsurface performance and economic evaluation. It now supports 31 user-editable parameters, with convective data and new editable parameters in the near future. Notably, the integration of the SBT model provides rapid simulations of deeper and more complex designs that previously would take hours to run a single case study. Additionally, by reducing memory usage by 97%, GeoCLUSTER v2.0 makes high-level geothermal simulations accessible for users with limited computational resources. With its ease of use and speed, this application is well positioned to facilitate early scoping of closed-loop projects, supporting both researchers and industry stakeholders in their efforts to commercialize these systems.

4. ACKNOWLEDGEMENTS

This work was authored in part by the Pacific Northwest National Laboratory, operated by Battelle Memorial Institute, for the U.S. Department of Energy (DOE) under contract GT0100000-05450-1005612 (2022) and the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, under contract DE-AC36-08GO28308. Funding was provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Geothermal Technologies Office. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. 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 U.S. Government purposes.

The authors of this paper would like to acknowledge Alex Buchko for his notable data engineering in significantly reducing GeoCLUSTER's memory. Furthermore, we would like to thank Koenraad Beckers for leading this paper's modeling methods and Anastasia Bernat for curating and programming the interface.

REFERENCES

Alimonti, C., Soldo, E., Bocchetti, D., and Berardi, D.: The Wellbore Heat Exchangers: A Technical Review, Renewable Energy, 123, (2018), 353. https://doi.org/10.1016/j.renene.2018.02.055.

Aljubran, M. J., and Horne, R. N.: Techno-economics of Enhanced Geothermal Systems Across the United States Using Novel Temperature-at-Depth Maps, 49th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2024).

Beckers, K., Vasyliv, Y., Bran-Anleu, G.A., Martinez, M., Augustine, C., and White, M., 2023. Tabulated Database of Closed-Loop Geothermal Systems Performance for Cloud-Based Technical and Economic Modeling of Heat Production and Electricity Generation Proceedings, 48th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2023).

Beckers, K. F., Koch, D. L., and Tester, J. W.: "Slender-body theory for transient heat conduction: theoretical basis, numerical implementation and case studies." Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 471(2184), p. 20150494 (2015).

Beckers, K. F., Rangel-Jurado, N., Chandrasekar, H., Hawkins, A. J., Fulton, P. M., and Tester, J. W., 2022. "Techno-economic performance of closed-loop geothermal systems for heat production and electricity generation." Geothermics 100, p. 102318.

Blankenship, D., Gertler, C., Kamaludeen, M., O'Conner, M., and Porse, S.: Pathways to Commercial Liftoff: Next-Generation Geothermal Power, Department of Energy, (2023). https://liftoff.energy.gov/next-generation-geothermal-power/.

Bláquez,C. S., et al.: GES-CAL: A new computer program for the design of closed-loop geothermal energy systems, Geothermics, 87 (2020), 101852. https://doi.org/10.1016/j.geothermics.2020.101852.

Budiono, A., Suyitno, S., Rosyadi, I., Faishal, A., and Ilyas, A.X.: A Systematic Review of the Design and Heat Transfer Performance of Enhanced Closed-Loop Geothermal Systems, Energies, 15, (2022), 742. https://doi.org/10.3390/en15030742.

Budiono, A., Suyitno, S., Rosyadi, I., Faishal, A., and Ilyas, A.X.: A Systematic Review of the Design and Heat Transfer Performance of Enhanced Closed-Loop Geothermal Systems, Energies, 15, (2022), 742. https://doi.org/10.3390/en15030742.

Computer Modelling Group Ltd.: GEM—Advanced Reservoir Simulation Software. Available online: https://www.cmgl.ca/solutions/software/gem/.

Eavor Technologies Inc.: Eavor Suite 2024 Partner Summit, (2024). Available online: https://www.eavor.com/wp-content/uploads/2024/09/Eavor-Suite-2024-Parnter-Summit-Local.pdf#page=9.00.

Fervo Energy: Fervo Energy Breaks Ground on the World's Largest Next-Gen Geothermal Project, Fervo Energy, (2023), 27 Nov. https://fervoenergy.com/fervo-energy-breaks-ground-on-the-worlds-largest-next-gen-geothermal-project/.

Fowler, M. T.: A Feasibility Study on Three Geothermal Designs, (2021), ResFrac. Available online: https://www.resfrac.com/wp-content/uploads/2024/06/Fowler-2021A-Feasibility-Study-on-Three-Geothermal-Designs.pdf.

Idaho National Laboratory: Impact of Enhanced Geothermal Systems (EGS) on the United States in the 21st Century, MIT Energy Initiative, (2006). https://energy.mit.edu/research/future-geothermal-energy/.

Kliphuis, T., Markowitz, A., Yang, R., and Vesselinov, V.: GeoDAWN To GeoTGo: From Complex Data To Decisions Related To Geothermal Prospectivity, Proceedings, 50th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2025).

Liu, S., Taleghani, A., and Ji, K.: An Advanced Closed-Loop Geothermal System to Substantially Enhance Heat Production, Energy Conversion and Management, 322, (2024), 119168. https://doi.org/10.1016/j.enconman.2024.119168.

Los Alamos National Laboratory: FEHM: Finite Element Heat and Mass Transfer Code. Available online: https://fehm.lanl.gov/.

National Renewable Energy Laboratory (NREL): Introducing GEOPHIRES v2.0: Updated Geothermal Techno-Economic Simulation Tool (2018). Available online: https://www.nrel.gov/docs/fy18osti/70856.pdf.

OpenGeoSys Community: OpenGeoSys—an open-source project for thermo-hydro-mechanical-chemical (THMC) processes in porous media. Available online: https://www.opengeosys.org/.

Pacific Northwest National Laboratory (PNNL): Subsurface Transport Over Multiple Phases (STOMP) Simulator. Available online: https://www.pnnl.gov/projects/stomp.

Pan, L., and Oldenburg, C.M.: T2Well—An Integrated Wellbore–Reservoir Simulator, Computers & Geosciences, 65, (2014), 46–55. https://doi.org/10.1016/j.cageo.2013.06.005.

PFLOTRAN Development Team: PFLOTRAN. Available online: https://pflotran.org/.

Project InnerSpace: Project InnerSpace News and Geomap Asia. Available online: https://projectinnerspace.org/news/geomap-asia.pdf.

U.S. Department of Energy: FORGE: Geothermal Data Repository. Available online: https://gdr.openei.org/forge.

Weers, J., Taverna, N., Podgorny, S., and Huggins, J.: Lessons Learned from AskGDR: Usage and Impact Analysis of the Geothermal Data Repository's AI Research Assistant, Proceedings, 50th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2025).

White, M., et al.: Numerical Investigation of Closed-Loop Geothermal Systems in Deep Geothermal Reservoirs, Geothermics, 116, (2023), 102852. https://doi.org/10.1016/j.geothermics.2023.102852