Bulletin of Geophysics and Oceanography
Vol. XX, n. X, pp. XXX-XXX; Xxxxx 20xx
DOI 10.4430/bgo00533

Probabilistic neural network-based seismic attribute prediction of pore pressure: a case study in the Timimoun Basin, Algeria

K. Zabel[1], R. Ketteb[1], M.C. Berguig[1], S. Amoura[1] and Y. Kiche[2]

[1] *Geophysics Laboratory, Universite des Sciences et de la Technologie Houari Boumediene, Algiers, Algeria*
[2] *Datageo Solutions, Pau, France*

(Received: 29 September 2025; accepted: 25 March 2026; published online: 31 August 2026)

ABSTRACT This study aims to estimate the subsurface pore pressure (P_p) within the Frasnian shale in the Algerian Timimoun Basin, where there have been no prior studies. Seismic and well-log data were incorporated to assess the lateral distribution of P_p. A probabilistic neural network (PNN) model was implemented based on its capabilities to better integrate the diverse data sources and adapt to complex geological conditions. The model is trained using data from four wells, using the P_p log, and the composite traces of seismic inversion outputs, validated on one blind well and compared to a baseline method. Then, the model was extended from the boreholes to the entire study area. The application of the PNN algorithm exhibited a robust Pearson coefficient of correlation of 0.94 and 4 MPa as root-mean-square error, signifying a substantial alignment between the predicted and actual P_p values. This correlation emphasises that seismic data effectively encapsulates the physical implications of predicted volumes. The PNN-predicted P_p aligns with the structural patterns of the Frasnian shale in the subsurface, suggesting that the lateral distribution of P_p is influenced by these structural characteristics. The assessment demonstrates an abnormal subsurface P_p which poses a challenge to drilling operations for future well exploration or exploitation.

Key words: seismic inversion, pore pressure, machine learning, probabilistic neural network, shale formation.

1. Introduction

The accurate forecasting of gas content in shale reservoirs has emerged as a focal point of research within the fields of geology and geophysics. To enhance shale gas recovery, it is essential to accurately determine the gas content of shale reservoirs (Luo *et al.*, 2022). Pore fluids are unable to escape when they are trapped in impermeable rocks, such as shales, which are formed through sediment compaction. This results in an elevated formation pressure due to the weight of the entire overlying rock column, which is supported by these fluids. Pore pressure (P_p) analysis involves the study of pressure variations with depth, inferring essential insights into subsurface conditions, and facilitating the reliable determination of the safety mud weight window to maintain wellbore stability, thus ensuring the safety of drilling operations and offering critical inputs for well design (Wang *et al.*, 2022; Das and Maiti, 2023; Weng *et al.*, 2025). The inversion of seismic reflection data into subsurface properties such as P_p is the standout of a powerful geophysical method named the pre-stack simultaneous seismic inversion process.

Seismic inversion methods are widely utilised to estimate attributes such as P impedance (*Ip*), S impedance (*Is*), density (*d*), P-wave and S-wave velocities (*Vp* and *Vs*), and other elastic properties. These attributes help to understand subsurface lithology and fluid content (Kushwaha *et al*., 2020).

Pandey and Sain (2021) evaluated the lateral variation of porosity on the basis of two-dimensional (2D) seismic reflection data, using two different approaches based on acoustic impedance (AI) transformation and direct seismic inversion. The latter approach provides satisfactory results along seismic lines compared to acoustic impedance transformation. Li *et al*. (2016) developed a strategy to infer the properties of subsurface geology (such as elastic and petrophysical properties) on the basis of pre-stack seismic data obtained with Gassmann equations with deterministic optimisation techniques. Charkhi *et al*. (2019) proposed a straightforward and applied methodology to estimate reservoir properties from seismic inversion, demonstrating an alternative approach of rock physics requiring a considerable number of predefined constant parameters, which makes their use complicated in practice. Moreover, the effective pressure within geological formations is closely related to seismic velocities (Carcione and Tinivella, 2008). By analysing these velocities, particularly through acoustic impedance, geophysicists can infer P_p levels. Moreover, due to the reduced effective stress on the rock matrix, the P_p increases while the seismic velocity decreases. Consequently, using traditional methods for predicting P_p is related to geological conditions and empirical relationships (Hottmann and Johson, 1965; Liu *et al*., 2018). These relations are often determined from historical data and can vary significantly by region. In addition, these methods may lack the capabilities of advanced technologies in terms of accuracy. The incorporation of artificial intelligence in oil and gas domains, particularly machine learning (ML), has reflected a significant advancement by promising potential in overcoming traditional methods and enhancing various aspects of research and applications. In other words, by leveraging the performance of ML, researchers aim to address challenges, enhance predictive accuracy, and optimise operational processes, thus contributing to improved efficiency and safety in the extraction and utilisation of hydrocarbon resources (Oshim *et al*., 2023; Stefanou and Darve, 2024).

Huang *et al*. (2022) integrated the effective stress theorem into a new method using ML algorithms to predict P_p, where the suggested ML models provide reliable accuracy. Xinyu *et al*. (2025) developed a preliminary formation P_p model by improving the accuracy of pre-drilling P_p in carbonate rocks, considering the abnormal pressure. Radwan *et al*. (2022) predict an accurate subsurface P_p from a limited range of well logs using ML tools instead of empirical relationships. Jin *et al*. (2024) predicted P_p by integrating rock physics models and deep neural networks (DNNs) into the Eaton method. Seismic-inverted data was used to calculate the normal compaction velocity, taking into account the effects of porosity, fluid types, and lithology. This innovative approach marked the capability and the reliability of deep learning compared to empirical relationships. Furthermore, Krishna *et al*. (2024) evaluated several regression algorithms, including decision tree, extreme gradient boosting, random forest (RF), recurrent neural network, and convolutional neutral network for P_p prediction using a dataset from five wells. Their analysis demonstrated that RF is the best-performing predictive model with a Pearson coefficient of correlation (*r*) of 0.905 and a root-mean-square error (*RMSE*) of 6.48 MPa. Kushwaha *et al*. (2020) estimated acoustic impedance, porosity, density, gamma-ray volume, and P-wave velocity by implementing model-based inversion and geostatistical methods including probabilistic neural network (PNN), multi-layer feed-forward neural network (MLFFN) and radial basis function neural network (RBFNN). The analysis shows that among the geostatistical methods, the PNN provides slightly better results.

Meanwhile, several studies have shown that PNNs achieve high correlation coefficients between predicted and actual reservoir properties, indicating superior predictive performance over other neural network types such as MLFFNNs or RBFNNs (Haris *et al*., 2017; Akbari and Riahi, 2020). This highlights that, although RF performs well in certain contexts, PNNs provide a powerful alternative due to their probabilistic approach and strong accuracy in reservoir property prediction. Previous studies in Algerian basins have used conventional methods to estimate P_p at well locations. In the Illizi Basin, in the Takouazet Field**,** the analysis indicates an abnormal P_p induced by disequilibrium compaction in Silurian marine shales. The P_p drops sharply from Silurian to Ordovician sediments (Baouche *et al*., 2020a, 2020b). However, in the Tin Fouye Tabankort Field in the same basin, a sub-normal pressure regime was identified in the Ordovician reservoir (IV-3 unit) caused by reservoir depletion due to prolonged hydrocarbon production (Laouini *et al*., 2025). In the Ahnet Basin, an overpressure regime against the hot shale unit with a gradient of 0.58 psi/ft was caused by the high in situ temperature (Bouchachi *et al*., 2022). Furthermore, the same regime was observed in the Berkine Basin (Saadaoui *et al*., 2025).

The objective of the current work is to establish the extension of the subsurface P_p to prevent drilling issues within the Frasnian-aged shale in the Timimoun Basin, where there have been no prior studies, based on seismic-driven data and ML methods. This approach is intended to evaluate well-log data and the acoustic parameters given by the pre-stack seismic inversion process in predicting P_p and enhance the accuracy using the capabilities of ML techniques, particularly those of the PNN. The research seeks to improve the prediction of subsurface characteristics and enable the extension of P_p predictions from boreholes to the entire volume of the study area, by integrating seismic cubes (*Ip*, *Is*, *Vp*, *Vs*, *Vp*/*Vs*, and *d*).

2. The geologic setting of the study area

The city of Timimoun, located within the Gourara region of northern Algeria, is situated within a distinct geomorphological framework. It is limited by the Grand Erg Occidental, an extensive sand sea characteristic of the Saharan Desert, to the north. The south-western margin is defined by the Saoura Valley and Taout regions, whereas the south-eastern boundary is marked by the Tademait Plateau, a Palaeozoic sedimentary uplift (Ben Charif and Belakehal, 2024). The inclusive geological setting of the area includes a series of regional structural trends that segregate sedimentary basins. These structural features (Figs. 1 and 2), which formed during major tectonic episodes such as the Hercynian orogeny, influence the spatial distribution of lithostratigraphic units and basin architecture (Klett, 2002).

In the study area, the stratigraphic series spans a Palaeozoic sequence, which overlies the Precambrian basement and exceeds 4,000 m in thickness, to Cenozoic intervals (Fig. 3). The Palaeozoic strata were folded Cambrian to Carboniferous strata deformed during the Hercynian orogeny (Cochran and Petersen, 1999). Structural succession controls segmentation and anisotropic stress, which are vital to determining the P_p distribution. Lower Silurian marine clays and Upper Devonian (Frasnian) organic-rich mudstones, deposited during major transgressive events, serve as regional seals. These low permeability units restrict vertical fluid migration, promoting overpressure development in underlying reservoirs. During thermal maturation, the Frasnian mudstone, which is related to organic-rich sediment, was associated with hydrocarbon generation. Thus, it is a potential driver of elevated P_p (Daniels and Emme, 1995; Makhous *et al.*, 1997). Spatial variability in the organic richness and thermal maturity of source rocks reflects palaeogeographic controls, with burial history and lithologic heterogeneity directly impacting

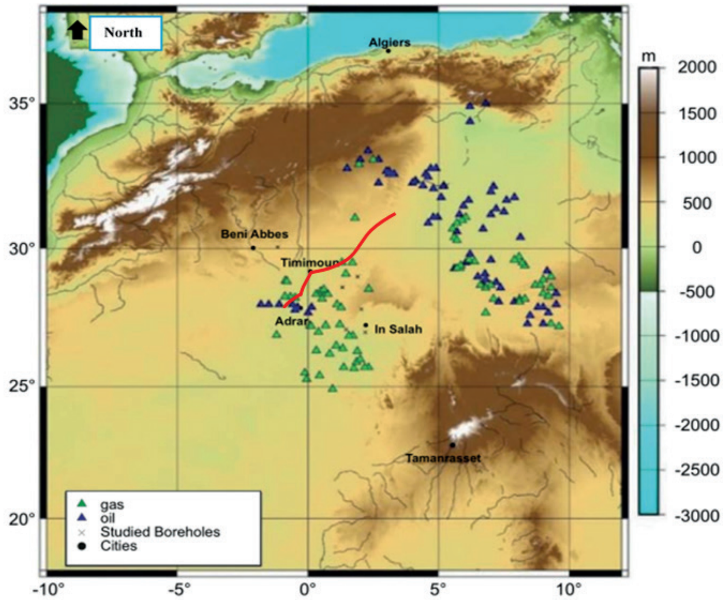

Fig. 1 - Situation of Timimoun, and the trace of the NE–SW cross-section (Askri *et al.*, 1995).

compaction disequilibrium (Daniels and Emme, 1995, Green and Vernik, 2020). Reservoir-seal pairs that influence pressure sectionalisation are created by interbedded Devonian sandstone–mudstone sequences. Tectonic features derived from Hercynian deformation, including fault systems and structural highs, further modulate (Fig. 2) fluid migration pathways (Klett, 2002). Within Silurian–Devonian units, the total organic carbon (TOC) content and vitrinite reflectance data show that thermal maturation is sufficient for hydrocarbon generation, which is correlated with regional trends documented in analogous basins (Makhous *et al.*, 1997; Lüning *et al.*, 2003). In structurally complex settings, these processes coupled with lithologic variability underpin predictive models for the distribution of P_p.

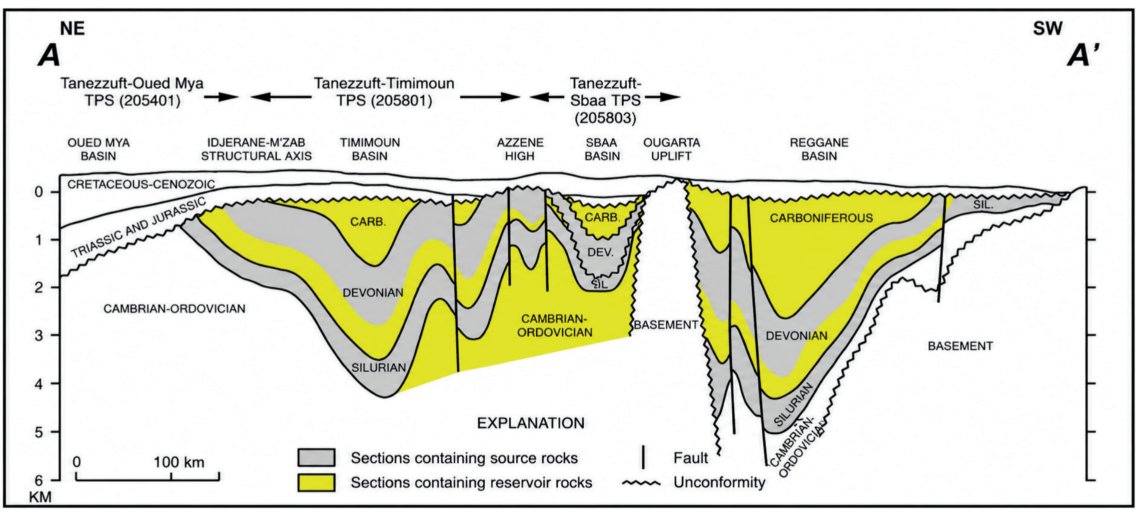

Fig. 2 - NE–SW stratigraphic cross-section through the Oued Mya, Timimoun, Sbaa, and Reggane basins (modified from Aliev *et al.*, 1971; Makhous *et al.*, 1997).

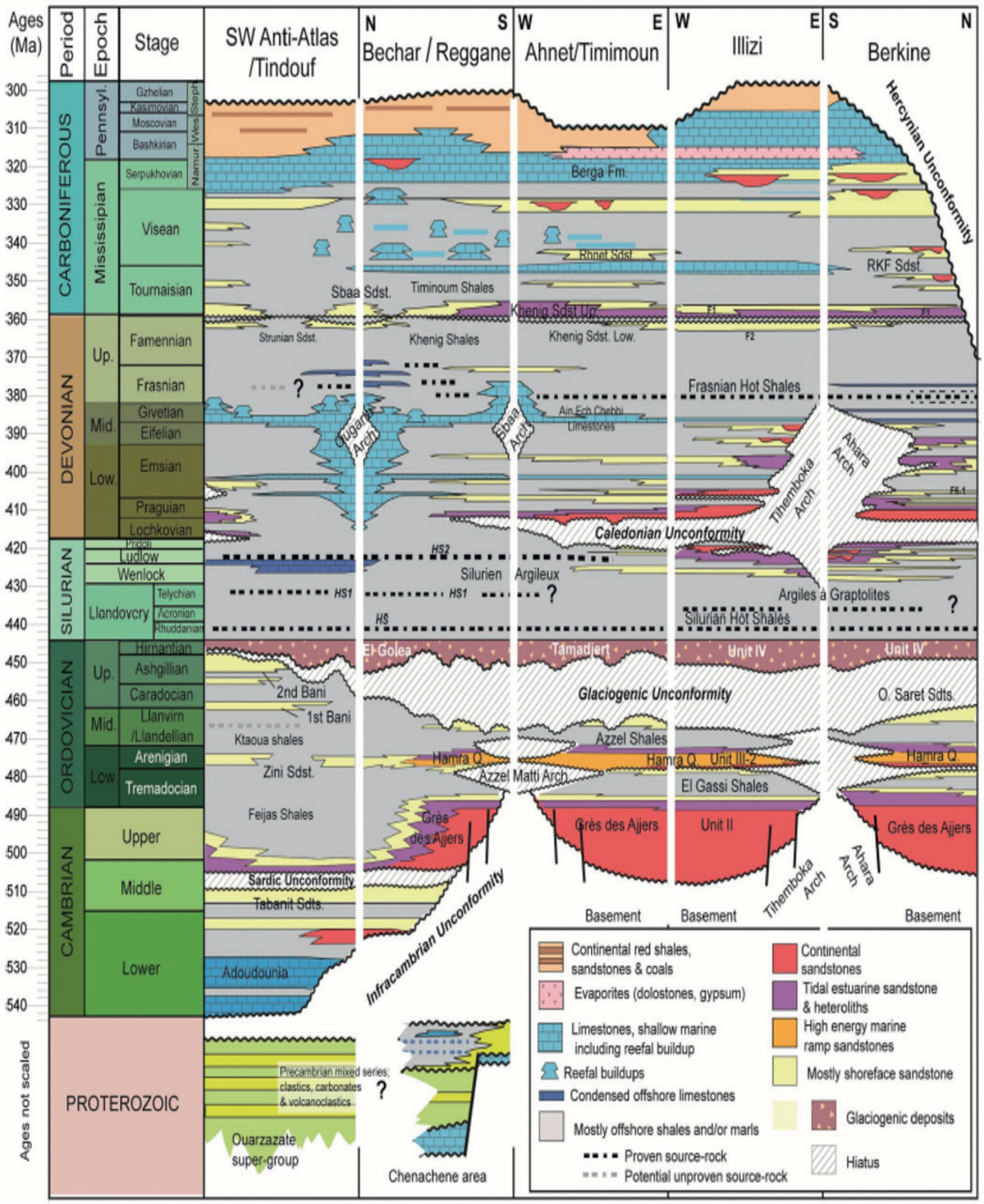

Fig. 3 - Chronostratigraphic chart on the scale of the Algerian Saharan basins (Lopez, 2025).

3. Methodology

In this research, a three-dimensional subsurface P_p model was predicted by analysing well-log data with pre-stack seismic inversion outputs including *Ip*, *Is*, *Vp*, *Vs*, *Vp*/*Vs*, and *d* using Emerge via Hamson Russel software. Fig. 4 represents the proposed workflow. Emerge is a program that analyses well-log and seismic data by finding a relationship between them at the well locations.

The P_p logs for four wells were estimated using the Eaton method, [Eq. (1)]. Then, these logs, along with the extracted seismic traces from each seismic inversion output, served as training data for the PNN model.

When a probabilistic network is training, it is finding the set of sigma, which refer to the smoothing parameters used in the radial basis function kernels at the pattern layer of the network. The training process aims to find the set of sigma values that minimise the validation error (Hampson *et al.*, 2001). This is a nonlinear optimisation problem, which Emerge attempts to solve in two stages.

The first stage is to determine the best single sigma, assuming that all the different sigmas are equal. Once this stage is completed, the second stage is to determine the individual sigma using a conjugate-gradient analysis with the first global sigma as the starting point. Then, the optimum set of sigma is determined and used for prediction. Furthermore, cross-validation is a good, preventative check for overtraining. It requires reserving some data out of the training process, losing what this data could add to the training process to gain what it can correct through validation. These skipped points would be a well log. The well that is skipped is called the validation well (or blind well). This process is usually carried out separately from the original training (which can, then, use all the data). Following successful validation, the model was applied across the entire field to predict subsurface P_p using seismic data. On the other hand, a baseline method was used to validate the model by extracting P_p traces from the predicted P_p seismic volume and comparing them to the P_p log-based density method.

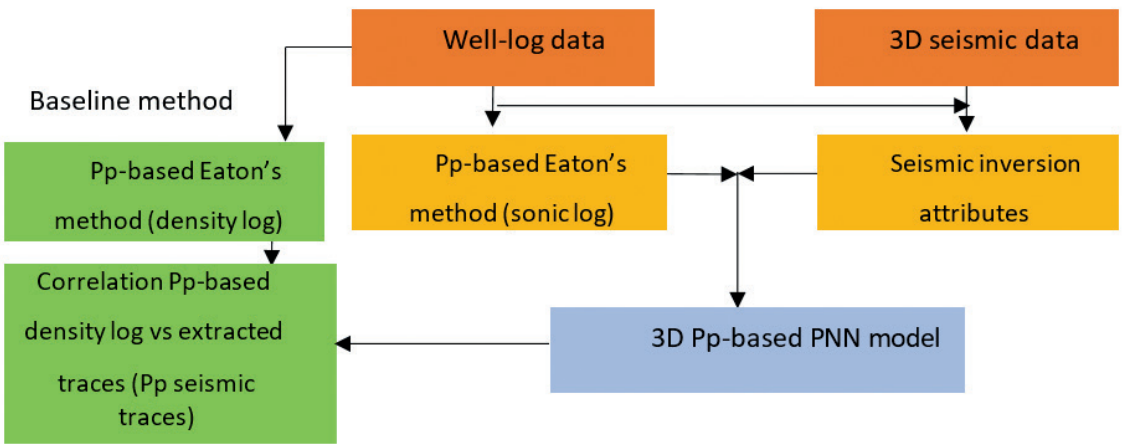

Fig. 4 - Proposed workflow of estimation subsurface P_p.

3.1. Overview of model based seismic inversion

Reflected seismic traces (acoustic signals) are inverted into rock properties by the pre-stack seismic inversion process. This process involves analysing the amplitude and phase of these signals to derive parameters such as *Ip*, *Is*, *Vp*, *Vs*, *Vp*/*Vs*, and *d*. Using algorithms, geophysicists can gain insights into the subsurface characteristics of geological formations (Banerjee and Chatterjee, 2022). Using Zoeppritz (1919) equations, particularly the Fatti *et al.* (1994) approximation, seismic response of geological layers can be calculated with varying elastic properties, such as velocities and densities, across multiple acquisition angles. These elastic properties provide details of reservoir characterisation (Banerjee and Chatterjee, 2022). As illustrated in Fig. 5, the standard workflow in model-based pre-stack inversion includes seismic and well data as input data. The processing of required pre-stack seismic data is essential to enhance the signal-to-noise ratio for improving the overall quality of the data. The required processing involves correcting irregularities by applying trim static process and eliminating random noise using a Radon filter (Bancroft *et al.*, 2000). The next step in the seismic inversion process is seismic-to-well tie. The primary goal of this process is to ensure that the seismic data accurately represent the geological features indicated by the well logs, thereby enhancing the reliability of subsurface interpretations (Russell, 2014). The seismic-to-well tie involves aligning events from a synthetic seismogram generated from well data with corresponding events observed in a seismic trace at the well location. This procedure is performed after generating angle stacks (near stack range 02°–20°, middle stack range 18°–30°, and far stack range 30°–45°) and extracting wavelets (reverse statistical wavelets were initially extracted using angle stacks followed by another extraction using wells) (Figs. C1 and C2 in Appendix C). The next step involves well data and structural interpretations (horizons) to build a low frequency model which aims to supplement the low frequency component for enhancing the initial volumes of acoustic impedance, shear impedance, and density (Barclay *et al.*, 2007). During the iterative process of adjustment, an optimal fit between the observed seismic data and the modelled predictions has to be achieved and various parameters have to be fine-tuned to ensure that the model accurately represents the geological features indicated by

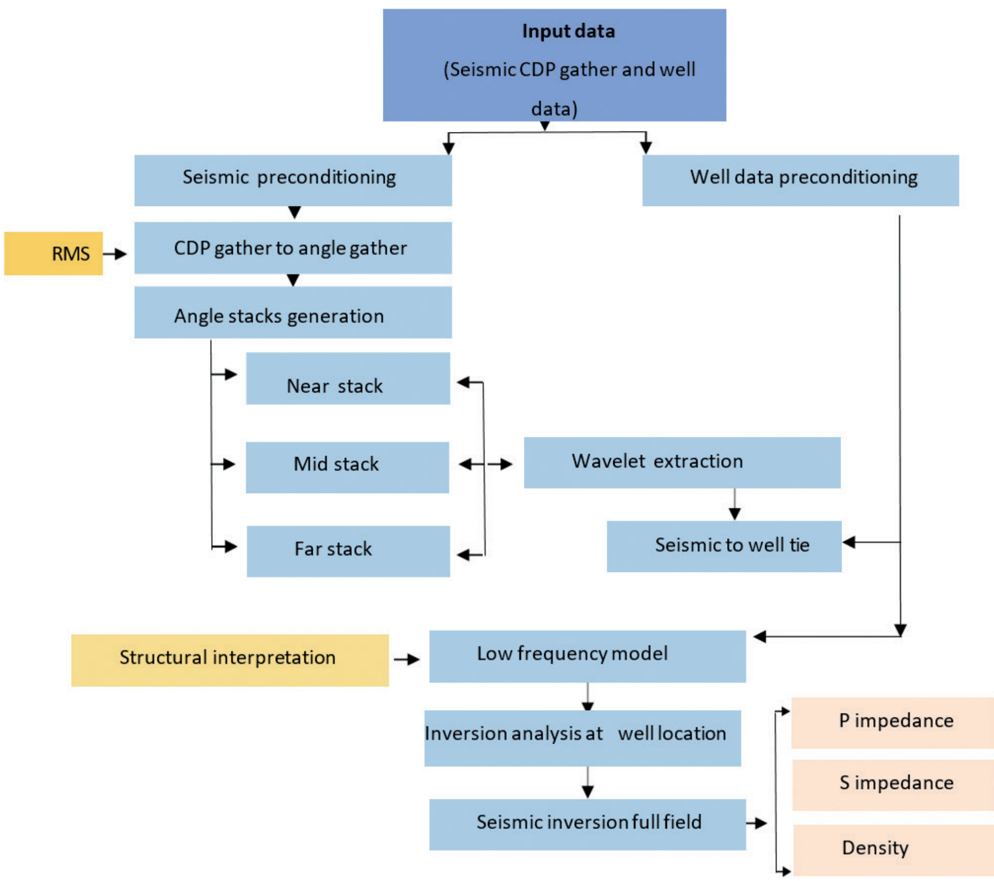

Fig. 5 - Pre-stack seismic inversion workflow.

the well logs (regularisation) (Di Luca Vingelli, 2016). Then, the optimal parameters are selected and applied into the entire seismic volume, providing subsurface properties such as the acoustic parameters cubes (*Ip*, *Is*, *Vp*, *Vs*, *Vp*/*Vs*, and *d*) (Betancourt *et al*., 2016).

3.2. Pore pressure log estimation

P_p is fluid pressure trapped in pore spaces of a porous formation [Eq. (1)]. It can range from hydrostatic pressure [Eq. (2)] to extreme overpressure. Abnormal pressure refers to P_p that deviates significantly from the hydrostatic pressure gradient. It can be classified either as abnormal low pressure or abnormal high pressure (Wang and Ma, 2024).

Various methods are used to estimate P_p, with the Eaton method [Eq. (1)] being the most commonly employed and chosen as it is a reliable approach that uses vertical stress *Sv* to determine P_p [Eq. (3)] (Das and Maiti, 2023):

$$P_p = S_v - (S_v - P_h) \times \left(\frac{DT_n}{DT}\right)^3 \tag{1}$$

$$P_h = \int_0^Z gz \tag{2}$$

$$S_v = \int_0^Z \rho gz \tag{3}$$

where *DT* represents the observed values of the sonic transit time log, DT_n is the value of a normal compaction line for sonic transit time log, *Ph* is the normal hydrostatic pressure, *P* is the density at a depth *z*, and *g* is the gravitational constant. However, the Eaton method can also be applied to density logs for P_p estimation by using the deviation of the measured formation bulk density from a normal compaction trend of density versus depth. The principle is similar to the sonic log application but uses density data to infer abnormal P_p:

$$P_{p_{based\ density}} = S_V - (S_V - P_h) \times \left(\frac{d_{trend}}{d_{measured}}\right)^{1/2} \tag{4}$$

where d_{trend} is the density of the formation expected under normal pressure and $d_{measured}$ is the recorded log. This method was used as a baseline model affirming the model's robustness and validity.

3.3. Probabilistic neural network

PNN is defined as a type of ML algorithm used for classification and pattern recognition tasks (Specht, 1991). The PNN estimates the joint probability density function of input-output pairs and derives the conditional expectation of the output given the input. Given a training dataset with input-output pairs (x_i, y_i), where $x_i \in R^d$ and $y_i \in R^d$ are the conditional expectation (regression output) for a new input, *x* is estimated as (Nadaraya, 1964; Watson, 1964):

$$\hat{y}(x) = E(y|x) = \sum_{i=1}^{N} y_i\, K(x, x_i) / \sum_{i=1}^{N} K(x, x_i) \tag{5}$$

where *N* is the number of training samples, $K(x, x_i)$ is the kernel function measuring similarity between input x and training sample x_i. Typically, K is a Gaussian kernel and σ is the smoothing parameter (bandwidth) controlling the kernel width:

$$K(x, x_i) = e^{(-|x-x_i|^2/2\sigma^2)}. \quad (6)$$

The numerator in Eq. (5) sums outputs y_i weighted by the similarity between x and each training point x_i. The denominator normalises the weights. Thus, the output is a weighted average of the training outputs. This approach is also known as the Nadaraya–Watson kernel regression and is a non-parametric regression method. The architecture of the PNN regression model consists of an input layer which receives input vector x, a pattern layer, where each neuron corresponds to a training sample and computes kernel function $K(x, x_i)$, and a summation layer which computes the weighted output sum $\sum_{i=1}^{N} y_i K(x,x_i)$ and kernel weights sum $\sum_{i=1}^{N} K(x,x_i)$.

The output layer that divides the weighted sum of outputs by the sum of kernel weights to produceis illustrated in Fig. 6.

3.4. Input data description and preparation

Before constructing the model, it is crucial to evaluate the impact of each parameter on the target variable, specifically P_p. This involves selecting the input data in relation to the desired output. The selection of training log data is based on their strong or weak correlation with the actual P_p log and is assessed through r, which assesses the strength of the linear relationship between two variables:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (7)$$

where n is the number of paired data points, x and y are the individual sample points for variables X and Y, $\sum xy$ is the sum of the product of paired values, $\sum x$, $\sum y$ are the sums of the x and y values,

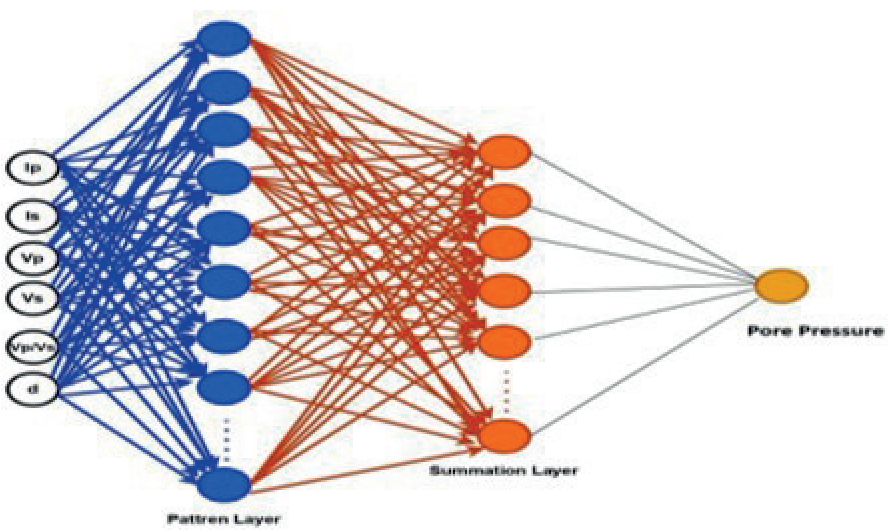

Fig. 6 - PNN model architecture.

respectively, and $\sum x^2$, $\sum y^2$ are the sums of the squares of *x* and *y* values.

This statistical metric quantifies the degree of relationship between two variables and ranks the importance of features. Furthermore, in this paper, 1,167 datasets are collected from four wells as training dataset (Well A, Well B, Well C, and Well E). Fig. 7 illustrates a scatter plot comparing actual P_p logs with the mentioned logs. The analysis showed a positive correlation and identified outliers that should be removed. Moreover, the correlation coefficients are further detailed in Fig. 8, which includes a heatmap demonstrating that the input features exhibit a good correlation with the P_p log. The correlation between P_p and Vp is $r(P_p, Vp)$ = 0.85, $r(P_p, Ip)$ = 0.81, $r(P_p, Vs)$ = 0.72, $r(P_p, Is)$ = 0.7, $r(P_p, Vp/Vs)$ = 0.69, and $r(P_p, d)$ = 0.52. In addition, the contribution of each well is depicted in same Fig. 8, where Well A and Well C represented 22.15% of the training dataset and 27.85% for both Well B and Well E. One unseen well (Well D) with 327 data is used to validate the model. Table 1 summarises key statistical metrics for the training dataset measured over 1,167 samples. The depth ranges from 1,378 to 1,580 m, with an average of 1,477 m representing what is a typical value of the depth within the studied interval. Relatively consistent lithology is indicated by the average value of density (2.374 g/cc), showing low variability. Moreover, *Ip* reflects substantial heterogeneity in rock elastic properties, as it has a relatively high standard deviation (STD) of 1366.58 (m/s) × (gr/cc).

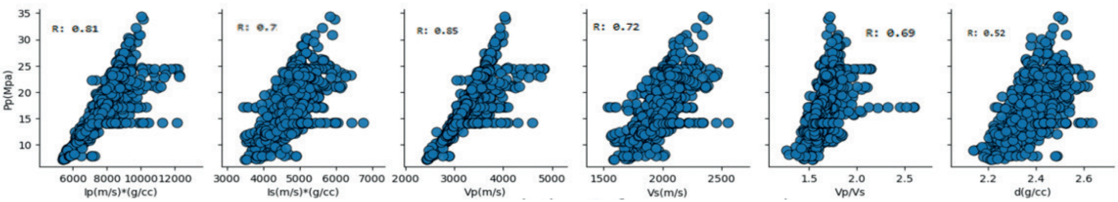

Fig. 7 - Cross-plot of P_p log versus log data before preprocessing.

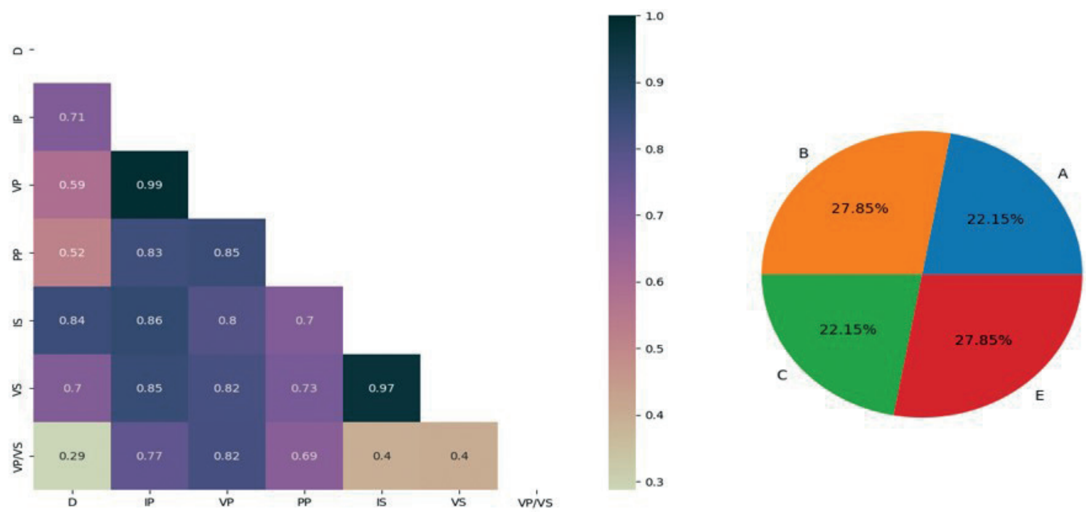

Fig. 8 - Bar plot and heatmap of feature correlation with P_p log and the contribution of each well in training the model.

Table 1 - Statistical description of the training dataset.

	d (gr/cc)	*Ip*(m/s)×(gr/cc)	*Vp*(m/s)	P_p (Mpa)	*Is*(m/s)×(gr/cc)	*Vs* (m/s)	*Vp*/Vs	*d* (gr/cc)
Count	1167.000	1167.000	1167.000	1167.000	1167.000	1167.000	1167.000	1167.000
Mean	2.374	8074.345	3389.810	30.818	4752.606	1997.036	1.694	2.374
STD	0.083	1366.583	490.500	2.290	551.938	176.537	0.153	0.083
Min	2.126	5369.271	2433.776	25.172	3431.216	1534.482	1.260	2.126
25%	2.319	7071.186	3013.075	28.691	4389.634	1879.440	1.594	2.319
50%	2.373	7928.483	3327.914	32.273	4714.557	1982.565	1.672	2.373
75%	2.431	8881.014	3681.416	34.681	5054.311	2097.736	1.766	2.431
Max	2.626	14245.677	5452.916	40.339	6930.854	2647.924	2.621	2.626

Conversely, lower STD values, such as 0.083 (g/cc) for density log, suggest more uniformity. The distribution of P_p can be characterised using percentile statistics. The 25th percentile (25 MPa) represents the lower-pressure zones, the median value (32 MPa) reflects the central tendency of the dataset, and the 75th percentile (40 MPa) identifies the higher-pressure zones. Before building the model, the dataset must be normalised to help better understand the relationship between input and output data and to expedite the training process. Then, the Isolation Forest algorithm is utilised to remove outliers from the dataset as depicted in Fig. 9. Isolation Forest is an unsupervised anomaly detection algorithm that isolates outliers by recursively partitioning data using random splits. The contamination key parameter used in this work is 0.08 after multiple trials and errors, which means the model assumes that 8% of the data are outliers. After

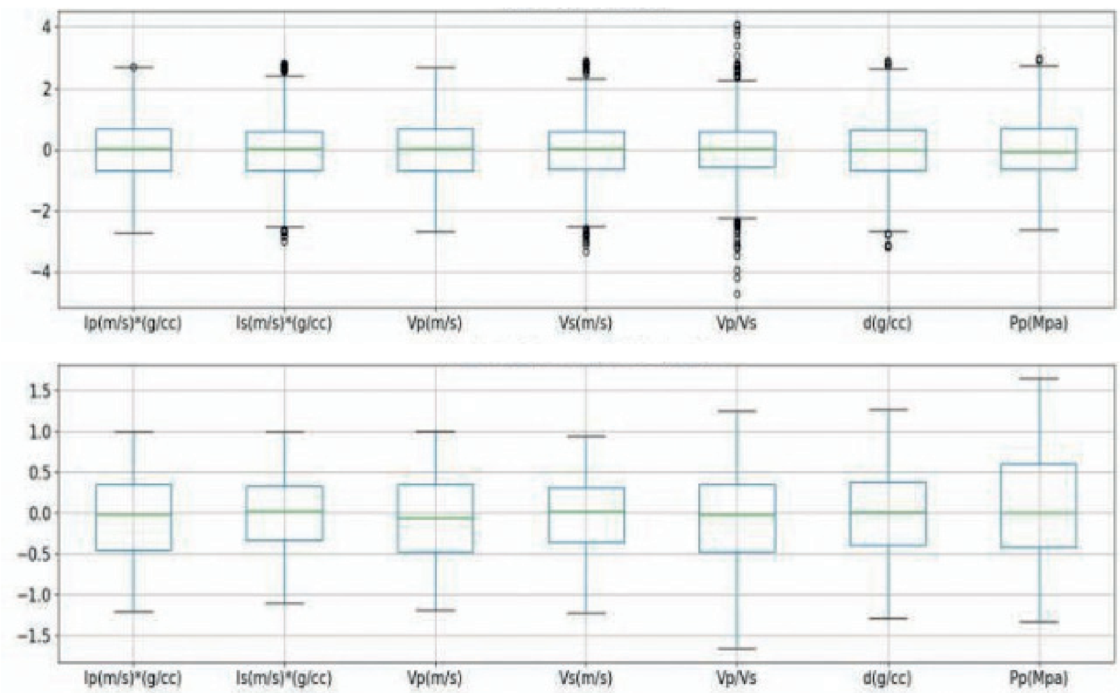

Fig. 9 - Box plots of the dataset before and after outlier removal.

removing outliers, the dataset was 1,080 samples. Fig. 10 illustrates a cross-plot of the prepared dataset after removing outliers and Table 2 summarises the statistical description of the training dataset after normalisation and outlier removal.

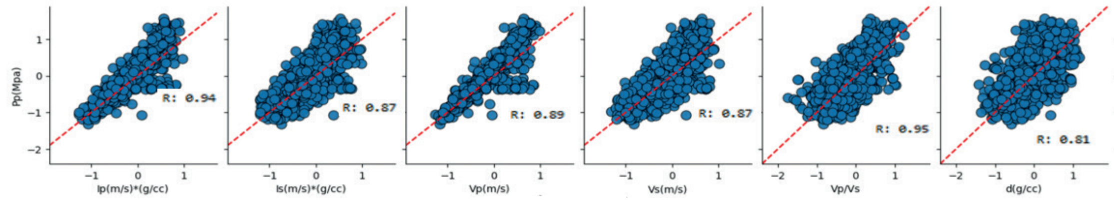

Fig. 10 - Cross-plot of P_p log versus input well-log data after outlier removal.

Table 2 - Statistical description of the training dataset after normalisation and outlier removal.

	d (gr/cc)	*Ip*(m/s)×(gr/cc)	*Vp*(m/s)	P_p (Mpa)	*Is*(m/s)×(gr/cc)	*Vs* (m/s)	*Vp*/Vs
Count	1080.000	1080.000	1080.000	1080.000	1080.000	1080.000	1080.000
Mean	-0.055	-0.083	-0.092	-0.063	-0.076	-0.042	0.058
STD	0.555	0.511	0.520	0.503	0.506	0.517	0.659
Min	-1.354	-1.204	-1.241	-1.221	-1.259	-1.397	-1.257
25%	-0.477	-0.492	-0.516	-0.413	-0.425	-0.453	-0.447
50%	-0.046	-0.049	-0.082	0.004	-0.015	-0.010	-0.013
75%	0.362	0.330	0.323	0.315	0.304	0.327	0.546
Max	1.176	1.033	1.049	0.967	0.988	1.186	1.771

4. Results and discussion

As mentioned above, in this study, acoustic parameters (seismic attributes), including *Ip*, *Is*, *Vp*, *Vs*, *Vp*/*Vs*, and *d*, were inverted, specifically for a shale reservoir, using the pre-stack simultaneous seismic inversion process. At well locations, composite traces were initially extracted from each attribute.

During the PNN training phase, 25 sigma were tried within a specified range. Since the input data are normalised to a standard deviation of 1.0 before analysis, the best sigma is expected to be around 1.0 (ranging from 0.5 to 1.5). The number of iterations for the standard conjugate-gradient algorithm used to optimise the individual sigma around the global value (0.220) was 20 iterations, which is the stopping criteria. The reliability of the model was rigorously evaluated against a single unseen well from the same field to assess the performance of the PNN in predicting P_p. This validation was used after multiple trials and errors. However, leaving-one-well-out cross-validation (LOWO-CV) across the five wells yielded per-fold and aggregate *r*, *RMSE*, and mean average error (*MAE*) that indicated a weak performance (Figs. C1, C2, and C3 in Appendix C), showing *r* of 0.78, *RMSE* of 3.80 Mpa, and *MAE* of 2.58 Mpa (Table 3). The *RMSE* is used to recognise the outliers in the dataset and to evaluate the accuracy of the model:

$$RMSE = \sqrt{\sum_{i=1}^{n} \left(\frac{\hat{y}_i - y_i}{n}\right)^2} \quad (8)$$

$$MAE = \frac{1}{n}\sum_{i=0}^{n}|y_i - \hat{y_i}|. \quad (9)$$

Additionally, the average absolute median/percentile errors and percentage of dynamic range (Table C1 in Appendix C), suggested it did not provide the best estimation compared to using a single blind well (Well D). Thus, validation based on one blind well (Well D) revealed a robust correlation between the PNN-predicted P_p and the actual P_p log from the well with an R^2 of 94% and an *RMSE* of 4 MPa.

Table 3 - LOWO-CV across the five wells.

	Well A	**Well B**	**Well C**	**Well D**	**Well E**	**Average**
r	0.89	0.56	0.72	0.94	0.78	0.78
RMSE (Mpa)	4.50	3.83	3.41	4.00	3.28	3.80
MAE (Mpa)	4.09	3.12	2.25	1.09	2.33	2.58

Quality control of the predicted P_p was conducted by exporting a single trace at the well location. This trace was, then, compared to P_p logs estimated using the Eaton method, as illustrated in Fig. 11, where the exported trace of the PNN-predicted P_p is depicted by the red log (P_pressure trace), while the blue log represents the P_p log (P_pressure_f) after application of a frequency filter with a cutoff frequency of 20–60 Hz. A band-pass frequency filter is applied to address discrepancies in frequency domains between well-log data (high frequency) and seismic data (low frequency). In other words, the applied filter matched the frequency content (bandwidth) of the well-log data to that of the seismic data at the same location.

Additionally, the extracted seismic traces were compared to the baseline model (P_p-based density log) as depicted in Fig. 12. The results show a similar trend between the seismic traces and the P_p-based density log. On the other hand, this baseline method yielded a good correlation with the Eaton-based sonic P_p log (Fig. B1 and Table B1 in Appendix B).

Once validated, the PNN-predicted P_p was extrapolated across all seismic sections based on seismic inversion outputs. Fig. 13 showed the PNN-predicted P_p along an arbitrary line with the P_p-based density logs, which indicated that the PNN-predicted P_p aligned with structural patterns of the Frasnian-aged shale in subsurface (Horizon B to Horizon C). This correlation emphasises that seismic data effectively encapsulates the physical implications of predicted volumes. As a result, lateral P_p distribution is described by an overpressure, as it is given by a high value along the structural pattern.

The analysis revealed abnormally high predicted subsurface P_ps in Frasnian-age shale in the Timimoun Basin. The achieved accuracy highlights the PNN strength in making reliable P_p extrapolation even in areas lacking well control by capturing complex nonlinear relationships between well-log data and seismic data. The overpressure identified in the Frasnian-age shale within the Algerian Timimoun Basin can be attributed to several geological factors, such as compaction mechanisms, which indicate that P_p is affected by additional hydrocarbon generation processes (Daniels and Emme, 1995; Green and Vernik, 2020).

Greater overburden (vertical) stress compresses pore fluids more intensely as depth increases. High TOC values (around 10%) are another factor indicating a strong potential (up to 52 kg HC/t) for hydrocarbon generation (Daniels and Emme, 1995; Makhous *et al*., 1997; Lüning *et al*., 2003). However, the Frasnian shale is described as an excellent source rock (high TOC value obtained

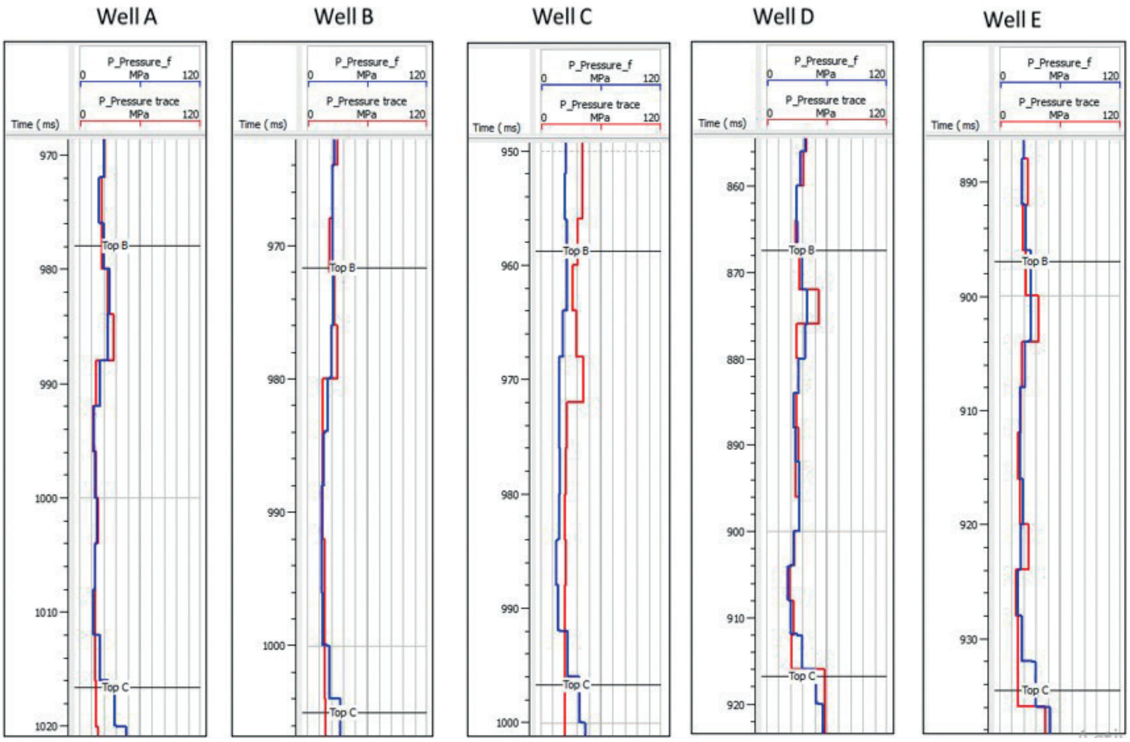

Fig. 11 - Comparison of the P_p trace trend (the red log) with the P_p filtered log (blue log) at well level.

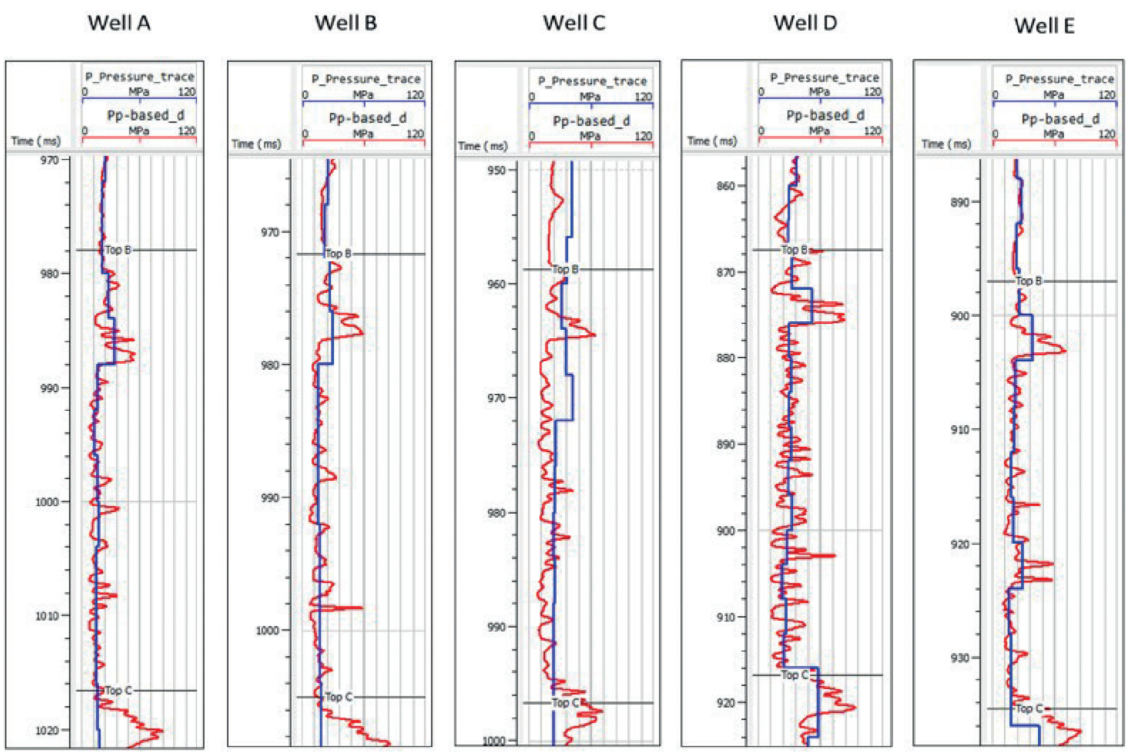

Fig. 12 - Comparison of the P_p trace trend (blue log) with the P_p-based density log (red log) at well level.

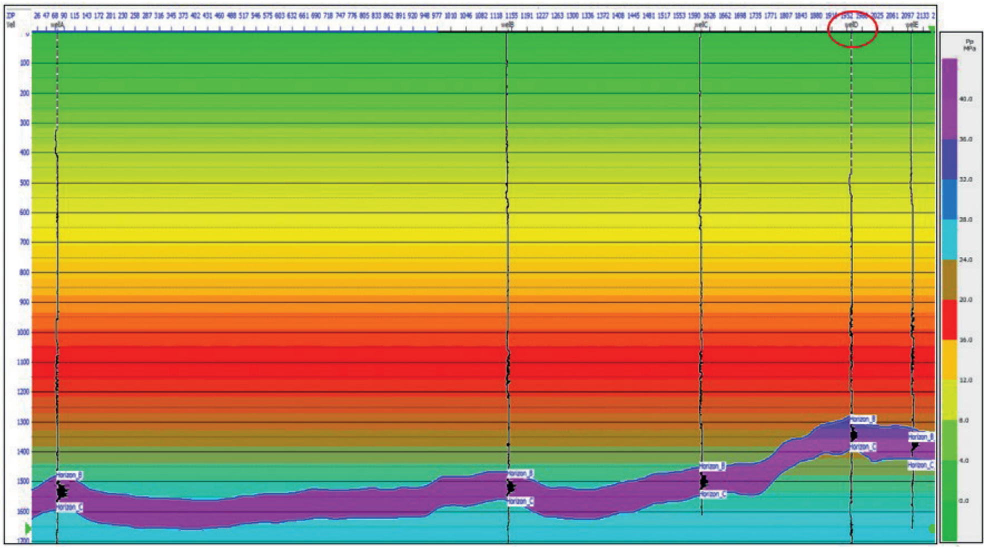

Fig. 13 - The PNN-predicted P_p along the arbitrary line in Frasnian-aged shale and the blind well (circled well).

using the Passey method, as shown in Table 1). These high values explicate the low density and sonic velocity log values (high sonic slowness log value) as illustrated in Figs. B2 and B3.

The TOC content in the shale exhibited a lower permeability, which made it harder for fluids to escape as the rock was buried and compacted. The trapped fluids contribute to retained P_p or overpressure. In other words, it has a considerable effect on the P_p by affecting the rock's ability to accommodate fluids due to the higher compressibility and lower density of the organic matter compared to that of the mineral matrix. Moreover, the ductile nature of organic matter can retain porosity, allowing higher P_p preservation.

These outcomes are affirmed by the strong positive correlation between the P_p seismic trace and the P_p-based density log (Fig. 13). The regional consistency of these findings is validated in adjacent Algerian basins, such as the Berkine, Illizi, and Ahnet, that showed similar overpressure trends (Baouche *et al*., 2020a; Bouchachi *et al*., 2022; Saadaoui *et al*., 2025). Fig. 14 shows the STD of the predicted P_p across the study area. Findings highlight a lower STD value (under 0.6

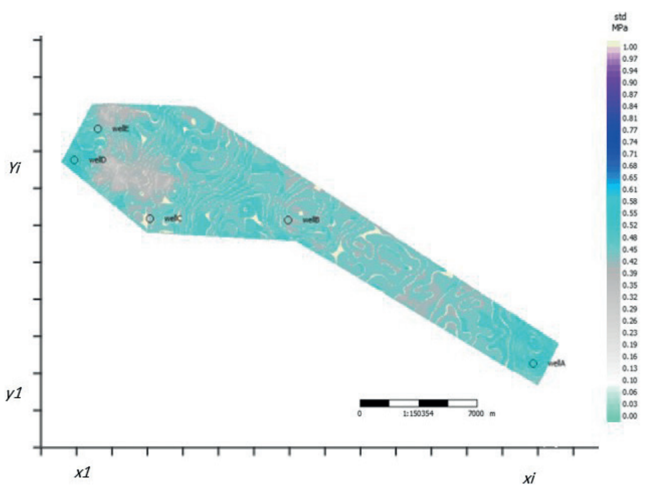

Fig. 14 - STD map of predicted P_p (Mpa).

Mpa), indicating high confidence and low uncertainty in the model's output. Such maps guide well placement and inform geomechanical, reservoir management decisions, highlighting where P_p predictions are reliable and where further data or model refinements may be necessary.

5. Conclusions

Abnormal pressure trends have been documented in Algerian basins such as the Berkine, Illizi, and Ahnet. These studies have employed conventional methods to estimate P_p at well locations. Far from the wells, P_p evaluation lacks spatial coverage and cannot detect other abnormal zones to mitigate drilling risks. The main purpose of the current work is to extend P_p prediction across block X in Frasnian age shale within the Timimoun Basin, where there had been no prior studies. The predicted P_p was achieved using the PNN model, which incorporated multi-source inputs, including well-log data and seismic inversion outputs. Firstly, the model was trained using four wells and validated on an unseen fifth well within the same field, achieving high prediction accuracy (r = 0.94, $RMSE$ = 4 MPa). The model was, then, compared with a baseline method and subsequently extended to the entire study area. Thus, an abnormally high predicted subsurface P_p was identified, and the regional consistency of the findings was validated in adjacent Algerian basins. These findings indicate that PNN techniques have significant potential to enhance P_p predictions, enabling ML models to provide rapid and reliable insights using readily available well-log data from multiple boreholes and seismic data, showcasing the transformative power of these techniques in the oil and gas industry.

Acknowledgments. The authors would like to thank the responsible members of DEGG for their valuable insights during the data interpretation phase. They also wish to express their appreciation for the support and assistance provided by the Geophysics Laboratory team at USTHB University. Special thanks to the anonymous reviewers for their constructive comments that improved the quality of this manuscript. This research received no specific financial support. The data used in this study are confidential and were handled in accordance with the relevant agreements.

REFERENCES

Akbari M. and Riahi M.A., 2020: *Estimation of reservoir quality from multi-attribute analysis by using a probabilistic neural network: case of Sarvak formation in an offshore oil field*. Boll. Geof. Teor. Appl., 61, 539–554, doi: 10.4430/bgta0330.

Aliev M., Aït Laoussine N., Avrov V., Aleksine G., Barouline G., Lakovlev B., Korj M., Kouvykine J., Makarov V., Mazanov V., Medvedev E., Mkrtchiane O., Moustafinov R., Oriev L., Oroudjeva D., Oulmi M. and Saïd A.; 1971: *Geological structures and estimation of oil and gas in the Sahara in Algeria*. Altamira-Rotopress, S.A., Spain, 265 pp.

Askri H., Belmechri A., Benrabeh B., Boudjema A., Boumendjel K., Daoudi M., Drid M. , Ghalem T., Docca A.-M., Ghanbriche H., Ghomari A., Guellati N., Khennous M., Lounici R., Naili H.,Takherist A. and Terkmani M.; 1995: *Geology of Algeria*. In: Well evaluation conference Algeria, Schlumberger, Houston, TX, USA, 93 pp.

Banerjee A. and Chatterjee R.; 2022: *Mapping of reservoir properties using model-based seismic inversion and neural network architecture in Raniganj Basin, India.* J. Geological Society of India, 98(4), 479–486, doi: 10.1007/s12594-022-2005-2.

Bancroft J.C., Richards A. and Ursenbach C.P.; 2000: *Statics: the abusive power of trimming*. CREWES Research Report, 12, 7 pp.

Baouche R., Sen S. and Sankar Ganguli S.; 2020a: *Pore pressure and in-situ stress magnitudes in the Bhiret Hammou hydrocarbon field, Berkine Basin, Algeria*. J. African Earth Sciences, 171, doi: 10.1016/j.jafrearsci.2020.103945.

Baouche R., Sen S. and Boutaleb K.; 2020b: *Distribution of pore pressure and fracture pressure gradients in the paleozoic sediments of Takouazet field, Illizi basin, Algeria.* J. African Earth Sciences, 164, doi: 10.1016/j.jafrearsci.2020.103778.

Barclay F., Bruun A., Rasmussen K.B., Alfaro J.C., Cooke A., Cooke D. and Roberts R.; 2007: S*eismic inversion: reading between the lines.* Oilfield Review, 20(1), 42–63.

Ben Charif H. and Belakehal A.; 2024: *Documenting earthen construction processes and techniques. Case study of Ksar Timimoun, Algeria.* In: Joffroy T., Crosby T., Gandreau D., Hubert A., Marchi S. and Spencer J., Nile's Earth 2023 International Conference: Study and Conservation of Earthen Archaeological Sites in Ancient Egyt and Sudan, CRAterre, Villefontaine, France, pp. 96–103.

Betancourt Sigmonin J.G., Arminio J.F., García E. and Yoris F.; 2016: *Modelo petrofísico integrado: Herramienta clave para la detección de yacimientos de gas condensado en capas delgadas y de baja permeabilidad*. *Cuenca del Valle inferior del Magdalena, Colombia*. Ingeniería Petrolera, 56(5), 256–276.

Bouchachi Y., Boudella A., Bourouis S., Boukhallat S. and Harbi A.; 2022: *In-situ stress analysis of Ahnet Basin, south western Algeria: a 1D geomechanical approach*. J. African Earth Sciences, 196, 104678, doi: 10.1016/j.jafrearsci.2022.104678.

Carcione J.M. and Tinivella U.; 2008: *The seismic response to overpressure: a modelling study based on laboratory, well and seismic data*. Geophysical Prospecting, 49, doi: 10.1046/j.1365-2478.2001.00273.x.

Charkhi A.H., Javaherian A. and Nabi-Bidhendi M.; 2019: *A reliable regression-based approach for seismic reservoir characterization.* Exploration Geophysics, 50(1), 86–93, doi: 10.1080/08123985.2019.1570710.

Cochran M.D. and Petersen L.E.; 1999: *Hydrocarbon exploration in the Berkine Basin, Grand Erg Oriental, Algeria*. AAPG Bulletin, 83(12), 1–32, doi: 10.1306/E4FD4689-1732-11D7-8645000102C1865D.

Daniels R.P. and Emme J.J.; 1995: *Petroleum system model, eastern Algeria, from source rock to accumulation; when, where, and how?* In: Proceedings of the Seminar on Source Rocks and Hydrocarbon Habitat in Tunisia; Entreprise Tunisienne d'Activités Petrolieres Memoir 9, pp. 101–124.

Das G. and Maiti S.; 2023: *A machine learning approach for the prediction of pore pressure using well log data of Hikurangi Tuaheni zone of IODP expedition 372, New Zealand.* Energy Geoscience, 5, 100227. doi: 10.1016/j.engeos.2023.100227.

Di Luca Vingelli M.; 2016: *Caracterización geofísica de yacimientos heterogéneos de la formación Porquero: area piloto Bloque A, cuenca del Valle Inferior*. Magister thesis, Universidad Nacional de Colombia, Facultad de Ciencias, Departamento de Geociencias, Bogotá, Colombia, 150 pp.

Fatti J.L., Smith G.C., Vail P.J., Strauss P.J. and Levitt P.R.; 1994: *Detection of gas in sandstone reservoirs using AVO analysis; a 3-D seismic case history using the Geostack technique.* Geophysics, 59, 1362–1376, doi: 10.1190/1.1443695.

Green S. and Vernik L.; 2020: *Correcting density/sonic logs for total organic carbon to reduce uncertainty in pore pressure prediction*. Geophysical Prospecting, 69, doi: 10.1111/1365-2478.12968.

Guo X., Tang M., He S., Zou L., Zhang G., Zhang, L. and Zhang Y.; 2025: *A new pore pressure model for pre-drilling prediction in carbonate rocks considering the abnormal pressure.* Geoenerg. Science and Engineering, 252, doi: 10.1016/j.geoen.2025.213912.

Hampson D., Schuelke J.S. and Quirein J.; 2001: *Use of multiattribute transforms to predict log properties from seismic data*. Geophysics, 66, 220-236.

Haris A., Sitorus R.J. and Riyanto A.; 2017: *Pore pressure prediction using probabilistic neural network: case study of South Sumatra Basin.* IOP Conference Series: Earth and Environmental Science, 62, doi: 10.1088/1755-1315/62/1/012021.

Hottmann C.E. and Johnson R.K.; 1965: *Estimation of formation pressures from log-derived shale properties.* Journal of Petroleum Technology, 17, 717-722.

Huang H., Li J., Yang H., Wang B., Gao R., Luo M., Li W., Zhang G. and Liu L.; 2022: *Research on prediction methods of formation pore pressure based on machine learning.* Energy Science & Engineering, 10, 1886–1901. doi: 10.1002/ese3.1112.

Jin H., Liu C. and Guo Z.; 2024: *Estimating pore pressure in tight sandstone gas reservoirs : a comprehensive approach integrating rock physics models and deep neural networks.* J. Applied Geophysics, 230, 105526, doi: 10.1016/j.jappgeo.2024.105526.

Klett T.R.; 2002: *Total Petroleum Systems of the Grand Erg/Ahnet Province, Algeria and Morocco - The Tanezzuft-Timimoun, Tanezzuft-Ahnet, Tanezzuft-Sbaa, Tanezzuft-Mouydir, Tanezzuft-Benoud, and Tanezzuft-Béchar/Abadla.* USGS Bulletin 2202(B), USGS, Denver, CO, USA, 144 pp., http://pubs.usgs.gov/bul/b2202-b/.

Krishna S., Irfan S.A., Keshavarz S., Thonhauser G. and Ilyas S.U.; 2024: *Smart predictions of petrophysical formation pore pressure via robust data-driven intelligent models.* Multiscale and Multidisciplinary Modeling, Exp. and Des., 7, 5611–5630, doi: 10.1007/s41939-024-00542-z.

Kushwaha P.K., Maurya S.P., Rai P. and Singh N.P.; 2020: *Estimation of subsurface rock properties from seismic inversion and geo-statistical methods over F3-block, Netherland.* Exploration Geophysics, 52(3), 258–272, doi: 10.1080/08123985.2020.1815528.

Laouini R., Hacini M., Merabti H., Medjani F. and Mahmou, O.; 2025: *Mud loss analysis through predictive modeling of pore pressure and fracture gradients in Tin Fouye Tabankort Field, western Illizi Basin, Algeria.* Energies, 18(7), doi: 10.3390/en1807183.

Li Z., Song B., Zhang J. and Hu G.; 2016: *Joint elastic and petrophysical inversion using prestack seismic and well log data*. Exploration Geophysics, 47(4), 331–340.,doi: 10.1071/EG14074.

Liu L., Shen G., Wang Z., Yang H. Han H. and Cheng Y.; 2018: *Abnormal formation velocities and applications to pore pressure prediction.* Journal of Applied Geophysics, 153, doi: 10.1016/j.jappgeo.2018.02.013.

Lopez R.; 2025: *Sedimentary basins of Algeria: their role in Europe's energy security and emerging green hydrogen future.* EGU Blogs, SSP, https://blogs.egu.eu/divisions/ssp/2025/12/02/sedimentary-basins-of-algeria-their-role-in-europes-energy-security-and-emerging-green-hydrogen-future/.

Lüning S., Adamson K. and Craig J.; 2003: *Frasnian organic-rich shales in North Africa: regional distribution and depositional model.* Geological Society, London, Special Publications, 207, 165-184, doi: 10.1144/GSL.SP.2003.207.9.

Luo S., Xu T. and Wei S.; 2022: *Prediction method and application of shale reservoirs core gas content based on machine learning*. J. Applied Geophysics, 204, 104741. doi: 10.1016/j.jappgeo.2022.104741.

Makhous M., Galushkin Y. and Lopatin N.; 1997: *Burial history and kinetic modeling for hydrocarbon generation, Part II: applying the GALO model to Saharan Basins*. American Association of Petroleum Geologists Bulletin, 81, 1679–1699.

Nadaraya E.A.; 1964: *Estimating regression.* Theory of probability and Its applications, 9(1), 141-142.

Oshim F., Ezeonyema C., Modekwe D.,and Dunu A.; 2023: *The application of deep learning in pore pressure prediction and reservoir optimization: a brief review*. Asian J. Geological Research, 6, 160-171, https://journalajoger.com/index.php/AJOGER/article/view/141.

Pandey L. and Sain K.; 2021: *Porosity mapping of shallow subsurface sediments: a case study in Mahanadi basin, eastern margin of India*. Exploration Geophysics, 53(5), 517–531, doi: 10.1080/08123985.2021.1994352.

Radwan A.E., Wood D.A. and Radwan A.A.; 2022: *Machine learning and data-driven prediction of pore pressure from geophysical logs: a case study for the Mangahewa gas field, New Zealand*. Journal of Rock Mechanics and Geotechnical Engineering, 14(6), 1799–1809, doi: 10.1016/j.jrmge.2022.01.012.

Russell B.; 2014: *Prestack seismic amplitude analysis: an integrated overview*. Interpretation, 2(2), SC19– SC36, doi: 10.1190/INT-2013-0122.1.

Saadaoui A., Ahmadi R., Khalfi C., Hadji R. and Laaouar M.S.; 2025: *Pore pressure study in the berkine basin, south-east Algeria; characterisation of overpressure intervals and implications for hydrocarbon accumulation and exploration.* J. African Earth Sciences, 226, 105587, doi: 10.1016/j.jafrearsci.2025.105587.

Specht D.F.; 1991: *Probabilistic neural networks*. Neural Networks, 3(1), 109-118.

Stefanou I. and Darve F.; 2024: *Machine learning in Geomechanics 1: overview of machine learning, unsupervised learning, regression, classification and artificial neural networks*. Wiley, Hoboken, NJ, USA, 241 pp., doi: 10.1002/9781394325634.

Wang H. and Ma J.; 2024: *Pore pressure prediction based on rock physics theory and its application in seismic inversion*. J. Applied Geophysics, 229(1), doi: 10.1016/j.jappgeo.2024.105494.

Wang J., Cao J.X. and Zhou X.; 2022: *Reservoir porosity prediction based on deep bidirectional recurrent neural network*. Prog. Geophys., 37, 267–274.

Watson G.S.; 1964: *Smooth regression analysis.* Sankhyā: The Indian Journal of Statistics, 26(4), 359-372.

Weng C., Li J., Yang H., Long Z., Zhang G., Wang B. and Zhao Y.; 2025: A *hybrid model of formation pore pressure prediction based on geological sequence matching*. Geoenergy Science and Engineering, 252, 213972, doi: 10.1016/j.geoen.2025.213972.

Zoeppritz K.; 1919: *The reflection and propagation of seismic waves*. Gottinger Nachrichten, 1(5), 66–84.

Corresponding author: Khalda Zabel
Geophysics Laboratory, Universite des Sciences et de la Technologie Houari Boumediene
Bab-Ezzouar, BP 32 El Alia, 16111 Algiers, Algeria
Phone: +213 0541354840; e-mails: khalda.zabel@usthb.dz, zabelkhalda@gmail.com

Appendix A

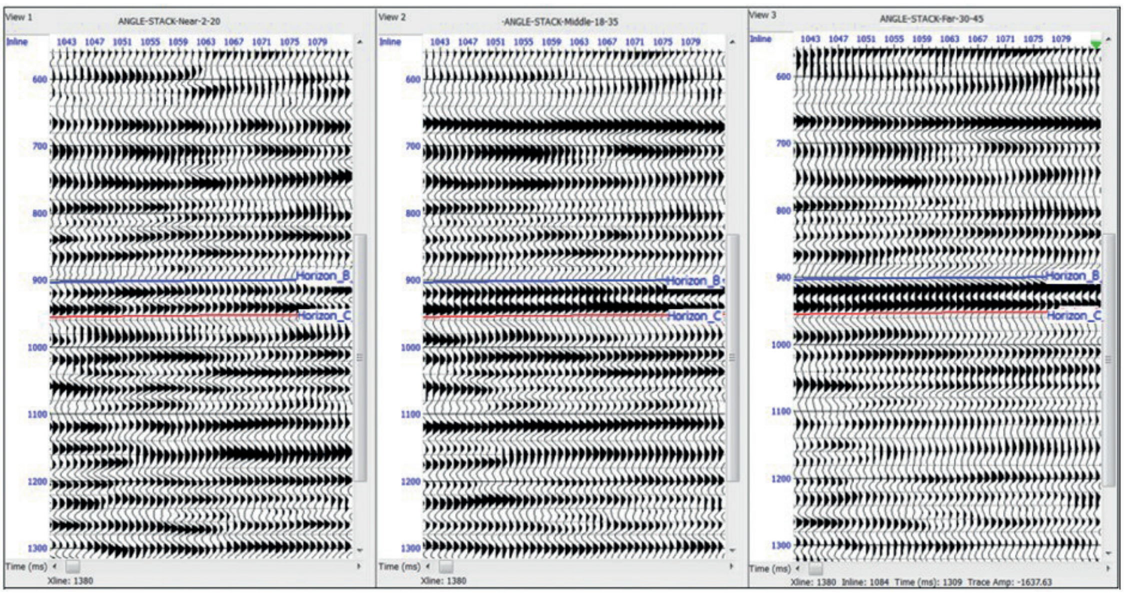

Fig. A1 - Angle stacks: near stack 02°-20°, middle stack 18°-35°, and far stack 30°-45°.

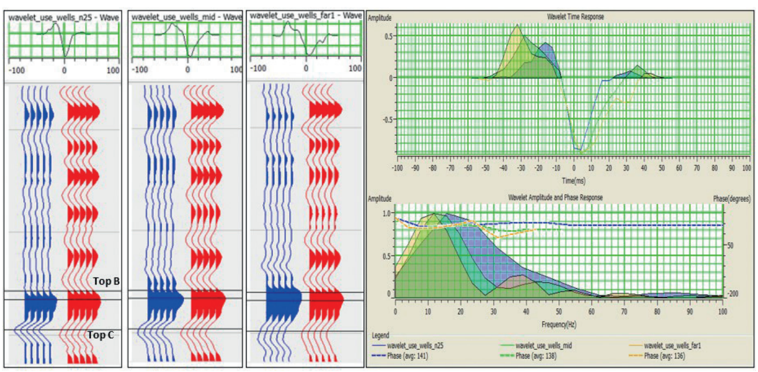

Fig. A2 - Well tie (synthetic in blue, composite traces in red) and the extracted wavelet use well (near phase = 141°, middle phase = 138°, and far phase = 136°).

Appendix B

Fig. B1 shows the results of P_p estimation using two different well-log approaches for five wells (Well A, Well B, Well C, Well D, and Well E). The track of each well highlights a specific depth interval, Frasnian interval, or study interval, where the black and green logs show a positive correlation between the P_p-based density log and P_p-based sonic log. Table B1 summarises the statistical description of the results obtained.

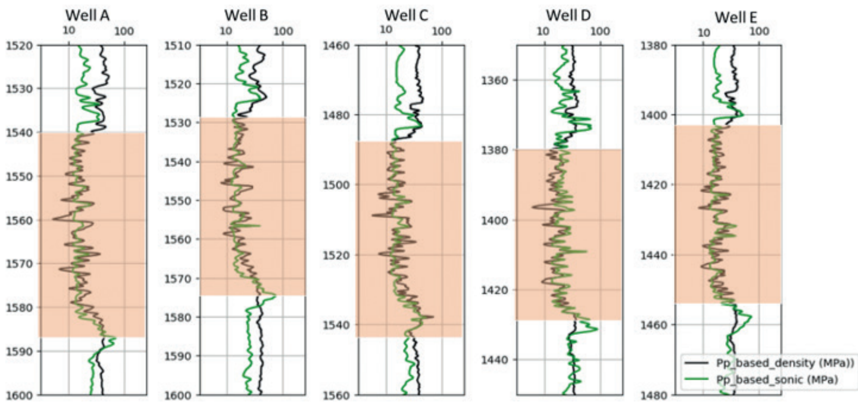

Fig. B1 - Strong positive correlation between the P_p-based density log and P_p-based sonic log at the target reservoir.

Table B1 - Statistical description of the P_p-based density, P_p-based sonic, and total organic carbon (TOC) log.

	P_p-based density (Mpa)	**P_p-based sonic (Mpa)**	**TOC (%)**
Count	1080.000	1080.000	1080.000
Mean	29.954	30.818	7.064
STD	2.277	2.290	3.499
Min	24.198	25.172	0.342
25%	28.060	28.691	4.854
50%	30.245	32.273	7.517
75%	33.871	34.681	9.650
Max	41.867	40.339	14.115

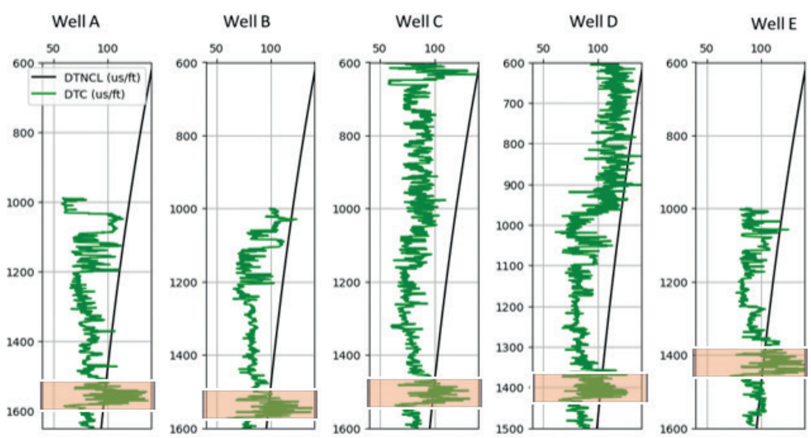

Fig. B2 - Measured transit time (DTC) and the normal compaction trend (DTNCL) of transit time versus depth.

Appendix C

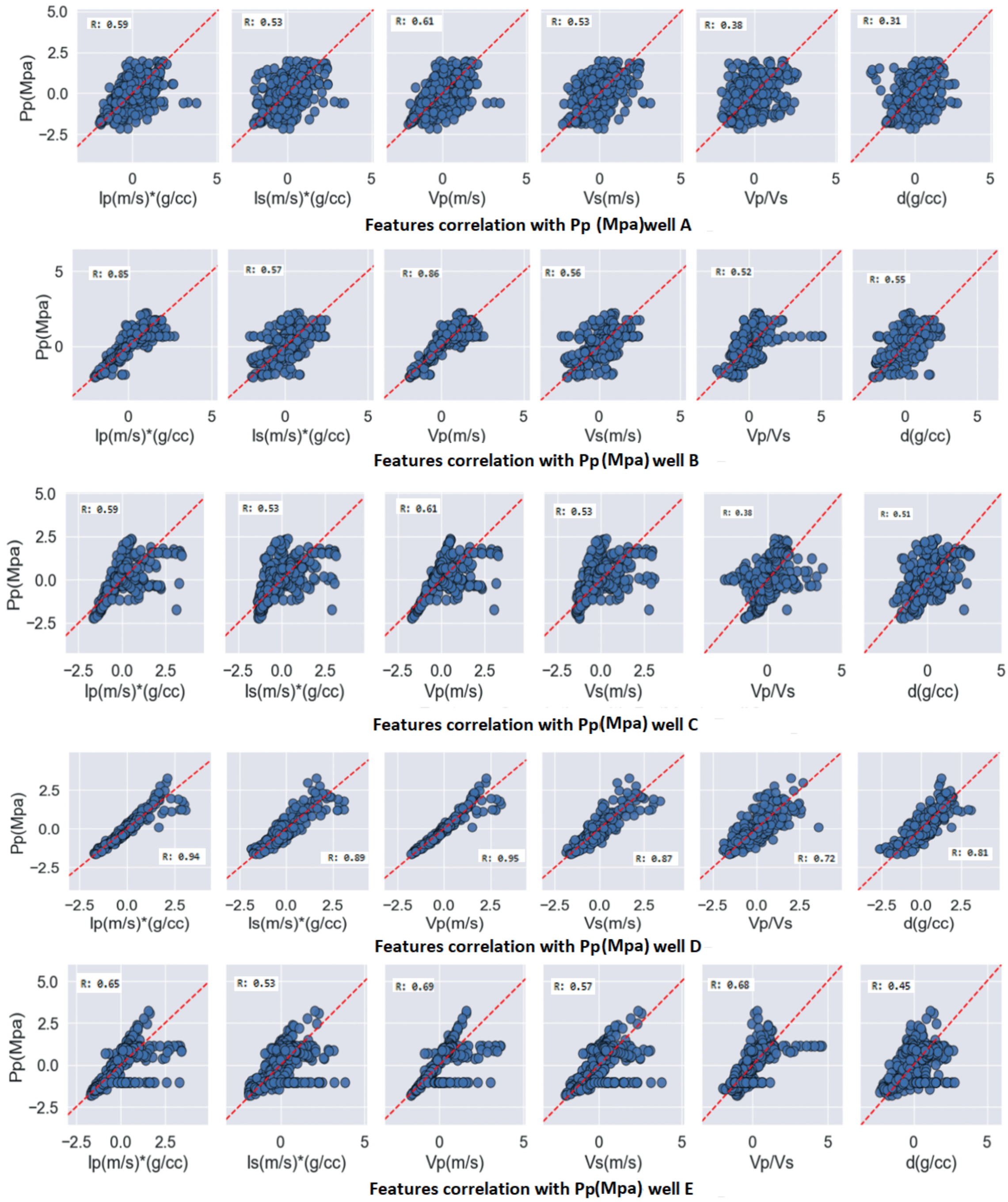

Fig. C1 - Feature correlation with P_p (Mpa) for each well.

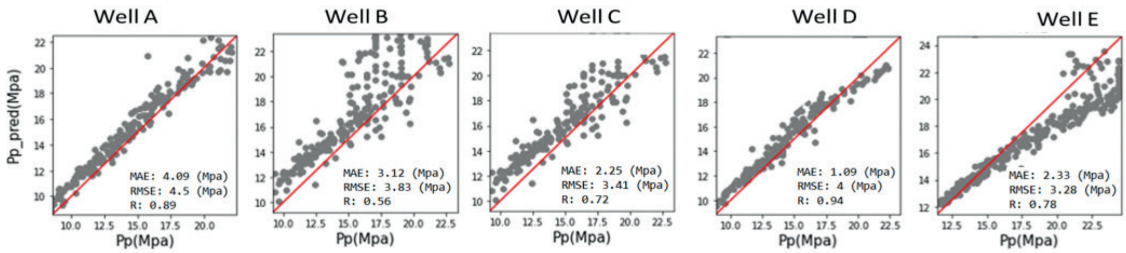

Fig. C2 - LOWO-CV across the five wells with *MAE*, *RMSE*, and R^2.

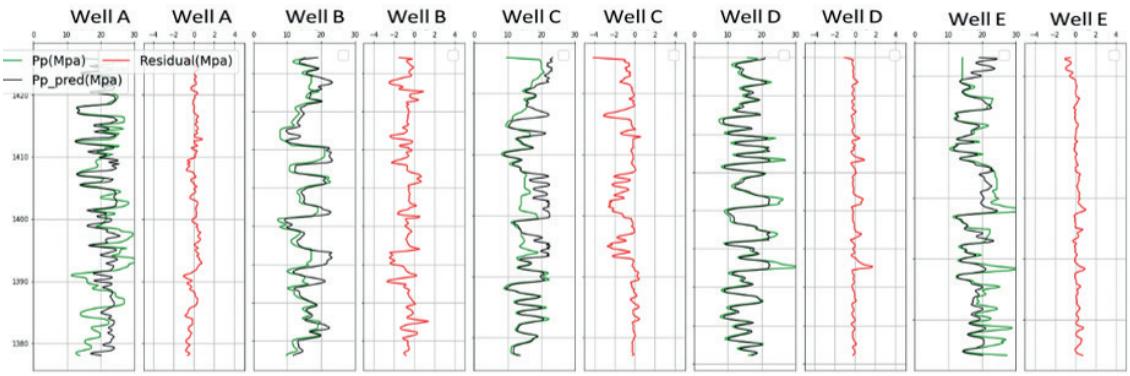

Fig. C3 - Composite log of P_p-based sonic log P_p (Mpa), P_p _pred (Mpa), and the residual log (Mpa) of each well after LOWO-CV.

Table C1 - Median/percentile errors and percentage of dynamic range for each well comparing P_p (Mpa) and predicted P_p (Mpa).

Well	Median error (MPa)	Median error (% of range)	25th percentile error (MPa)	25th percentile error (% of range)	75th percentile error (MPa)	75th percentile error (% of range)
Well A	0.62	3.29%	-2.83	-14.95%	2.55	13.47%
Well D	-0.77	-4.08%	-1.09	-5.76%	-0.13	-0.67%
Well B	-1.57	-8.30%	-2.28	-12.04%	-0.94	-4.97%
Well C	-0.85	-4.51%	-3.25	-17.19%	-0.40	-2.11%
Well E	1.09	5.74%	-0.18	-0.95%	2.80	14.82%
Avg of abs	0.98	5.18	1.92	10.17%	1.36	7.20%