ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/352519185

Machine Learning to Characterize the State of Stress and its Influence on Geothermal Production

Conference Paper · June 2021

CITATIONS
0

READS
9

4 authors:

Velimir Vesselinov
Los Alamos National Laboratory
225 PUBLICATIONS **1,887** CITATIONS
SEE PROFILE

Luke Frash
Los Alamos National Laboratory
36 PUBLICATIONS **336** CITATIONS
SEE PROFILE

Bulbul Ahmmed
Los Alamos National Laboratory
21 PUBLICATIONS **20** CITATIONS
SEE PROFILE

Maruti Mudunuru
Pacific Northwest National Laboratory
50 PUBLICATIONS **253** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Edge Computing View project

Quantum computing View project

All content following this page was uploaded by Bulbul Ahmmed on 18 June 2021.

The user has requested enhancement of the downloaded file.

Machine Learning to Characterize the State of Stress and its Influence on Geothermal Production

Velimir Vesselinov[1], Luke Frash[2], Bulbul Ahmmed[1], Maruti K. Mudunuru[3]

[1]Earth & Environmental Science Division, Computational Science Group, Los Alamos National Laboratory, Los Alamos, NM 87545

[2]Earth & Environmental Science Division, Earth Systems Observations Group, Los Alamos National Laboratory, Los Alamos, NM 87545

[3]Watershed & Ecosystem Science, Pacific Northwest National Laboratory, Richland, WA 99352

Keywords

unsupervised machine learning, geothermal, geothermal power production, state of stress, and permeability

ABSTRACT

Geologic subsurface conditions, including the state of stress, provide key controls on hydrothermal fluid flow. Characterizing the specific aspects of the geologic structure that facilitate fluid flow in geothermal systems is of particular interest. It is even more critical to understand how the subsurface conditions impact decisions regarding well design for optimized energy production. We have explored the impact of geologic conditions on energy production using the unsupervised machine learning method of non-negative matrix factorization with k-means clustering (NMF*k*). NMFk is a part of our open-source machine learning and artificial intelligence framework called SmartTensors (http://tensors.lanl.gov; https://github.com/SmartTensors). The results of our machine-learning studies identified the key geologic and well-design parameters that have the greatest impact on energy production. We have extracted specific energy production signatures that relate to reservoir properties, state of stress, and well construction.

1. Introduction

Fractures dominate the subsurface flow system and provide the mechanism for heat extraction for Enhanced Geothermal Systems (EGS). Hydraulic fracture and hydraulic shear stimulation are the key methods being used to manipulate the in-situ fracture network by creating or enhancing the subsurface flow pathways that are facilitated by fractures. Predicting when and where fractures will be stimulated is of the utmost importance when designing wells to efficiently and effectively mine heat from the subsurface. However, there are many unknowns that hinder our ability to make reliable predictions for where hydraulically conductive fractures are likely to be, such as the unknown locations of pre-existing natural fractures, heterogeneities, and weak layers. To design optimized EGS systems, we must address this uncertainty head-on by leveraging all the available information that we can feasibly obtain and by applying our fundamental understanding of the behavior of fractures and fluid flow in natural rock systems. Knowledge of

the stress state and trends for the orientations and geometry of natural fractures and faults at a site offers some of the most important information for estimating where fractures are likely to exist or to be created by hydraulic stimulation.

The common approach adopted by the oil and gas industry is to drill lateral wells in the direction parallel to the minimum principal stress. This encourages hydraulic fracture growth perpendicular to the well so that a maximum volume of rock can be stimulated per length of well (i.e., stimulated reservoir volume). However, oil-bearing rock is sedimentary, weak, and does not typically contain significant natural fracturing. In contrast, the crystalline rock of geothermal hot spots can be much stronger, higher stress, and densely naturally fractured. In this case, it becomes increasingly favorable and likely that natural fractures will be stimulated by fluid injection. With shear known to be the dominant mechanism for fracturing in high-stress environments, knowledge of the stress state and natural fracture strengths becomes crucial for predicting the orientations of natural fractures that will be prone to shear slip induced by increasing pore fluid pressure. Owing to geologic data analysis combined with diagnostic fracture injection tests (DFIT), estimates of the magnitude and orientation of in-situ stresses can be feasibly obtained for most sites. With knowledge about the stress state, the next objective is to identify optimal well orientations that can best leverage both hydraulic fractures and natural shear fractures to maximize the heat extraction efficiency and longevity of the geothermal system. However, this is a complex three-dimensional problem where many cases must be considered due to the chronic lack of sufficiently high-fidelity data for accurate *a priori* models.

In this study, we use a recently developed fast physics-based model (GeoDT) (Frash, 2021) to generate a database of possibilities for a modestly characterized geothermal system (UtahFORGE). This data is mined using non-negative matrix factorization with k-means clustering (NMF*k*) to identify the dominant parameters and relationships that link site data and well design to ultimate energy production. GeoDT investigates three-dimensional stochastic fracture networks that are subjected to variable well designs, tensile and shear hydraulic stimulation, and long-term heat extraction to generate electrical power. The GeoDT/NMF*k* analysis reveals the relative importance of each piece of site information (e.g., in-situ stress state) and the consequences/benefits of design decisions (e.g., well spacing and orientation) as they relate to ultimate power production. This offers a new, rapid, means to consider these complex interactions for better optimized new EGS site designs. Crucially, our approach is founded on physical measurements of coupled fracture properties so as to not underestimate the influence of real-world complexity.

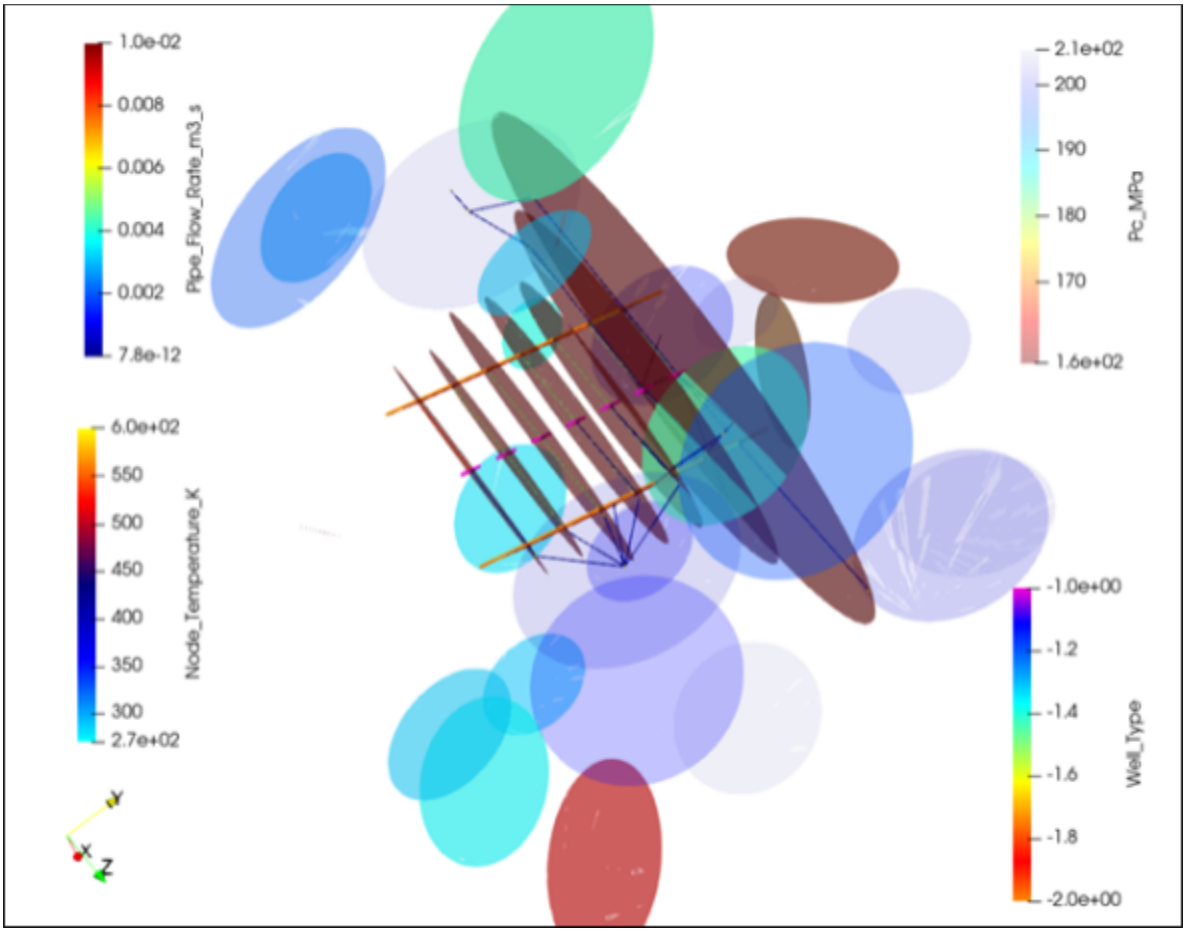

Figure 1. Example stochastically generated fracture and well scenario with injection into one well across seven isolated intervals and production from two bounding wells. The parallel hydraulic fractures propagated from each injection interval are shown in red, the color indicating that these fractures require relatively low pressure for activation (*Pc*). Note that most, but not all, of the scattered natural fractures require significantly higher pressures to activate.

2. Methods

2.1 GeoDT

GeoDT is a multi-physics code for rapid prediction of the performance of geothermal energy systems, dependent on design decisions and site data uncertainty. The GeoDT model, described by Frash, (2021), was used to generate a library of over 2000 geothermal production scenarios based loosely on the UtahFORGE site's parameters. This model includes three-dimensional natural fracture networks, mechanisms for tensile and shear fracture growth, fracture flow with roughness and tortuosity effects, long-term heat extraction, stress-dependent fracture properties, and well design and control parameters. Embedded scaling laws link fracture dimensions to mechanical, seismic, and hydrological parameters (Frash et al., 2021). This coupled approach enables valuation that considers the interplay between general site parameters (e.g., depth and thermal gradient), in-situ stress attributes (e.g., stress anisotropy), rock mechanical attributes (e.g., elastic moduli), natural fracture strength and permeability characteristics (e.g., hydraulic

aperture and friction angle), natural fracture intensity (e.g., number, orientation, and spacing for fractures), fracture complexity (e.g., roughness), and site design decisions (e.g., well spacing and well orientation). As a byproduct of the scaling relations, GeoDT predicts maximum induced seismic magnitudes from the hydraulic stimulation process. The site-specific parameters from the UtahFORGE site used for the model are given in Table 1. Each modeled scenario included stochastically generated natural fractures. An example system is visualized in Figure 1. To solve this system, GeoDT completes the following computational sequence:

1. Natural fracture placement, well placement, and calculation of fracture activation pressure (Pc) based on the far-field stress state, mechanical properties, and orientations.
2. Hydraulic stimulation by simultaneous injection into all of the intervals. This ignores sequencing and staging which sacrifices some accuracy in preference for obtaining a solution faster.
3. Long-term flow calculation with consideration of continued stimulation, far-field leakoff, and 3D connectivity issues through the well and fracture network.
4. Long-term transient heat-extraction and production simulation where heat from the rock transfers to the injected fluid.
5. Electrical power output calculation via the Single-flash Rankine steam cycle. More advanced and higher efficiency cycles are not evaluated at this time, so the estimates will be lower than what is achievable by the best available technologies.

Using the above sequence, this tool simulates the whole geothermal development cycle from the initial well design to the end of production. Since GeoDT includes geomechanical coupling between fracture properties and stress, results allow us to probe the influence of stress data on geothermal production potential at a given site. GeoDT also enables us to investigate links to seismicity and the benefits or consequences of key design attributes such as well spacing, orientation, and diameter.

2.2 NMFk

We applied NMF*k*, which can analyze hidden properties in data that are not visible with open eyes or traditional data analytic. Our group has successfully applied NMF*k* to understand several geothermal systems (Ahmmed et al., 2020a; Ahmmed and Vesselinov, 2021; Ahmmed et al., 2020b,c; Vesselinov et al., 2020a, 2021, 2020b). Also, there are various similar unsupervised ML tools such as principal component analysis (PCA) (Wold et al., 1987), independent component analysis (ICA) (Common, 1994), non-negative matrix factorization with customized *k*-means clustering (NMF*k*) (Alexandrov and Vesselinov, 2014; Vesselinov et al., 2019,2018). PCA, ICA, and NMF*k* learn to represent a data matrix using linear combinations of basis matrices by reducing their dimensions. PCA and ICA are not additive and parts-based techniques; therefore, interpretation of outcomes becomes challenging. Alternatively, NMF*k* is an additive and parts-based technique so it has better interpretability. NMF*k* discovers dominant signals/latent features, finds dominant attributes in signals, and in

some cases clusters spatial similarities in a dataset (Alexandrov and Vesselinov, 2014; Vesselinov et al., 2019,2018). Because of the preceding advantages of NMF*k* and its successful application in geothermal data, we selected this technique for the complex dataset.

NMF*k* combines non-negative matrix factorization (NMF) and customized *k*-means clustering. NMF learns from parts of an object (Lee and Seung, 1999) while *k*-means clustering partitions similar data points together and discovering underlying patterns (Hartigan and Wong, 1979). Customized *k*-means clustering estimates the similarity of an object (derived from *k*-means clustering) to its own cluster compared to other clusters (Alexandrov and Vesselinov, 2014; Vesselinov et al., 2019,2018). In this study, first, a data matrix, X of size (n,m), is built with non-negative elements, where n and m are numbers of locations and attributes/variables, respectively.

NMF of NMF*k* algorithm factorizes X to decompose into two matrices W (mixing matrix) and H (attribute matrix) of size (n,k) and (k,m), respectively as:

$$X \cong W \times H \tag{1}$$

where k is an unknown/hidden number of signals present in the data. The W matrix represents how locations are related to the hidden signatures. The H matrix depicts the relationship between features and the hidden signatures. The number of hidden signatures is unknown that is optimized by performing a series of NMF for $k = 2, 3, \cdots, d$, where $d \leq m$. The optimization process minimizes the following objective function, L, based on Frobenius norm for specified k as:

$$L(k) = ||X - W \times H||_F \tag{2}$$

For each k, NMF is solved for 1,000 random initial guesses for W and H matrices. The least value of L for a given k is assumed as the best value or the reconstruction error, $E(k)$. After completion of the NMF process, the 1,000 initial guesses of H are clustered into k clusters using customized k- means clustering. However, k is also unknown in k-means clustering. The algorithm consecutively examined specified k by obtaining 1,000 H for each attribute. During clustering, the similarity between two variables was assessed using the cosine distance.

Next, the Silhouette value (Rousseeuw, 1987) is calculated and used to estimate a particular choice of k using 1,000 H matrices. The Silhouette value quantifies how similar an object is to its own cluster compared to other clusters and varies from −1 to +1; high values indicate that the object is well matched to its own cluster and poorly matched to neighboring clusters. The combination of $L(k)$ and the Silhouette value were used to determine the number of hidden signals. The best estimate for k is a number that optimizes a low $L(k)$ and the high Silhouette value.

3. Results and Discussion

Results from GeoDT (Figure 1) predict the time series of geothermal power production for each of over 2000 subsurface scenarios. For each scenario, GeoDT estimates when thermal

breakthrough (i.e., produced fluid cooling) begins, and the thermal and electrical power outputs over time. Initial inspection of the results from our GeoDT analysis shows an apparent link between the well spacing and the electrical power output of the system after 20 years of production. There also appears to be a strong link between the number of injection intervals and power output. However, the dataset consists of >90 input attributes and >16 result outputs for 2000 realizations, and these links are only a small portion of what can be identified using ML methods developed by our team.

Applying NMF*k* reveals four constitutive multi-attribute input signatures that control the time series of the produced fluid enthalpy (i.e., geothermal fluid energy) and the related electrical power potential. The structure of these signatures is shown in Figure 3. The temporal contribution of four signatures varied for production enthalpy (i.e., thermal power output) and electrical power output. The contribution of Signature A of thermal power output gradually increases until 10 years and flattens out after 10 years. The contribution of Signatures B and C gradually decreases from the very beginning in a similar fashion. The contribution of Signature D also decreases; however, it sharply decreased during the first 2 years.

The contribution of Signature A of electrical power output gradually increases until ~2.5 years and decreases thereafter. The contribution of Signatures B and C is low compared with the enthalpy output signatures; however, they decrease initially; interestingly, they increase after 10 years indicating the dominance of this signal after a significant portion of the operation cycle. The contribution of Signature D is higher during the early stage; however, it precipitously decreases during the first ~1.5 years and continues to decreases till the end of the production cycle.

The contribution of each signature indicates that some attributes have better control over enthalpy output and others have better control over electrical power output. Also, the temporal contribution of each signature demonstrates that some attributes have more/less control during the early stage but their contribution may decrease/increase over time. In the next paragraph, we will provide a high-level interpretation of how the attributes are related to each signature and their temporal contribution.

Based on our previous experience, the weight (contribution) of an attribute range from 0 to 1 and any attribute is dominant if its weight value is >0.5. The weight values of critical parameters are listed in Figure 4. *Well spacing* is the one and only dominant attribute of Signature A. However, Signature B captures several attributes to be dominant and those attributes are *geothermal gradient*, *rock thermal conductivity*, *rock volumetric specific heat capacity*, *rock elastic modulus*, *rock Poisson's ratio*, *rock shear modulus*, *minimum lateral pressure coefficient*, *intermediate pressure coefficient*, *fracture count set 1*, *fracture count set 2*, *fracture count set 3*, *well spacing*, *well length*, *well azimuth*, *borehole/casing radius*, *cohesion*, and *friction angle*. The contribution of these dominant attributes to Signature B varies and *borehole/casing radius* is the most critical among all these attributes. *Reservoir depth*, *reservoir pore pressure*, *reservoir temperature*, *overburden stress*, and *intermediate stress* are dominant attributes of Signature C. Among these

five attributes, *reservoir depth*, *reservoir temperature*, and *overburden stress* are the most dominant (i.e., they got perfect weight value). *Well dip* is the dominant attribute of Signature D.

We can use these dominant attributes to categorize the signatures into (1) well spacing, (2) stress, geothermal, and fracture attributes, (3) system (i.e., site conditions) attributes, and (4) well dip (i.e., orientation). These attributes and temporal variation of signatures provide a better understanding of the temporal control of attributes on both thermal and electrical power outputs. For example, *well spacing* has a steady control over thermal power production; alternatively, it has higher control over electrical power output during the early stage but its control slowly dissipates over time. Effect of stress, geothermal, fracture attributes, and system attributes (Signature B and C) slowly decreases over time on thermal production output; however, they gain control at the later stage of electrical power output. In other words, increased in situ stress causes decreased production over time. Here it is important to note that increased stress will cause fracture closure after stimulation, which will likely reduce production, but this stress increase will also provide for more shear stress. Shear stress is a prerequisite for shear stimulation of fractures to increase reservoir performance, but it is also a driver for induced seismicity. *Well dip* (Signature D) has a high effect during the early stage, but its effect gradually decreases.

Also, NMF*k* allows investigation of the effects of the input attributes on other outputs such as maximum induced seismic magnitude, far-field leakoff, and the number of fractures that interlink the injection and production wells (Figure 5). Interestingly, there appears to be a link between the well-spacing dominated signature and the maximum induced seismic event magnitude. It is not yet clear what underlying mechanism drives this connection. Less surprisingly, the system attributes (e.g., natural fractures, well length, well diameter, and rock properties) have a strong influence on the amount of fluid loss (i.e., boundary outflow rate) from the system. Stress effects on the GeoDT results are clearly evident, but a clear causal pattern is not immediately apparent. Instead, stress appears to associate with mixed effects, some positive and others negative. Another surprising result is the importance of well dip and azimuth (i.e., well orientation). The cause of this importance is suspected to be linked to the natural fracture orientations, especially Joint Set 3, which is northeast striking and southeast dipping, making it a prime target for shear slip. At the UtahFORGE site, the nearby Opal Mound Fault is also northeast striking and southeast dipping. The planned well orientation at UtahFORGE is nearly perpendicular (i.e., face on) to this fault.

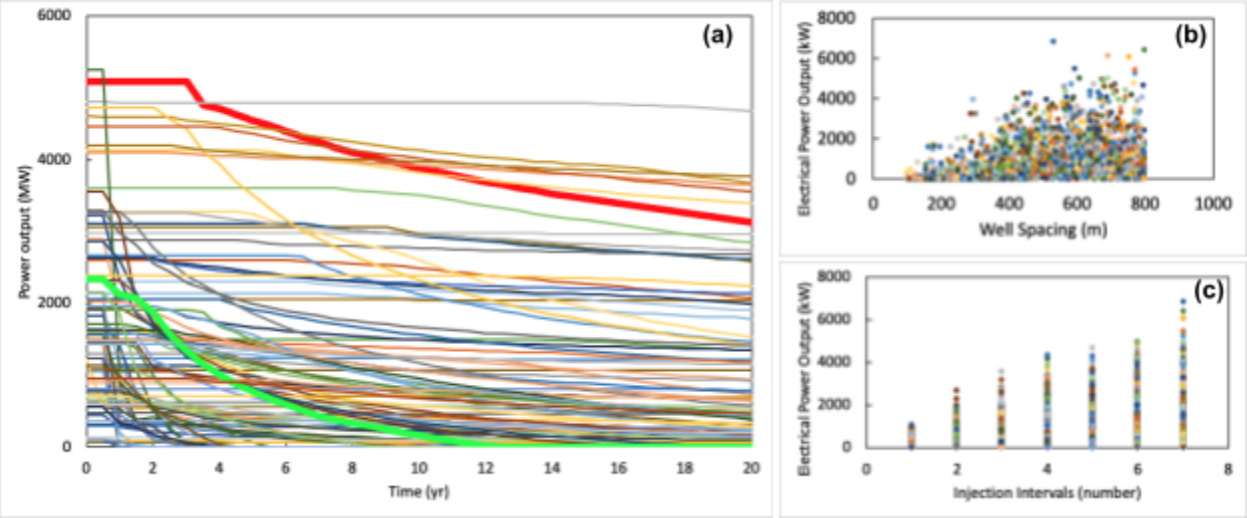

Figure 2: Compiled results from more than 2000 geothermal power production simulations based on the parameters described in Table 1. In the time series plot (a), a high-performing case is highlighted in red, and a poor performer is highlighted in green. Also, the well spacing (b) and injection intervals (c) demonstrate a clear correlation to electrical power outputs.

Table 1: Partial list of properties for GeoDT modeling scenarios.

Attribute	Minimum	Mean	Maximum
Domain size (m)	3000	3000	3000
Depth (m)	4000	6000	8000
Geothermal gradient (C/km)	40	45	50
***Well orientation & dip (deg AzN & deg down)**	89 & -15	104 & 0	119 & 15
Well length & spacing (m & m)	1500 & 100	1500 & 450	1500 & 800
Well diameter (m)	0.15	0.28	0.41
Number of injection zones (#)	1	5	8
***Joint set 1 strike & dip (deg AzN & deg down)**	-15 & -7	96 & 80	+15 & +7
***Joint set 2 strike & dip (deg AzN & deg down)**	-15 & -7	185 & 48	+15 & +7
***Joint set 3 strike & dip (deg AzN & deg down)**	-15 & -7	35 & 64	+15 & +7
Minimum horizontal stress anisotropy (%)	30%	60%	90%
Intermediate horizontal stress anisotropy (%)	60%	90%	120%
Rock density (kg/m3)	2700	2700	2700
Thermal conductivity (W/m-K)	2.1	2.45	2.8
Specific heat capacity (kJ/m3-K)	1900	2050	2200
Rock Young's modulus & Poisson's ratio (GPa & ratio)	30 & 0.15	60 & 0.25	90 & 0.35
Production lifespan (yr)	20	20	20
Injection temperature (C)	95	95	95
Production pressure drawdown (MPa)	2.0	2.0	2.0
Stimulation rate & volume (m3/s & m3)	0.08 & 100000	0.08 & 100000	0.08 & 100000
Long term injection rate per zone (m3/s)	0.01	0.01	0.01
Fracture frictional angle & cohesion (deg & MPa)	25 & 5	35 & 10	45 & 15
Fracture roughness factor (no unit)	0.7	0.85	1.0
Gutenberg-Richter b-value (no unit)	1.0	1.0	1.0

*data based on UtahFORGE

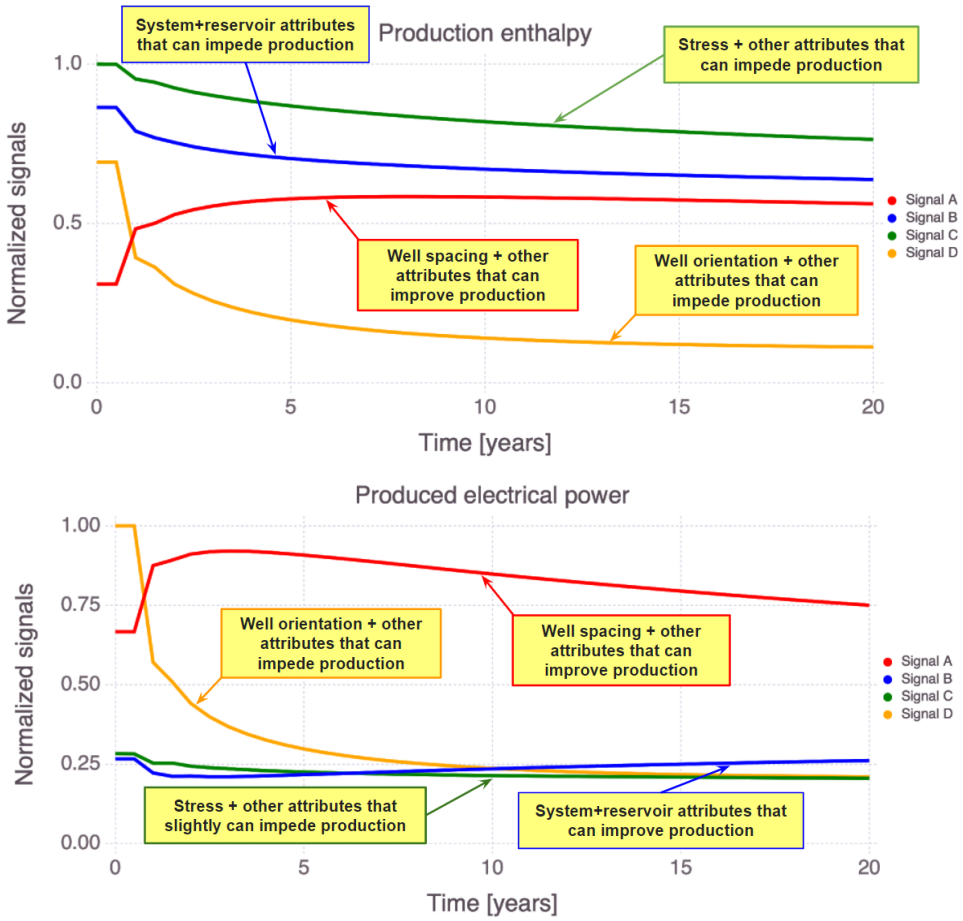

Figure 3. ML identifies the signature structure of the enthalpy and power production time series predicted by GeoDT. The primary physical components of each mixed signature are provided to aid interpretation. Only one of the signatures (red) shows inputs that associate with increased production over time.

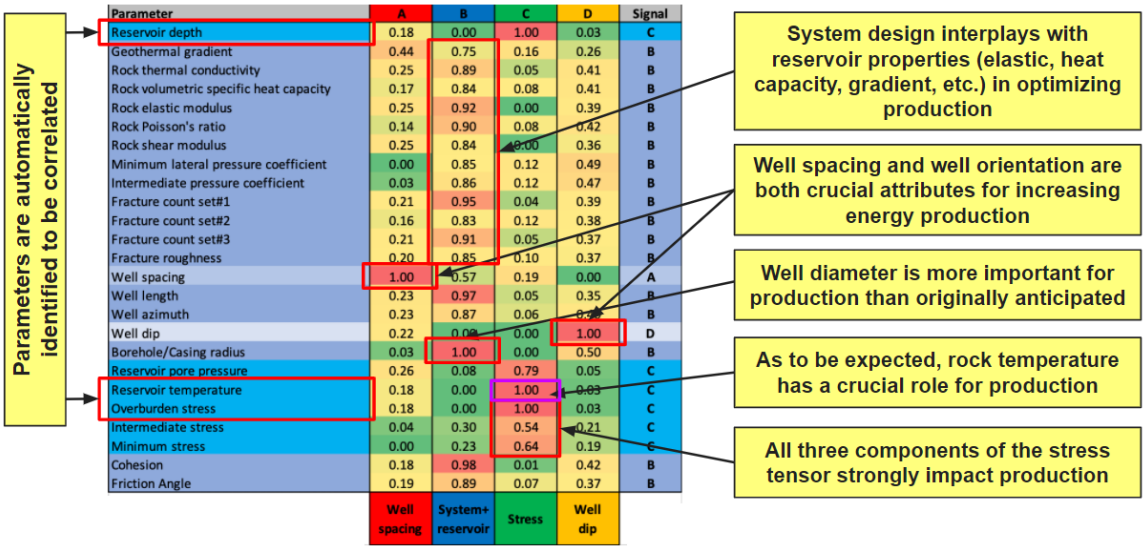

Figure 4. Combined inputs of the four ML identified signatures that control geothermal power production. Callouts are included to highlight the primary physical components of each signature. We categorize each signature by its most dominant component. Red colors indicate parameters with high importance with that particular signature, green colors show that a parameter has a low weight with that specific signature.

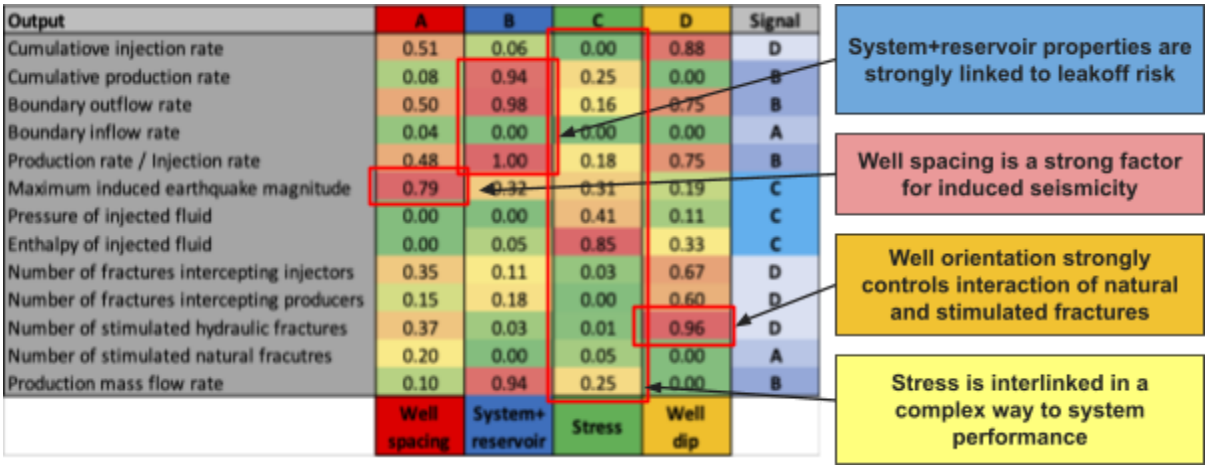

Figure 5. ML predicted the effects of the identified multi-attribute signatures on various outputs from GeoDT. Callouts are included to highlight significant links that aid understanding of subsurface geothermal processes coupled with system design. Red colors indicate outputs with high importance with that particular signature, green colors show that a specific output has a low weight with that specific signature.

3. Conclusions

ML analyses of GeoDT simulations focused on the influence of stress state, geothermal reservoir properties, and natural fractures on geothermal well drilling and well production. Using GeoDT, we simulate the time series (up to 20 years) of geothermal power production for each of over

2000 subsurface scenarios. For each scenario, GeoDT estimates when thermal breakthrough (i.e., produced fluid cooling) begins, and the thermal and electrical power outputs over time. The dataset consists of >90 input attributes and >16 result outputs. Later, we apply an unsupervised machine learning method called NMFk to the dataset to find signatures characterizing effects of stress state, geothermal reservoir properties, and natural fractures on well drilling and on thermal power output and electrical power output.

NMFk found four signatures called Signatures A, B, C, and D. The temporal contribution of four signatures varied for production enthalpy (i.e., thermal power output) and electrical power output. We also found dominant attributes of each signature. *Well spacing* is the dominant attribute of Signature A. The dominant attributes of Signature B are *Geothermal gradient*, *rock thermal conductivity*, *rock volumetric specific heat capacity*, *rock elastic modulus*, *rock Poisson's ratio*, *rock shear modulus*, *minimum lateral pressure coefficient*, *intermediate pressure coefficient*, *fracture count set 1*, *fracture count set 2*, *fracture count set 3*, *well spacing*, *well length*, *well azimuth*, *borehole/casing radius*, *cohesion*, and *friction angle*. The dominant attributes of Signature C are *reservoir depth*, *reservoir pore pressure*, *reservoir temperature*, *overburden stress*, and *intermediate stress*. The dominant attribute of Signature D is *Well dip*.

We also discovered the dominance of attributes over thermal and electrical power outputs. For example, *well spacing* has a steady control over thermal power production. Effect of stress, geothermal, fracture attributes, and system attributes (Signature B and C) are dominant in the later stage of electrical power output. In other words, increased in situ stress causes decreased thermal power production over time. *Well dip* has a high effect during the early stage for both kinds of power outputs, but its effect gradually decreases.

We also found a correlation between the well-spacing dominated signature and the maximum induced seismic event magnitude. Moreover, the system attributes (e.g., natural fractures, well length, well diameter, and rock properties) have a strong influence on the amount of fluid loss (i.e., boundary outflow rate) from the system. Furthermore, NMFk elucidated that stress appears to associate with mixed effects, some positive and others negative. Finally, we demonstrated that the importance of well dip and azimuth (i.e., well orientation) is suspected to be linked to the natural fracture orientations.

ML analyses identified well spacing and well orientation as critical parameters impacting energy production and induced seismicity. Our results also support the idea that “fracture caging” and “well caging” can limit induced seismic event magnitudes. “Caging“ aims optimized the drilling of injection and production wells so that they can contain the circulated fluids within a portion of the reservoir where fracture-dominated flow occurs.

Also, we got plenty of results and additional work is needed to parse out the meaning of these signatures and implications for site-specific geothermal energy production. NMF*k* coupled with GeoDT provides a good platform for this future work, owing to its ability to rapidly model the effect of complex interactions and design decisions on production for an extensive range of site conditions. Note that the presented results are preliminary, and the GeoDT model was only just completed in 2021. More validation of GeoDT is needed to gain confidence in these model predictions and their importance to guide field exploration and drilling decision-making. Further investigation of the identified signatures is required to more clearly understand the links and implications. This additional work is our future plan.

Acknowledgment

This research is based upon work supported by the U.S. Department of Energy's (DOE) Office of Energy Efficiency and Renewable Energy (EERE) under the Geothermal Technology Office (GTO) Machine Learning (ML) for Geothermal Energy funding opportunity, Award Number DE-EE-3.1.8.1. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy (Contract No. 89233218CNA000001). Additional information regarding the datasets and codes can be obtained from Velimir V. Vesselinov (Monty) (vvv@lanl.gov) and Bulbul Ahmmed (ahmmedb@lanl.gov).

This paper was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

REFERENCES

Ahmmed, B., Lautze, N., Vesselinov, V., Dores, D., and Mudunuru, M. (2020a). Unsupervised machine learning to extract dominant geothermal attributes in Hawaii Island Play Fairway data. Geothermal Resources Council, Reno, NV, October 18–23.

Ahmmed, B. and Vesselinov, V. (2021). Prospectivity Analyses of the Utah FORGE Site using Unsupervised Machine Learning. In *Geothermal Rising, San Diego, CA*.

Ahmmed, B., Vesselinov, V., and M.K., M. (2020b). Machine learning to characterize regional geothermal reservoirs in the western USA. Fall Conference, Geological Society of America, Abstract T185-358249, October 26–29.

Ahmmed, B., Vesselinov, V., and M.K., M. (2020c). Non-negative matrix factorization to discover dominant attributes in Utah FORGE Data. Geothermal Resources Council, Reno,

View publication stats

NV, October 18–23.

Alexandrov, B. and Vesselinov, V. (2014). Blind source separation for groundwater pressure analysis based on nonnegative matrix factorization. *Water Resources Research*, 50(9):7332–7347.

Comon, P. (1994). Independent component analysis, a new concept? *Signal processing*, 36(3):287– 314.

Frash, L. (2021). Geothermal Design Tool (GeoDT). In *Proceedings of the 46th Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, February 15-17*.

Frash, L., Welch, N., Meng, M., Li, W., and Carey, J. (2021). A Scaling Relationship for Frac- ture Permeability after Slip. In *Proceedings of the 55th US Rock Mechanics/Geomechanics Symposium, Houston, TX, June*.

Hartigan, J. and Wong, M. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1):100–108.

Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65.

Vesselinov, V., Ahmmed, B., and M.K., M. (2020a). Unsupervised machine learning to discover at- tributes that characterize low, moderate, and high-temperature geothermal resources. Geothermal Resources Council, Reno, NV, October 18–23.

Vesselinov, V., Ahmmed, B., Mudunuru, M., Karra, S., and Middleton, R. (2021). Hidden geothermal signatures of southwest New Mexico. In *Proceedings of the World Geothermal Congress, Reykjavik, Iceland*.

Vesselinov, V., Mudunuru, M., Ahmmed, B., S., K., and R.S, M. (2020b). Discovering signatures of hidden geothermal resources based on unsupervised learning. *45th Annual Stanford Geothermal Workshop*.

Vesselinov, V., Mudunuru, M., Karra, S., O'Malley, D., and Alexandrov, B. (2019). Unsupervised machine learning based on non-negative tensor factorization for analyzing reactive mixing. *Journal of Computational Physics*, 395:85 – 104.

Vesselinov, V. V., Alexandrov, B. S., and O'Malley, D. (2018). Contaminant source identification using semi-supervised machine learning. *Journal of contaminant hydrology*, 212:134–142.

Wold, S., Esbensen, K., and Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52.