Estimating geomechanical parameters from hydraulic fracturing tests using a soft computing-based methodology

Rafael Abreu[1,2], Cristian Mejia[2], Deane Roehl[1,2], Leonardo Cabral[3]

[1]*Department of Civil and Environmental Engineering, Pontifical Catholic University of Rio de Janeiro*
225 Marquês de São Vicente Street, Gávea, 22451-900, Rio de Janeiro, RJ, Brazil
rafaeloabreu@tecgraf.puc-rio.br, deane@tecgraf.puc-rio.br
[2]*Tecgraf Institute, Pontifical Catholic University of Rio de Janeiro*
225 Marquês de São Vicente Street, Gávea, 22451-900, Rio de Janeiro, RJ, Brazil
crisms@tecgraf.puc-rio.br
[3]*Petrobras - Petróleo Brasileiro S.A, Brazil*
leonardocabral@petrobras.com.br

Abstract. The discovery of naturally fractured reservoirs in the Brazilian pre-salt has attracted considerable attention for a better understanding of reservoir characterization and fluid flow inside fracture channels. Predicting the hydromechanical behavior of these reservoirs is a cumbersome task, which requires the identification of their geomechanical parameters. In this scenario, a soft computing-based methodology is implemented to estimate geomechanical parameters from borehole injection pressure in hydraulic fracturing tests. Based on artificial intelligence techniques, this approach integrates a proxy model and an optimization algorithm to match the field measurements and the borehole pressure curve predicted by a finite element model. Considering a multistep-ahead strategy to predict time series, a multilayer perceptron-based proxy model computes the borehole pressure curves, substituting the numerical simulation of a minifrac test. The adoption of a proxy model substantially reduces the computational effort of the parameter identification task. Therefore, a genetic algorithm can efficiently estimate the reservoir geomechanical parameters by solving a nonlinear least squares problem. The application to field-measured data from a minifrac test confirms the capability of the proposed methodology to estimate geomechanical parameters from hydraulic fracturing tests.

Keywords: hydraulic fracturing, geomechanical parameter identification, neural networks, genetic algorithm.

1 Introduction

Information about geomechanical properties and in-situ stresses is essential for petroleum engineering in conducting several operations, analyses, and estimations. Some techniques have been developed to estimate in-situ stresses [1–3]. Nonetheless, these methods may require substantial investments to acquire reliable data in deepwater applications. In this scenario, various empirical relationships have been proposed to estimate the elastic properties, permeability, and porosity of rock formations using well logs [4,5]. However, these estimated parameters are only representative near the wellbore [6]. While hydraulic fracturing tests are the most direct means to estimate the minimum in-situ stress in the oil and gas industry, they are inaccurate to determine the maximum horizontal in-situ stress [7,8].

In this context, inverse analysis can be applied to estimate geomechanical parameters in hydraulic fracturing tests. Inverse problems arise when observations are available, but the underlying causes are unknown. Inverse analysis serves as a powerful tool to identify parameters in mathematical models, replacing the trial-and-error process. In a modern approach, stochastic optimization methods are employed to solve inverse problems, particularly due to the occurrence of complex topologies in objective functions to be minimized [9]. Therefore, several studies utilize optimization methods inspired by natural processes to solve inverse analysis problems. It is important to emphasize that these methods do not guarantee obtaining the optimal solution; however, these methods often compute solutions satisfactorily close to the global minimum.

In cases where the mathematical model under study lacks an analytic expression, evaluating the objective function during a parameter identification task may be computationally expensive. To overcome this, several researchers have employed proxy (surrogate) models of numerical simulations, often utilizing artificial neural networks. Considering hydraulic fracturing simulations, Zhang and Yin [10] and Zhang et al. [11] applied neural networks and genetic algorithms to estimate geomechanical parameters based on borehole pressure data. Additionally, Zhang and Yin [12] and Zhang et al. [13,14] employed these techniques to compute horizontal stresses using information from leak-off tests. Abreu et al. [15,16] introduced approaches for parameter identification in hydraulic fracturing tests based on time series recursive predictions.

Therefore, this study presents a soft computing-based methodology for parameter identification in finite element simulations of hydraulic fracturing. This methodology uses artificial intelligence methods to match numerical and experimental borehole pressure curves. In this scenario, a genetic algorithm minimizes the difference between numerical and experimental data. In addition, a multilayer perceptron is trained to approximate the outcomes of numerical models, reducing the computational cost of the optimization problem by decreasing the number of required finite element analyses. The presented methodology is heavily based on the proof of concept developed by Abreu et al. [15]. The methodology is then applied to identify parameters in a field minifrac test.

2 Methodology

For the numerical simulation of minifrac tests, the finite element method is employed, utilizing the Abaqus® software. Thus, a proxy model is defined to reduce the computational effort of the parameter identification task. According to the definitions presented in this section, this proxy model is built based on machine learning concepts.

2.1 Artificial neural networks

In this study, the proxy model is defined using a multilayer perceptron, as shown in the following equation:

$$\boldsymbol{a}_{n+1} = f_{n+1}(\boldsymbol{W}_{n+1}\boldsymbol{a}_n + \boldsymbol{b}_{n+1}) \quad (1)$$

where $\boldsymbol{a}_{n+1}$ and $\boldsymbol{a}_n$ respectively represent the outputs of the current and previous layers; f_{n+1} denotes the activation function of the current layer; $\boldsymbol{W}_{n+1}$ represents the weight matrix that connects the current and previous layers; and $\boldsymbol{b}_{n+1}$ defines the biases of the current layer. Moreover, $\boldsymbol{a}_0$ represents the input values of the neural network. In this context, the weights and biases are determined using the Levenberg-Marquardt algorithm [17]. Additionally, the weights are initialized according to the strategy proposed by Glorot and Bengio [18].

2.2 Multistep-ahead prediction

Predicting time series, such as the borehole pressure curves studied in this work, can be a cumbersome task due to the accumulation of errors, increased uncertainty, and deteriorating accuracy [19]. Nevertheless, various approaches, particularly the recursive strategy, have been successfully adopted in diverse time series applications [19]. When training a machine learning model using the recursive strategy, the one-step-ahead prediction is used:

$$\hat{y}_{t+1} = \hat{f}(y_t, y_{t-1}, \ldots, y_{t-h+1}, x_1, x_2, \ldots, x_m) \quad (2)$$

in which $\hat{y}$ is a predicted value, $\hat{f}$ is the machine learning model, y is a known value from the series, t is the time step, h is the moving window size, x_k is an exogenous variable, and m is the number of exogenous variables. In this study, the exogenous variables are the geomechanical parameters and a variable that regulates the injection flow. Once the model $\hat{f}$ is established, the next value $\hat{y}_{t+2}$ is predicted according to Equation (3).

$$\hat{y}_{t+2} = \hat{f}(\hat{y}_{t+1}, y_t, \ldots, y_{t-h+2}, x_1, x_2, \ldots, x_m) \quad (3)$$

Therefore, the values $\hat{y}_{t+3}, \hat{y}_{t+4}, \ldots, \hat{y}_{t+n}$ are predicted utilizing the same idea presented in Equation (3). In this work, the initial values y_t to y_{t-h+1} are defined based on the initial pore pressure p_o. During the training of $\hat{f}$, all pressure data are normalized respecting the same interval.

2.3 Genetic algorithm

A genetic algorithm computationally represents the biological evolution, simulating the Darwinian natural selection mechanisms. This algorithm is extensively applied in the solution of global optimization problems. This work adopts the real-coded genetic algorithm implemented by Abreu et al. [15] to identify the geomechanical parameters. The applied algorithm solves the following optimization problem:

$$\min_{\boldsymbol{x} \in \boldsymbol{\Omega}} F(\boldsymbol{x}) \tag{4}$$

in which $F(\boldsymbol{x})$ denotes the objective function, $\boldsymbol{x}$ denotes the vector of continuous components, and $\boldsymbol{\Omega}$ denotes the problem domain. To characterize the parameter identification problem, $F(\boldsymbol{x})$ is defined in Section 2.4. As this algorithm is population-based, the global minimum is the individual with the lowest objective function.

2.4 Inverse analysis procedure

Estimating parameters of mathematical models based on a set of observations is an inverse problem. Optimization methods are often used to solve such problems. This approach aims to minimize an objective function, which directly depends on observed and predicted data. Once the optimization problem (4) is defined, the objective function is specified based on the nonlinear least squares, as shown in Equation (5).

$$F(\boldsymbol{x}) = \boldsymbol{r}(\boldsymbol{x})^T \boldsymbol{r}(\boldsymbol{x}) \tag{5}$$

with

$$\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{y}_{obs} - \boldsymbol{y}_{pred}(\boldsymbol{x}) \tag{6}$$

considering that $\boldsymbol{x}$ is a set of parameters to be estimated, $\boldsymbol{y}_{obs}$ is of the observed data, and $\boldsymbol{y}_{pred}$ is of the predicted data (outputs of a mathematical model). In the presented methodology, the model output $\boldsymbol{y}_{pred}(\boldsymbol{x})$ is calculated by applying an approximation of the objective function. Consequently, the objective function directly relies on a machine learning model that has been previously developed to replace a numerical simulator. Figure 1 presents a flowchart of the soft computing-based inverse analysis procedure.

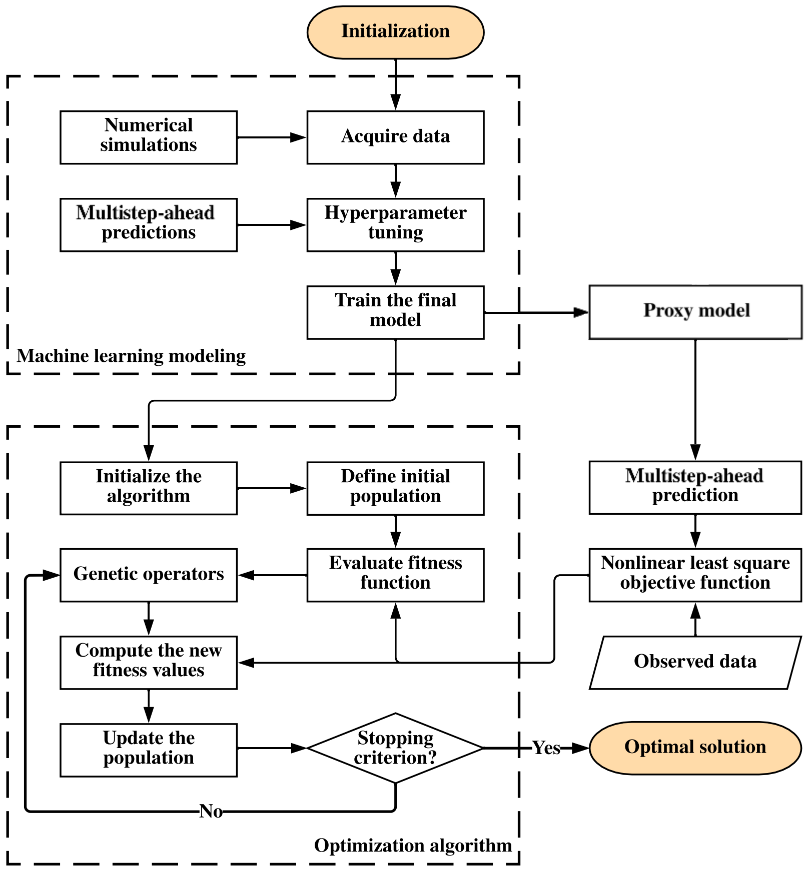

Figure 1. Flowchart of the adopted inverse analysis procedure

3 Application

A parameter identification study for a minifrac test was conducted based on the described methodology. The primary aim is to estimate certain properties of the fractured medium and the injected fluid. To accomplish this, field data are utilized. The minifrac test was carried out in a vertical well of a carbonate reservoir situated in Santos Basin (Brazil). The finite element model of this test is illustrated in Figure 2. Further details about the finite element model and formulation can be found in the works of Abreu et al. [15,16] and Rueda et al. [20].

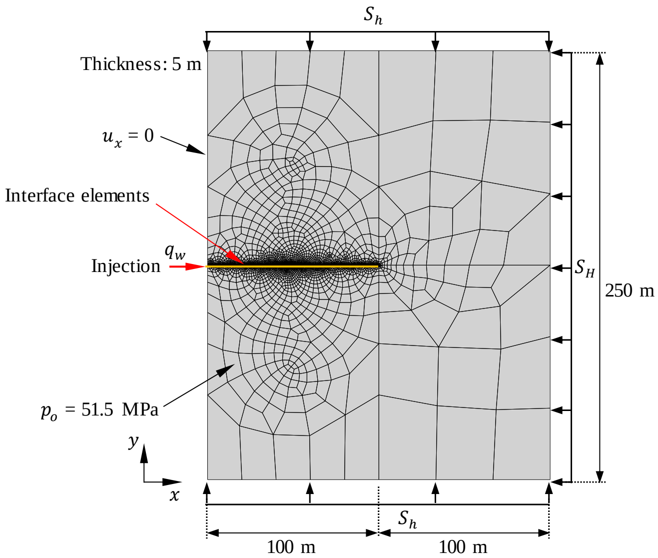

Figure 2. Geometry, finite element mesh, and boundary conditions of the minifrac test

The proxy model for the problem was built from various numerical simulations. The Latin hypercube sampling method was employed to generate 1000 sample points (approximately 307 hours of execution time) consisting of the geomechanical parameters investigated in this inverse problem: Young's modulus E, rock tensile strength σ_t, minimum horizontal stress S_h, ratio of horizontal stresses $K = S_H / S_h$, rock permeability k, and dynamic fluid viscosity u_f. Table 1 shows the upper and lower bounds adopted for the sampling process. The generated dataset comprises equally spaced data points for time intervals of 30 seconds. For simplification reasons, the time simulation was limited to 1200 seconds, despite the experimental test extending over a longer period.

Table 1. Upper and lower bounds for sampling and optimization procedures

Parameter	Upper bound	Lower bound
Young's modulus E (GPa)	40	10
Rock tensile strength σ_t (MPa)	5	0.5
Minimum horizontal stress S_h (MPa)	81.5	56.5
Ratio of horizontal stresses K	1.2	1
Rock permeability k (md)	500	5
Dynamic fluid viscosity μ (cp)	1000	500

Exhaustive search is employed to evaluate certain predefined sets of hyperparameters of neural networks, namely the number of neurons and hidden layers. Subsequently, the set presenting the best validation performance determines the final model. This study adopts a 10-fold cross-validation process to quantify the performance for each set of hyperparameters. The performance evaluation utilizes the Root Mean Squared Error (RMSE) metric. Moreover, the error calculated through multistep-ahead prediction is utilized in performance calculations since the final application considers this type of prediction.

Once the optimal set of hyperparameters is determined, the final model is built. This involves training the same model multiple times and selecting the one with the lowest validation error. To mitigate overfitting, models displaying significant disparities among training, validation, and test errors are excluded. Importantly, the neural network modeling was developed using the MATLAB® platform.

Initially, multilayer perceptrons were modeled using 70%, 15%, and 15% of all collected data for the training, validation, and test datasets, respectively. Figure 3 presents the mean and best model performances (multistep-ahead prediction) for the predetermined number of neurons and hidden layers employed in the search. The number of neurons and hidden layers investigated in the exhaustive search is listed in Table 2. In the adopted notation, 10-20-20-1 indicates a model with ten and one neurons in the input and output layers, respectively, and two hidden layers with twenty neurons each. It should be noted that the RMSE was calculated using normalized outputs. As indicated in the graph, the best set of hyperparameters was selected based on the average validation error.

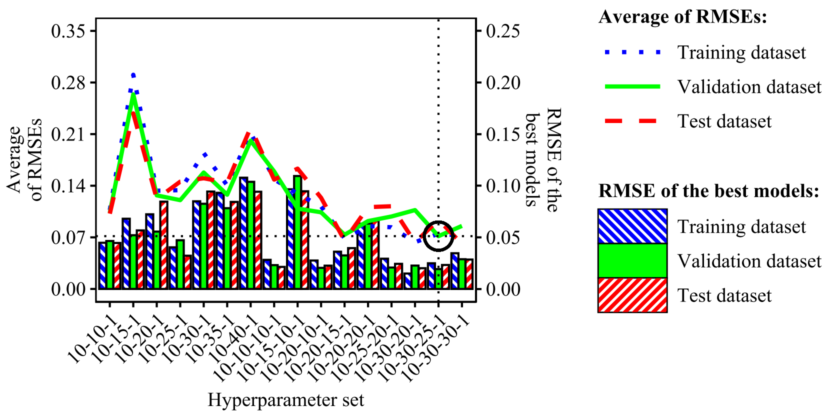

Figure 3. Multilayer perceptron performance computed for the exhaustive search

Table 2. Investigated neural networks structures

Number of hidden layers	Number of neurons in each layer
1	10-10-1, 10-15-1, 10-20-1, 10-25-1, 10-30-1, 10-35-1, and 10-40-1
2	10-10-10-1, 10-15-10-1, 10-20-10-1, 10-20-15-1, 10-20-20-1, 10-25-20-1, 10-30-20-1, 10-30-25-1, and 10-30-30-1

Considering the chosen number of neurons and hidden layers determined for the exhaustive search, the final proxy model is trained. This model was employed to predict the borehole pressure curves of 150 simulations from the test dataset. Therefore, the multistep-ahead prediction was applied to compute the results depicted in Figure 4, which presents a linear regression between the normalized predicted data and the normalized observed ones. Furthermore, the Pearson correlation coefficient R is provided. It is evident that the predicted and observed data exhibit a strong correlation. Consequently, the final model shows a good performance, being appropriate to the parameter identification task.

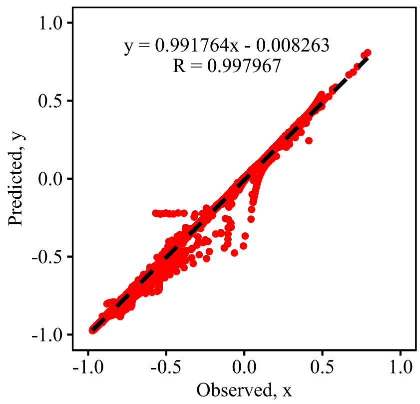

Figure 4. Correlation between the observed and predicted test dataset considering multistep-ahead predictions

Using the genetic algorithm, the aforementioned geomechanical parameters are estimated via inverse analysis. The domain $\boldsymbol{\Omega}$ of the optimization problem is defined by the intervals specified in Table 1. Due to the stochastic nature of the genetic algorithm, it is considered a best practice to execute it multiple times, as this may yield various solutions to be assessed. Nevertheless, the multiple executions of the genetic algorithm consistently yielded the following estimations of the geomechanical parameters: $E = 10$ GPa, $\sigma_t = 0.5$ MPa, $S_h = 60.69$ MPa, $K = 1.20$, $k = 133.56$ md, and $u_f = 500$ cp. In terms of performance, Figure 5 illustrates the average and best values of the objective function computed throughout the optimization procedure. Note that the algorithm required approximately 20 generations to reach the vicinity of the optimal solution and 45 generations to achieve convergence. It is worth mentioning that the number of evaluations of the objective function during the optimization process (8100 evaluations) is greater than the number of numerical models used to construct the neural network. This suggests that approximately eight times the number of numerical simulations would be necessary to identify the geomechanical parameters without using the multilayer perceptron-based proxy model. Moreover, given the importance of executing the optimization algorithm multiple times, proxy modeling becomes a worthy tool to reduce the computational effort involved in the parameter identification tasks.

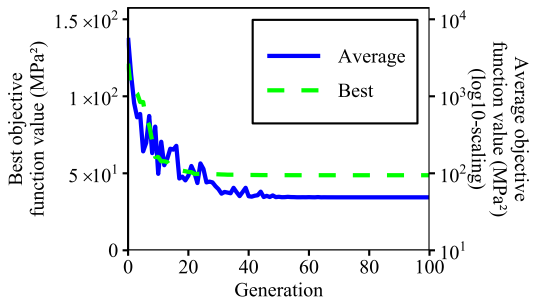

Figure 5. Average and best objective function values computed during the optimization process

Finally, Figure 6 illustrates the borehole pressure curve obtained through the parameter identification process. The results of the numerical simulation and the neural network, utilizing the aforementioned parameter set, are compared with the field-measured borehole pressure data. It is worth emphasizing that the neural network accurately predicts the numerical response, including the abrupt pressure change resulting from the interruption of fluid injection.

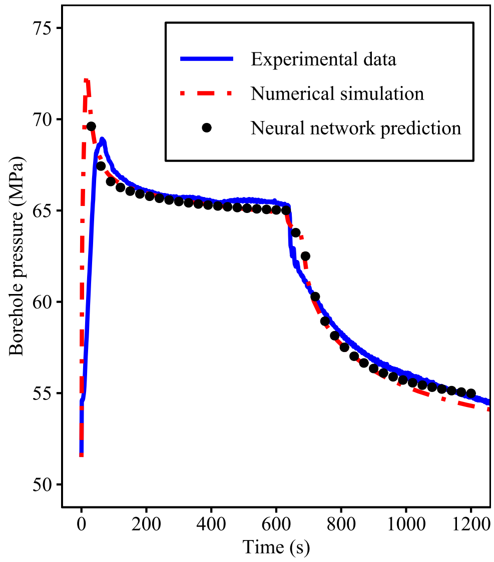

Figure 6. Comparison between experimental data, numerical simulation, and neural network prediction

4 Conclusions

To identify geomechanical parameters from a minifrac test, a methodology combining artificial intelligence techniques was applied. This soft computing-based approach performed the parameter identification process of a hydraulic fracturing problem under a feasible computational cost. A multilayer perceptron was trained as a proxy model, generating an approximate objective function to be minimized through a genetic algorithm. The hyperparameters of the multilayer perceptron were successfully computed by exhaustive search. Remarkably, the proxy model could satisfactorily represent the numerical simulation outcomes, as evidenced by the strong correlation between observed and predicted data. As a consequence of the adopted methodology, a reduced number of finite element analyses were required to solve the parameter identification problem.

Notably, the current proxy model was trained under specific conditions related to the present minifrac test, such as flow rate and injection time. Therefore, the model is tailored to this particular application and may not be directly transferable to different scenarios or conditions. Nonetheless, this study emphasizes the potential of integrating machine learning models and genetic algorithms to provide a powerful and efficient strategy to estimate geomechanical parameters. Future research should continue to explore the boundaries and applicability of this approach in a wider spectrum of geomechanical settings.

Acknowledgements. The authors gratefully acknowledge Petrobras, Brazilian National Council for Scientific and Technological Development (CNPq) Grants 142461/2020-2, 308056/2022-1, and 309384/2019-2, Carlos Chagas Filho Foundation for Supporting Research in the State of Rio de Janeiro (FAPERJ) Grants E-26/201.391/2021 and E-26/202.928/2019, and Tecgraf Institute/PUC-Rio for their support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] B. Amadei, O. Stephansson, Rock Stress and Its Measurement, Springer Netherlands, Dordrecht, 1997.

https://doi.org/10.1007/978-94-011-5346-1.
[2] G. Hareland, L.L. Hoberock, Use of Drilling Parameters To Predict In-Situ Stress Bounds, in: SPE/IADC Drill. Conf., SPE, Amsterdam, Netherlands, 1993: pp. 457–471. https://doi.org/10.2118/25727-MS.
[3] M.D. Zoback, C.A. Barton, M. Brudy, D.A. Castillo, T. Finkbeiner, B.R. Grollimund, D.B. Moos, P. Peska, C.D. Ward, D.J. Wiprut, Determination of stress orientation and magnitude in deep wells, Int. J. Rock Mech. Min. Sci. 40 (2003) 1049–1076. https://doi.org/10.1016/j.ijrmms.2003.07.001.
[4] J.C. Jaeger, N.G.W. Cook, R.W. Zimmerman, Fundamentals of Rock Mechanics, 4th ed., Blackwell Publishing, Oxford, 2007.
[5] M.D. Zoback, Reservoir Geomechanics, Cambridge University Press, Cambridge, 2007. https://doi.org/10.1017/CBO9780511586477.
[6] S. Rafieepour, S. Baldino, S.Z. Miska, Determination of in-situ elastic properties and reservoir boundary conditions, J. Nat. Gas Sci. Eng. 81 (2020) 103397. https://doi.org/10.1016/j.jngse.2020.103397.
[7] J.W. Puller, K.W. Mills, R.G. Jeffrey, R.J. Walker, In-situ stress measurements and stress change monitoring to monitor overburden caving behaviour and hydraulic fracture pre-conditioning, Int. J. Min. Sci. Technol. 26 (2016) 103–110. https://doi.org/10.1016/j.ijmst.2015.11.017.
[8] H. Xu, S. Sang, J. Yang, J. Jin, Y. Hu, H. Liu, P. Ren, W. Gao, In-situ stress measurements by hydraulic fracturing and its implication on coalbed methane development in Western Guizhou, SW China, J. Unconv. Oil Gas Resour. 15 (2016) 1–10. https://doi.org/10.1016/j.juogr.2016.04.001.
[9] N. Moreira, T. Miranda, M. Pinheiro, P. Fernandes, D. Dias, L. Costa, J. Sena-Cruz, Back analysis of geomechanical parameters in underground works using an Evolution Strategy algorithm, Tunn. Undergr. Sp. Technol. 33 (2013) 143–158. https://doi.org/10.1016/j.tust.2012.08.011.
[10] S. Zhang, S. Yin, Determination of in situ stresses and elastic parameters from hydraulic fracturing tests by geomechanics modeling and soft computing, J. Pet. Sci. Eng. 124 (2014) 484–492. https://doi.org/10.1016/j.petrol.2014.09.002.
[11] S. Zhang, J. Lv, X. Yuan, S. Yin, BP Neural Network with Genetic Algorithm Optimization for Prediction of Geo-Stress State from Wellbore Pressures, Int. J. Comput. Intell. Appl. 15 (2016) 1650013. https://doi.org/10.1142/S1469026816500139.
[12] S. Zhang, S. Yin, A hybrid ANN-GA method for analysis of geotechnical parameters, in: 2016 12th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov., IEEE, Changsha, China, 2016: pp. 1–6. https://doi.org/10.1109/FSKD.2016.7603141.
[13] S. Zhang, S. Yin, F. Wang, H. Zhao, Characterization of In Situ Stress State and Joint Properties from Extended Leak-Off Tests in Fractured Reservoirs, Int. J. Geomech. 17 (2017) 04016074. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000757.
[14] S. Zhang, Y. Yuan, H. Fang, F. Wang, An application of soft computing for the earth stress analysis in hydropower engineering, Soft Comput. 24 (2020) 4739–4749. https://doi.org/10.1007/s00500-019-04542-x.
[15] R. Abreu, C. Mejia, D. Roehl, Inverse analysis of hydraulic fracturing tests based on artificial intelligence techniques, Int. J. Numer. Anal. Methods Geomech. 46 (2022) 2582–2602. https://doi.org/10.1002/nag.3419.
[16] R. Abreu, C. Mejia, D. Roehl, Parameter identification of minifrac numerical tests using a gradient boosting-based proxy model and genetic algorithm (in press), Int. J. Numer. Anal. Methods Geomech. (2023).
[17] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the Marquardt algorithm, IEEE Trans. Neural Networks. 5 (1994) 989–993. https://doi.org/10.1109/72.329697.
[18] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: Y.W. Teh, M. Titterington (Eds.), Proc. 13th Int. Conf. Artif. Intell. Stat., Proceedings of Machine Learning Research, Sardinia, 2010: pp. 249–256. https://proceedings.mlr.press/v9/glorot10a.html.
[19] S. Ben Taieb, G. Bontempi, A.F. Atiya, A. Sorjamaa, A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition, Expert Syst. Appl. 39 (2012) 7067–7083. https://doi.org/10.1016/j.eswa.2012.01.039.
[20] J. Rueda, C. Mejia, R. Quevedo, D. Roehl, Impacts of natural fractures on hydraulic fracturing treatment in all asymptotic propagation regimes, Comput. Methods Appl. Mech. Eng. 371 (2020) 113296. https://doi.org/10.1016/j.cma.2020.113296.