Rio Oil & Gas Expo and Conference 2020

ISSN 2525-7579

Conference Proceedings homepage: https://biblioteca.ibp.org.br/riooilegas/en/

Technical Paper

Digital subsurface transformation: challenges and perspectives towards an AI-assisted G&G workflow

Félix Gonçalves [1]
Francois Lafferriere [2]
Mathieu Ducros [3].

1. KOGNITUS, COO, . RIO DE JANEIRO - RJ - BRASIL, felix@kognitus.com.br
2. KOGNITUS, CEO, . RIO DE JANEIRO - RJ - BRASIL, francois@kognitus.com.br
3. KOGNITUS, R&D, . RIO DE JANEIRO - RJ - BRASIL, mathieu@kognitus.com.br

Abstract

The O&G industry is being disrupted at multiple levels by changes in the global energy landscape, business and operating models, and geopolitical order. Such a confluence of elements occurs at the moment when a technological revolution is also underway. This work focuses on discussing what Artificial Intelligence (AI) means for the subsurface data analysis and G&G disciplines. Despite all the enthusiasm around the topic and the proliferation of AI pilot projects, an effective insertion of these technologies in the mainstream of G&G workflow still poses significant challenges. Some of them are related to the unique characteristics of subsurface data, such as marked data sparsity, a high degree of uncertainty, strong spatial dependence, and complicated physics background. Other challenges are more related to the business and operational context of the O&G industry, such as domain silos in data infrastructure, difficulties to assemble teams with the required mix of skills, obstacles to deploy ML and DL solutions in a timely and reliable way and availability of domain-specific AI platforms. This work will demonstrate some examples of how a cloud-native AI platform for G&G can be employed to automate the analysis of large-volume datasets at scale and to solve existing subsurface problems such as prediction of missing well log curves, seismic inversion, reservoir properties prediction, seismic interpretation and seismic data compression.

Keywords: E&P. Subsurface. Artificial Intelligence. Geology. Geophysics

Received: March 15, 2020 | **Accepted:** Jun 06, 2020 | **Available online:** Dec 01, 2020
Article Code: 455
Cite as: Rio Oil & Gas Expo and Conference, Rio de Janeiro, RJ, Brazil, 2020 (20)
DOI: https://doi.org/10.48072/2525-7579.rog.2020.455

1. Introduction

The O&G industry is being disrupted at multiple levels by changes in the global energy landscape, business and operating models, and geopolitical order. Such a confluence of elements occurs at the moment when a technological revolution is also underway. Digital technologies now comprise an integral part of our day to day lives and become a critical success factor for businesses across all industries.

For the O&G companies, embracing digital transformation through the widespread adoption of technologies such as Big Data, Cloud Computing, Artificial Intelligence (AI), Internet of Things and Blockchain, has enormous potential for enhancing operational efficiency and capturing new insights that can unlock higher value.

This work focuses on discussing what AI – a critical enabling technology of the digital transformation revolution – means for the subsurface data analysis and the geologic and geophysical disciplines. The authors highlight the foremost technical and managerial challenges to ensure a pervasive adoption and a successful implementation of AI-assisted methods. The work also provides case studies in areas in which it is already possible to see the real value of AI to better harness legacy data and transform the geologic and geophysical (G&G) workflow.

2. Challenges and perspectives of AI in subsurface

In the subsurface domain, when AI is mentioned, geologists and geophysicists are usually referring to machine learning (ML) and deep learning (DL) techniques. Some of the methods that cause great excitement among subsurface professionals nowadays are not actually new. What has changed over the last decade is the emergence and commoditization of technologies and architectures that made data access and processing much easier and faster.

Despite all the enthusiasm around the topic and the proliferation of initiatives for the application of ML and DL technologies in practical use cases, an effective insertion of these technologies in the mainstream of G&G workflow still poses significant challenges.

2.1. Challenges related to the nature of the subsurface domain

Some of the challenges are related to the very nature of subsurface data. When compared to more conventional business and technical domains, subsurface presents several unique characteristics, such as marked data sparsity, a high degree of uncertainty, strong spatial dependence, and complicated physics background.

Regarding data scarcity, one of the main challenges in applying ML/DL is to obtain enough data to train the algorithms. A few methods are conventionally applied to remedy this issue in classical ML/DL workflows, such as data augmentation or neural transfer style. More recently, a complementary solution is to use simulations to generate physically-plausible synthetic data.

Regarding the physical complexity, ML models that entirely disregard physical constraints can eventually produce unrealistic and non-interpretable results. By introducing physics-inspired features or using physics-constrained optimization functions in the models, it is possible to achieve physically-realistic insights and losing the benefits of the data-driven approach.

2.2. Challenges related to the context of the O&G industry

Other challenges are more related to the business, operational, and technology context of the O&G industry. It is crucial, for instance, to overcome domain silos that have traditionally prevailed in E&P data infrastructure. Overcoming such silos requires the implementation of cloud infrastructure and technologies that can meet companies' access control and data privacy requirements, and scale to support the performance needed by AI solutions.

The widespread adoption of AI technologies is also highly dependent on assembling teams with the right mix of skills (G&G, ML/DL, and software development). Unfortunately, the current combination of an aging workforce and a strong skills shortage faced by the O&G industry is a roadblock that can slow the incorporation of AI technologies into E&P workflows.

An additional challenge that O&G shares with all other industries is the difficulty of deploying ML and DL solutions in a timely and reliable way. Most of the oil companies seem stuck in a never-ending cycle of AI pilot projects that never progress to full deployment. Building a continuous deployment process will require a high degree of standardization, automation, and integration with existing systems and procedures.

Finally, pervasive incorporation of AI technologies into the E&P decision-making process will also rely on the availability of domain-specific data science platforms that can allow asset teams to collect and prepare data, build models, automate workflows and deploy ML/DL solutions without necessarily being able to code.

2.3. Perpectives on the application of AI in G&G workflows

This work shows some examples of the application of mechine learning techniques to solve existing subsurface problems such as (i) prediction of missing well log curves using RNN and LSTM algorithms, (ii) seismic inversion using CNN algorithms trained with synthetic data generated by Markov-Chain simulation, (iii) reservoir properties prediction using numerous ML algorithms, (iv) automated seismic interpretation using CNN algorithms, (v) seismic data compression using autoencoders, and (vi) petroleum system undertaities and risk assessment.

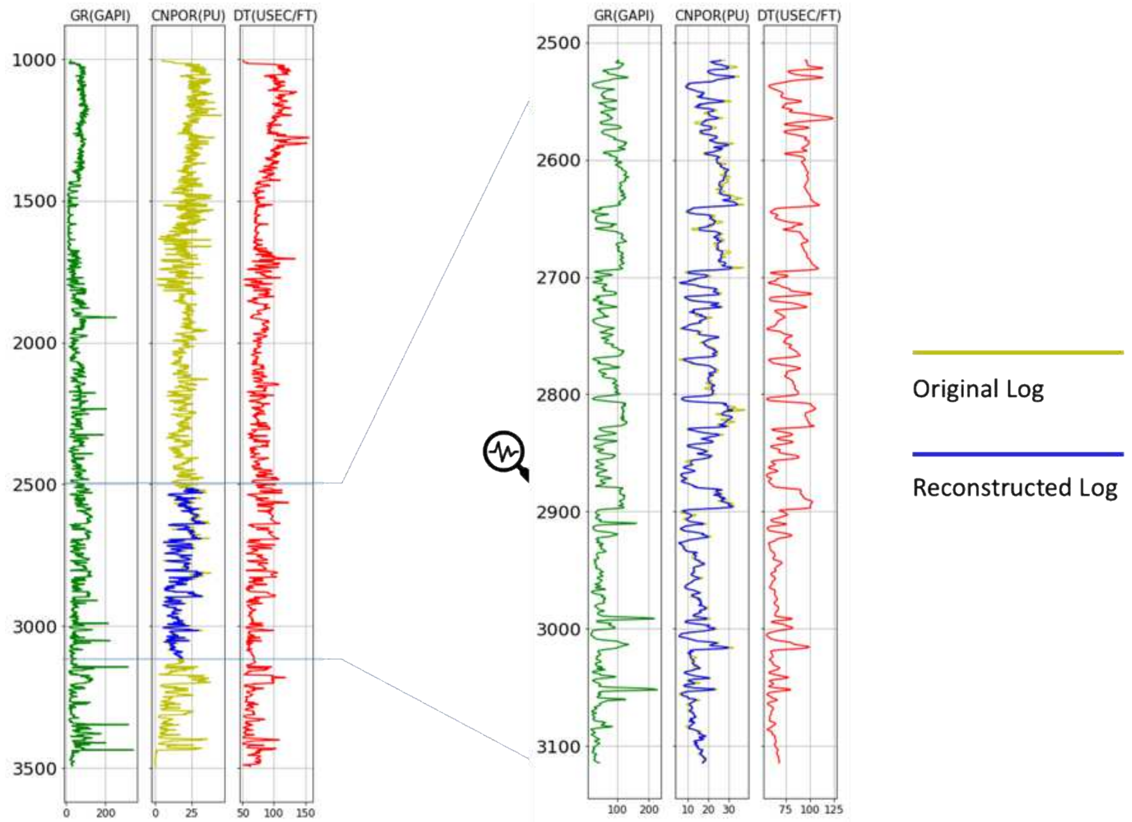
Prediction of missing well logs

Acquisition issues or economical restriction frequently leads to incomplete logset from well data. Still, reservoir characterization workflows would greatly benefit from having the entire logset as input for every well to get optimum results. Time/spatial-series algorithms, such as recurring neural network (RNN) or long short-term memory (LSTM) are well suited to account for the continuity of the geological signal, as aneeeded for the prediction of missing log curves. In a case study conducted using public data from North American basins, the network was first optimized, trained and tested on the available complete sets of logs. The trained network was then applied on the wells presenting missing curves. By using only one well data log it was possible to predict different log curves with high accuracy (Figure 1).

Seismic Elastic Inversion using ML & Pseudo-Wells Training

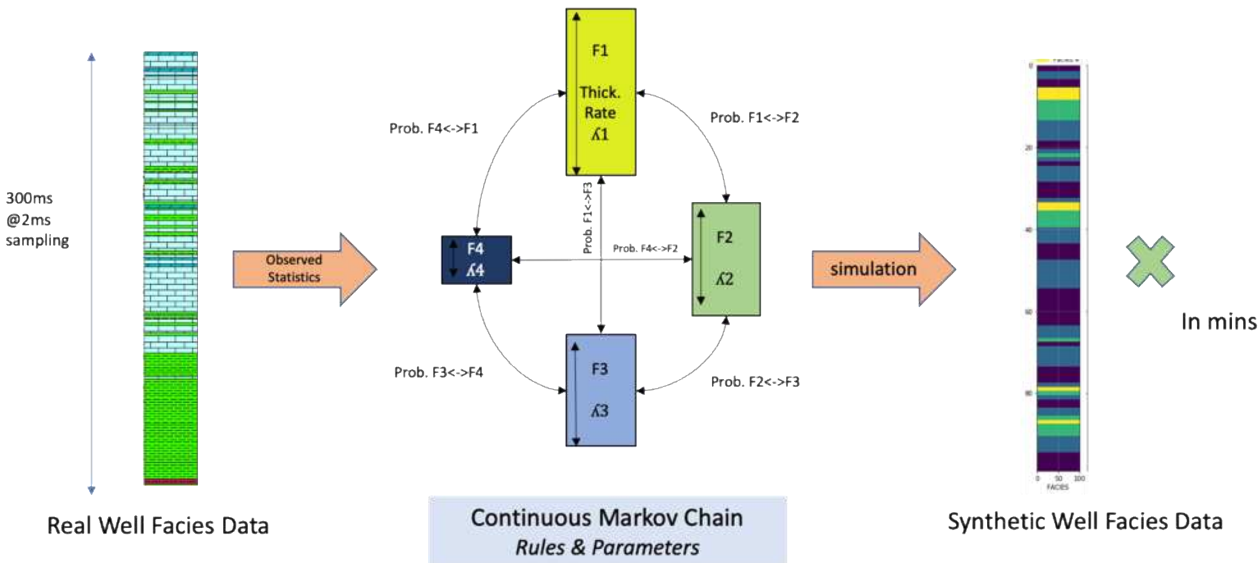
Seismic inversion is a typical problem which still did not benefit from all the recent developments in artificial intelligence as algorithms usually require to be fed with more than the available wells. To solve this issue, it was developed a workflow to quickly generate millions of realistic pseudo-wells (Zambrini et al., 2019). They are based on statistical information from the local geology and on Markov-Chain simulation (Figure 2). An associated synthetic seismic, including random noise, is generated (Figure 3) and the coupled synthetic-facies log is then used, in addition to the original dataset, to train the machine learning algorithms. For seismic inversion, ML algorithms applied to time-series, such as RNN and CNN-1D, were selected due to their capacity to invert directly the post-stack or angle-stack traces to the original facies, with a high quality of prediction. The method achieved an accuracy of more than 93% on a dataset of 40Hz peak frequency for the prediction of four facies in a study case on pre-salt data (Figure 4).

Figure 1 – Prediction of missing CNPOR log curve using RNN and LSTM algorithms.



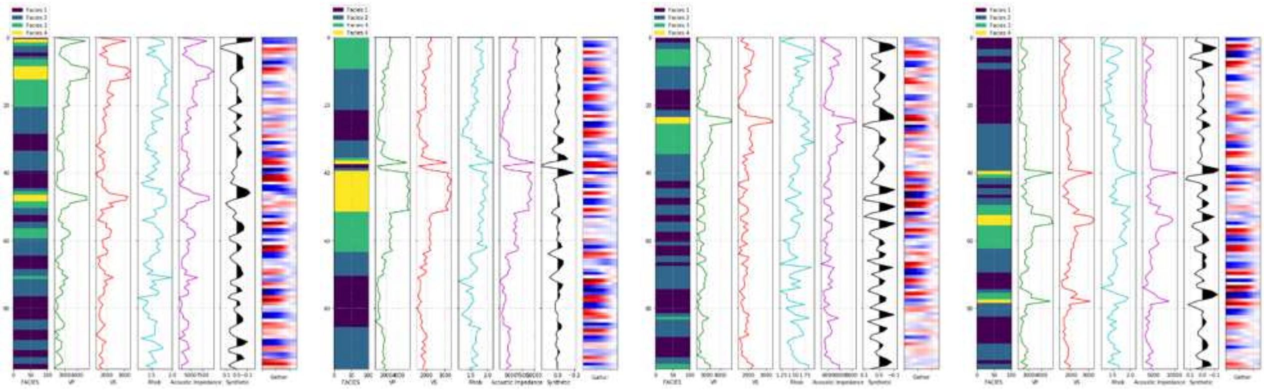
Source: from the authors.

Figure 2 – Markov-Chain simulation used to generated realistic pseudo-wells base on actual well data.



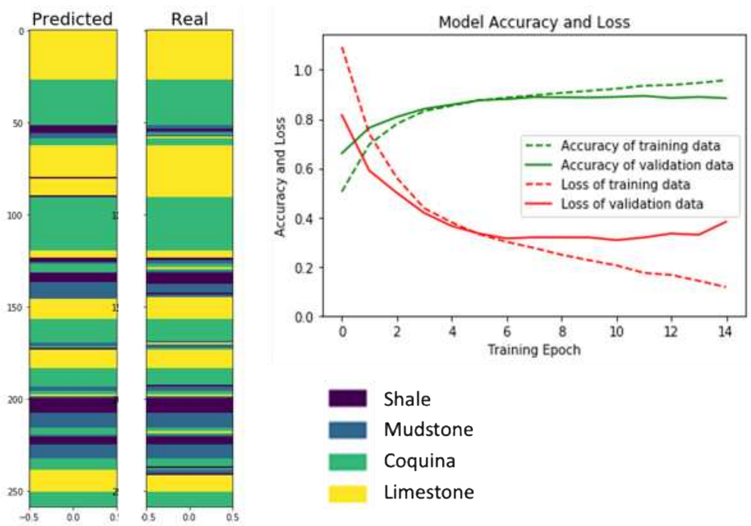
Source: Zambrini et al. (2019)

Figure 3 – Synthetic seismic (including random noise) generated from pseudo-well data. Coupled synthetic-facies log were used, in addition to the original dataset, to train the ML algorithms used for seismic inversion.



Source: Zambrini et al. (2019)

Figure 4 – Seismic inversion results in a study case fo the pre-salt section using RNN and CNN-1D algorithms.

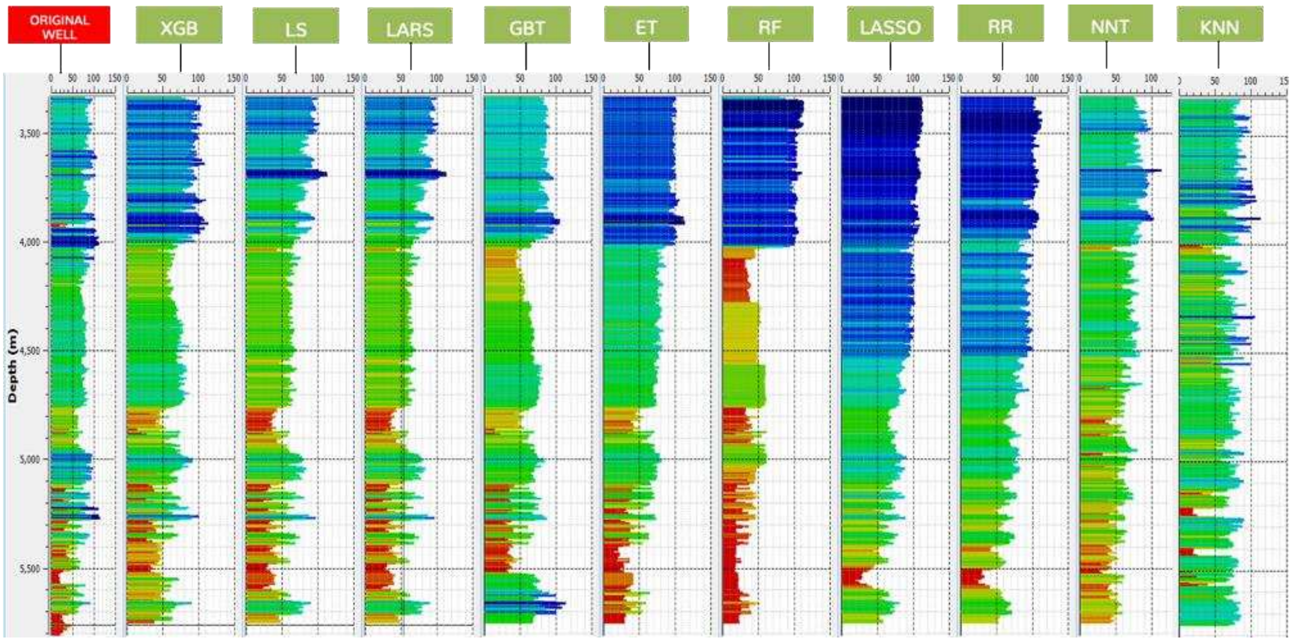


Source: Zambrini et al. (2019)

Reservoir Property modeling

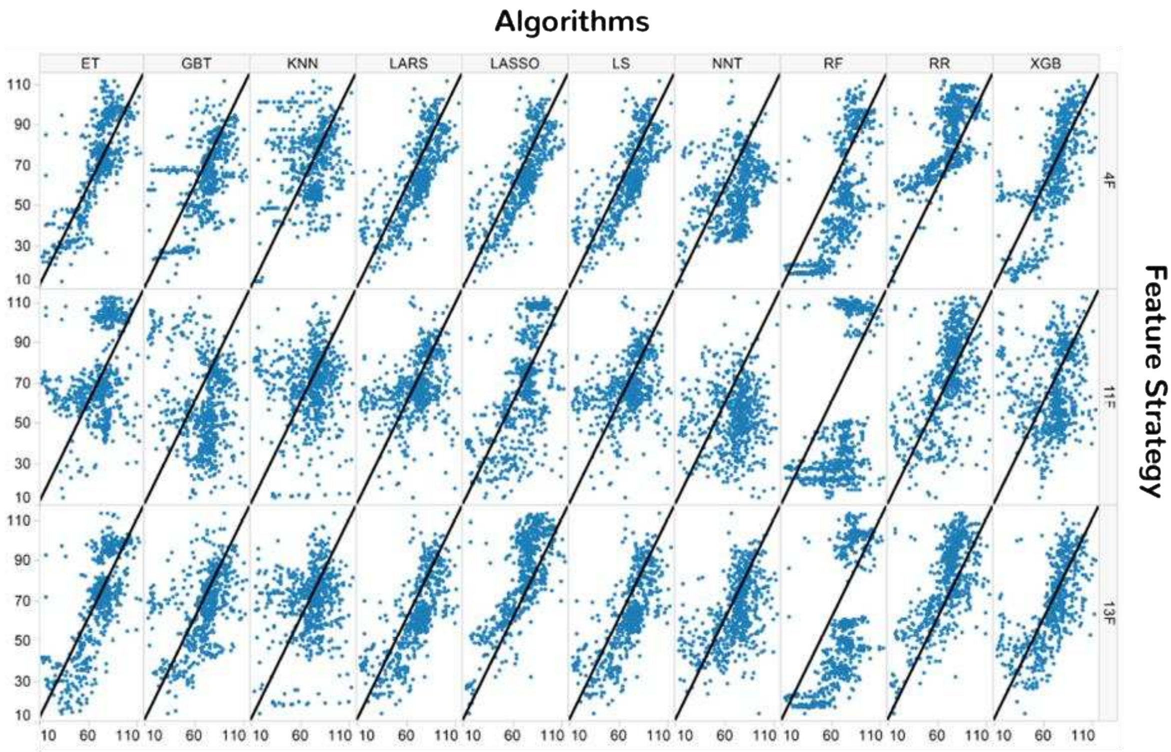
Reservoir property modeling has historically been driven by geostatistics, based on averaged distributions of properties and trends, to extrapolate information away from the well data. These approaches were designed to build models correct in average but with low prediction capacity. Recent developments of artificial intelligence allows to better capture the non-linear and complex inter-relationships between the different reservoir properties and improve the predictions capabilities of geological models. A myriad of algorithms and “feature engineering” strategies have been tested and ranked in term of success of predictability for area of interest in northern Santos Basin (Lafferriere et al., 2018). A key advantage of the Machine Learning algorithms is to be able to incorporate several input features into the prediction, and to validate a posteriori the performance of the prediction on a testing dataset. Figures 5 and 6 shows the results for a 1D facies prediction usinf differente ML algorithms and features combinations.

Figure 5 – 1D facies prediction results using different machine learning algorithms.



Source: Lafferriere et al. (2018)

Figure 6 – Data corretation obtained with different ML algorithms and features engineering strategies.

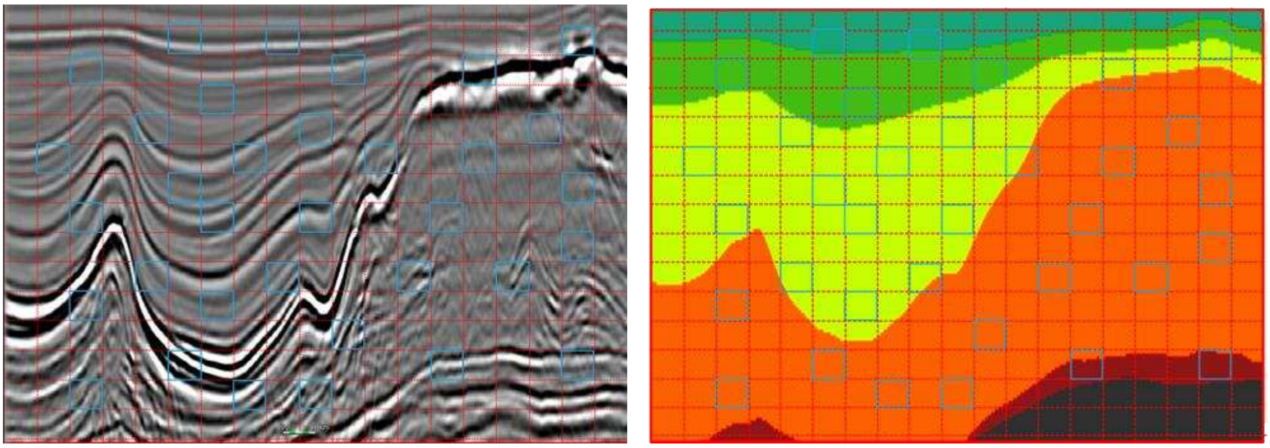


Source: Lafferriere et al. (2018)

Automated Seismic Interpretation with Convolutional Neural Networks

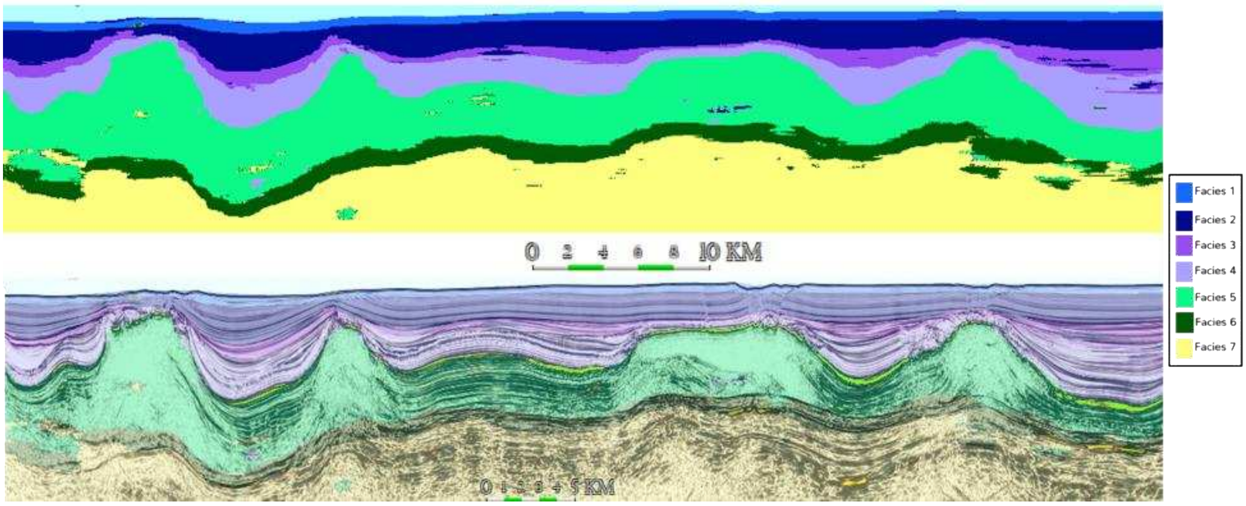
Seismic interpretation usually relies on heavy manual or semi-automatic processes of picking which are very time-consuming for geophysicists especially on 2D datasets. A machine-learning based workflow was developed to use prior interpretations made by geophysicists to consistently and automatically extend it to adjacent seismic 2D lines. After being loaded, preexisting interpretations were pre-processed to generate a discrete property model segmented into aleatory small squared images (Figure 7). These images fed a 2D Convolutional Neural Network, specially parametrized for seismic interpretation. Once accurately trained, the model was applied to a new 2D datasets still not interpreted. Depending on the quality of the original dataset, prediction accuracy of more the 97% can be achieved on nearby 2D lines (Figure 8), saving considerable time to geophysicist. This kind of approach also paves the way to automated velocity modeling at basin scale.

Figure 7 – Preexisting interpretations after being loaded and pre-processed to generate a discrete property model segmented into aleatory small squared images, which fed a CNN specially parametrized for seismic interpretation.



Source: Lafferriere et al. (2019)

Figure 8 –Seismic interpretation result obtained with the CNN (above), compared to the manual interpretation (bellow).

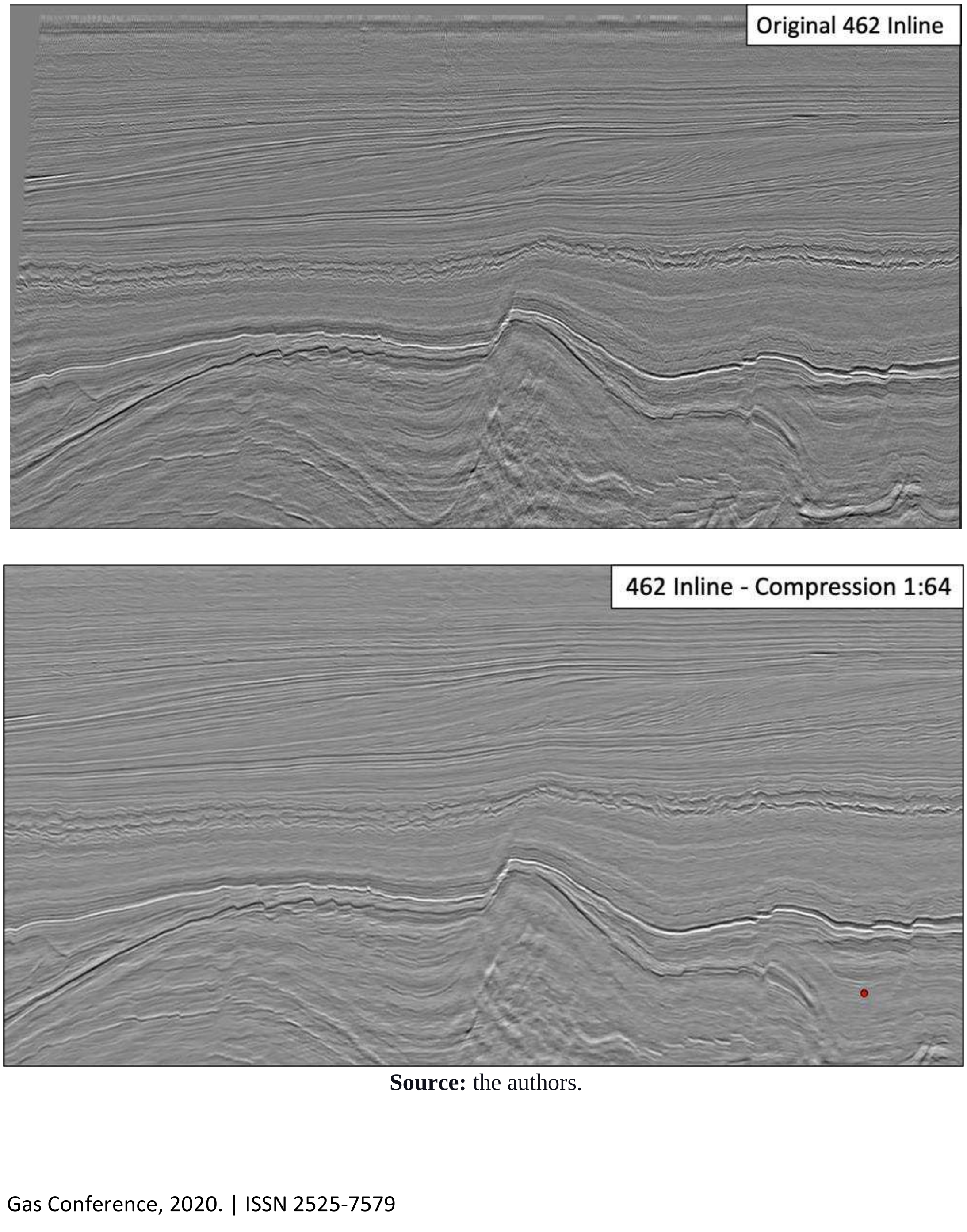


Source: Lafferriere et al. (2019)

Seismic Data Lossy compression

Seismic surveys size has been continuously increasing over the past decades. With higher resolution and multitude of processing vintages, the storage of such information and its access have become challenging for the E&P companies and services providers. A dedicated workflow to compress post-stack SEGY datasets using state-of-the-art Machine Learning algorithm was tested. This algorithm performs a lossy compression with very high Compression Rate (CR) and minimal quality loss (Figure 9) that could be applied to different steps of E&P workflow: for fast-track transmission, storage and access or any geoscience application that could support some data loss. The algorithm is based on an encoding/decoding strategy with a Convolutional Neural Networks (CNN) to minimize quality loss in each SEGY.

Figure 9 –Comparison between original (above) seimic section and compressed (1:64; bellow) using CNN.

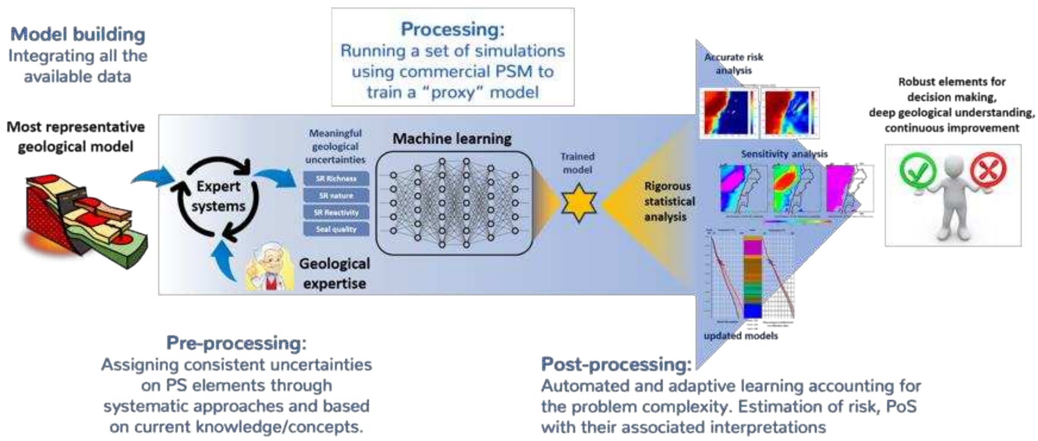


Source: the authors.

Petroleum system risk assessment

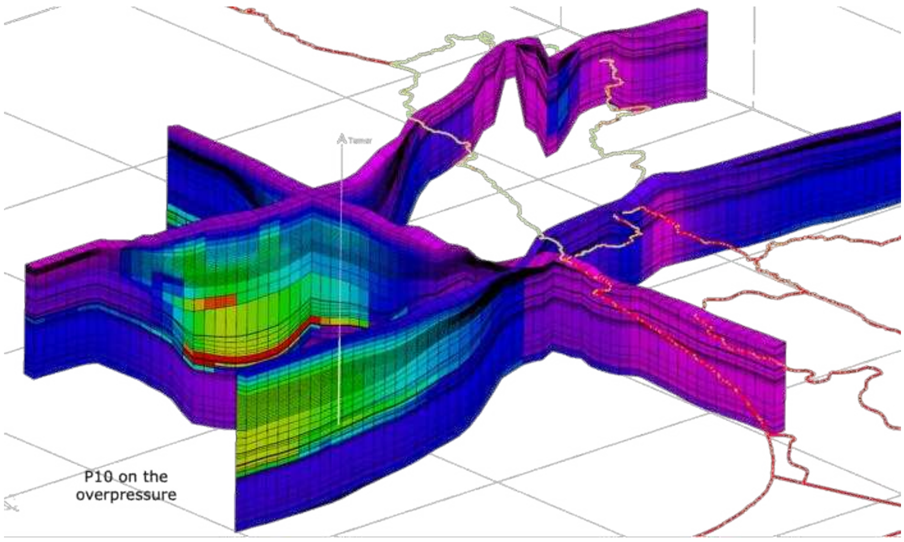
Petroleum system modeling (PSM) plays a key role in the assessment and mitigation of exploration risks by integrating most of the geologic data. However, such powerful method require labor and computing intensive efforts to integrate uncertainties coming from seismic data and to keep the pace of geologic knowledge evolution along the prospect's life cycle, hampering an accurate assessment of the risks. A novel approach based on machine learning techniques and expert systems was developed to overcome these limitations (Figure 10; Mathieu & Gonçalves, 2020). The application of the method in the still poorly known Levant basin gave access to the identification of sweet spots and to clues on how to further reduce the risks related to HC charge, seal efficiency and overpressure development (Figure 11). The results suggest that the method could enhance the role of PSM by fastening the reevaluation of risks and resources while acquiring new data.

Figure 10 – Machine learning and expert systems workflow used to integrate geologic uncertainties into petroleum system modeling and to improve the assessment of exploration risk.



Source: Mathieu & Gonçalves (2020).

Figure 11 – Assessement of overpressure risk usin the ML-assisted petroleum system modeling workflow.



Source: Mathieu & Gonçalves (2020).

3. Final considerations and conclusions

Despite the proliferation of ML and DL pilot projects in most of the oil companies, an effective insertion of these technologies in the mainstream of E&P workflow still poses significant technical and managerial challenges, such as natural complexity of subsurface data, domain silos in data infrastructure, difficulties to assemble teams with the required mix of skills, obstacles to deploy ML and DL solutions in a timely and reliable way and availability of domain-specific AI platforms.

Notwithstanding the foregoing, case studies presented herein demonstrate that artificial intelligence techniques can be used as a powerful tool for automating the analysis and interpretation of large-volume datasets at scale and improving the quality of complex E&P decision-making.

Referências

François Lafferriere, João Zambrini, Félix Gonçalves, João Victor Anjos, & Bruno Carvalho. (2019). Automated seismic interpretation using CNN: a 2D case study in the Santos Basin (Vol. 1, p. 1). Presented at the I Seminario de Inteligencia Artificial Aplicado a Indústria do Petróleo da ABGP, Rio de Janeiro. Retrieved from https://abgp.com.br/wp-content/uploads/2019/10/a7.02-12.-Automated-seismic-interpretation-using-CNN-a-2D-case-study-in-the-Santos-Basin.pdf

François Lafferriere, Félix Gonçalves, Juliana Araújo, Alexandre Gonçalves, João Victor dos Anjos, Bruno Carvalho, … Luis Neves. (2018). Integrating 3D seismic and well log data for improved estimation of lithological characteristics in the northern Santos Basin using Machine Learning techniques (Vol. 1, p. 1). Presented at the Joint SBGf/SEG Workshop on Machine Learning, Rio de Janeiro. Retrieved from https://seg.org/events/ML18

João Zambrini, François Lafferriere, Félix Gonçalves, & Edson Alonso. (2019). CNN seismic inversion using geological realistic scenarios derived from Markov-Chain synthetics (Vol. 1, p. 1). Presented at the I Seminario de Inteligencia Artificial Aplicada a Industria do Petroleo da ABGP, Rio de Janeiro. Retrieved from https://abgp.com.br/wp-content/uploads/2019/10/p8.04-39.-CNN-Seismic-inversion-using-geological-realistic-scenarios-derived-from-Markov-Chain-synthetics.pdf

Mathieu Ducros, & Félix Gonçalves. (2020, March 2). AI and Petroleum Risk Assessment.*GeoExpro Magazine*, *17*(1), 26–28. Retrieved from https://www.geoexpro.com/articles/2020/03/artificial-intelligence-and-petroleum-system-risk-assessment