# UNISIM ON-LINE

Year 18, Number 2 | 153rd edition | May, 2023


# All-in-one proxy to replace 4D seismic forward modelling with machine learning algorithms


Shahram Danaei


*"Faster 4D seismic forward modeling through machine learning to accelerate 4D seismic data assimilation procedure."*

## Introduction

The oil and gas industry has seen several paradigm shifts through centuries. Machine learning (ML) techniques perhaps mark the most important milestone in the history of this industry which may change the traditional methods and approaches. In this context, ML algorithms could be seen as an alternative 4D seismic forward model. Quantitative applications of 4D seismic data need a forward model to build a bridge between simulated rock and fluid properties and synthetic seismic responses. The 4D seismic forward modeling is a complex process which includes two back-to-back models. The first is a petro-elastic model (PEM) and the second is a seismic model. Therefore, the 4D seismic forward modeling is composed of a combination of PEM and the seismic model. This combination brings some problems in 4D seismic quantitative applications such as 4D seismic history matching. For example, its multidisciplinary nature is a problem especially when the forward model is used in history matching process. Moreover, it is a time-consuming process within iterative ensemble-based data assimilation scheme where hundreds or even thousands simulation runs are needed. Finally, the forward model is a step-by-step combo of models. To mitigate these problems, ML models could be seen as a proxy to substitute the traditional 4D seismic forward model. In this research, we propose a methodology to develop the proxy model (we call it, S4D-Proxy) and apply it to a post salt Brazilian offshore field.

## Methodology

The proposed method to develop the S4D-Proxy model has three main steps as follows:

**The first step (datasets preparation):** The first step for our method is data preparation. This step is important as the ML algorithms learn from the data and find hidden patterns between the input features (such as saturation-pressure changes, porosity) and the target (for our application, dRMS or time-lapse difference in root mean square amplitude). We use an ensemble of 3D reservoir simulation models to prepare the datasets and train the ML algorithms. First, each 3D reservoir model is simulated until 2016 (monitor seismic survey time for our application). Knowing the baseline (2013) and monitor (2016) times, the time-lapse property changes, such as saturation and pressure changes, are extracted from the simulation model. The simulation outputs are then transformed to the seismic attribute RMS using petro-elastic and seismic models and the synthetic dRMS is generated. Map-based input features such as porosity, NTG, and initial saturation and pressure are extracted. Lastly, time-lapse changes in saturation-pressure and dRMS are also extracted as 2D maps. The above process is repeated for all the models in the ensemble. Eventually, the prepared dataset for our application is divided into the training (70% of the models in the ensemble, or 140 models), the validation (10%, or 20 models), and the test dataset (20%, or 40 models). For more details, interested readers are referred to Danaei et al. (2023).

**The second step (training the ML algorithms):** Two ML algorithms are considered for our research. The first is an Extreme Gradient Booster (XGBoost) that is described in Chen and Guestrin (2016) and the second is a tailored Deep Neural Networks (DNN) architecture which its details could be found in Simonyan and Zisserman (2015). Moreover, there are two characteristics for our application to train the ML algorithms. The first is the use of an ensemble of reservoir simulation models (for our case, prior ensemble) and the second is map-based inputs and output. For the training phase, two strategies are adopted. The first is a standard training in which input features are related pointwise to the desired output and the second is 3X3 neighborhood strategy where in the input features, a point with its neighbors within a 3x3 window are related to the desired output.

**The third step (test):** The performance of the trained ML models is evaluated using a test dataset. This dataset is not used in the training and is completely unknown to the ML models. It is worth noting that we evaluate the ML models based on a quantitative measure (R-squared) and a visual comparison between the predicted dRMS from ML models and the results from a full-fledge PEM and seismic model.

## Application

The proposed methodology to develop the proxy model was applied to a post-salt offshore field located in the Campos Basin. The field is composed of unconsolidated (soft) sandstone and there are seven producers and four injectors. Figure. 1a shows a random simulation model in the ensemble of models (prior ensemble of models before data assimilation) and Figure. 1b illustrates dRMS map of observed seismic data. Main 4D signals of the mentioned field are located around water injectors (hardening signals) and a noticeable softening 4D signal in the middle of the reservoir because of gas coming out of solution. There are some small-scale 4D signals which are considered for our application as well.

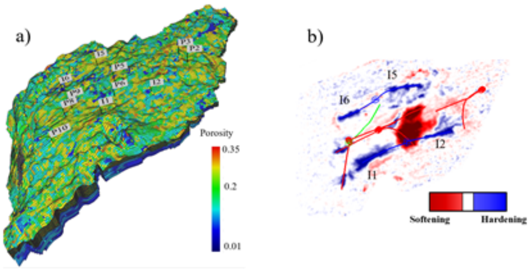


***Figure 1**: (a) a random porosity model in the ensemble of reservoir models and (b) dRMS attribute to define softening (red) and hardening (blue) signals.*

Three ML models are considered for our application (Table 2). For the first model, we used XGBoost (XGB) with a standard training strategy, for the second model, a 3x3 neighborhood was considered with the XGBoost algorithm (XGB-3x3) and finally, the DNN algorithm was trained with a 3x3 neighborhood strategy (DNN-3x3).

***Table 2**: Different ML models as S4D-Proxy.*

| ML model | Name | Machine learning algorithm | Training strategy | Color code |
|---|---|---|---|---|
| 1 | XGB | XGBoost | Standard | |
| 2 | XGB-3x3 | XGBoost | 3x3 neighbourhood | |
| 3 | DNN-3x3 | Deep Neural Network | 3x3 neighbourhood | |

## Results

This section is divided into two parts. The first part shows the performance of the ML models based on the quantitative measure (R-squared), and the second is a visual comparison between ML predictions and the traditional approach.

**Quantitative measure (R-squared):**
We compared the responses of the ML models on the test dataset. For each ML model result, a boxplot was

## 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.*

*"Our approach reduces turnaround time in data assimilation and enhances the use of 4D seismic data in model-based decision making."*

drawn for the group of the R-squared in the test dataset. Figure. 2 illustrates the R-squared measure for each ML model. Aside from the boxplots, the group of the R-squared in the test dataset was represented with a histogram. The results demonstrate that the DNN-3x3 had the best predictability.

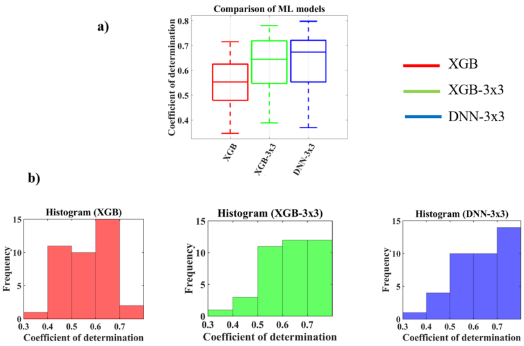


***Figure 2**: R-squared measure for different ML model: (a) boxplots of the R-squared measures and (b) histograms of the quantitative measure.*

An interesting observation in Figure. 2a concerns the median value of the DNN-3x3 and XGB-3x3. The median for the DNN-3x3 model was 0.67 and 0.64 for the XGB-3x3. This indicates that when comparing the group of R-squared, the DNN-3x3 had more test models (20 test models) with R-squared higher than 0.67 compared to the XGB-3x3 (only 14 test models). The comparison of the XGB and the XGB-3x3 showed that the XGB-3x3 provided better predictability compared to the XGB.

**Visual comparison of the ML models prediction**

Figure. 3 shows the results for the ML models and the traditional PEM and seismic model (reference solution). From the figure, it is clear that all three ML models were able to predict the main 4D signals. For example, all ML models were able to predict hardening signals around all injectors. One could compare the ML predictions with the reference solution in the location of injector in the southwest (I1). The softening signals due to the gas out of the solution in the center and the right flank of the reservoir were also predicted by all three ML models. Aside from prediction of the main 4D signals with all the model, the DNN-3x3 was able to estimate more details in the 4D map. For instance, the DNN was successful in recovering details as shown with blue arrows in the figure. For some softening signals especially around injectors, the DNN-3x3 model was able to predict them as shown in Figure. 4 where the softening signals around Injector 5 were estimated by the model. The reason could be in the DNN architecture used for our research, where different convolutional layers could be able to capture more features in the input data to train the DNN algorithm.

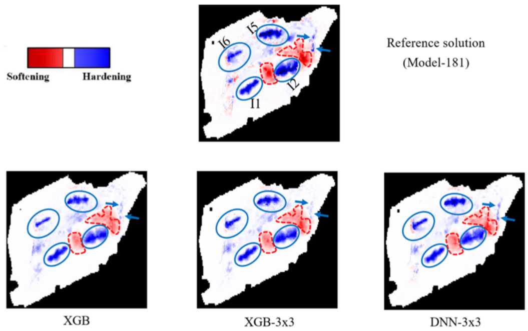


***Figure 3**: dRMS predictions with different ML models.*

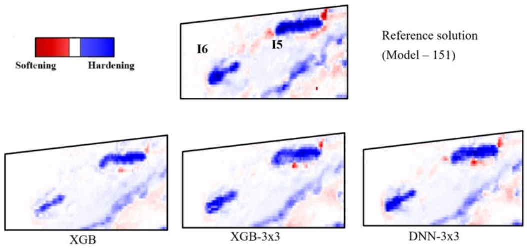


***Figure 4**: DNN-3x3 model predicted small-scaled softening signals especially around injectors.*

## Conclusions

Quantitative applications of 4D seismic data require a 4D seismic forward model with a petro-elastic model then followed by a seismic model. This paper proposed a very fast alternative 4D seismic forward model to replace the traditional approach with machine learning models (S4D-Proxy). The specific conclusions and applications of the S4D-Proxy are as follows:

1. The DNN-3x3 model performed better than the XGBoost models. Therefore, we recommend this model as an alternative fast forward model.
2. The immediate application of the S4D-Proxy is inside 4D seismic history matching to fasten the process.
3. S4D-Proxy model could also be considered as an alternative forward model to invert saturation-pressure changes from 4D seismic data.

## References

Danaei, S., Cirne, M., Maleki, M., Schiozer, D. J., Rocha, A., Davolio, A., 2023. All-in-one proxy to replace 4D seismic forward modeling with machine learning algorithms. Geoenergy Science and Engineering 222, 211460. https://doi.org/10.1016/j.geoen.2023.211460

## About the author:

Shahram Danaei has a MSc by research in Geosciences from PETRONAS University and a MSc degree in Petroleum Engineering from University of Tehran. His research interest includes: Reservoir modeling, 3D/4D seismic data analysis, and Machine learning. He is currently a postdoc researcher at the University of Campinas (UNICAMP).

## *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 and Petroleum Studies
University of Campinas
Campinas - SP

Phone.: 55-19-3521-1220

unisim@cepetro.unicamp.br