UNISIM ON-LINE
Year 18, Number 1 | 152nd edition | May, 2023

"The main challenge inherent in proxy models is the inclusion of 3D geostatistical realizations (block-to-block grid properties such as porosity and permeability) as uncertain attributes in the proxy construction."

Specials interests:

- UNISIM
- UNISIM Publications
- Reservoir Simulation and Management Portal
- Previous Issues

Links:

- UNICAMP
- Cepetro
- Petroleum Engineering Division
- School of Mechanical Engineering
- Petroleum Sciences and Engineering

Graduate:

Petroleum Sciences and Engineering: interested in Masters and PhD in the Simulation and Oil Reservoir Management area click here.

Data Assimilation Using Principal Component Analysis and Artificial Neural Network

Célio Maschio

Introduction

Data assimilation (DA) for uncertainty reduction using reservoir simulation models normally demands high computational time; it may take days or even weeks to run a single reservoir application, depending on the reservoir model characteristics. Therefore, it is important to accelerate the process to make it more feasible for practical studies, especially those requiring many simulation runs. One possible alternative is to use proxy models to replace the reservoir simulator in some time-consuming parts of the procedure. However, the main challenge inherent in proxy models is the inclusion of 3D geostatistical realizations (block-to-block grid properties such as porosity and permeability) as uncertain attributes in the proxy construction. In most cases, it is impossible to treat the values of all grid properties explicitly as input of the proxy building process due to the high dimensionality issue.

This text presents a summary of the paper published by Maschio et al. (2023), which proposes a new methodology for data assimilation combining principal component analysis (PCA) with artificial neural networks (ANN) to solve this problem. The PCA technique is applied to reduce the dimension of the problem, making it possible and feasible to use grid properties in proxy modelling. The trained ANN is used as a proxy to the reservoir simulator, with the final goal of reducing the total computational time spent in the application.

Methodology

The general flowchart of the methodology is shown in Figure 1, and the main steps are described in the following:

1) Using a set of geostatistical realizations previously generated using geological/geostatistical modeling software, apply the PCA technique for dimensional reduction and parameterization.
2) Validate the PCA parameterization running the data assimilation process with reduced dimensions and comparing with a reference solution from the conventional process.
3) Once the PCA parametrization is validated, a training data set is created to train the ANN, which is composed of the PCA variables besides other scalar attributes such oil water contact, relative permeability exponents, etc.
4) Train and validate the artificial neural networks. A batch mode training process is applied in this step.
5) Apply a DA method using the trained ANN to replace the simulator.
6) Validate the results with the reservoir simulator.

To validate the methodology, we run three data assimilation processes (DA1, DA2 and DA3), described as follows:

Data assimilation 1 (DA1): DA1 is the conventional process in which the DA method (in our case, the ES-MDA, Emerick and Reynolds, 2013) updates the entire grid properties (all properties values, such as porosity and permeability, for example) during the data assimilation. The procedure starts with the simulation of the prior numerical models (ensemble of models generated from the prior uncertainties, including the grid properties and the other scalar attributes), followed by an analysis step. From this point, there is an iterative process in which the ES-MDA updates the attributes according to the maximum number of iterations defined by the user. The results from the process DA1 are used as a reference solution to processes DA2 and DA3.

Data assimilation 2 (DA2): before applying the image parametrization via PCA in the ANN training process (and further in DA, replacing the simulator with the ANN), it is necessary to confirm if the data assimilation works properly with the reduced-dimension problem. To do so, we run the process DA2. In DA2, instead of updating the entire grid property values, the ES-MDA updates the PCA coefficients. It is noteworthy that DA2 is an intermediate validation step, i.e., it is only run to confirm and validate the image parametrization via PCA.

Data assimilation 3 (DA3): the objective of the DA3 is to speed up the DA process, the final goal of the methodology. Once the PCA parameterization is validated with the DA2, we create the training data set, composed of the PCA coefficients plus the other scalar attributes. The DA3 parametrization is exactly the same as DA2; that is, the uncertain attributes are composed of the PCA coefficients plus the other scalar attributes. The key difference in relation to DA1 and DA2 is that, in DA3, the trained ANN replaces the reservoir simulator, while in DA1 and DA2, the reservoir simulator is used in the entire process. In DA3, it was necessary to couple the trained ANN to the data assimilation algorithm (ES-MDA) in such a way that, instead of calling the reservoir simulator, the ES-MDA calls the trained ANN. After the data assimilation in DA3, the results are validated submitting the posterior ensemble models to the reservoir simulator.

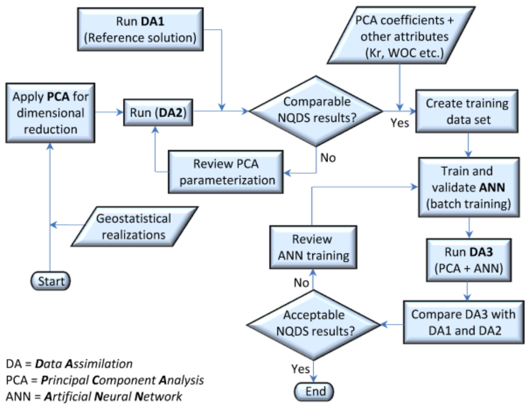

Figure 1*: General flowchart of the methodology.*

Application and results

The proposed methodology was applied in a Brazilian offshore field named S-Field, located in Campos Basin. There are 2,359 days of history data from which 2,206 days (from 0 to 2,206, denominated history period) were used in the data assimilation, and the rest (from 2,207 to 2,359, denominated validation period) was used to validate the models regarding its predictive capacity. The prior uncertainties are composed of 200 geostatistical realizations plus 59 other scalar attributes. The PCA reduced the number of variables from 81,398 to 209 (150 PCA coefficient plus other 59 scalar attributes).

Figure 2 shows the history matching quality (NQDS, Normalized Quadratic Deviation with Sign) for oil rate and bottom-hole pressure for the producer wells, comparing the performance of the three processes (DA1, DA2 and DA3). Analyzing the process DA2, it is possible to note, in general, good similarity when compared with DA1 (reference solution). The results are similar for both prior and posterior ensembles. First, this means that the simulation results using the original prior images (DA1 prior) and using the prior images generated via PCA (DA2 prior) are consistent. Second, the similarity between the posterior results (after the data assimilation) means that the DA process using the images generated via PCA (DA2) was successfully applied, i.e., the image parameterization using PCA (dimensional reduction) was validated.

Analyzing the process DA3, although some specific differences have occurred, the DA3 provided overall good results. For most local OF, we can see that it converges to NQDS distributions close to the reference solution (DA1).

Figure 3 shows the oil rate curves for one producer well (P6), comparing DA2 and DA3. In both processes, the data assimilation reduces the variability of the prior curves, but maintains a range of models encompassing the observed data.

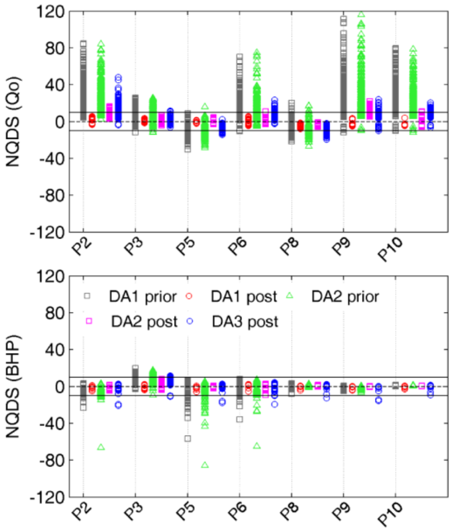

***Figure 2**: NQDS plots comparing the history matching quality of three DA processes.*

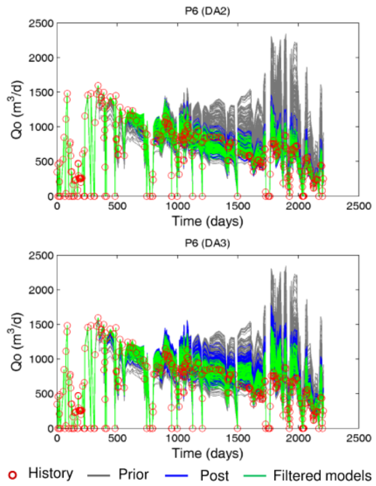

***Figure 3**: Oil rates curves for one producer well (P6) comparing DA2 and DA3.*

Figure 4 shows the NQDS computed in the validation period. We can see an effective uncertainty reduction comparing the prior and posterior ensembles for the three DA processes. These plots also show that the DA1 tends to reduce the uncertainty more than DA2 and DA3 in some wells.

It is worth mentioning that from DA1 to DA3, two major simplifications were introduced in the process: the strong dimensional reduction (by applying the PCA) and the substitution of the reservoir simulation by a proxy model (trained ANN). Thus, a small loss of accuracy would be naturally expected. Nevertheless, the global results from the proposed approach (DA3) are consistent.

Conclusions and final remarks

- The results from DA2 (using reduced dimensions) were similar to the results from DA1 (using full dimensions), meaning that the PCA parameterization applied in the data assimilation process worked properly, that is, the validation purpose was successfully accomplished.

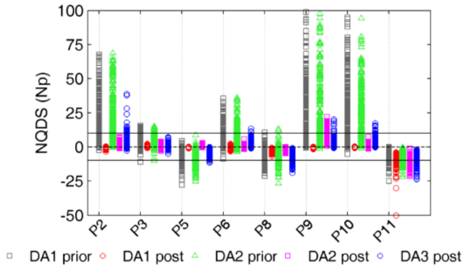

***Figure 4**: NQDS plot comparing the validation period (Np) of three DA processes.*

- The proposed methodology using dimensional reduction (applying PCA) allowed the effective inclusion of 3D spatial properties in the construction and application of ANN-based proxies in the data assimilation process, contributing to a practical approach and covering a gap in the literature, especially in real-field applications.
- The process DA3 (using the training ANN) enabled 76.4% of computational time reduction, providing suitable results.
- The main contribution of this work is the significant reduction in computational time of the data assimilation process, preserving the global quality of the results.

More details and results can be found in the full version of the paper.

Acknowledgments

This work was conducted with the support of Projects (1) "MFM - Developing a Multi-Fidelity Modelling Approach for Uncertainty Reduction & Improved Production Forecasting" (UNICAMP/Shell Brazil/ANP) funded by Shell Brazil, under the ANP R&D levy as "Compromisso de Investimentos com Pesquisa e Desenvolvimento", (2) "Development of integration between reservoir simulation and seismic 4D - Phase 2", also funded by Shell Brazil, for providing us with the S-Field model (MFM), and (3) Energi Simulation Chair program. We also thank UNISIM, FEM-UNICAMP, CEPETRO, and CNPq for supporting this work and CMG, Emerson and Schlumberger for the software licenses. The authors would also like to thank Alexandre A. Emerick (from Petrobras) for providing the EHM tool to UNISIM.

References

Emerick, A. A., Reynolds, A. C.; 2013. "Ensemble smoother with multiple data assimilation", Computers & Geosciences, 55, 3-15.

Maschio, C., Avansi, G. D., Schiozer, D. J.; 2023. "Data Assimilation Using Principal Component Analysis and Artificial Neural Network". SPE Reservoir Evaluation & Engineering. v. 26, pp. 1-18, 2023. https://dx.doi.org/10.2118/214688-PA

About the author:

Célio Maschio is graduated in Mechanical Engineering from UNESP, holds a MS and a PhD degree in Mechanical Engineering from UNICAMP. He is a researcher at UNISIM/CEPETRO/UNICAMP since 2001 developing research and leading projects focused on data assimilation for uncertainty reduction.

"The main contribution of this work is the significant reduction in computational time of the data assimilation process, preserving the global quality of the results."

UNISIM opportunities:

If you are interested in working or developing research in the UNISIM Group, please contact us.
For further information, click here.

Research in Reservoir Simulation and Management Group

Petroleum Engineering Division - Energy Department
School Of Mechanical Engineering
Center for Energy & Petroleum Studies
University of Campinas
Campinas - SP

Phone.: 55-19-3521-1220

unisim@cepetro.unicamp.br