Introduction

The geological provinces on the Pacific Coast of Peru have a long and complex geological history. The petroleum system has undergone more than a hundred years of exploration and exploitation (de Souza *et at.*), however its full potential remains to be understood. Integration of seismic and well data is key for building a consistent geological framework that can be used to explore its prospectivity. A case study with a robust petrophysics and rock physics workflow implemented in eleven wells from the North-West Pacific region of Peru is presented here. This workflow allowed us to predict the in-situ elastic response of the well logs, as well as investigate, in real time, how potential geological scenarios, such as changes in porosity, mineral volume and fluid properties, can affect the response in elastic well logs and by extension in the seismic amplitudes.

Additionally, we propose a method for predicting shear velocity log (Vs) from other logs using a machine learning algorithm. The two key advantages of this method over other traditional methods in the industry is that it doesn't require a petrophysical interpretation as an input, and that it can be deployed over hundreds of wells in a matter of minutes.

Method and theory

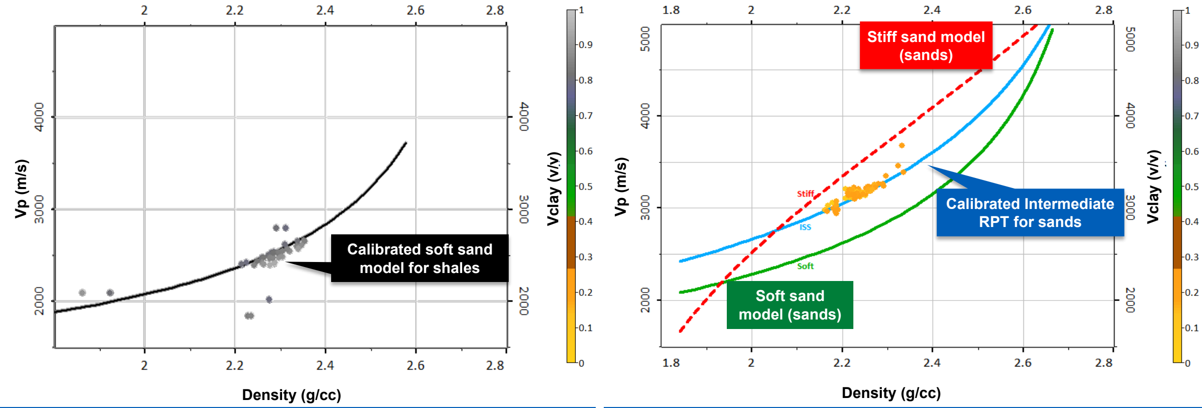
The first step was to perform a complete petrophysical analysis of well-log information to determine the volume of minerals and fluids present in the rock at in situ well-log conditions for each of the eleven wells. Then, a rock physics analysis was performed to derive rock physics models (RPMs) describing relationships between the elastic response and the geological properties of the rock such as mineralogy, porosity, pore shape, permeability, fluid saturation, pore pressure, laminations, fractures, etc. These calibrated RPMs were also used to edit the compressional and shear wave velocity logs and density where issues where found, and to predict shear wave velocity in three wells where this log was not acquired.

In addition to the traditional rock physics approach for prediction of shear log, a tree-based machine learning algorithm was implemented. Tree-based models, derive their prediction from the recorded target properties of a group of observations that are similar to the current sample, where the similarity rules are defined during the training process. The work described here uses boosted trees (BT), as these are less prone to overfitting the training data than other tree-based algorithms. The models were implemented using the XGBoost library (Chen and Guestrin, 2016), an open-source software library providing a regularizing gradient boosting framework for many languages.

Results

Results indicated a strong relationship between the derived RPM and the geological setting. Post Paleozoic shales overall exhibited a response consistent with the soft sand model (Dvorkin,1991), i.e. a slow increase in the elastic logs (acoustic-wave and shear-wave velocities and density) with decreasing porosity. This can be associated with porosity reduction through deposition of shale in the pore spaces or compaction. Non-shale lithologies within this interval tended to exhibit a faster increase of the elastic logs with decreasing porosity, which could be related to cementation or changes in the cement type within the grain contacts, so an intermediate stiff sand model was implemented. The intermediate stiff model is a model that is halfway between the soft sand model and the stiff sand model (Dvorkin et at. 1994, Dvorkin et al 1996).

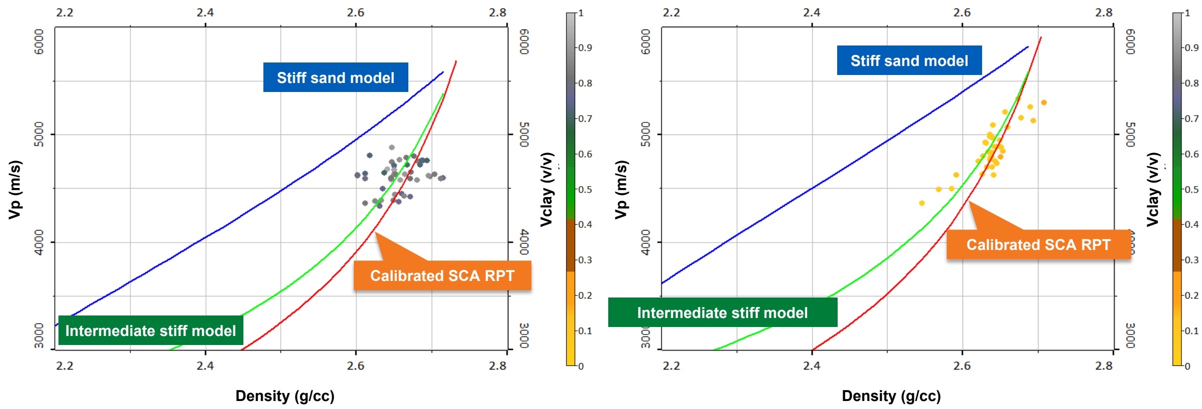
Figure 2 shows results corresponding to rocks from the Palaeozoic sequence. In this interval, fractured quarzites present an exceptionally stiff response in terms of elastic logs, alongside very low porosities. Moreover, the porosity system in these rocks is quite complex and two types of porosities have been identified: primary porosity or porosity preserved from deposition through lithification and burial history, and secondary porosity, which is mainly associated to fractures. Based on this observation, an inclusion model and more specifically, a self-consistent approach (SCA) model (Budiansky, 1965; Hill, 1965; Wu,1966) which allows the interpreter to account not only for the presence of porous space but also for the different pore shapes present in the rock was required to accurately describe the elastic response of these rocks. Furthermore, this model can be used to predict how changes in rock properties such as porosity and clay content, can impact the ability to discriminate between brine and hydrocarbon sands.



Density vs Vp crossplot for sands (Vclay ≤ 0.3 & Sw =1) and shales (Vclay ≥ 0.8), over Tertiary sediments with calibrated Rock Physics Template (RPT)

Density vs Vp crossplot for Vclay ≤ 0.3 & Sw =1, over Palaeozoic section with calibrated Rock Physics Template (RPT)

Figure 1 *Rock physics diagnostics for Tertiary shales (left) and sands (right). Different calibrated rock physics templates are presented for reference. The soft sand model is a good fit for the shales, however the elastic behaviour of the sands is somewhat in between the soft sand model and the stiff sand model, therefore the use of an intermediate stiff sand model.*



Density vs Vp crossplot for Vclay ≥ 0.7 over Palaeozoic section, with calibrated Rock Physics Template (RPT)

Density vs Vp crossplot for Vclay ≤ 0.3 & Sw =1, over Palaeozoic section with calibrated Rock Physics Template (RPT)

Figure 2 *Rock physics diagnostics for Palaeozoic shales (left) and wet sands (right). Different calibrated rock physics templates are presented for reference. Palaeozoic rocks have a more complex elastic response compared to sediments from the Tertiary. Both sands and shales are very proximal to the theoretical density of the mineral and the SCA seem to be the best model for describing their elastic behaviour.*

Once the rock physics atlas for eleven wells in the Talara-Progreso Basin had been built, it was possible to assess if ML models could be trained using this regional knowledge database for performing shear wave prediction up to a standard comparable to that of what a specialist could produce. For this, the data was randomly split on a well-by-well basis, out of the eleven wells available, 9 wells (81%) were used for training and 2 wells (19%) were kept aside for blind testing of the model. This ensured that the ML algorithm predicts on wells that it has never seen, allowing us to estimate the accuracy of the predictions and the generalization power of the model. Three different sets of input well logs were used to train the models:

a) **Base logs** - a model with GR, NPHI, logarithm of deep resistivity, RHOB_RAW (unconditioned density), Vp_RAW(unconditioned V) and TVDBML
b) **Processed logs** - model trained using GR, NPHI, logarithm of deep resistivity, conditioned Vp, conditioned RHOB and TVDBML
c) **Minimal logs** - a model that only considers conditioned Vp, RHOB and TVDBML

Figure 3 shows the results from the ML prediction using three different sets of input log. All models produce an excellent shear prediction, there are no significant differences between the true Vs and the BT predicted logs. The R^2 score (Kramer, 2005) was above 0.98 for all models, and on average it took the algorithm less than a second to generate the prediction of a Vs log once the model was trained. The accuracy and swiftness of this method could mean that predicting Vs log for hundreds of wells in the area could be done in a matter of minutes.

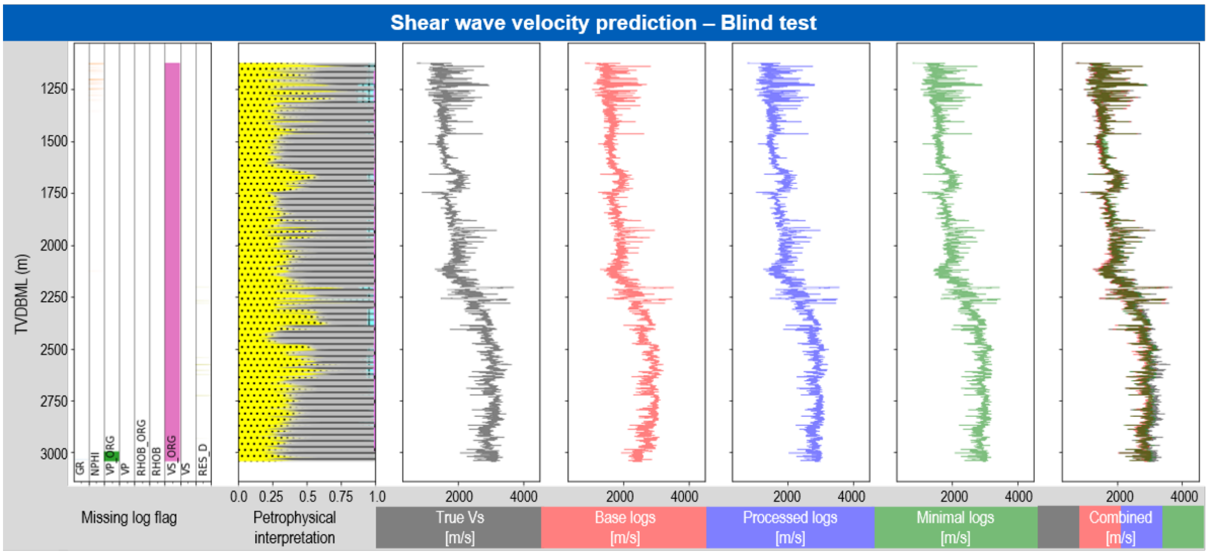


Figure 3 *Vs (m/s) prediction using BT algorithm in a well from the Talara-Progreso Basin with three different suites of logs as inputs. This well samples Tertiary and Cretaceous sediments and has not acquired shear, but it has been predicted by a specialist using Greenberg-Castagna relationship with standard coefficients validated by nearby wells. The three BT models predict quite well along the entire length of the well and no significant difference has been found between the specialist generated shear and the predictions. The petrophysical interpretation is presented for reference only and has not been used as input for any of the BT predictions.*

Conclusions

A comprehensive petrophysical and rock physics workflow was applied on eleven wells from the Talara-Progreso Basin. Observations showed that changes in lithology, porosity and fluid properties in the area can be described using a soft sand model and an intermediate stiff sands model in the Post-Paleozoic sequences, and a self-consistent approximation model, that accounts for different aspect ratios of the pores, in the Paleozoic sequence.

Nine wells from the current rock physics atlas have been used to build ML models for shear-wave prediction from different input logs. The high R^2 score between predicted and true Vs from the blind tests confirms that any of the proposed BT models can be used to predict Vs logs. Given the high quality and the short turnaround of this method, this workflow represents a robust and fast alternative for predicting Vs log in the hundreds of wells without shear in the area. Finally, increasing the number of wells with shear could positively impact the implementation of integrated Amplitude Versus Offset (AVO) and Quantitative Interpretation studies for the screening of opportunities in the region.

Acknowledgements

The authors would like to thank the PGS MultiClient for providing the data (well CPIs and rockAVO Atlas) and technical collaboration, and the EPCC for the technical collaboration over the development of the machine learning codes.

References

Budiansky, B., [1965] On the elastic moduli of some heterogeneous materials. *J. Mech. Phys. Solids*, **13**, 223-227.

Chen, T. and Guestrin, C. [2016] XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM and SIGKDD international conference on knowledge discovery and data mining*, August, 785-794.

De Souza, P. H., Bustamante, E. and Castillo, J. A., [2012] Seismic Interpretation in an Ultra-Mature Area: Talara Basin, Peru. *SEG Technical Program Expanded Abstracts:* 1-5

Dvorkin, J., Mavko, G., and Nur, A., [1991] The effect of cementation on the elastic properties of granular material. *Mechanics of Materials*, **12**, 207-217.

Dvorkin, J., Nur, A., and Yin, H., [1994] Effective properties of cemented granular material. *Mechanics of Materials*, **18**, 351-366.

Dvorkin, J., and Nur, A., [1996] Elasticity of high-porosity sandstones: Theory for two North Sea datasets. *Geophysics*, **61**, 1363-1370.

Greenberg, M.L., and Castagna, J.P. [1992] Shear-wave velocity estimation in porous rocks: Theoretical formulation, preliminary verification and applications. *Geophysical Prospecting*, **40**, 195-209.

Hill, R., [1965] A self-consistent mechanics of composite materials. *J. Mech. Phys. Solids*, **13**, 213-222.

Kramer, M. [2005] R2 statistics for mixed models. *15th Annual Conference on Applied Statistics in Agriculture.*

Krief, M., Garat, J., Stellingwerff, J., and Ventre, J. [1990] A petrophysical interpretation using the velocities of P and S waves (full-waveform sonic). *The Log Analyst*, **31**, November, 355-369.

Wu, T. T., [1966] The effect of inclusion shape on the elastic moduli of a two-phase material. *Int. J. Solids Struct*., **2**, 1-8.