Facies Classification using Seismic Inversion and Machine Learning Methodologies

Ana Sofia Tita Martins Moutoso

anamoutoso@tecnico.ulisboa.pt

DECivil, Instituto Superior Técnico, Universidade de Lisboa, Portugal

December, 2020

Abstract

An accurate and precise seismic interpretation is vital to understand the geology of a target area. However, when performed in the conventional way, by human interpreters, it is considered to be a difficult and time-consuming task. Over the years several alternative and complementary approaches have been proposed and recently machine learning methods have been successfully implemented in the facies classification process.

In this thesis deterministic and geostatistic seismic inversion methods will be explored and compared to the results of an automated workflow performed by the GeoScience Advisor (GSA), a system developed both by Galp and IBM, that assists experts with decisions involving geological characterization based on seismic data through the use of multiple machine learning methods.

Keywords: Machine Learning, Facies Classification, Seismic Inversion, Automated Seismic Interpretation.

Introduction

Understanding the rock properties that constitute a petroleum system is a crucial factor in the exploration and production stages of any field. This work proposes not only the estimation of elastic properties associated with a hydrocarbon reservoir but also the attempt to take a step further and predict lithofacies through the use of innovative techniques.

Seismic inversions have been used for several decades in the petroleum industry, both for exploration and production purposes. Nowadays, predicting spatial distribution of reservoir properties is leveraged by inverting 3D seismic data calibrated to elastic properties obtained from well-log data. These methodologies can be used to look inside a geobody, by providing volumes of the elastic properties of its interior and a quantitative interpretation of porosity, lithology and litho-fluid facies. The inversion methods are either deterministic or stochastic and these approaches can be applied to post and/or pre-stack seismic [1].

Over the years, ML has attracted attention in several areas of petroleum engineering and geosciences. Some of the areas in which ML methods have been successfully implemented in the petroleum industry include reservoir

characterization, seismic pattern recognition, reservoir properties prediction, lithofacies identification and well production prediction [2].

Case Study – Kupe Field

The Kupe gas condensate field is located in the Tasman Sea at 35 km offshore of the town Manaia, New Zealand. This field has an area of 256 km^2, an aerial extent of 100.000 km^2, and it is situated in the southern Taranaki Basin [3] [4], represented in Figure 1.

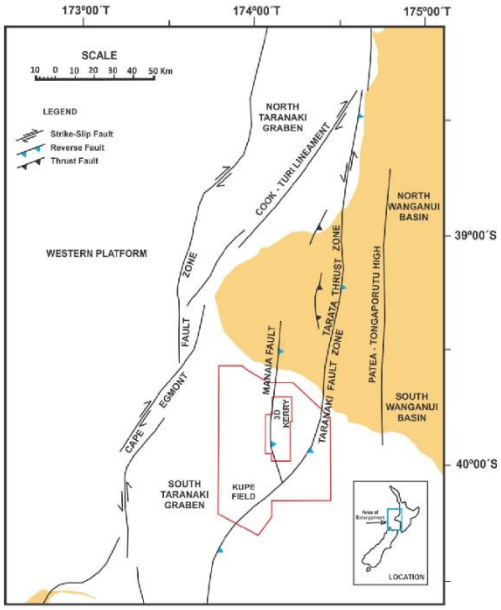

Figure 1. Location of Kupe field, New Zealand, and major faults [5].

Geological Setting and Stratigraphy

The evolution process of the Taranaki Basin is divided in three different phases. The evolution starts with a rifting phase in the mid-Cretaceous–Paleocene. This rifting is associated with the breakup of the supercontinent Gondwana, which originated the Tasman Sea and numerous extensional basins on the New Zealand subcontinent, accompanied by coal measure deposition in restricted, fault-controlled basins. The second phase consists in the evolution of an Eocene–Early Oligocene passive margin, post-rift thermal contraction and regional subsidence, originating the deposition of coal measures in marginal marine settings and on transgressive coastal plains. The third and current phase started in Oligocene and comprises the evolution of the basin to an active marginal due to the development of the Australia-Pacific convergent plate boundary [6] [4].

In terms of stratigraphy, the age range of the basin goes from the Late Cretaceous until nowadays, as it is illustrated in Figure 2.

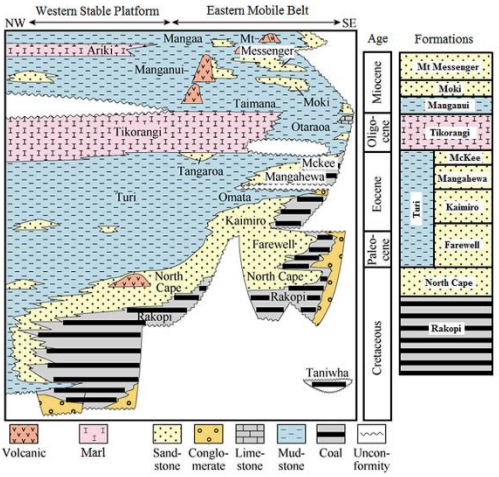

Figure 2. Stratigraphic sequence and lithostratigraphic column of the Taranaki Basin, New Zealand [6].

The Taranaki Basin is subdivided in two distinct structural zones [4]:

- Western Stable Platform: explicitly distressed by Post-Eocene tectonic activities;
- Eastern Mobile Platform: influenced by several tectonic regimes since the Late Miocene until nowadays and composed by:
    - Northern Graben;
    - Central Graben;
    - Southern Graben.

The extensional regime is much more dominant in the Northern and Central Graben while the Southern Graben is mostly characterized by structural inversion, in other words, compressional reactivation of already existing extensional structures [4].

The Kupe field is located in an area within the Eastern Mobile Platform of the southern Taranaki Basin, which comprises the offshore part of the Manaia Anticline. The north-south Taranaki fault placed in the east of the basin was active since mid-Oligocene. The main fault in the region of the Kupe field is a north-south oriented Manaia fault, which crosses the central part of the area [3].

Petroleum System

The source rocks from the Taranaki basin are typically coals from the Late Cretaceous Pakawau Group and Paleogene Kapuni Group. These source rocks are most likely capable to generate gas instead of oil. Regarding the seal rock, mudstones deposited during the Oligocene are common topseals or internal seals in the basin [7].

The oldest reservoir sequence in the basin is the Paleocene Farewell formation (Figure 2), which comprises shoreline systems and coastal plains. This formation was deposited in the transgressive regime after the rift subsidence. In the Kupe field, this sandstone unit has good reservoir qualities and comprises porosities that exceed 20% [7].

Taranaki basin has a complex tectonic structure and therefore structural traps occur in the area. Distinctive features like horsts, folds, overthrusts and thrusted anticlines are expected throughout the basin [8].

Data Description

The Kerry 3D marine seismic data used was obtained from NZP\&M online data base and it was acquired in 1996 and reprocessed to improve the quality in 2004. This survey contains inlines ranging from 2042 to 3188, xlines ranging from 300 to 3100, a total two-way time recorded length of 4996 ms and a resample rate of 4 ms. In order to facilitate the seismic interpretation, a sub-volume was created, having inlines ranging from 2050 to 3180 and xlines ranging from 715 to 2255.

Within the seismic sub-volume created, there are six wells available. Three of these wells are classified as exploration wells: Kupe-1, Kupe South-1 and Momoho-1. The remaining wells are classified as appraisal wells: Kupe South-2, Kupe South-4 and Kupe South-5. Each well has available three types of well-logs: density, gamma-ray and P-wave velocity.

Methodology and Workflow

The methodology adopted can be subdivided in:

- Seismic Interpretation;
- Well-to-Seismic Tie and Wavelet Estimation;
- Seismic Inversions;
- GSA.

Seismic Interpretation

Three different horizons were interpreted in Petrel ® and they correspond approximately to the seafloor, top of the reservoir formation Farewell and another in between these last two, corresponding to a significant unconformity inserted in the Matemateaonga formation.

Regarding the structural interpretation, multiple faults can be identified throughout the seismic

cube. As explained before, the evolution process of the Taranaki basin included both extensional and compressional regimes, which justifies the presence of normal and inverse faults.

Concerning the seismic stratigraphy, multiple seismic features can be identified throughout the seismic cube as well. These features can be a great help to interpret depositional regimes.

Well-to-Seismic Tie / Wavelet Estimation

Two different wavelets were used in the inversions performed, since the stochastic inversion only considers a portion of the seismic volume. Consequently, the well-to-seismic tie was done separately too. Both wavelets were generated using Hampson Russel ®.

Stochastic Seismic Inversion

This thesis incorporated the Global Stochastic Inversion (GSI), an iterative approach based on the simulation and co-simulation of acoustic and elastic impedances. The simulation algorithm used is the Direct Sequential Simulation (DSS), which is based on the resample of the global distribution conditioned by the local conditional mean and variance [9] [10]. The GSI methodology is illustrated in Figure 3.

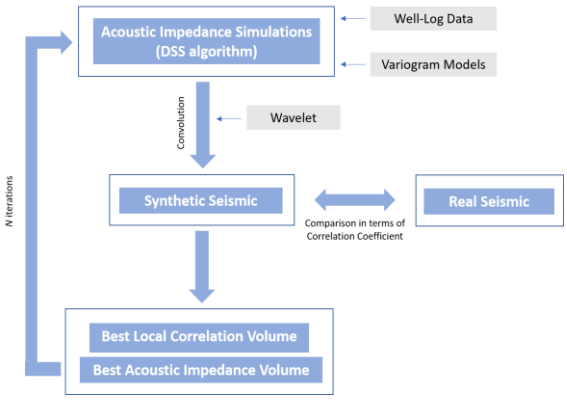

Figure 3. Scheme summarizing the GSI methodology. Modified from [9].

In a first stage, a study area was delimited and divided in two different zones: one above the reservoir unit and another corresponding to the reservoir itself. This division allowed a separate assessment of the continuity of the study area. Seismic data was used to assess the continuity and develop variogram models of the two zones. The upper zone, above the reservoir unit, has a higher continuity while the second zone has a much lower continuity. Both variograms were adjusted using exponential structures.

In order to complement the well-log information and later determine the input P-Impedance for the seismic inversion, a Gardner's relation was used as an approximation of density.

A low frequency model, developed for the deterministic inversion, was incorporated in the stochastic inversion process to add a spatial constraint.

The inversion was performed using a computational code developed in Centro de Recursos Naturais e Ambiente (CERENA) at Instituto Superior Técnico, using 6 iterations with 32 simulations each.

Deterministic Seismic Inversion

A model-based inversion was also performed. This methodology uses an iterative forward modelling and comparison procedure. Essentially, an initial impedance model is convolved with a certain wavelet to obtain a synthetic seismic response that is compared with the real seismic trace. The impedance model is modified in an iterative process until a certain correlation value is attained or a pre-defined number of iterations fulfilled, meaning that the final inversion result is a solution in which the impedance model has been checked

against the seismic traces and the errors calculated and minimized [11]. The model-based methodology is illustrated in Figure 4.

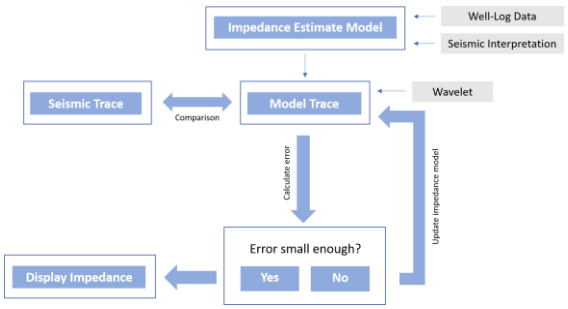

Figure 4. Generalized flow-chart for model-based inversion. Modified from [11].

This deterministic seismic inversion was performed on HampsonRussell ® considering a soft-constraint model. The initial low frequency (3-10 Hz) model was built in the time domain using the horizons interpreted and five of the wells available.

GeoScience Advisor

Seismic Facies Analysis (SFA), a module of GSA, applies cognitive visual computing in order to create an image-based machinery that assists experts with decisions involving geological characterization based on seismic data. The image processing and computer vision (CV) techniques will be used to approximate the geometry of geological structures, their topology and to classify textures. These features will empower the expert by feeding the knowledge base (KB) semi-automatically, retrieving structural information and guiding visualizations.

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM). The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, generating a GLCM, and then extracting statistical measures from this matrix. Each matrix is defined by two parameters, the angle Φ and the distance *d*, creating a mask to assemble the co-occurrence matrix. After the GLCMs are calculated, the formulations proposed by Haralick are used to obtain the imaging parameters based on each matrix.

The task template considered in this thesis used the parameters Φ (0, 45, 90 and 135) and *d* (1,2,3 and 4), as well as the image parameters: auto-correlation, correlation, energy and difference variance.

The result of the visual inspection combined with a Gaussian Mixture Model (GMM) methodology originates the clustering of seismic data. After the clustering, SFA suggests a classification in terms of seismic facies and lithofacies, based on a SVM algorithm.

Results and Discussion

Cross-Plots of Well-Log Data

The cross-plot of P-impedance vs. density (Figure 5) shows two main structures that can correspond to two different lithologies. The cross-section shown in Figure 5, exhibits the two structures and displays a clear distinction of lithologies, that match the formations identified in the well-log data.

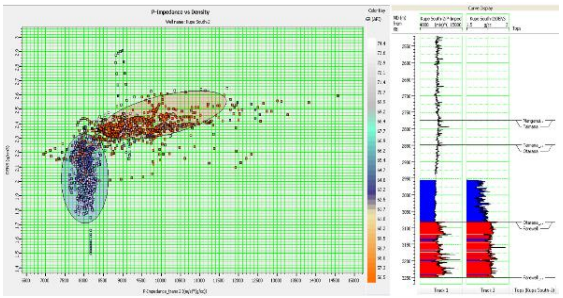

Figure 5. Cross-Plot of P-impedance vs. Density well-logs and corresponding cross-section, identifying two main structures.

The cross-plot of P-impedance vs. gamma ray (Figure 6) also displays two different structures. Similar to the previous cross-plot, the structures identified match the formations exhibited in the cross-section.

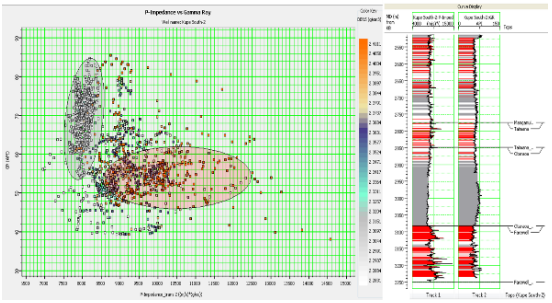

Figure 6. Cross-Plot of P-impedance vs. Gamma Ray well-logs and corresponding cross-section, identifying two main structures.

By analyzing the last two cross-plots it is possible to confirm the geology expected. The high gamma ray and low density responses that help identify the vertical structures are very compatible to a clay-rich rock formation like a compact mudstone. The low gamma ray and high density responses that characterize the horizontal structures are very compatible to a sandstone unit interspersed with clay, as shown in both cross-sections.

This cross-plot analysis will facilitate the interpretation of the seismic inversion results since it associates high values of P-impedance (> 8500 (m/s)*(g/cc)) to sands of interest and lower values of P-impedance (< 8500 (m/s)*(g/cc)) to the presence of shales.

Seismic Inversions

As described previously, two different types of seismic inversions were performed. In a first stage the results can be compared through a general look of the acoustic impedance models (Figure 7). Since the same low frequency model was used in both seismic inversions, the overall behavior shows some resemblances, such as the clear distinction of the two formations (Otaroa-Farewell), which validates de horizon interpreted. Although the models present comparable structures, the deterministic one is significantly more continuous and less detailed when compared to the stochastic inversion. The stochastic model displayed presents lower values of P-impedance comparing to the deterministic, which was expected, considering that the chosen image corresponds to the mean of the simulations from the last iteration performed.

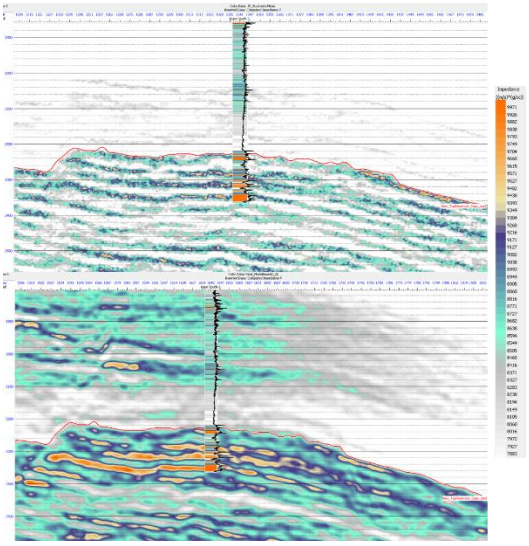

Figure 7. Acoustic impedance mean model (from the last iteration) of the stochastic seismic inversion and acoustic impedance model of the deterministic seismic inversion, respectively. This figure corresponds to the inline 2777 and exhibits the computed P-impedance of the well Kupe South-2.

Comparing the correlation coefficients in the inline of study, the results reveal the same resemblance between the synthetic seismic data produced by both models and the real seismic data, approximately 86%.

The cross-plot analysis performed identified the sands of interest in ranges of P-impedance superior to 8500 (m/s)*(g/cc). Figure 8 exhibits a color key edit that allows an easier interpretation of the sands in the reservoir

region, by applying an opacity filter above that P-impedance value.

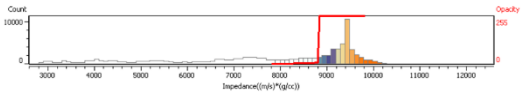

Figure 8. New color key edit with an opacity filter applied.

In Figure 9, lower values of P-impedance (represented in blue) correspond to sands that most likely have a small percentage of clay minerals, while the higher values of P-impedance (represented in orange) can correspond to "cleaner" sands. The deterministic inversion presents a more optimistic view of the reservoir area, while the mean model of the stochastic inversion presents a more conservative version.

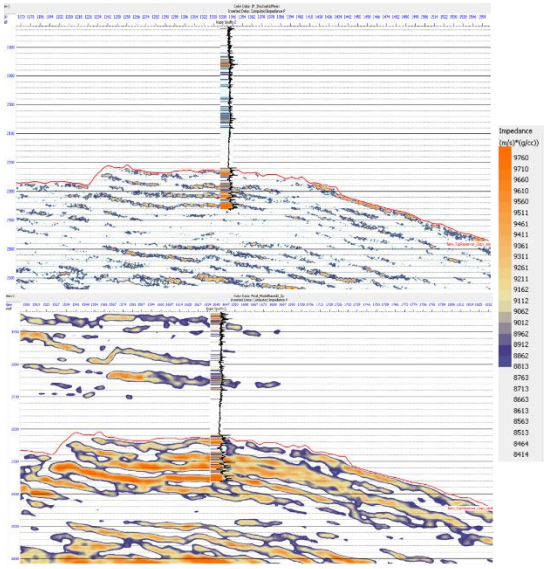

Figure 9. Same models displayed in Figure 9 in a different color key. This figure corresponds to the inline 2777 and exhibits the computed P-impedance of the Kupe South-2.

GeoScience Advisor

In order to study the area of interest on the chosen inline (2777), a 2D region was created around the interpreted horizon, similar to the region outlined previously for the stochastic inversion. As a result of the visual inspection explained and a GMM algorithm, the area of interest was divided into 8 clusters, as exhibited in Figure 10.

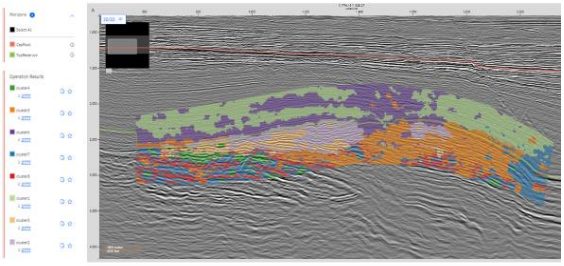

Figure 10. Clustering executed by the system using a GMM algorithm that considers the task template chosen (inline 2777).

The result of the clustering shows a very clear distinction between the upper and lower formations (Otaroa-Farewell), also validated by the cross-plots and both seismic inversions. Figure 11 displays a comparison between the results obtained through the clustering, stochastic inversion and deterministic inversion around the well location, as well as the computed P-impedance well-log.

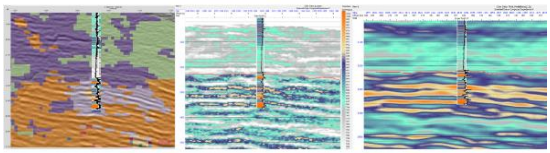

Figure 11. Comparison around the well location between the clustering performed by SFA, the stochastic inversion and the deterministic inversion, respectively, displaying the well-log P-impedance of the Kupe South-2 well (inline 2777).

Through Figure 11 it is possible to characterize clusters 2 and 5 as being the most likely to contain hydrocarbons since they are very compatible with the beginning of the reservoir area displayed in the well-log, although the size of the clusters is smaller when compared with the seismic inversions, specially the deterministic one that exhibits a larger reservoir area. The lower values of P-impedance in the well, associated to shales, match clusters 1 and

6, which suggests they can be associated to shale.

Concerning the classification of clusters, SFA has a classification system that suggests seismic facies and lithofacies (sand or shale). This classification gives the user a probability of finding a certain type of facies according to the CV exclusively or CV+KB when the system finds similar seismic data interpreted in the KB. Figure 12 displays the classification suggested for the clusters mentioned before: 1, 2, 5 and 6.

Figure 12. Classification suggested by the system for clusters 2, 5, 1 and 6, respectively.

Analyzing the classification displayed in Figure 12, it is possible to verify that, as expected, clusters 2 and 5 present a higher probability of containing sands, while clusters 1 and 6 most likely present higher contents of shales. This classification fits the geological description mentioned in the literature as well as the previous results.

Regarding the seismic facies, not every suggestion is correct, but the overall classification is reasonable. In some cases, the CV and CV+KB provide contradictory information, which should not happen. Those situations may reflect the need for more prior training.

After characterizing clusters 2 and 5 as the most likely to contain the sands of interest, it is important to study their location throughout the area of interest. Figure 13 displays a comparison of results concerning the presence of sands, through the application of an opacity filter similar to the one used for the seismic inversions.

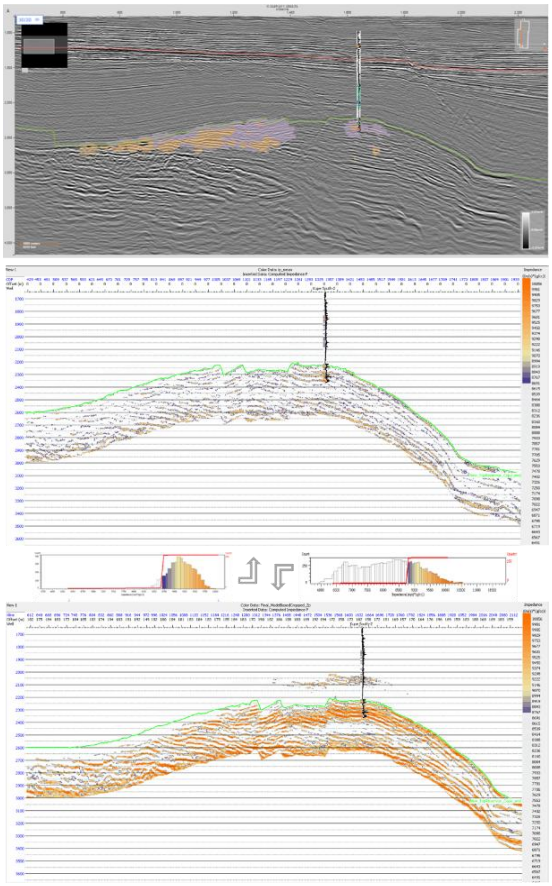

Figure 13. Comparison between clusters 2 and 5, the stochastic inversion and the deterministic inversion, respectively, displaying the well-log P-impedance of the Kupe South-2 well and the opacity filter applied on the color key (inline 2777).

The presence of clusters associated to the reservoir sands on the left side of the clustering model (Figure 13) is probably the most dubious aspect of these results. In the seismic

inversions, the presence of high P-impedances associated to sands does not dismiss the theory of having sands on the left side of the section, but they certainly do not highlight that region when compared to the well location. Once the literature only describes superficially the Farewell formation on the Kupe field and there are no wells available around that area, the presence of a reservoir on the left side cannot be confirmed nor dismissed by hard data.

Figure 14 displays clusters 1 and 6 in order to compare their location with the results of both seismic inversions with an opacity filter applied to enhance the lower P-impedances.

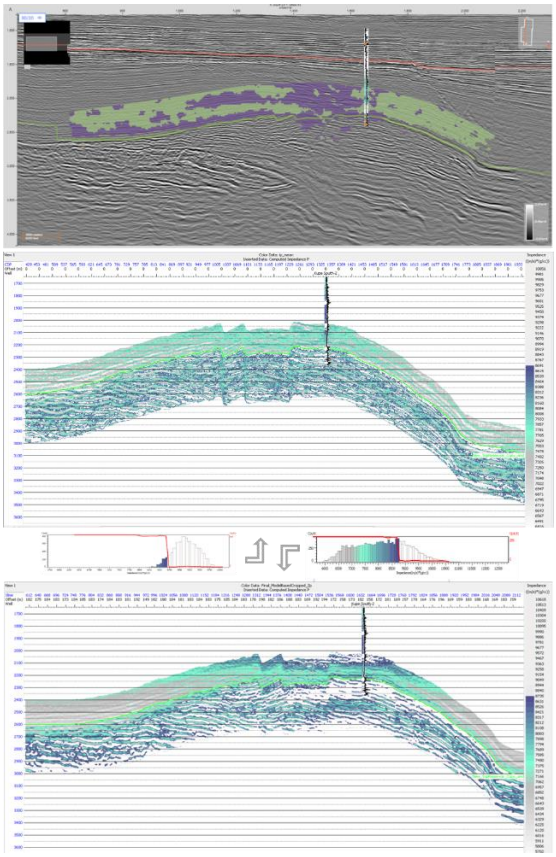

Figure 14. Comparison between clusters 1 and 6, the stochastic inversion and the deterministic inversion, respectively, displaying the well-log P-impedance of the Kupe South-2 well and the opacity filter applied on the color key (inline 2777).

Figure 14 compares clusters 1 and 6 to the lower P-impedance values of both models. This comparison validates the high contents of shale compatible to a mudstone in the zone above the interpreted horizon.

Conclusions

In this work two different seismic inversion methods were explored and compared to the results of an automated workflow.

The results obtained in the seismic inversions performed fulfilled the geological description mentioned in the literature and presented two valid models of the subsurface. Even though both exhibited good and identical correlations to the original seismic data, the stochastic model presented the most realistic prediction.

Despite the quality of the seismic data used, GSA presented results that reasonably fitted the seismic inversions as well as the geological information described in the literature. Some of the results generated by the GSA and seismic inversions lacked confirmation from hard data, such as the presence of sands of interest away from the existing wells.

The use of GSA had several positive impacts in this thesis, such as:

- Validated the seismic interpretation performed since the results portray the distinction of the Otaroa and Farewell formations, matching the horizon interpreted;
- Suggested a separation and classification of the area of interest that fitted the expected results;
- Provided new information about the area of interest and confirmed some of the results predicted by the seismic inversions, which should minimize the overall risks associated with the interpretation of data.

Future Work

The implementation of ML methodologies in seismic interpretation processes is a recent approach and it is still in the early stages, which means that there is a large room for improvement.

GSA is a prototype that uses several simple supervised and unsupervised ML methodologies, such as GMM, K-means and SVM. Currently, more complex techniques like programmable neural networks have been implemented in the area and the results are far more promising. In the future, GSA could benefit from incorporating these techniques, such as CNN, that have the potential to improve the results.

Another interesting approach for future work would be the incorporation of seismic inversions to the system as an input, to further analyze the results obtained.

Bibliography

[1] I. Priezzhev, P. Veeken, S. Egorov e U. Strecker, "Direct prediction of petrophysical and petroelastic reservoir properties from seismic and well-log data using nonlinear machine learning algorithms," *The Leading Edge,* vol. 38, pp. 949-958, 2019.

[2] F. Anifowose, A. Abdulraheem e A. Al-Shuhail, "A parametric study of machine learning techniques in petroleum reservoir permeability prediction by integrating seismic attributes and wireline data," *Journal of Petroleum Science and Engineering,* vol. 176, 02 2019.

[3] A. Almasgari e U. Hamzah, "Sequence stratigraphy of the pliocene deposits, Central Taranaki Basin, New Zealand," vol. 1784, p. 060001, 11 2016.

[4] S. T. Qadri, M. A. Islam, M. Shalaby e A. E. Haque, "Seismic interpretation and structural modelling of Kupe field, Taranaki Basin, New Zealand," *Arabian Journal of Geosciences,* vol. 10, 07 2017.

[5] J. Palmer e G. Bulte, "Taranaki Basin, New Zealand: Chapter 9," 1991.

[6] N. Jumat, M. Shalaby, A. E. Haque, M. A. Islam e L. Hoon, "Geochemical characteristics, depositional environment and hydrocarbon generation modeling of the upper cretaceous Pakawau group in Taranaki Basin, New Zealand," *Journal of Petroleum Science and Engineering,* vol. 163, 04 2018.

[7] W. Alotaby, "Fault interpretation and reservoir characterization of the Farewell Formation within Kerry Field, Taranaki Basin, New Zealand," Missouri University of Science and Tecnology, 2015.

[8] M. Webster, S. O'Connor, B. Pindar e R. Swarbrick, "Overpressures in the Taranaki Basin: Distribution, causes, and implications for exploration," *AAPG Bulletin,* vol. 95, pp. 339-370, 03 2011.

[9] L. Azevedo e A. Soares, Geostatistical Methods for Reservoir Geophysics, 2017.

[10] A. Soares, J. Diet e L. Guerreiro, "Stochastic Inversion with a Global Perturbation Method," 2007.

[11] R. Simm e M. Bacon, Seismic Amplitude: An Interpreter's Handbook, 2012, pp. 1-271.