POLITECNICO DI TORINO
Repository ISTITUZIONALE

Discrete Fracture Network insights by eXplainable AI

Original
Discrete Fracture Network insights by eXplainable AI / Berrone, S., Della Santa, F., Mastropietro, A., Pieraccini, S., Vaccarino, F.. - ELETTRONICO. - (2020). (Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS) December 11, 2020).

Availability:
This version is available at: 11583/2882079 since: 2022-12-29T01:13:32Z

Publisher:
Neural Information Processing Systems Foundation

Published
DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

29 November 2026

Discrete Fracture Network insights by eXplainable AI

Stefano Berrone
Politecnico di Torino
stefano.berrone@polito.it

Francesco Della Santa
Politecnico di Torino
francesco.dellasanta@polito.it

Antonio Mastropietro
Politecnico di Torino
Addfor S.p.A.
antonio.mastropietro@polito.it

Sandra Pieraccini
Politecnico di Torino
sandra.pieraccini@polito.it

Francesco Vaccarino
Politecnico di Torino
ISI Foundation, Turin
francesco.vaccarino@polito.it

Abstract

In the framework of flow simulations in Discrete Fracture Networks (DFN), we consider the problem of identifying backbones. Backbones of a DFN are sub-networks of fractures that preserve the main flow properties and can be fruitfully used to reduce the computational cost of simulations or to analyze clogging and waste storage problems. With a well-trained Neural Network for flux regression in a DFN at hand, we use the Layer-wise Relevance Propagation method to compute the expected relevance of each fracture to identify the backbone.

1 Introduction

Several applications related to underground analysis require to simulate flow in fractured media (e.g. in geothermal applications, oil & gas production, storage of CO_2). According to their features, underground fractures can represent preferential paths or even flow obstructions, having a strong impact on the flow intensity and directionality. For this reason, within this framework, we consider the Discrete Fracture Network (DFN) model [1, 2, 3], in which each fracture in the network is characterized by its own geometrical and hydro-geological features. A DFN is a discrete model that represents a network of n fractures as a set of 2D polygons $\mathcal{F}_1, \dots, \mathcal{F}_n$, in a 3D domain $\mathbb{D} \subset \mathbb{R}^3$ (see Figure 1), where each polygon $\mathcal{F}_i$ is endowed with its own size and orientation, i.e. the fracture's geometrical properties, and with its own transmissivity parameter κ_i, that is the hydro-geological parameter characterizing the flow facilitation through the fracture.

The demand for accurate flow simulations results in large networks is a challenging task. One of the major issues depends on the geometrical complexity of the computational domain in realistic networks, since it is often quite large and the relative position of the fracture can form very small and narrow angles. Dealing with these complexities typically calls for quite cumbersome mesh generation processes, unless using some specific approaches to overcome it; for the DFN simulations of this work, we follow the approach described in [4, 5, 6], consisting in a reformulation of the problem as a PDE-constrained optimization problem, in which the need for conforming mesh is completely overcome. However, the computational cost of simulations on a large DFN is still expensive, and it may be prohibitive for large numbers of simulations as it can be required, e.g., by Uncertainty Quantification (UQ) analyses. Another issue related to DFN flow simulations is the lack of full-

Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020), Vancouver, Canada.

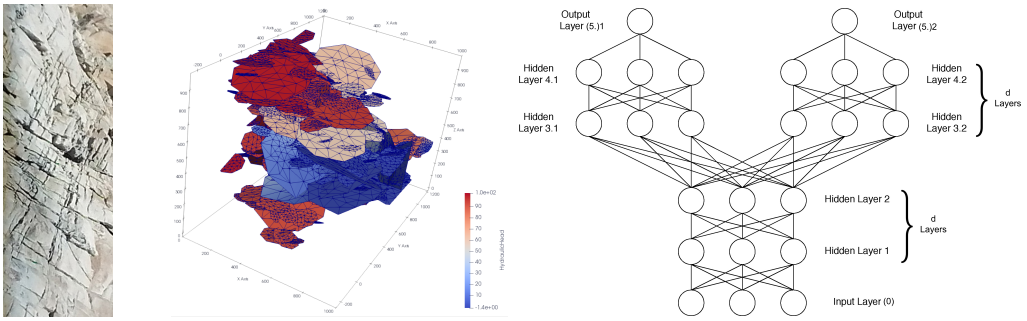

Figure 1: External surface of a natural fractured medium (left) and a DFN (center), NN built for vector valued regression concerning flux prediction (right)

deterministic information about geometrical and hydro-geological properties of fractures. Then, this information is only known in terms of probability distributions, sampling the data needed for actual simulations from the available ones.

In this paper we propose a novel strategy to identify backbones, namely, sub-networks of fractures with transport characteristics approximating the ones of the original DFN [7]. These sub-networks can be used in many applications and furnish precious and fundamental information for clogging problems and for waste storage problems simulated through DFNs. In literature, many backbone identification methods exist (see [8, 9, 10, 7, 11]); these methods use as criterion for identifying the backbones the preservation of the minimum time spent by particles to flow from the source to the sink of the network (also called "first passage time"). However, in the framework of UQ here we consider the total flux exiting the DFN as quantity of interest (QoI), varying the hydro-geological properties of the fractures; therefore, the method that we propose identifies backbones such that their exiting flux approximates the one of the full DFN, modelling transmissivities as random variables.

The new method is based on the Layer-wise Relevance Propagation (LRP) algorithm [12, 13] applied to Neural Networks (NNs) trained for flux regression of DFNs, computing an approximation of the *expected relevance scores* of all the features in the domain space. In this way, we use the LRP as a feature selection method and we identify the backbone fractures as the ones with higher expected relevance score. The validation of the LRP-based feature selection method is done comparing the flow simulations run on the subnetworks with the fluxes of the full DFN.

2 Neural Networks for Flux Regression in Discrete Fracture Networks

Let us briefly describe the Neural Networks (NN) solving a flux regression problem of a test case DFN. The DFN considered, called DFN158, consists of a fixed geometry with $n = 158$ fractures, immersed in a cubic domain $\mathbb{D}$ with a 1000 meters long edge. Fractures have been randomly generated using geological distributions of the geometrical features [14, 15]. The flow problem defined for DFN158 is characterized by setting fixed boundary Dirichlet conditions on those fracture edges obtained intersecting the DFN with the leftmost and rightmost faces of the domain (w.r.t. x axis). Dirichlet conditions impose a fixed pressure difference of 10 meters between the same two faces of $\mathbb{D}$, characterizing the flux directionality; therefore, the resulting outflow fractures are fixed independently of the fracture transmissivities: $\mathcal{F}_8, \mathcal{F}_{12}, \mathcal{F}_{14}, \mathcal{F}_{78}, \mathcal{F}_{90}, \mathcal{F}_{98}, \mathcal{F}_{107}$ ($m = 7$ outflowing fractures). The fracture transmissivities κ_i for each fracture $\mathcal{F}_i$ are modelled as random variables with log-normal distribution [16, 10] : $\log_{10} \kappa_i \sim \mathcal{N}(-5, 1/3)$.

We use a NN model to approximate by regression the $m = 7$ fluxes exiting from DFN158, for any vector of fracture transmissivities $\boldsymbol{\kappa} = [\kappa_1, \ldots, \kappa_n]^\top$. We build a dataset $\mathcal{D}$, given by $N = 10\,000$ pairs $(\boldsymbol{\kappa}, \boldsymbol{\varphi}) \in \mathbb{R}^n \times \mathbb{R}^m$ obtained with a random sampling of N vectors $\boldsymbol{\kappa}$ and computing the corresponding fluxes $\boldsymbol{\varphi}$ through DFN simulations. The test set $\mathcal{P}$ is created randomly picking the 30% of $\mathcal{D}$; the remaining elements are randomly split into the training $\mathcal{T}$ and validation $\mathcal{V}$ set, $|\mathcal{V}| \sim 20\%|\mathcal{D} \setminus \mathcal{P}|$. The NN model is a fully-connected multi-headed NN, characterized by a "tree-shape" structure as illustrated in Figure 1 (right) where the "trunk" and the m "branches" have all the same depth d. In the test case here illustrated, we consider a NN $\mathcal{N}^*$ characterized by this architecture, with a trunk/branch depth parameter $d = 3$ and all hidden layers given by n softplus units each; $\mathcal{N}^*$

Table 1: Dissimilarity between actual and predicted distributions for the outflux fractures

	$\mathcal{F}_8$	$\mathcal{F}_{12}$	$\mathcal{F}_{14}$	$\mathcal{F}_{78}$	$\mathcal{F}_{90}$	$\mathcal{F}_{98}$	$\mathcal{F}_{107}$
$D_{\mathrm{KL}}/\mathcal{E}$	0.0009	0.0003	0.0010	0.0002	0.0033	0.0379	0.0010

is trained on $\mathcal{T}$ using adam [17] and the early stopping regularization method (patience parameter $p^* = 150$). All the hyper-parameters have been chosen after a brief grid-search. As performance measure we assume the mean of the mean weighted error

$$\frac{1}{|\mathcal{P}|}\sum_{(\boldsymbol{\kappa},\boldsymbol{\varphi})\in\mathcal{P}}\frac{1}{m}\sum_{j=1}^{m}\frac{|\widehat{\varphi}_j-\varphi_j|}{\sum_{j=1}^{m}\varphi_j}\,, \tag{1}$$

on the test set, where $\widehat{\boldsymbol{\varphi}}\in\mathbb{R}^m$ is the prediction of $\mathcal{N}^*$ for any vector $\boldsymbol{\kappa}$. The value obtained for $\mathcal{N}^*$ is 0.0082, meaning that the mean error of the predicted flux is globally less than 1% of the total exiting flux. In the framework of UQ, we compare also the distributions of predicted and actual fluxes; then we introduce a dissimilarity measure defined as the ratio between the Kullback-Liebler divergence (D_{KL}) between the two distributions and the entropy ($\mathcal{E}$) of the actual one:

$$D_{\mathrm{KL}}(P_j\|Q_j)/\mathcal{E}(P_j)\,, \tag{2}$$

where P_j is the actual flux's probability distribution of the j-th output flux and Q_j is the one of the corresponding predictions. The ratio (2) can be interpreted as a "relative information error" when we describe P_j using a random sampling from Q_j. Also with respect to (2), $\mathcal{N}^*$ shows good performances (see Table 1), returning distributions Q_j very similar to P_j, for each $j = 1\,,\ldots,m$.

3 Layer-wise Relevance Propagation for Backbone Identification

Layer-Wise Relevance Propagation (LRP) is a subclass of the "eXplainable AI" (XAI) algorithms [18]. Given a NN model, an input $\boldsymbol{x}$ and a score $R(\boldsymbol{x})\in\mathbb{R}^m$ equal to the outflow flux predicted by the NN on each outflux fracture, the LRP algorithm here used propagates the score backward through the NN from the output layer to the input one redistributing it among all the input neurons. The propagation takes into account the NN's weights and architecture and it assumes to preserve the score sum between subsequent layers. The score assigned to each input neuron $\boldsymbol{x}_k$ indicates how much it has contributed in computing the prediction of the NN. In the literature, LRP is applied with respect to one input at a time: the relevance is considered as a characteristic of the input, therefore the most relevant components can vary changing the input taken into account.

The propagation rule comprises the definition of messages between neurons of subsequent layers, from neuron j of layer $(\ell+1)$ to neuron i of layer ℓ . The message $R_{i\leftarrow j}^{(\ell\,,\,\ell+1)}$ represents how much the output of i sent to j is relevant for the overall prediction. The relevance score $R_i^{(\ell)}$ is given by the sum of all the incoming messages:

$$R_i^{(\ell)} = \sum_{j\in(\ell+1)} R_{i\leftarrow j}^{(\ell\,,\,\ell+1)}\,. \tag{3}$$

Therefore, the relevance of the component x_i is given by the quantity $R_i^{(0)}$ computed starting from $R(\boldsymbol{x})$. We choose the LRP algorithm named $\alpha-\beta$ *rule* [12, 13], fixing $\alpha=1,\beta=0$. The idea behind the feature selection proposed is to compute the *expected relevance scores*:

$$\left[\mathbb{E}_{\boldsymbol{x}\sim p}[R_1^{(0)}]\,,\ldots,\mathbb{E}_{\boldsymbol{x}\sim p}[R_n^{(0)}]\right]^T \sim \bar{\boldsymbol{r}} \tag{4}$$

for a random input vector $\boldsymbol{x}\in\mathbb{R}^n$ with distribution p, approximated by its sampling mean $\bar{\boldsymbol{r}}$.

From the point of view of the regression problem in DFNs, the expected relevance score are computed with respect to the inputs $\boldsymbol{\kappa}\sim p$ and allows to order the input features according to their importance. Then, LRP tells us that the contribution of some fractures is negligible in the DFN flux simulations; eventually, by keeping the fractures with high scores and discarding the others, we obtain a subset of fractures that can be identified as a backbone preserving the total exiting flux of the full DFN. The validation of a backbone selected by the exposed procedure consist of running flow simulations on this sub-network of fractures and comparing its flux distributions with the ones of the full DFN. If the two simulated fluxes of the DFN and of the sub-DFN agree, we can say that the NN have captured the main characteristics of the simulated physical phenomenon and LRP have indicated the most important fractures for the flow propagation.

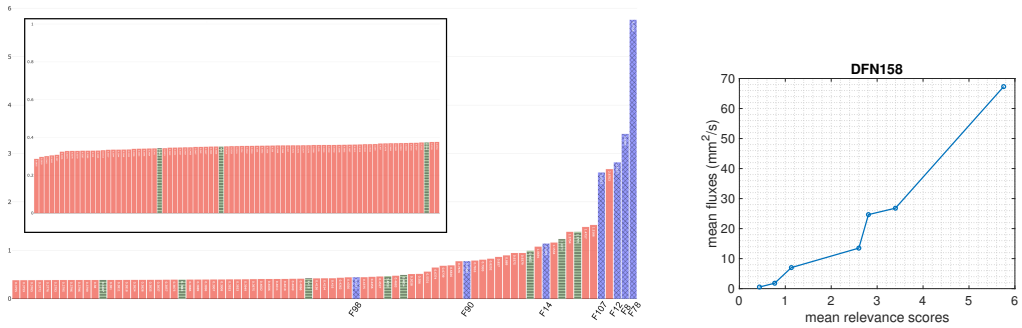

Figure 2: Left: Fractures ordered by ascending value of $\boldsymbol{r}$ (the box in the top-left corner contains the lowest 60%); in blue the labelled outflow fractures, in green the inflow fractures. Right: mean relevance scores versus mean fluxes (mm^2/s) of outflow fractures.

4 Results

Let us analyze the vector of mean relevance scores $\bar{\boldsymbol{r}} \in \mathbb{R}^{158}$ computed on the whole dataset $\mathcal{D}$. A first observation reveals that all the fractures with exiting flux belong to the set of fractures in the top 25% with highest relevance scores (see Figure 2-left). This observation has non-trivial consequences: it indicates that $\mathcal{N}^*$ learnt to approximate the fluxes coherently with the topology of the fractures network, even if it has no information about relationships between inputs and outputs; indeed, if $\mathcal{F}_i$ is the j-th boundary exit fracture, no information about the strict physical-based relationship between κ_i and φ_j has been given to $\mathcal{N}^*$. In addition, the mean relevance scores of the outflow fractures are characterized by a non-negligible dependence on the mean value of the fracture fluxes (see Figure 2-right). Focusing on the graphs that characterize the sub-networks given by the set of fractures in the top 25% , 50% , 75% relevance scores (Figure 3) we observe that: i) less relevant fractures are mainly those belonging to "dead-end" branches of the original graph, since they are the first fractures ignored when we keep the 75% most relevant ones; ii) keeping 50% and 25% of the most relevant fractures, fractures belonging to source-sink paths (i.e. paths from any inlet fracture to any outlet fracture) are also ignored, thus reducing the number of these paths but leaving always at least one of them; iii) $\mathcal{N}^*$ seems to have understood that some bottleneck nodes are fundamental for the existence of a source-sink path, since these fractures belong to the top 25% relevance scores.

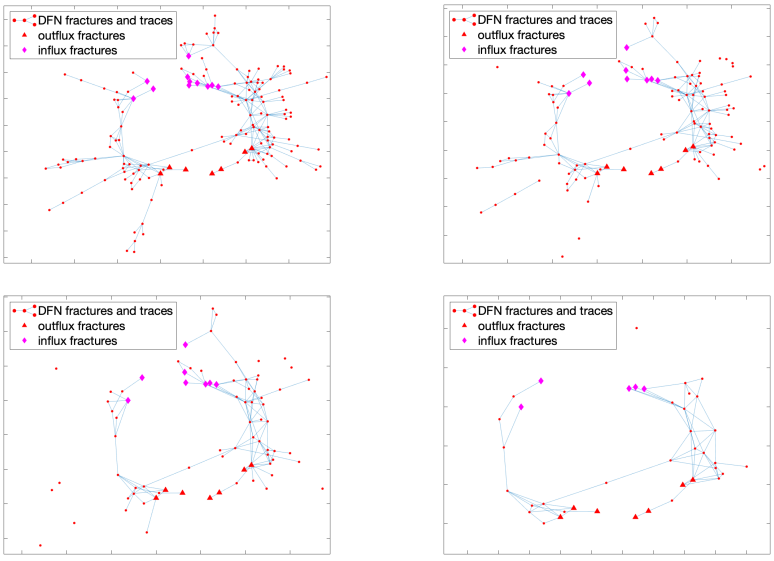

Figure 3: Graphs of fractures networks. Relevance scores: top 100%, 75%; bottom 50%, 25%.

Finally, in order to confirm that the detected sub-networks are actually backbones of DFN158, we run numerical simulations on all the sub-DFNs, assuming input transmissivities equal to the ones in the original $\mathcal{D}$ restricted to the remaining fractures. The results show a good preservation of both the mean and standard deviation of the original total flux for all the sub-DFNs; indeed, the sub-networks given by the $75\%\,, 50\%\,, 25\%$ most relevant fractures preserve the mean total exiting flux with the $96.61\%, 94.26\%, 85.62\%$ and a flux standard deviation with the $99.78\%, 99.01\%, 97.85\%$ of the original one, respectively.

Broader Impact

The method proposed can have useful applications in many geological clogging and storage problems, since it takes into account not only the graph connections of the DFN but also the stochasticity of the fractures hydro-geological properties. Going beyond the DFN context, an interesting aspect of this work is the novel application of the LRP algorithm as feature selector; indeed, we used an eXplainable AI algorithm to understand the most relevant input features of a numerical simulator, through a NN trained to approximate the simulator. This new application of XAI algorithms and NNs can have interesting implications, first of all concerning the dimensionality reduction of high dimensional problems. In a broader sense, we argue that the coupling between model-based simulations and data-driven machine learning algorithms could provide a powerful and insightful approach both to the explainability of AI, by integrating formalised field knowledge in the actual learning paradigm, and the integration of AI into simulable systems as the engineered ones. In addition the usage of machine learning and eXplainable AI algorithms allow to give a new meaning to go "forward", towards the prediction, or "backward", towards the input, giving insights about the main or negligible components of the global problem.

Acknowledgments and Disclosure of Funding

Research performed in the framework of the Italian MIUR Award "Dipartimento di Eccellenza 2018-2022" granted to the Department of Mathematical Sciences, Politecnico di Torino, CUP: E11G18000350001. The research leading to these results has also been partially funded by INdAM-GNCS and by the SmartData@PoliTO center for Big Data and Machine Learning technologies. S. Pieraccini also acknowledges support from Italian MIUR PRIN project 201752HKH8_003. F. Vaccarino acknowledges partial support from Intesa Sanpaolo Innovation Center. The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

References

[1] P. M. Adler, *Fractures and Fracture Networks*. Kluwer Academic, Dordrecht, 1999.

[2] G. Cammarata, C. Fidelibus, M. Cravero, and G. Barla, "The hydro-mechanically coupled response of rock fractures," *Rock Mechanics and Rock Engineering*, vol. 40, no. 1, pp. 41–61, 2007.

[3] C. Fidelibus, G. Cammarata, and M. Cravero, *Hydraulic characterization of fractured rocks. In: Abbie M, Bedford JS (eds) Rock mechanics: new research.* Nova Science Publishers Inc., New York, 2009.

[4] S. Berrone, A. Borio, and S. Scialò, "A posteriori error estimate for a PDE-constrained optimization formulation for the flow in DFNs," *SIAM J. Numer. Anal.*, vol. 54, no. 1, pp. 242–261, 2016.

[5] S. Berrone, S. Pieraccini, and S. Scialò, "On simulations of discrete fracture network flows with an optimization-based extended finite element method," *SIAM J. Sci. Comput.*, vol. 35, no. 2, pp. A908–A935, 2013. [Online]. Available: http://dx.doi.org/10.1137/120882883

[6] ——, "An optimization approach for large scale simulations of discrete fracture network flows," *J. Comput. Phys.*, vol. 256, pp. 838–853, 2014.

[7] S. Srinivasan, S. Karra, J. Hyman, H. Viswanathan, and G. Srinivasan, "Model reduction for fractured porous media: a machine learning approach for identifying main flow pathways,"

Computational Geosciences, Mar 2019, doi:10.1007/s10596-019-9811-7. [Online]. Available: https://doi.org/10.1007/s10596-019-9811-7

[8] G. Aldrich, J. D. Hyman, S. Karra, C. W. Gable, N. Makedonska, H. Viswanathan, J. Woodring, and B. Hamann, "Analysis and visualization of discrete fracture networks using a flow topology graph," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 8, pp. 1896–1909, 2017.

[9] J. D. Hyman, A. Hagberg, G. Srinivasan, J. Mohd-Yusof, and H. Viswanathan, "Predictions of first passage times in sparse discrete fracture networks using graph-based reductions," *Physical Review E*, vol. 96, no. 1, pp. 1–10, 2017.

[10] J. D. Hyman, A. Hagberg, D. Osthus, S. Srinivasan, H. Viswanathan, and G. Srinivasan, "Identifying Backbones in Three-Dimensional Discrete Fracture Networks: A Bipartite Graph-Based Approach," *Multiscale Modeling & Simulation*, vol. 16, no. 4, pp. 1948–1968, 2018.

[11] S. Srinivasan, E. Cawi, J. Hyman, D. Osthus, A. Hagberg, H. Viswanathan, and G. Srinivasan, "Physics-informed machine learning for backbone identification in discrete fracture networks," *Computational Geosciences*, no. Mc, 2020.

[12] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," *PloS one*, vol. 10, no. 7, p. e0130140, 2015.

[13] G. Montavon, W. Samek, and K.-R. Müller, "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1–15, 2018.

[14] Svensk Kärnbränslehantering AB, "Data report for the safety assessment, SR-site," SKB, Stockholm, Sweden, Tech. Rep. TR-10-52, 2010.

[15] J. D. Hyman, G. Aldrich, H. Viswanathan, N. Makedonska, and S. Karra, "Fracture size and transmissivity correlations: Implications for transport simulations in sparse three-dimensional discrete fracture networks following a truncated power law distribution of fracture size," *Water Resources Research*, 2016.

[16] X. Sanchez-Vila, A. Guadagnini, and J. Carrera, "Representative hydraulic conductivities in saturated grqundwater flow," *Reviews of Geophysics*, vol. 44, no. 3, pp. 1–46, 2006.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2015. [Online]. Available: http://arxiv.org/abs/1412.6980

[18] F. K. Dosilovic, M. Brcic, and N. Hlupic, "Explainable artificial intelligence: A survey," in *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, 2018, pp. 0210–0215.