12º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS

TÍTULO DO TRABALHO:

PGA-ITML - a Physics and Geostatistics-Aware, Information-Theoretic Machine Learning for Upscaling Heterogeneous Reservoir

AUTORES:

Ianyqui Falcão Costa – UFPE - ianyqui.costa@ufpe.br
Liliane de Allan Fonseca – UFPE - Liliane.fonseca@ufpe.br
Ramiro Brito Willmersdorf - UFPE - ramiro.willmersdorf@ufpe.br
Ézio da Rocha Araújo - UFPE - ezio@ufpe.br

INSTITUIÇÃO:

Universidade Federal de Pernambuco

PGA-ITML - a Physics and Geostatistics-Aware, Information-Theoretic Machine Learning for Upscaling Heterogeneous Reservoir

Resumo

Redes Neurais Profundas têm mostrado avanços significativos na solução de problemas de equações diferenciais parciais dependentes do tempo com coeficientes suaves, principalmente nos casos de altas dimensões e Big Data. Porém, a despeito disso, elas produzem resultados com acurácia pobre. Análise de incertezas em simulações de reservatórios oferecem grandes dificuldades já que são problemas de baixa dimensão (pequeno espaço de fases) e possuem propriedades heterogêneas, além de coeficientes variáveis no tempo. Não há claras vantagens de redes neurais profundas sobre algoritmos tradicionais quando as propriedades acima coexistem. Elas falham na solução de reservatórios heterogêneos, tornando necessários homogeneizá-los por subdomínios antes de submetê-los aos PINNs. As heterogeneidades mudam com o tempo, e assim, sucessivas homogeneizações devem ser feitas ao longo do tempo. Neste trabalho propomos o uso de Teoria da Informação em Aprendizado de Máquinas que leva em conta a física e a geoestatísticas a priori, para resolver os problemas acima descritos, fazendo primeiro uma redução de escalas. Essa mudança de escalas obedece à geoestatística e às físicas envolvidas, contrariamente aos PINNs que a impõe a física a posteriori. A solução do modelo substituto pode ser feita por métodos diretos com a vantagem de alta confiabilidade que não existe em redes neurais em regime de Small Data. Resultados que contemplam a variação dos coeficientes no tempo será descrito em outra publicação futura.

Palavras-chave: Simulação estocástica de reservatório, DNN, PINN, ITML, mudança de escala.

Abstract

Deep Neural Networks (DNNs) have shown significant advances in solving time-dependent partial differential equations with smooth coefficients, mainly in high-dimensional and Big-Data problems. However, despite this, they could have poor accuracy. Uncertainty analysis (UA) of reservoir simulations offers substantial difficulties because they are low-dimension problems and have heterogeneous reservoir properties and coefficients varying in time. Moreover, they belong to the Small-Data problem class. There is no clear advantage of DNNs over traditional algorithms with those collective properties. Physical Informed NNs (PINN) used to be the main driver of Machine Learning applications of scientific and engineering problems. However, it failed to solve reservoir heterogeneous problems, making it necessary to homogenize subdomains before submitting problems to PINNs. As heterogeneity changes with time, successive homogenization of selected regions should be done, introducing new difficulties. In this paper, we propose to use Physics and Geostatistics—Aware Information-Theoretic Machine Learning (PGA-ITML) to solve the above set of problems, firstly performing a stochastic upscaling that reduces the cardinality and the probabilistic dimension of the original reservoir models and then solving the upscaled model by traditional methods. The upscaling naturally obeys existing geostatistical properties and physical laws, which is more effective than the posterior verification as done by PINNs. Solving the surrogates by direct methods adds the advantage of more reliable solutions not guaranteed by DNNs solutions in the Small-Data regime. Results that contemplate time-varying coefficients will be described in further publications.

Keywords: Stochastic Reservoir Simulation, DNN, PINN, ITML, upscaling.

Introduction

Deep Neural Network (DNN) models have shown remarkable results when applied to high dimensional (hundreds of phase states), Chen et al [1], and Big-Data (hundreds of thousand steps) problems, those that operate far beyond their interpolation threshold, where they attain excellent generalization results and good enough accuracy when overparameterized and calibrated cross-entropy, Bornschein [2]. Stochastic scientific problems, however, should operate with about two hundred elements of the supervised set. Due to those well-known severe limitations, DNNs do not solve with good accuracy, do not generalize well, and sometimes make unexpected incorrect results, He and Jiang [3], Bornschein et al. [2], Wang et al. [4].

The Monte Carlo (MC) method is the most prominent technique to perform (UA) from which the statistical measures of the model response can be done through random sampling. Usually, a large number of geostatistical realizations of the stochastic input should be collected to guarantee accuracy and the model must be solved to estimate the statistics of the solution.

The usual discretization mesh is the simulation model obtained from the geocellular model by successive upscaling. Laval et al. [5] upscaled a geocellular model for a giant oil field with about 350 million cells to a simulation model with 900 thousand cells. There could be substantial errors in the process of upscaling. Milad et al. [6] encountered a cumulative production error of 74.22% in the history matching when upscaling the horizontal permeability from a geocellular model with 13 million cells to a simulation model where horizontal cell dimensions were reduced by 16 times to obtain an upscaled simulation model with 87 thousand cells in vertical upscaling introduced fewer errors, about 10%, in the history matching. Simulation models are too fine to be routinely used with the MC method for UA of hundreds or thousands of realizations because of computational cost, so it is strictly necessary to do another upscaling to a much coarser mesh, introducing additional errors in the solution. Higher production history matching errors associated with natural fracture size have been observed too. Computational time reduction errors are in simple exponential proportional to the upscaling rate, Milad et al. [6].

The high computational overhead of the MC method is usually overcome with surrogate modeling methods. The methods of nonlinear system identification like regression techniques long used in short-time deterministic reservoir simulation suffer significant drawbacks in the MC environment due to their incapacity to do safe medium and long-term forecasts. Unfortunately, the same can be said concerning pure regressions with DNN applications. Model upscaling is, currently, the unique method that can make safe and accurate predictions of medium- to long-term reservoir productions. This falling can be attributed to the absence of geostatistics and physical laws in the regression techniques.

DNN and other non-physical techniques have been putting together to produce efficient surrogate models for UA. Physical laws and practical engineering theories have been incorporated into PINN. Good results, however, are only obtained for constant mean permeability, as in Wang et al. [4], or deterministic local homogenization in large subregions within a concept of extended PINN (PINNX), Alhubail et al. [7]. Advances in DNN are related to the progress in computer hardware such as Graphical Processing Units (GPUs) that speed up calculations. Common surrogates are Polynomial Chaos expansion (PCE), Gauss Process (GP), Probabilistic Collocation Methods, and the classical Kosambi-Karhunen-Loève Expansion (KKLE) in combination with PINNs, Wang et al. [4]. Those transformations and expansions aim to truncate the input random field, preserve smoothing components, prepare for submission to PINN, and reduce the probabilistic dimension.

This work starts with a given microscale model X, either a geocellular or a simulation model, defined over a micromesh G_f, providing a cardinal reduction with macromesh $G_c \subset G_f$, and constructs a macroscale model Y, in a process of global probabilistic homogenization that takes into account all the microscale X. The microscale model $X = \{G_f, \lambda_f, \overline{u}_f\}$ where λ_f is the geostatistical random field, and $\overline{u}_f$ means the deterministically prescribed initial and Newman (rates of injection wells) and Dirichlet boundary conditions, and the macroscale model $Y = \{G_c, \lambda_c, \overline{u}_c\}$ where λ_c is the homogenized random field and $\overline{u}_c$ the equivalent deterministic initial and boundary conditions.

Instead of using a DNN approach in the form of PINN or PINNX, the physical laws and the geostatistics relations are imposed at first, on the verge of the generation of the reduced field λ_c so that, contrary to PINNX, the internal boundary conditions are automatically satisfied. The ML training is done with a labeled set $D_f = \{x_i, u_{fi}\}_{i=1}^{N_R}$ for each realization N_R where u_{fi} (rates of production wells, velocities, saturations) is the solution to each problem x_i.

The macroscale problem Y is obtained with Information-Theoretic ML (ITML). The $n_R < N_R$ problems can be solved directly; if high accuracy is desired, it is not recommended, yet. At this stage, the classical history matching should be performed but now with the upscaled model Y.

Original research in the direction of this work has been started with Koutsourelakis [7]. Some of the relevant works in that direction are Grigo and Koutsourelakis [8], Boso and Tartakovsky [9], and Costa et al. [10].

Methodology

Let the reservoir be described by a set of N_R microscale geostatistical realizations of a random field collectively aggregated into an ensemble X, identifying microscale meshes with N cells. In general, X contains more information than necessary for the accurate solution of a given reservoir simulation problem. The goal is to reduce this ensemble into an ensemble Y over $n < N$ coarse-scale meshes and possibly $n_R < N_R$ macroscale realizations. If that reduction is possible, the answer to the specific problem can be encountered with less computational effort and reduced accuracy. Ensemble X can be as detailed as a cellular mesh, containing microscale information that can be compressed into a desired macroscale for the specific objective.

If the properties described by X can be compressed into Y for the solution of a specific problem, what needs to be determined is what information in X must be statistically preserved in such a way that distortion in the desired response statistic can be controlled. At least two measures therefore need to be established. The first is the measure of the fidelity of the representation of X by Y, and the other is the statistical accuracy of the desired response. The physics at the microscale and that of the macroscale need not be the same.

The temporal scale will not be explicitly addressed in this text, without losing generality. This will be done in a future text when the methodology is applied to history matching as a component of stochastic control of reservoir management and development.

Mutual information

Replacing the X field with a smaller Y field implies the possibility of assigning the same value to large regions of the domain. A mutual information indicator measures the average information that knowledge of Y can provide about X and vice versa.

The hypothesis that the large regions of the heterogeneous domain of interest can be assigned the same values makes it possible to describe the random field X by the smaller field Y. A measure of the ability of Y to approximate X can be given by the mutual information indicator given by,

$$I(X,Y) = \iint p_{XY}(x,y) log \frac{p_{XY}(x,y)}{p_X(x)p_y(y)} dxdy, \tag{1}$$

where $p_{XY}(x,y)$ is the joint density of (X,Y), and $p_X(x)$ and $p_Y(y)$ are the respective marginal distributions. Its minimum value is zero when the two fields are independent, that is, $p(X,Y) = p_X(x)p_Y(y)$. Its maximum value is known as the entropy of X, which occurs when the two fields are identical $X \equiv Y$. Defining joint entropy as the indicator itself, and developing,

$$H(X,Y) = - \int p_X(x) \log log\, p_X(x)dx \ + \iint p_{XY}(x,y) \log log \frac{p_{XY}(x,y)}{p_X(x)p_y(y)} dxdy\,. \tag{2}$$

The first term in the second member of eq. (2) is the entropy of X,

$$H(X) = - \int p_X(x) log\, p_X(x)dx, \tag{3}$$

while the second is the conditional entropy.

In information theory, $I(X,Y)$ is also known as the information transmission rate R of a continuous channel.

Interscale response fidelity

The measure of fidelity is known as the measure of distortion of the response provided by the two fields. This measure can be scalar, vector, or matrix. Without loss of generality, it will be assumed that the measure is a scalar since the other measures must always be transformed into one or a set of scalar measures. For example, let the distance be given by,

$$d(X,Y) = (r(X) - r(Y))^2, \tag{4}$$

where $r(X)$ is the response due to the microscale and $r(Y)$ to the macroscale, usually functions implicitly defined by a numerical simulator. For each given stance defined as in eq. (4), there will be a need for its gradient relative to Y.There are no statistics in eq. (5), that is, the random functions $r(X)$: $R^N \rightarrow R$, e $r(Y)$: $R^N \rightarrow R$ must be understood as an ensemble, or set, of samples, or realizations, $r(X_i)$, $i = 1, ..., n_R$. In this sense, $d(X,Y)$: $R^N \times R^N \rightarrow R^+$ is a collection of $N_R \times n_R$ samples in one dimension. Furthermore, it will be admitted, without loss of generality, that the microscale is represented only by a random variable Y. This means that the macromesh has a single element, cell or block, with a single property that, essentially, is the best stochastic representation. The distance, in this case, will be in the application of $d(X,Y)$: $R^N \times R \rightarrow R^+$.

It is important to note, at this stage, that no reference is made to the physics that prevails in the geometric domains. It is possible, for example, to evaluate $r(X)$ in complex carbonate physics, using Darcy-Stokes or Stokes-Brinkman with Darcyan fractures, or even in compositional models, and $r(Y)$ with Black-Oil physics.

The objective is twofold. First, find a compression scheme from X to Y that leads to a minimum distortion of the desired response, and then determines its value. Determining the optimal transformation involves evaluating the posterior distribution of Y, $p_Y(y)$. For the stochastic scheme to be complete, it is necessary to evaluate the distortion in a probabilistic way. One possibility is that the distortion, D, be the distance's mathematical expectation, but more meaningful functions for certain applications, such as percentiles and risk measures on the density's tails, can be chosen,

$$D = E[d(X, Y)]. \tag{5}$$

By hypothesis, the field describing the microscale is given by the ensemble of N_R geostatistical realizations. Thus, evaluation of the distortion, its expectation, or any other convenient stochastic measure, must be done with the Monte Carlo technique, or some simplifying variant. Since X is already sampled it is possible to write, using the relation between the joint density and the conditional density,

$$D = E[d(X, Y)] = \sum_{X,Y} p_{XY}(X, Y) d(X, Y) = \sum_{X,Y} p_X(X) p_{XY}(Y|X) d(X, Y), \tag{6}$$

where $p_{XY}(Y|X)$ is the conditional density of Y. Its minimum, with respect to $p_{XY}(Y|X)$, will be a Dirac's delta function that assigns to each x of X the Y that minimizes $d(X, Y)$.

This is an optimization problem that requires more computational effort than desired. Instead, we try to solve the augmented Lagrangian with the restriction that the indicator $I(X, Y)$ is less than or equal to a pre-specified value $I(X, Y)$. The constrained minimization problem is used to find using the Deterministic Simulating Annealing algorithm in a parallel form (see Costa [3,4] for details). The algorithm carries a high computational cost, but it finds the $p_{XY}(Y|X)$ in a nonparametric form, does not require any assumption about its functional form. The distribution $p_{XY}(Y|X)$ has the form of a Gibbs distribution.

The ensemble Y, and the probability of each one of the n_R coarse-scale realizations can solve any uncertainty reservoir simulation problem in inexpensive, accurate, and correct predictions. The macro ensemble model achieves an exceptional generalization for Neumann boundary conditions (different number and disposition of wells and their rates).

Applications

The above methodology is applied to a two-dimensional reservoir with a two-phase flow of oil and water and depletion by water injection.

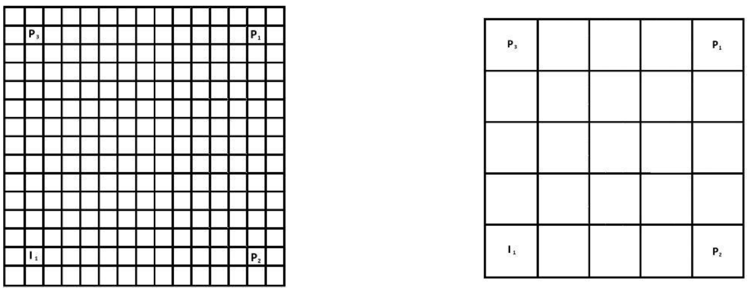

Figure 1 - Micromesh (15x15) and macromesh (5x5)

The reservoir has a 15×15 cells upscaled to 15×15 cells as shown in Figure 1. The permeability realizations in the X, Y, and XY directions were estimated from a $Lognormal(245.5896,\ 304.3432)$, and the correlation matrix is given by:

$$\sigma := \begin{bmatrix} 1.50 & -0.595 & 0.85 \\ -0.5 & 2.00 & 0.65 \\ 0.5 & -0.80 & 2.00 \end{bmatrix}$$

The porosity is constant and equal to 0.25, the viscosity of water is $1\ cP$ and that of oil is $110\ cP$, the gravity of water is $1000\ kg/m^3$ and that of oil is $700\ kg/m^3$. The initial pressure in the reservoir is $1\ bar$ and the fluid is incompressible. There is one injection well, and four production wells, one in each corner. The injection well operates with a rate of $0.5\ m^3/day$. The wells operate at bottom-hole pressure of $1\ bar$. The production period is $600\ days$, with measurements taken every $120\ days$.

A set of $N_R = 1000$ micro realizations are given. The procedure starts with simulating 1000 realizations of the fine grid. These are the only simulations of the fine grid. 25 upscaling operations were performed, one for each macrocell, with Dirichlet boundary conditions. The final number of microscale realizations used was $n_R = 10$. The computational cost was significantly reduced.

The problem is determining the new field of absolute permeabilities for the coarse grid given in Figure 1. Each set of 3×3 cells are upscaled to one microcell.

The reservoir simulations used in this example were performed with MRST-2017, and the upscaling was done with the values taken from the simulations in a single-phase (water) problem when the velocities were in a steady-state regime. The distance between the Darcy velocities fields were minimized in Eq. (4).

Figure 3 shows the cumulative oil and water of the original given models and the upscaled models. To compare the impact of the proposed methodology, the flow of water and oil from the wells was analyzed at both scales.

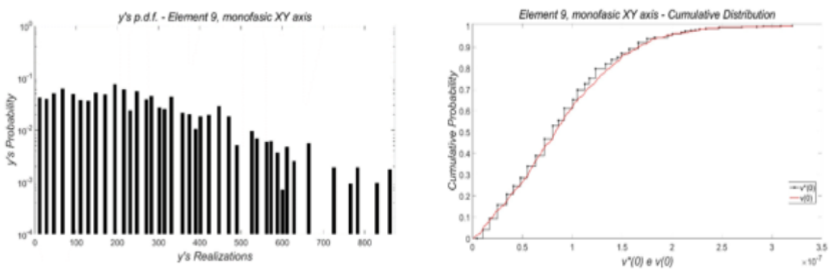

Figure 2: Quantized probability density and corresponding cumulative distribution of one macrocell.

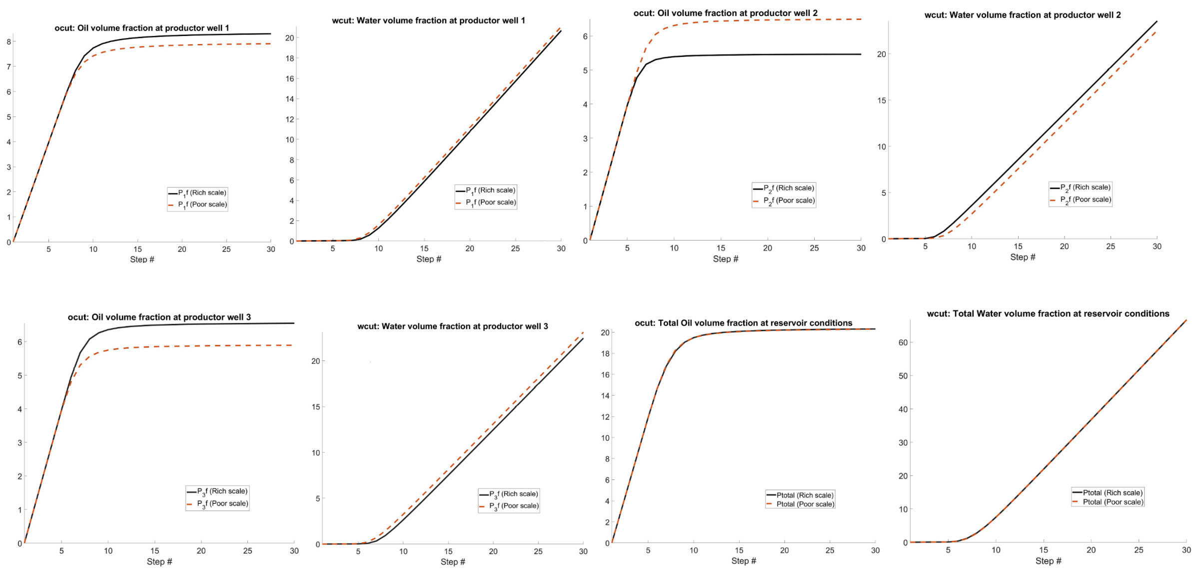

Figure 3 – Cumulative oil and water of the production wells.

It can be observed that the total cumulative macroscale production fits the microscale solution, while there are some errors in the individual well productions. Those productions can have better approximations if they are chosen to be minimized in Eq. (4). Figures 4 and 5 show the water saturations in the two scales. The upscaled results are in good agreement with microscale results.

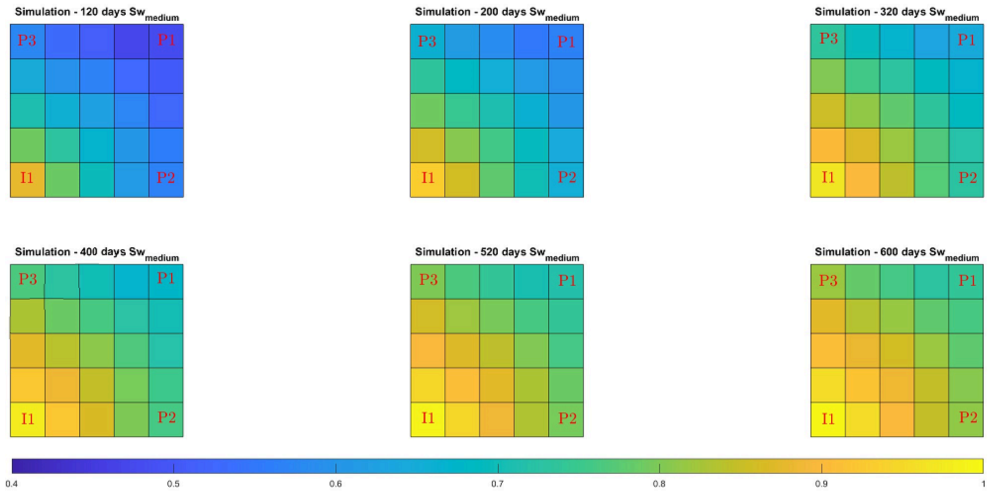

Figure 4 - Average water saturations in the reservoir in the coarse grid for time = 120, 200, 320, 400, 520, and 600.

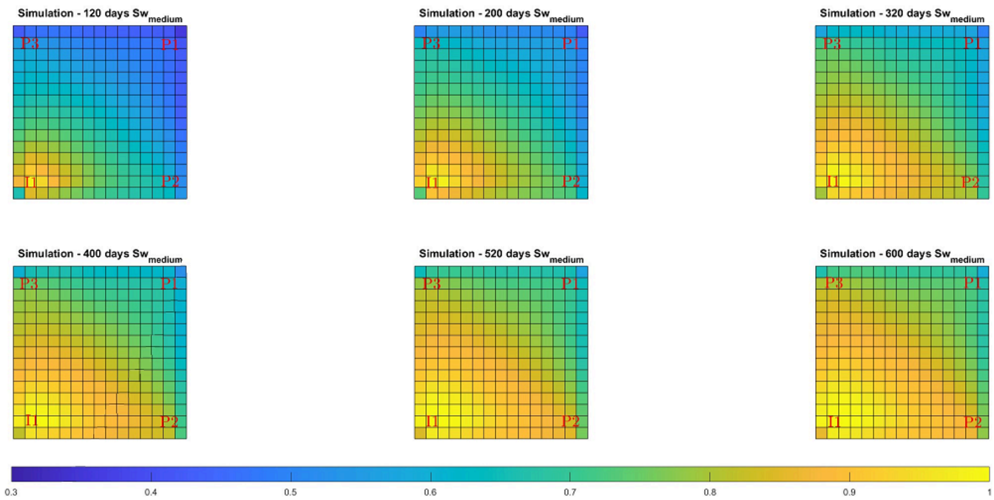

Figure 5 - Average water saturations in the reservoir in the fine grid for time = 120, 200, 320, 400, 520, and 600.

Conclusions

The stochastic upscaling works well as shown in the two-dimensional application above and in similar small examples not shown in this paper. It remains to solve the problem of downscale to capture fluid velocities, saturations, and pressures in the microcells. Discrepancies in the well's production suggest a careful choice for quantities to be minimized. The notion of goal-oriented upscaling should be introduced.

Acknowledgements

The Authors acknowledge the Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP) and Financiadora de Estudos e Projetos (FINEP), by the financial support provided by the Programa de Recursos Humanos da ANP – PRH-47.1 supported by investments by petroleum firms qualified in Cláusula de P, D&I of Resolução ANP nº 50/2015, with support to scholarships and complementary funds for realizations of research and professional formation in Geology and Civil Engineering for the Petroleum and Gas Industries.

Referências Bibliográficas

[1] J. Chen, S. X. Xi, Weinan E., Z. Young. “Bridging Traditional and Machine Learning-based Algorithm for Solving PDEs: The Randon Feature Method”, Journal of Machine Learning, 2022.
[2] J. Bornschein, F, Visin, S. Osindero. “Small Data, Big Decisions: Model Selection in the Small-Data Regime”, ArXiv, 2020.
[3] W. He and Z. Jiang. “ A Comprehensive Survey on Uncertainty Quantification for Deep Learning, ArXiv, 2024.
[4] N. Wang, H. Chang, D. Zhang. “Efficient Uncertainty Quantification for dynamic subsurface flow with surrogated by theory-guided Neural Network”, Comp Meth. In Applied Mech. and Eng., 373, 2021
[5] A. Laval, G. Auxiette, I. Echeverria. “3D Geocellular Modeling in a Complex Structural Setting: Furrial Giant Oil Field Venezuela”, SPE Latin American and Caribbean Petroleum Engineering Conference, 2003.
[6] B. Milad, S. Gosh, R. Slatt, K. Marfurt, M. Fahes. “Practical Aspects of Upscaling Geocellular Geological Models for Reservoir Fluid Flow Simulations: A Case Study in Integrating Geology, Geophysics, and Petroleum Engineering Multiscale Data from the Hunton Group”, Energies, 13, 1604, 2020.
[7] A. Alhubal, X. He, M. AlSinan, H. Kwajm, H. Hoteit. “Extended Physics-Informed Neural Networks for Solving Fluid Flow Problems in High Heterogenous Media”, 22IPTC-22163-MS, 2022.
[8] C. Grigo and PS. Koutsourelakis. “A physic-aware probabilistic machine learning framework for coarse-grained high-dimensional systems in the Small Data regime”, J. of Computational Physics, 397, 2019.
[9] Francesca Boso and D. M. Tartakovsky. “Information-Theoretic Approach to Bidirectional Scaling”, Water Resources Research, 54, 2018.
[10] I. F. Costa, L. A. Fonseca, E. R, Araújo. “Heterogeneous Stochastic Multiscale,” CILAMCE2017, XXXVIII Ibero-Latin American Congresso on Computational Methods in Engineering held in Florianópolis, Santa Catarina, Brasil, from November 5th to 8th, 2017.