Machine Learning for Geologically Consistent Flow Analysis in Fractured Geothermal Reservoirs

A Case Study

Kamel Targhi, E.; Bruna, P.; Daniilidis, A.; Rongier, G.; Geiger, S.

DOI
10.3997/2214-4609.2025101257

Publication date
2025

Document Version
Final published version

Citation (APA)
Kamel Targhi, E., Bruna, P., Daniilidis, A., Rongier, G., & Geiger, S. (2025). *Machine Learning for Geologically Consistent Flow Analysis in Fractured Geothermal Reservoirs: A Case Study*. Paper presented at 86th EAGE Annual Conference & Exhibition 2025 , Toulouse, France. https://doi.org/10.3997/2214-4609.2025101257

Machine Learning for Geologically Consistent Flow Analysis in Fractured Geothermal Reservoirs: A Case Study

E. Kamel Targhi[1], P. Bruna[1], A. Daniilidis[1], G. Rongier[1], S. Geiger[1]

[1] Delft University Of Technology (tu Delft)

Summary

Characterising fractures in geothermal reservoirs is crucial for understanding heat and fluid flow, as fractures control reservoir permeability. Due to data scarcity, estimating fracture network properties remains uncertain. Dynamic data, such as well tests, provides indirect insights into subsurface properties and workflows have been developed to illustrate how uncertainty in fracture data affects flow behaviour. However, they use simplified, randomly generated fracture geometries limiting their applicability to real-world scenarios.

This study presents a machine learning workflow for characterizing fractured reservoirs using transient data, focusing on geothermal reservoirs. A comprehensive dataset of 5000 geologically consistent Discrete Fracture Networks (DFNs) was generated using GeoDFN and directly linked to MRST for simulations. The workflow then applies a k-medoids clustering approach, using dynamic time warping (DTW) as a distance metric, to cluster pressure responses with similar transient behaviour. We identified 18 distinct pressure behaviour. Linking clusters to fracture properties reveals that fracture intensity, aperture, and length have the most significant impact on pressure behaviour, while fracture set type was found to be the least important factor. Future work will extend this workflow to temperature transient data and apply advanced machine learning techniques for both forward and inverse modelling of fractured geothermal reservoirs.

Machine Learning for Geologically Consistent Flow Analysis in Fractured Geothermal Reservoirs: A Case Study

A thorough understanding of heat and fluid flow in geothermal reservoirs requires capturing subsurface geological. Fracture characterization is fundamental for assessing geoenergy reservoirs, as natural fractures can enhance or even dominate reservoir permeability. However, directly measuring key fracture properties (e.g., intensity, aperture, and orientation) is challenging and the scarcity of data introduces significant uncertainty in estimating connectivity and permeability of the fracture network. This uncertainty extends to heat and flow simulations, requiring significant computational effort to quantify its impact on the results.

Dynamic data, such as well test data, have been extensively used for characterizing reservoirs by providing indirect insights into subsurface properties. However, traditional models for interpreting dynamic data in fractured reservoirs, such as Warren and Root (1963), rely on oversimplified assumptions, including fully connected fracture network and uniform fracture properties through the reservoir. This leads to inaccuracies when compared to real-world observations. Workflows have been developed to demonstrate how uncertainty in fracture data impacts flow predictions in geothermal reservoirs (Lepillier et al., 2019, 2020). Recently, Freites et al. (2023) developed a machine learning-based workflow to classify well-test pressure responses in naturally fractured reservoirs and link them to fracture network properties. While these approaches aim to improve reservoir characterization by linking dynamic responses to fracture network properties, they rely on simplified and randomly generated fracture geometries. This constrains their ability to capture the heterogeneous spatial organization of natural fracture systems, and limits their generalization to complex real-world systems.

This study offers a machine learning-aided workflow to classify a wide range of dynamic responses generated from a comprehensive dataset of geologically consistent Discrete Fracture Networks (DFNs). While the workflow is demonstrated using pressure transient data, it can also be applied to thermal transient data in fractured reservoirs, enabling systematic characterization of diverse dynamic responses.

A machine learning workflow for characterizing fractured reservoirs using transient data

The DFN dataset is generated using GeoDFN, which combines mechanical and statistical methods to produce geologically plausible 2D fracture networks at minimal computational cost (Kamel Targhi et al., 2024). A 1 km × 1 km domain was selected to ensure adequate spatial representation and to allow the pressure front sufficient distance to propagate. A Latin Hypercube Sampling (LHS) approach was used to sample 1000 parameter combinations, including fracture intensity, length, orientation, and aperture. This sample size was selected by analyzing the stability of the mean and variance of key parameters across varying sample sizes, to ensure sufficient coverage of the parameter space. For each parameter set, five realizations were generated, resulting in a total of 5000 DFNs, ensuring a comprehensive dataset for subsequent analysis.

Flow and heat transport simulations are conducted using Embedded Discrete Fracture Model (EDFM) in MRST that explicitly represents individual fractures within the network and automatically accounts for interactions between fractures and the surrounding rock matrix (Lie et al., 2012; Lie, 2019; Wong et al., 2021). A single production well with constant rate of 10 m^3/day is placed at the center of the model to simulate drawdown. To ensure that only the transient pressure response is captured, the simulation is terminated once the pressure front has reached the boundaries.

To systematically analyze the variability in the pressure responses and relate distinct flow behaviors to geological characteristics, we apply an k-medoids clustering approach. This clustering process allows us to group pressure responses with similar transient behavior (Liao, 2005). Since the duration and shape of pressure responses vary across different fracture networks, dynamic time warping (DTW) was employed as a distance metric to account for temporal misalignment in the curves and identify their similarity (Sakoe and Chiba, 1978).

We then analyze the fracture networks associated with each cluster to understand how fracture properties influence flow behavior. By comparing the distribution of fracture properties across clusters, we identify which fracture characteristics exhibit the greatest variability that implies they have a more significant impact on determining the cluster group of a pressure response curve. To further quantify their influence on flow response, we train a random forest classifier using the fracture properties as inputs and the cluster labels as outputs, with feature importance highlighting the key fracture properties in predicting dynamic flow behavior (Breiman, 2001).

Figure 1 presents a summary of the workflow used in this research. This workflow can be easily adapted to analyze temperature transient data by replacing drawdown simulations with thermal simulations (Collignon et al., 2021). An example with the injection well located on the west boundary is depicted in Figure 2.

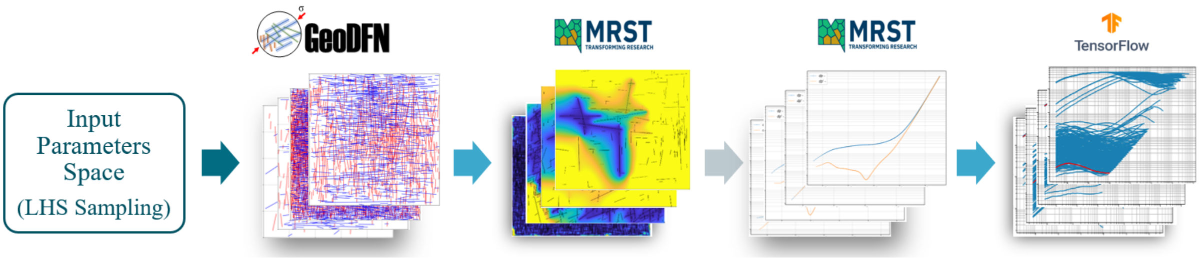
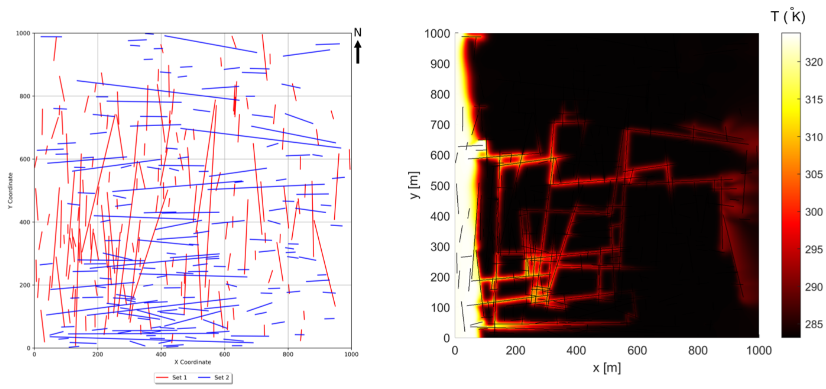


Figure 1 *Illustration of the machine learning workflow for characterizing fractured reservoirs using pressure transient data.*



(a) (b)

Figure 2 *DFN consisting of two orthogonal sets (a) and simulated temperature distribution(b). The west boundary is the injection boundary with $T_{inj} = 323.15$ K and $P_{inj} = 200$ bar, the east boundary is the production boundary with $T_{res} = 283.15$ K and $P_{res} = 100$ bar.*

Results - Clustering of pressure transient data and correlation with fracture properties

We identified 18 clusters, which we determined by analyzing the objective function and selecting the point where a further increase in number of clusters resulted in negligible reduction in the value of the objective function. The distinct patterns in mediods across different clusters indicate that the k-medoids has effectively captured the different flow regimes (Figure 3). The clustering yields acceptable performance, with an average silhouette index of 0.276, indicating moderate cluster separation (Rousseeuw, 1987). Only 11.3% of the samples were misclassified.

We then correlated the distinct pressure responses identified through clustering with the properties of the fracture networks. First, we visually analyzed how fracture properties vary across clusters and observed that fracture intensity, fracture aperture, and fracture length exhibited significant variability across clusters, suggesting their influence on the pressure response (Figure 4(a)). Then, we employed a Random Forest classification algorithm to further quantify the significance of these fracture properties (Figure

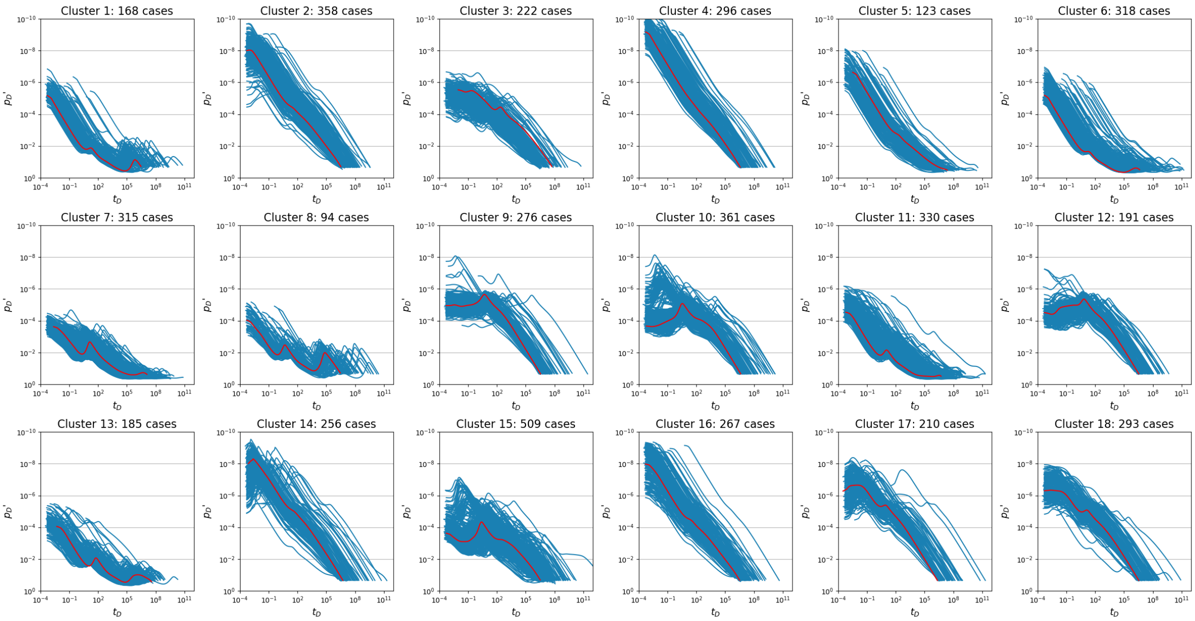
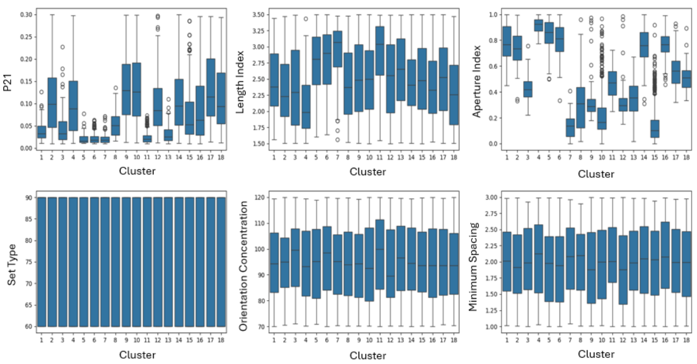
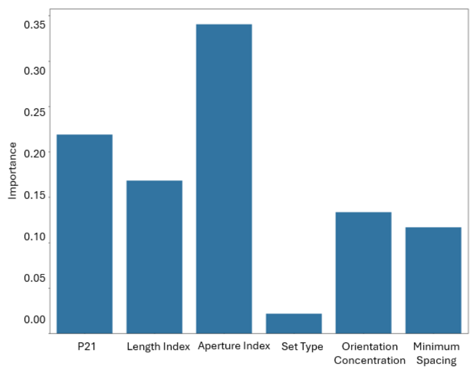


Figure 3 *Visualization of the first derivative of the dimensionless pressure, p_D', clusters for the entire dataset. Red curves indicate the medoid of each cluster, while blue curves show the other pressure response curves in each cluster.*

4(b)). The resulting feature importance analysis confirmed the visual observations by highlighting fracture intensity, fracture aperture, and fracture length as the most influential parameters in predicting cluster labels, with feature importance values of approximately 0.21, 0.33, and 0.18, respectively. In contrast, the type of fracture set was found to be the least significant parameter, with an feature importance value below 0.05.



(a)



(b)

Figure 4 *Box and whisker plots showing the median, P25, and P75 for fracture properties across all clusters, with circles representing the outliers (a), and and feature importance from the random forest classifier showing the influence of fracture properties on the clustering result (b). The fracture set type is a categorized variable representing either orthogonal (90°) or conjugate (60°) fracture sets.*

Conclusions

The work presented here demonstrates a machine learning approach for characterizing fractured reservoirs, where fractures control the permeability and connectivity of the reservoir, thereby influencing flow behavior. Using K-medoids clustering, we identified 18 distinct flow behavior for a dataset of pressure transient curves obtained through simulations on 5000 geologically consistent DFNs. By correlating the clustered curves to the fracture properties, we found that fracture intensity, fracture aperture, and fracture length have a significant impact on pressure behavior.

Future work aims to apply this workflow to temperature transient data and identify distinct heat flow behavior in fractured geothermal reservoirs. This extension is made possible by the flexibility provided through linking GeoDFN with MRST, which enables the generation of geologically consistent DFNs and directly simulating wide range of flow peocesses for them. These datasets will enable both forward and inverse modeling, meaning predicting dynamic flow and heat responses for a new set of fracture properties and inferring probable ranges of fracture properties from a given pressure or temperature response.

Acknowledgements

The authors acknowledge Energi Simulation for scholarship support for Elahe's PhD and for funding this research.

References

Breiman, L. [2001] Random forests. *Machine Learning*, **45**(1), 5–32.

Collignon, M., Klemetsdal, S. and Møyner, O. [2021] Simulation of Geothermal Systems Using MRST. In: *Advanced Modeling with the MATLAB Reservoir Simulation Toolbox*, Cambridge University Press, 491–514.

Freites, A., Corbett, P., Rongier, G. and Geiger, S. [2023] Automated Classification of Well Test Responses in Naturally Fractured Reservoirs Using Unsupervised Machine Learning. *Transport in Porous Media*, **147**(3), 747–779.

Kamel Targhi, E., Bruna, P.O., Daniilidis, A., Rongier, G. and Geiger, S. [2024] GeoDFN, A Flexible, Open-Source Software for Generating Geologically Consistent Discrete Fracture Networks.

Lepillier, B., Bruna, P.O., Bruhn, D., Bastesen, E., Daniilidis, A., Garcia, , Torabi, A. and Wheeler, W. [2020] From outcrop scanlines to discrete fracture networks, an integrative workflow. *Journal of Structural Geology*, **133**, 103992.

Lepillier, B., Daniilidis, A., Doonechaly Gholizadeh, N., Bruna, P.O., Kummerow, J. and Bruhn, D. [2019] A fracture flow permeability and stress dependency simulation applied to multi-reservoirs, multi-production scenarios analysis. *Geothermal Energy*, **7**(1), 1–16.

Liao, T. [2005] Clustering of time series data—a survey. *Pattern Recognition*, **38**(11), 1857–1874.

Lie, K.A. [2019] An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST). *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave*.

Lie, K.A., Krogstad, S., Ligaarden, I.S., Natvig, J.R., Nilsen, H.M. and Skaflestad, B. [2012] Open-source MATLAB implementation of consistent discretisations on complex grids. *Computational Geosciences*, **16**(2), 297–322.

Rousseeuw, P.J. [1987] Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, **20**(C), 53–65.

Sakoe, H. and Chiba, S. [1978] Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **26**(1), 43–49.

Warren, J. and Root, P. [1963] The Behavior of Naturally Fractured Reservoirs. *Society of Petroleum Engineers Journal*, **3**(03), 245–255.

Wong, D.L.Y., Doster, F. and Geiger, S. [2021] Embedded Discrete Fracture Models. In: Lie, K.A. and Møyner, O. (Eds.) *Advanced Modeling with the MATLAB Reservoir Simulation Toolbox*, Cambridge University Press, 375–408.