Journal of Petroleum Exploration and Production Technology (2024) 14:705–726
https://doi.org/10.1007/s13202-023-01731-1

ORIGINAL PAPER-EXPLORATION GEOLOGY

Geochemical evaluation and source rock zonation by multi-layer perceptron neural network technique: a case study for Pabdeh and Gurpi Formations-North Dezful Embayment (SW Iran)

Abolfazl Jamshidipour[1] · Mohammad Khanehbad[1] · Maryam Mirshahani[2] · Ali Opera[3]

Received: 27 August 2023 / Accepted: 20 November 2023 / Published online: 20 December 2023

Abstract

In this study, using a multi-layer perceptron neural network (MLPNN) model, total organic carbon (TOC) and hydrogen index (HI) values for Pabdeh and Gurpi Formations in the oil fields of Naft Sefid (NS-13), Kupal (KL-36, KL-38, and KL-48) and Palangan (PL-2) were calculated in the North Dezful Embayment located in the southwest of Iran. To build the MLPNN model, the geochemical data calculated by the Rock–Eval pyrolysis method (TOC and HI) and the conventional petrophysical well log data, including sonic transit time log (DT), formation density log (RHOB), total resistivity log (RT), spectral gamma-ray log, computed gamma-ray log and neutron porosity log from the NS-13 well were used. The log data were the input layer, and the geochemical data were the output layer of the model. Twenty-four datasets were used for MLPNN training, and seven datasets were used for MLPNN testing. Two hidden layers were considered in this technique. Each hidden layer has an activation function (tanh) and a solver parameter (lbfgs). The accuracy of measurement of TOC and HI indices of Pabdeh and Gurpi Formations in terms of R^2 was 0.93 and 0.90, respectively. This model has higher accuracy than the ΔlogR technique (R^2: 0.28). Considering the relationships between the input data and other wireline logs is an advantage of this technique. These two formations have five source rock zones. Pabdeh Formation has three zones. The middle zone of the Pabdeh Formation (Pz. II) has the highest TOC (2.6 wt%) and source rock potential. Pabdeh Formation has kerogen type II. Gurpi Formation has a weaker source rock potential than Pabdeh Formation due to its low TOC content (< 1%). Both source rock zones of this formation have low TOC, but in some layers of the lower zone of the Gurpi Formation (Gz. II), high values for TOC were predicted. Gurpi Formation has Kerogen types II and III.

Keywords Neural network · ANN · Rock–Eval · TOC · Well logs · Pabdeh formation · Gurpi formation · North Dezful Embayment

Abbreviations

AC	Acoustic log
ANN	Artificial neural network
BPANN	Back propagation artificial neural network
CGR	Computed gamma-ray log
CNL	Compensated neutron log
CNN	Convolutional neural network
DEN	Density log
DT	Sonic transit time log
Fig	Figure
Gu	Gurpi Formation
Gz. I	Upper zone of Gurpi Formation
Gz. II	Lower zone of Gurpi Formation
HI	Hydrogen index
KI	Kerogen type I
KII	Kerogen type II
KIII	Kerogen type III
KL	Kupal
L. Cr	Late Cretaceous
LOM	Level of maturity
MFF	Mountain frontal fault
MLPNN	Multi-layer perceptron neural network
MLPRegressor	Multi-layer perceptron Regressor tool
NISOC	National Iranian South Oil Company

✉ Mohammad Khanehbad
mkhanehbad@ferdowsi.um.ac.ir

[1] Department of Geology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran

[2] Research Institute of Petroleum Industry (RIPI), Tehran, Iran

[3] Department of Geology, National Iranian South Oil Company (NISOC), Ahwaz, Iran

NPHI	Neutron porosity log
NS	Naft Sefid
OM	Organic matter
Pb	Pabdeh Formation
PL	Palangan
POTA	Potassium
Pz. I	Upper zone of Pabdeh Formation
Pz. II	Middle zone of Pabdeh Formation
Pz. III	Lower zone of Pabdeh Formation
RBF	Radial basis function
RHOB	Formation density log
RIPI	Research Institute of Petroleum Industry
RT	Total resistivity log
SGR	Spectral gamma-ray log
SGS	Sequential Gaussian Simulation
skLearn	Scikit learn
SVM	Support vector machine
SW	Southwest
THOR	Thorium
TOC	Total organic carbon

List of symbols

gr/cm^3	Gram per cubic centimeter
gapi	American Petroleum Institute gamma ray units
km	Kilometer
km^2	Kilometer squared
m	Meter
mg HC/gr	Milligram hydrocarbon per gram
R	Coefficient of correlation
R^2	Coefficient of determination
tanh	Hyperbolic tangent function
µs/ft	Microsecond per feet
v/v	Volume per volume
wt%	Weight percent
Ω m	Ohm meter
°C	Degree celsius

Introduction

Shale is capable of sedimentation in various environments, including marine (such as turbidity currents, post-pelagic transport, contour currents, slump flow-subaqueous slide, and hyperpycnal flow), transitional, and terrestrial (Stow et al. 2001; Liu 2016; Awan et al. 2021; Mahdi et al. 2022; Wu et al. 2022; Wu et al. 2022; Fathy et al. 2023; Li et al. 2023). Typically, source rocks consist of organic-rich shales and calcareous mudstones (Hakimi et al. 2023; Li et al. 2023). Source rock evaluation is a crucial stage in hydrocarbon exploration dependent on numerous parameters. The amount of total organic carbon (TOC) and hydrogen index (HI) are the most crucial parameters (Vega-Ortiz et al. 2020; Lai et al. 2020).

Geochemical techniques are frequently time-consuming and costly despite providing valuable and beneficial information. Since the physical properties of organic matter (OM), primarily bulk density, differ from the mineral components of its host rock, petrophysical data can be used to interpret source rocks (Wang et al. 2018; Vega-Ortiz et al. 2020; Awan et al. 2021; Abubakar et al. 2023).

Numerous studies, such as Schmoker (1981), Schmoker and Hester (1983), and Passey et al. (1990), cited the significant potential of petrophysical logs for source rock properties. However, the primary objectives of these studies were to distinguish source rocks from non-source rocks by predicting TOC. In the past 2 decades, intelligent systems such as artificial neural networks (ANN), extreme learning machines, sequential Gaussian simulation (SGS), support vector machine (SVM), and neuro-fuzzy techniques have also been utilized to determine OM content (Kamali and Mirshady 2004; Kadkhodaie-Ilkhchi et al. 2009; Tabatabaei et al. 2015; Shi et al. 2016; Alizadeh et al. 2018; Bolandi et al. 2017; Wang et al. 2019; Zheng et al. 2021; Khalil Khan et al. 2022; Abubakar et al 2023; Diab et al. 2023; Hassan et al. 2023).

ANN is a mathematical model that simulates the behavior of animal neural networks and processes data by altering the relationships between many internal nodes (Zheng et al. 2021; Liu et al. 2023). Multi-layer perceptron neural networks (MLPNN) constitute a significant subclass of neural networks. A network typically consists of sensory units or input nodes comprising the input layer, one or more concealed layers of neurons or computation nodes, and an output layer. The advantages of MLPNNs include hidden unit outputs (essential functions) that adapt during training, eliminating the need for users to select them in advance (Ouadfeul and Aliouane 2014; Zheng et al. 2021).

This research aims to estimate and evaluate several geochemical indexes (TOC and HI) using conventional petrophysical logs in Pabdeh and Gurpi Formations in Naft Sefid (NS-13), Kupal (KL-36, KL-38, and KL-48) and Palangan (PL-2) oilfields located in the North Dezful Embayment in SW Iran (Fig. 1) based on the ΔlogR and MLPNN techniques and comparing their performance with each other. Also, this study trained and tested an MLPNN model using the geochemical indexes obtained from the Rock–Eval analysis and logging data in one well. It was used to predict TOC and HI in some wells with logging data but no geochemical data. Finally, using the geochemical data predicted by the MLPNN model, kerogen type and source rock zoning have been done in Pabdeh and Gurpi Formations.

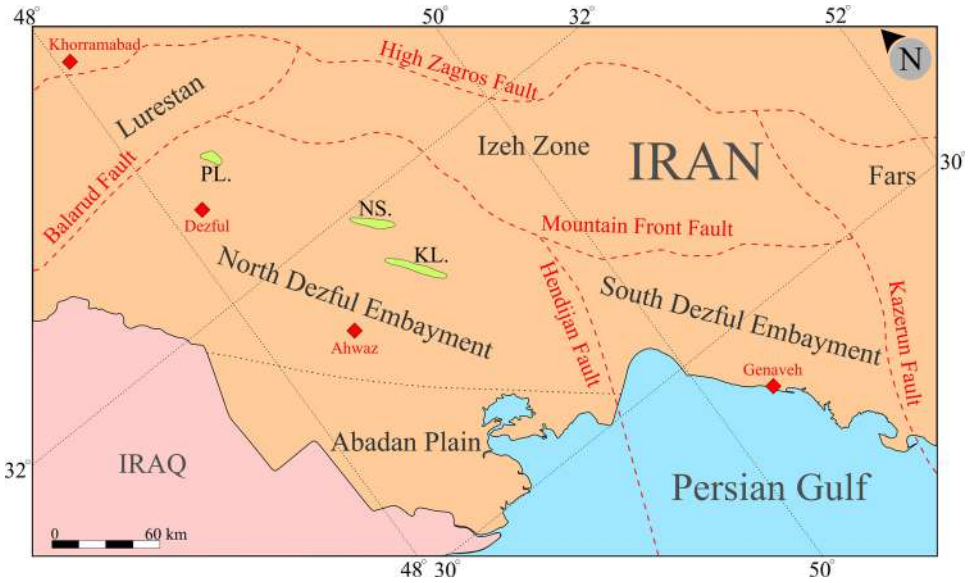

Fig. 1 Location map of North Dezful Embayment, SW Iran (PL: Palangan; NS: Naft Sefid; KL: Kupal)

Geological setting

The folded-thrust belt of Zagros, with a length of about 1375 km and a width between 120 and 250 km, extends from southeastern Turkey to northern Syria and Iraq to western and southern Iran and from the continuous and long-term convergence between the Arabian Plate and Eurasia Plate was created during the closing of the Neo-Tethys Ocean basin (Al-Husseini 2000; Hessami et al. 2001; Alavi 2004, 2007).

This belt is considered one of the wealthiest fold-thrust belts in the world, with substantial hydrocarbon fields. This belt was formed due to the change in the shape of the primary foreland system (pro-foreland) in the margin of the Zagros (Pirouz et al. 2017). The first comprehensive report on the Zagros fold-thrust belt was presented by James and Wynd (1965). Zagros can be structurally divided into Fars, Khuzestan, and Lurestan sub-basins (Aghanabati 2004). Regarding the structural pattern from the northeast to the southwest, Zagros includes thrust zones, folded belts, Dezful Embayment, and Abadan plain.

The Dezful Embayment is a part of the folded Zagros, which contains most of Iran's oil fields (including Ahvaz, Ab-Timur, Masjid Suleiman, Marun, Kupal, Aghajari, Karnaj, Parsi, Zilaei fields) (Bordenave and Burwood 1994; Bordenave and Hegre 2010). The northeastern range of Dezful Embayment is determined by the mountain frontal fault (MFF). Its southwestern range is located almost along the anticline parallel to the northwest–southeast structures of the Zagros front. Its eastern border is surrounded by the Kazerun Fault, and its northern border by the Balarud Fault (McQuarrie and van Hinsbergen 2013). As mentioned, this Embayment is divided into two regions of Dezful North and South due to the Handijan fault zone with a north–south direction almost in the middle. Dezful Embayment is tectonically more stable and less folded than the neighboring areas. This area has an area of about 40,000 km^2 (Aghanabati 2004).

The youngest Cretaceous stratigraphic unit in Zagros is the Gurpi Formation. The type section of this formation is located in Tang-e-Pabdeh, north of Masjid Sulaiman (Laeli oil field), with a thickness of 320 m. This formation has marl and gray calcareous shales and gray shaly limestones (Aghanabati 2004). In the studied area, this formation is 123–150 m thick in the Kupal oil field, 6 m in the Naft-Sefid oil field, and 176 m in the Palangan oil field. In the North Dezful Embayment, the Gurpi Formation is characterized by a sedimentary discontinuity from the Ilam and Pabdeh Formations. In Inner Fars and the north Dezful Embayment, the limestone facies of the Tarbor Formation replace the Gurpi Formation. Campanian to Maastrichtian age has been reported in Dezful Embayment (Aghanabati 2004) (Fig. 2).

In Lurestan and Dezful Embayment, the Imam Hassan limestone member is 114 m thick and has gray, thick-layered clayey limestone with interlayers of marl. Brown limestones characterize the Seymareh limestone member, and compared to the Imam Hassan limestone member, it is a shallower facies. The Mansouri limestone member is a neritic limestone belonging to the deep parts of the basin, which has an outcrop in western Khuzestan but disappears toward the northeast (Aghanabati 2004). The TOC level of Gurpi Formation in the Dezful Embayment is around 0.5–1.5 wt%. According to Ala et al. (1980), the hydrocarbon generation potential of this formation in the Dezful Embayment is weak and insignificant because of its low TOC. This formation is a source rock in oil fields such as Marun, Kupal, Haftkel, and Naft Sefid.

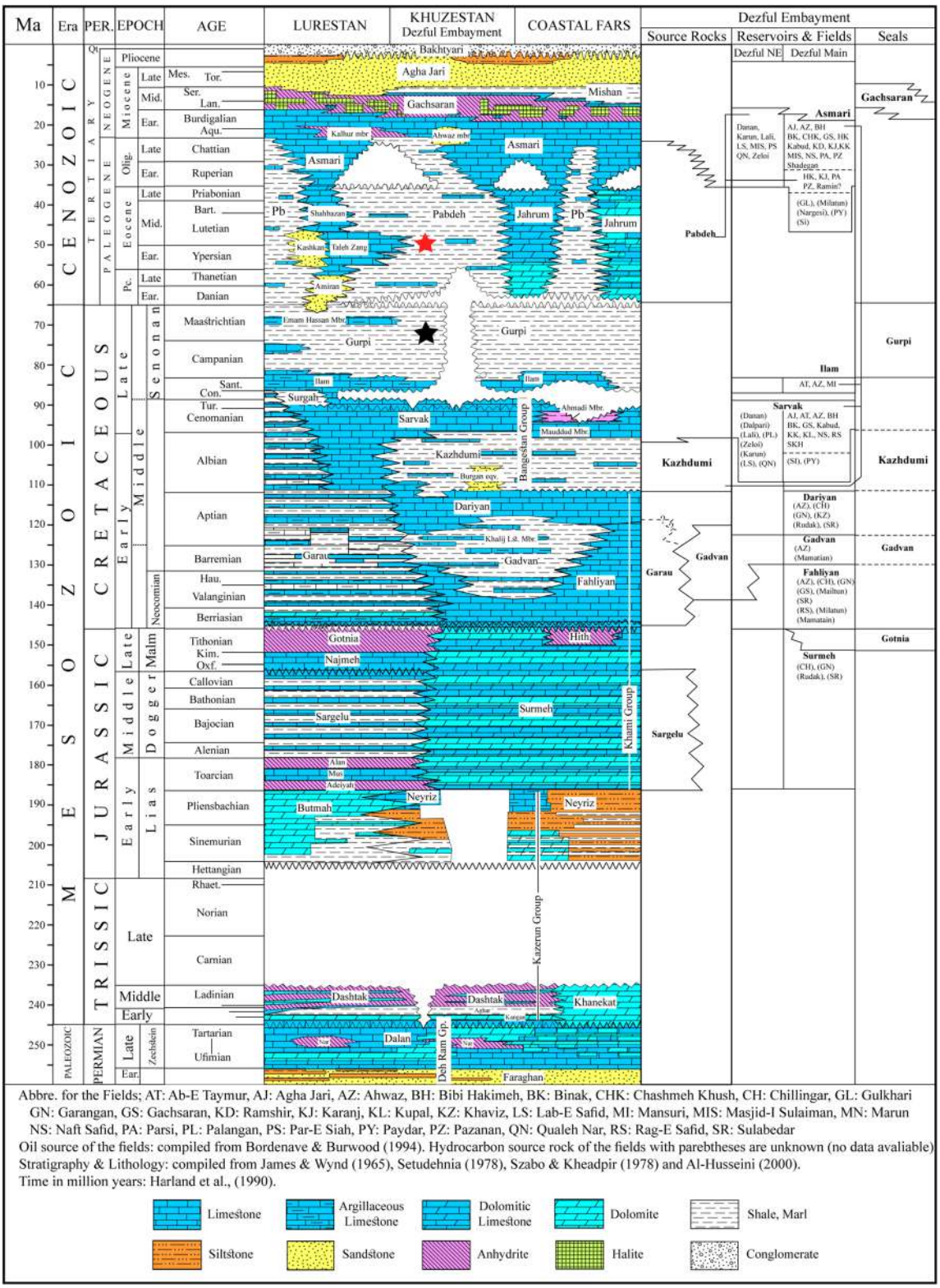

Fig. 2 Stratigraphic chart of the Dezful Embayment and adjacent areas (After Setudehnia 1978; Harland 1990; Bordenave 2002). The Pabdeh and Gurpi Formations are shown with red and black stars, respectively

The Pabdeh Formation type section in Tang-e-Pabdeh is 798 m thick. It consists of marl and gray shales and marine clayey limestone, divided into two informal parts: purple shales and cherty limestones (James and Wynd 1965). In the North Dezful Embayment, the lower boundary of the Pabdeh Formation with the Gurpi Formation is discontinuous, and the upper boundary of this Formation with the limestones of the Asmari Formation is gradational. The age of this formation is reported as the Late Paleocene—Late Oligocene (Aghanabati 2004). According to several sedimentological studies that have been carried out, the Pabdeh Formation is composed of three facies: pelagic, hemipelagic, and carbonate turbidite; which indicates deposition on the deep parts of a carbonate ramp (Hosseini Asgarabadi et al. 2019). In the studied area, the Pabdeh Formation is mainly gray and brown shales with interlayers of limestone and sometimes marl with a thickness of 156–160 m in the Kupal oil field, 256 m in the Naft-Sefid oil field, and 553 m in the Palangan oil field. Pabdeh Formation is a vital source rock and cap rock in Zagros. According to reports, its total organic matter (TOC) is between 2 and 6 wt%, and the kerogen is mostly type II with limited amount of type III continental organic matter (Rabbani and Kamali 2006; Karimi et al. 2016; Alizadeh et al. 2020) (Fig. 2). Alizadeh et al. 2020 reported that a brown shale unit was in the Pabdeh Formation in the Dezful Embayment. This brown shale unit (BSU) is rich in organic matter. Due to the presence of oil shale, this formation is also considered an unconventional source of hydrocarbon production (Alizadeh et al. 2020).

Methods and data

Espitalie et al. (1985) cite Rock–Eval pyrolysis as the most prevalent technique for determining the petroleum generation potential, thermal maturation, quantity, and quality of organic matter (OM) in sedimentary rocks. A Rock–Eval 6 instrument (manufactured by Vinci Technologies) was used to analyze 31 samples at the Petroleum Laboratory of Shahid Chamran University of Ahvaz (Fig. 3). Tissot and Welte (1984) describe this method in more detail. Following the standard procedure outlined by Behar et al. (2001), 70–80 mg aliquots of pulverized samples were inserted into crucibles, and various parameters (such as S1, S2, S3, and Tmax) were measured. Other parameters (such as TOC and HI) were calculated from these measurements (Table 1).

The most common and practical method for calculating TOC is the ΔlogR technique (Passey et al. 1990). This procedure involves superimposing a suitably scaled sonic transit time log on a resistivity (RT) log with a specified overlap coefficient. The two curves are adjusted to each other at a known organic-lean shale interval, forming the baseline for zero TOC. TOC is estimated based on the distance between

Fig. 3 Rock–Eval IV analyzer at the Petroleum Laboratory of Shahid Chamran University of Ahvaz

the two logs within other shale units (potential source rocks). The ΔlogR equation for calculating TOC from records of sonic velocity and resistivity is as follows (Passey et al. 1990):

$$\Delta \mathrm{Log}R = \mathrm{Log}_{10}\left(R/R_{\mathrm{base}}\right) + P * \left(\Delta t - \Delta t_{\mathrm{base}}\right) \quad (1)$$

where Δlog$R$ represents the difference between the two logs, R is the measured resistivity (Ω m), and t is the measured sonic velocity (microsecond/ft.). R baseline is the resistivity corresponding to zero TOC. Based on the ratio of 50 microsec/ft per one cycle of resistivity, the overlay coefficient is 0.02. Δlog$R$ has a linear relationship with organic carbon. The equation is Passey et al. (1990):

$$\mathrm{TOC} = (\Delta \mathrm{Log}R) *^{10(2.297-0.1688*\mathrm{LOM})} \quad (2)$$

where TOC is the organic carbon content calculation and LOM is the level of maturity. LOM for these samples is 12.

The first step in training an MLPNN is to collect the input data. The data required to train a neural network that can calculate TOC and HI values include sonic transit time log (DT), formation density log (RHOB), total resistivity log (RT) typically from the deep resistivity tool, spectral gamma-ray log (SGR), computed gamma-ray log (CGR) and neutron porosity log (NPHI) (Table 2). The TOC and HI were calculated by the Rock–Eval pyrolysis method (Table 3).

This study used a four-layer MLP neural network for the NS-13 well. Six neurons were included in wireline data in the first or input layer (Table 2). Seven neurons were used in the second layer (hidden layer), and three neurons were used in the third layer (hidden layer). Using two or more hidden layers in data with medium or low correlations can, in addition to measuring the relationships between input and output

data, evaluate the relationships of input data with each other and provide more accurate output data. In the fourth layer (output layer), one neuron was placed.

The first step is to remove null or invalid data (such as negative, zero, and exaggerated values) from the petrophysical logs in order to reduce the possibility of errors in the MLP neural network.

In the next step, the input and output data were normalized by the Gaussian method so that the data was 0–1 (Wang et al. 2019; Liu et al. 2023). The importance of this step is determined because not normalizing the data can lead to errors in the output and unexpected values.

To build up the structure of this MPLNN, Python software and the skLearn library were used. In this model, two hidden layers were considered, and 7 and 3 neurons were used for each layer (The number of neurons is calculated and suggested by MLPRegressor in the skLearn library.), respectively. The tanh activation function and the lbfgs solver parameter were used for hidden layers. The learning rate of the neural network was initially set to 0.01, which can be changed adaptively. The tolerance rate for the optimization here is 1e−5 (Fig. 4).

All input and output data at each step are controlled and corrected by some factors, such as log corrections based on the well environment and lithology. The output data have also been compared and controlled with other works on these formations, such as Alizadeh et al. (2020). In the output data, according to the geochemical data obtained from the Rock–Eval pyrolysis, values greater than 6 and less than 0 were considered uncertain and unrealistic values for TOC predicted by the MLPNN method. On the other hand, for the predicted HI based on the data from the Rock–Eval pyrolysis, values greater than 500 and less than 0 were considered unreliable and unrealistic.

Finally, the performance of the MLPNN was evaluated using the coefficient of determination (R^2). After MLPNN building-up, this neural network has been used in other wells to evaluate geochemical indexes.

In general, with the increase of the depth of burial (the density of the formation increases), the time to pass through the distance of mudstone decreases. However, in organic-rich intervals, the sonic transit time is more extensive than in adjacent organic-lean intervals at similar depths of burial because of the lower density of organic matter compared to the mineral matrix (Liu et al. 2013; He et al. 2016; Khalil Khan et al. 2022). TOC content tends to increase the apparent value of DT (correlation coefficient: 0.19). DT log with HI has a correlation coefficient of 0.66 (Figs. 5, 6a, 7a). DT log has a positive correlation coefficient with SGR, CGR, and NPHI logs and is negative with RT and RHOB logs (Fig. 8).

Since the bulk density of organic matter (1.1–1.4 g/cm^3) is lower than that of quartz (2.65 g/cm^3) and clay (2.77 g/cm^3) (He et al. 2016), the logarithm value of the density decreases. This log measures the apparent density of the formation. This density consists of the combined effects of matrix density and fluid density. Water saturation should be equal in shales with a similar degree of compaction and

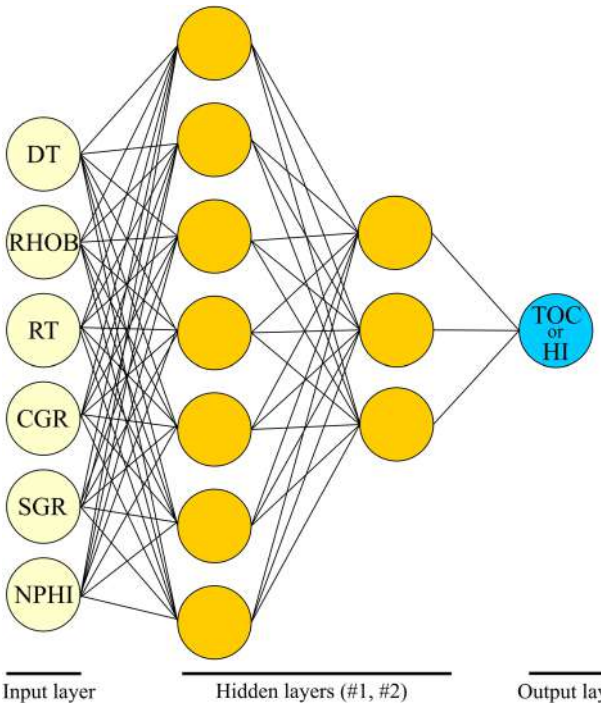

Fig. 4 Schematic diagram of MLPNN structure for this study. This MLPNN has two hidden layers were considered, and 7 and 3 neurons used for each layer, respectively. Input data are petrophysical logs. TOC and HI are the output data of this model

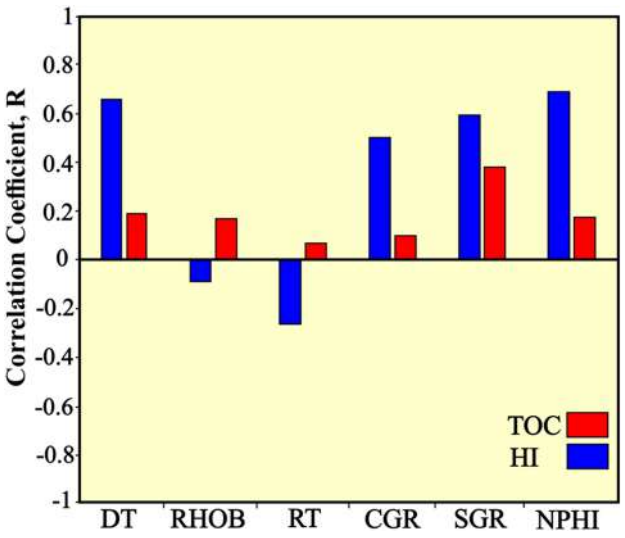

Fig. 5 Comparison of correlation coefficient of laboratory measurements (TOC and HI) and the logging data for well NS-13

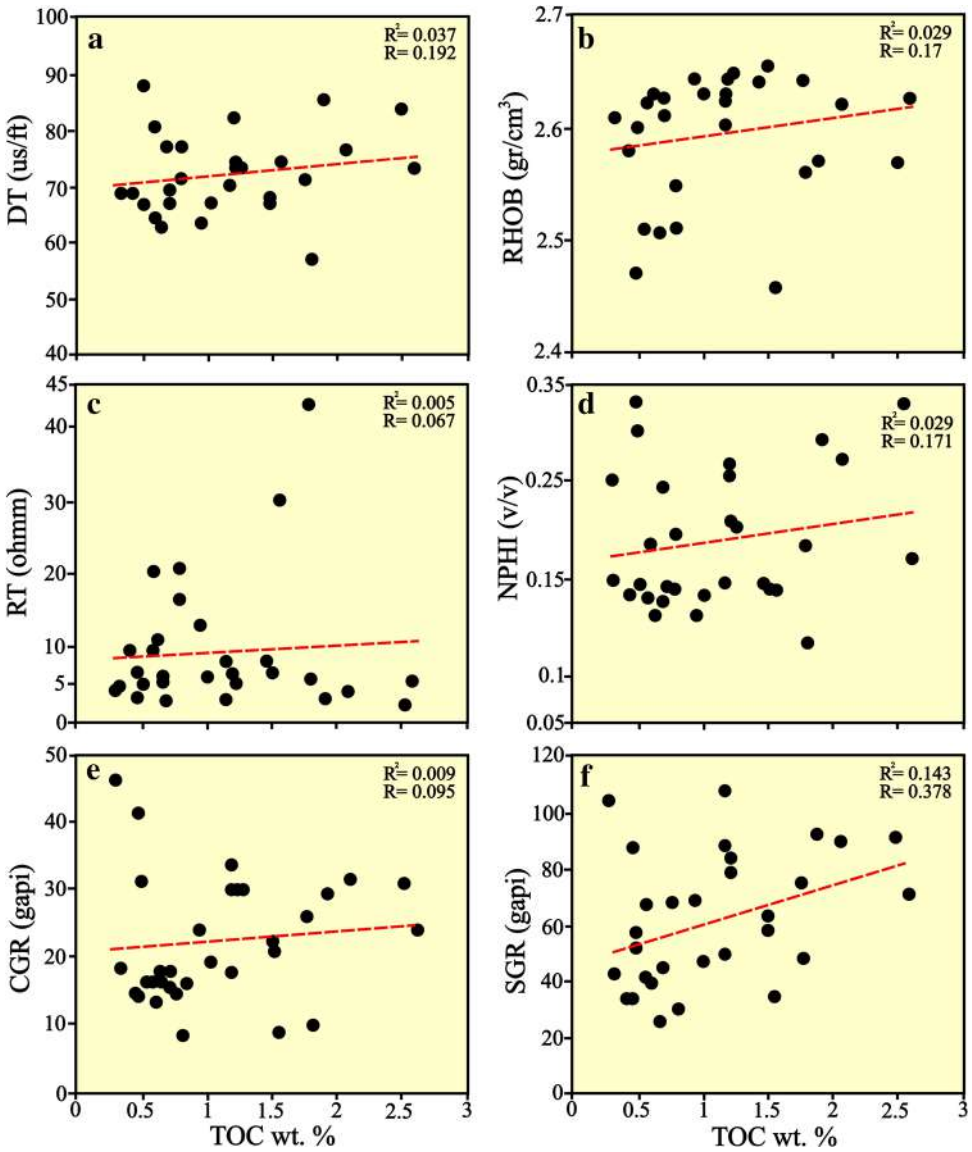

Fig. 6 Cross-plots of TOC laboratory-measurements and the logging data for well NS-13 **a** DT-TOC; **b** RHOB-TOC; **c** RT-TOC; **d** NPHI-TOC; **e** CGR-TOC; **f** SGR -TOC

matrix and similar fluid density. Solid OM has a density similar to that of water (approximately 1.0 g/ml) and, thus, less than the density of the surrounding rock matrix (Kadkhodaie-Ilkhchi et al. 2009). If the density values in source rocks are lower than in typical shales, it must be a function of the amount of OM present. Because solid organic matter has a density close to the water and is substantially less dense than the surrounding rock, formations with high TOC are often of low density. As a result, the density diminishes as the TOC matures. However, in addition to biological materials, a fall in rock density may also be brought on by increased porosity and clay minerals. Therefore, TOC estimates using density logs are favored for reservoirs with comparable fluid phases and constant mineralogy (He et al. 2016; Khalil Khan et al. 2022). The density log against the discussed geochemical indices has a negative correlation coefficient (−0.09) for HI (Fig. 7b) and a correlation coefficient of 0.17 for TOC (Figs. 5, 6b). The correlation of petrophysical logs with this log is negative; only the gamma-ray logs have a positive correlation coefficient (Fig. 8).

Shales and mudstones (including source rocks) are generally layered and electrically anisotropic. Mudstone intervals characteristically exhibit low resistivity due to the excellent conductivity of mudstone (both the rock skeleton and the water in the pores are conductive) (Passey et al. 1990; He et al. 2016). The RT log negatively correlates with other logs and geochemical indexes (except TOC) (Figs. 5, 6c, 7c, 8).

Because uranium or concentrated uranium ions in the shale layers may create rather intense radioactivity, shale layers rich in organic materials, particularly those from marine environments, produce significant gamma rays. A spectroscopic gamma-ray device may extract thorium, potassium, and uranium enrichment levels from radiation measurements. It can provide more reliable proof of the existence

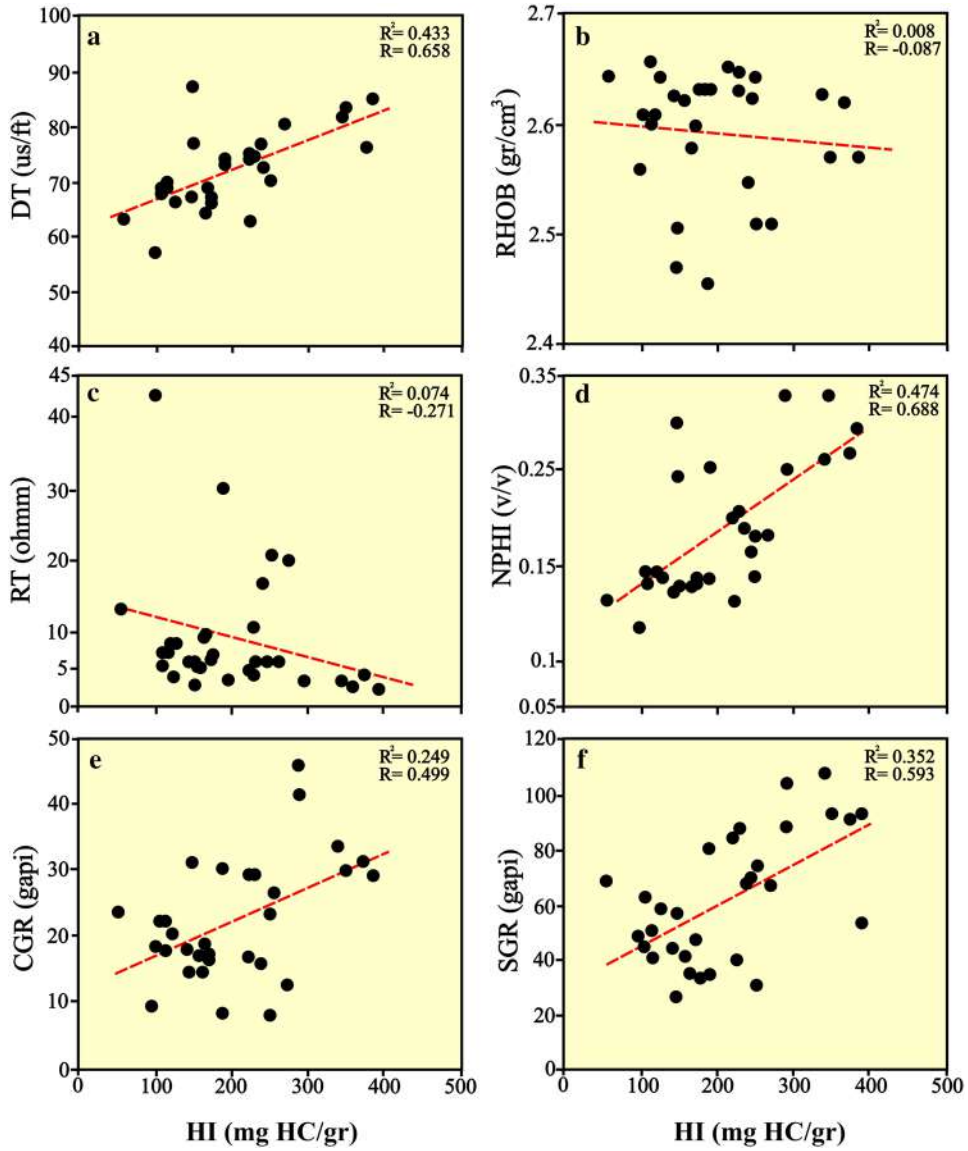

Fig. 7 Cross-plots of HI laboratory-measurements and the logging data for well NS-13 **a** DT-HI; **b** RHOB-HI; **c** RT-HI; **d** NPHI-HI; **e** CGR-HI; **f** SGR-HI

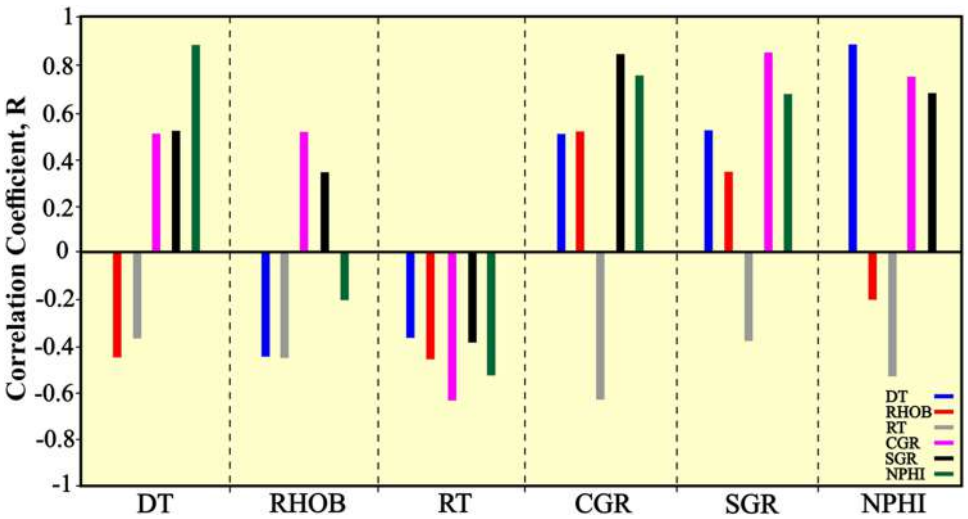

Fig. 8 Comparison between the correlation coefficient of different logs together for well NS-13

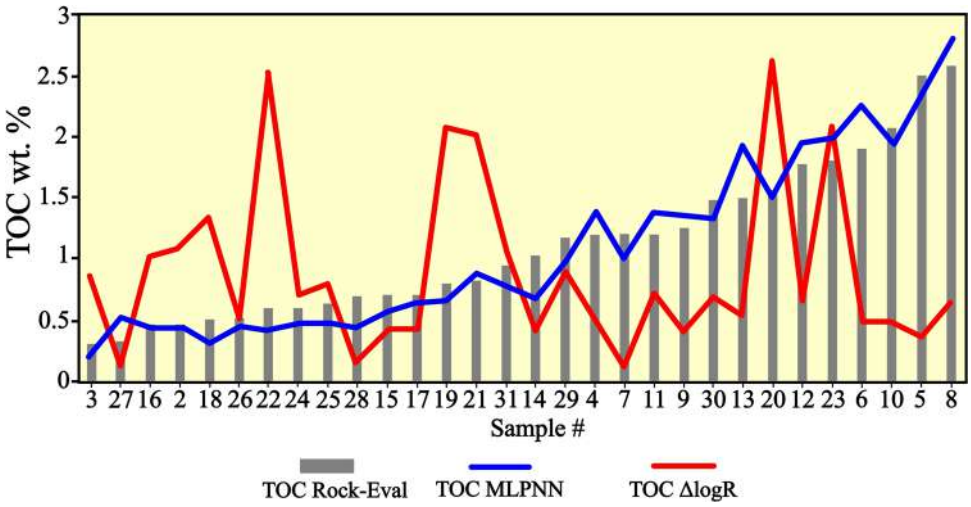

Fig. 9 Comparison between TOC values from Rock–Eval, MLPNN, and Δlog*R* techniques. This graph shows a good correlation between the laboratory-measurement data and the MLPNN technique

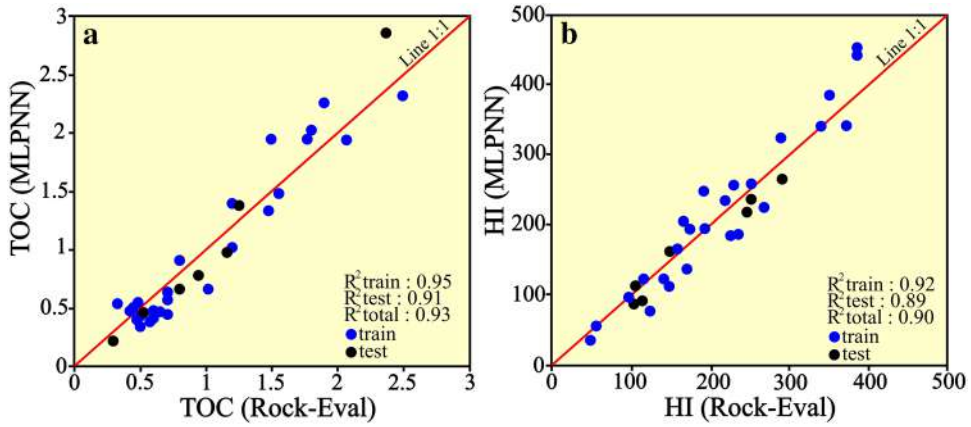

Fig. 10 Geochemical and wireline log for Pabdeh and Gurpi Formations in well NS-13 (L. Cr: late Cretaceous). Pabdeh and Gurpi Formations have 3 and 1 source rock zones, respectively

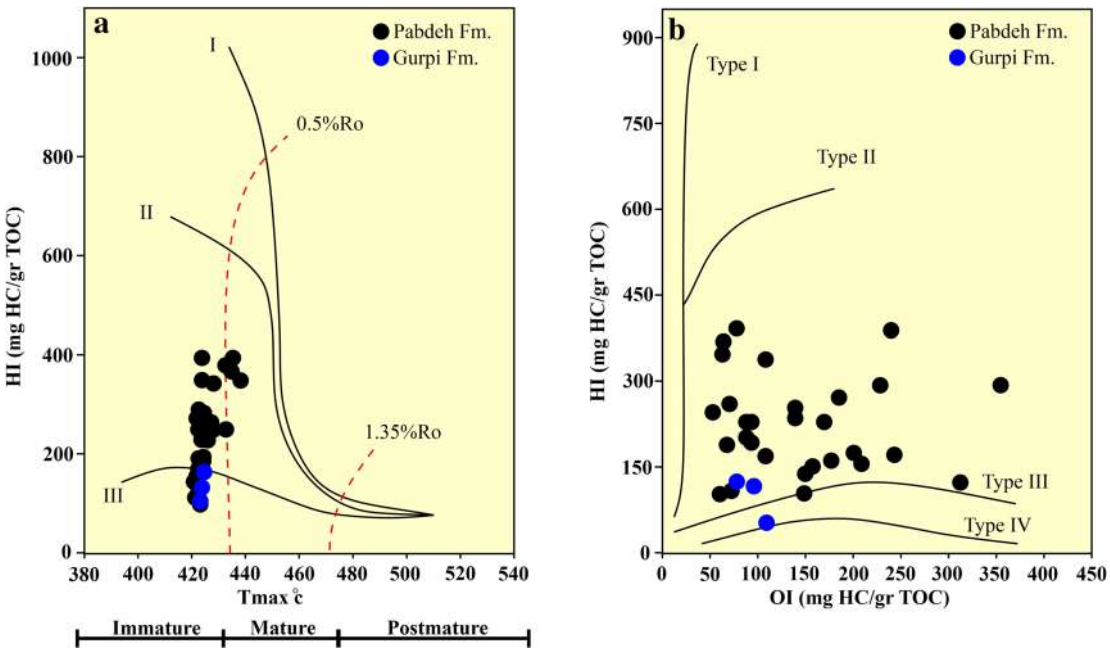

Fig. 11 **a** TOC (MLPNN) versus TOC (Rock–Eval) from train and test dataset. **b** HI (MLPNN) versus HI (Rock–Eval) from the train and test dataset

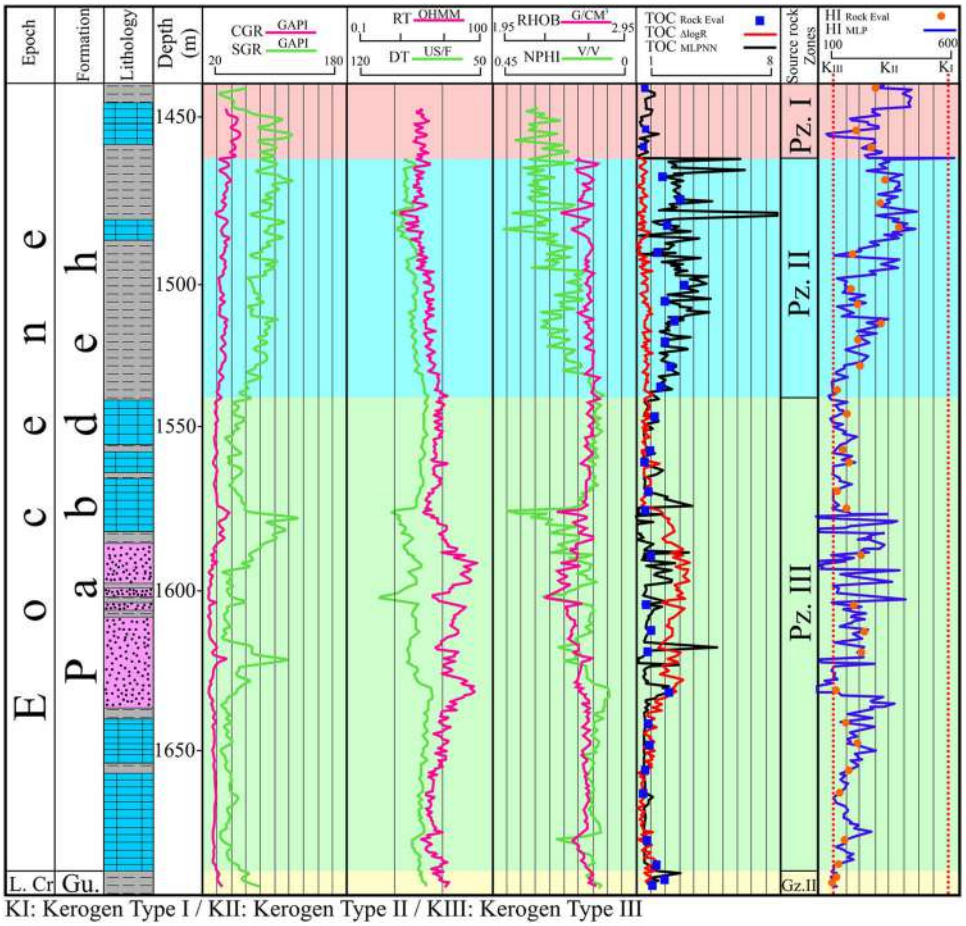

Fig. 12 Geochemical and wireline log for Pabdeh and Gurpi Formations in well KL-36. Pabdeh and Gurpi Formations have 3 and 2 source rock zones, respectively

of TOC than total gamma-ray data (Kadkhodaie-Ilkhchi et al. 2009). Uranium is often found in sedimentary salts and organic debris. It is conceivable that uranium ions, often present in saltwater and other trace elements, are adsorbable by plankton, which concentrates uranium in the source rock. Clay minerals are related to thorium and potassium (He et al. 2016; Khalil Khan et al. 2022). The gamma-ray log correlation coefficient (SGR and CGR) for TOC values based on the Rock–Eval pyrolysis does not show high values; it is about 0.1 for CGR and about 0.37 for SGR (Figs. 5, 6c, f, 7c, f, 8).

The correlation coefficient of HI with CGR is about 0.5, and for SGR is about 0.58. Since MLPNN is a vast and complex network, relationships between petrophysical logs are also important (Fig. 8). Gamma logs show a positive correlation coefficient with all logs (except the RT log), as shown in Fig. 8. The NPHI log recorded the highest positive correlation coefficient (CGR: 0.75 and SGR: 0.68).

High HI of organic matter leads to a high neutron registration value (Kadkhodaie-Ilkhchi et al. 2009). In this study, it has the highest correlation with the HI (0.69) and DT (0.89), and on the other hand, it has a negative correlation with the RHOB and RT logs (Figs. 5, 6d, 7d, 8).

Results and discussion

Performance of Δlog*R* technique

After using the Δlog*R* technique, TOC values were calculated using the differences between the DT and RT logs. The TOC values based on this technique were between 0.35 and 2.64 wt% (average: 0.88 wt%) for the samples, according to the values in Table 2. Based on Figs. 5 and 8 and the correlation of TOC values with the Δlog*R* technique against TOCs calculated based on the Rock–Eval pyrolysis, it was 0.28. This technique in previous studies also does not show a good correlation coefficient with the data obtained from laboratory measurements (e.g., Alizadeh et al. 2018; Khalil Khan et al. 2022). It was only a practical method for the overall evaluation of the source rock. The main reason for this problem is that the RT and DT logs in the calibration well NS-13 do not show a proper trend and a strong correlation with the TOC data calculated based on the Rock–Eval pyrolysis (Figs. 9, 12).

This technique is unsuitable for source rocks, especially those with the possibility of porosity and the presence of in-situ oil and oil shales. In such cases, it is better to use

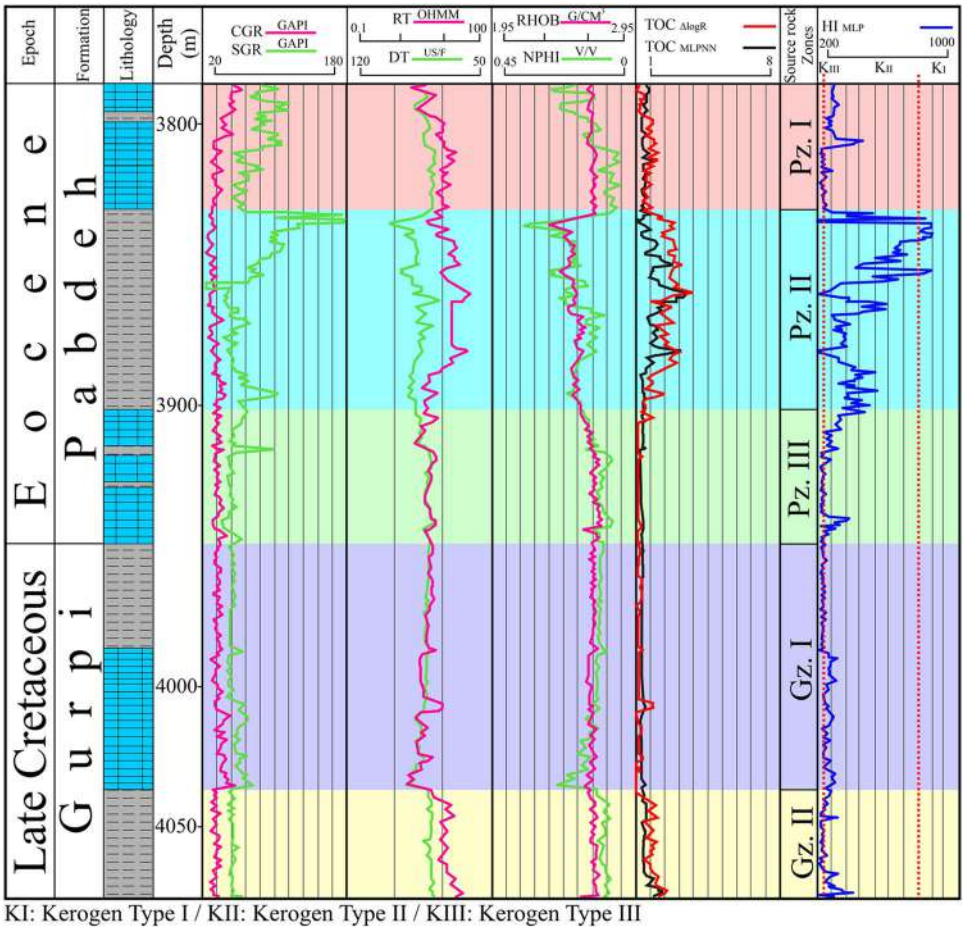

Fig. 13 Geochemical and wireline log for Pabdeh and Gurpi Formations in well KL-38. Pabdeh and Gurpi Formations have 3 and 2 source rock zones, respectively

other techniques to consider the relationships between the wireline data in source rock studies and the evaluation of organic carbon values. One of these techniques is artificial neural networks (ANN), especially MLPNN, which can be considered by relationships between all data (input and output data).

MLPNN model

Calculation of TOC and HI values based on DT, RT, ROHB, NPHI, SGR, and CGR logs for 24 datasets in the NS-13 well was performed. In Fig. 10a, predicted data from the train and measured TOC data are plotted against each other. The data are plotted near line 1:1 and shows R^2: 0.95. On the other hand, the same method was performed for the HI data based on the input wireline data in the same datasets, which are plotted near the line 1:1 and show R^2: 0.92 (Fig. 10b).

After MLPNN training, the trained neural network should be tested using a series of datasets. For this purpose, seven datasets from the same NS-13 well were selected, including DT, RT, ROHB, NPHI, SGR, and CGR logs. After running the neural network predicting the TOC and HI values and plotting them against the calculated values of these geochemical indexes, it was observed that the R^2 for TOC and HI was 0.91 and 0.89, respectively. All data are plotted near line 1:1 in Fig. 10.

Considering the high accuracy of MLPNN compared to single and multi-type wireline data evaluation techniques such as $\Delta \log R$ in the studied sample and other studies that have investigated neural networks and relationships of geochemical indexes, MLPNN is a high-accuracy evaluation technique. It is based on log input for source rocks in oil fields.

Comparing this method with previous studies for neural networks shows that the technique has higher accuracy even with less input data (Table 4). This higher accuracy in predicting geochemical parameters (TOC and HI) is due to using the solver function (lbfgs), which can train the neural network better in small datasets. Two hidden layers and the optimized learning rate made the neural network quickly fix its errors (Fig. 11).

The MLPNN trained and tested for other wells KL-36, KL-38, KL-48, and PL-2 was also implemented using the input data, including DT, RT, ROHB, NPHI, SGR, and

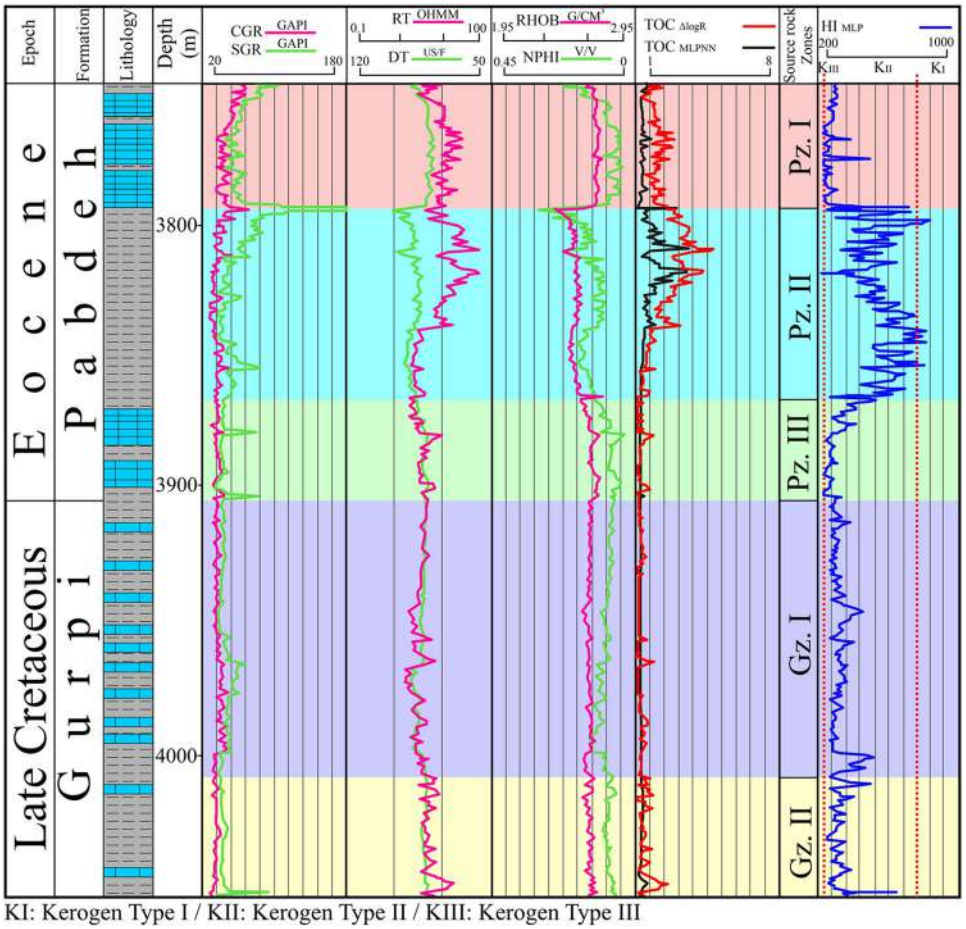

Fig. 14 Geochemical and wireline log for Pabdeh and Gurpi Formations in well KL-48. Pabdeh and Gurpi Formations have 3 and 2 source rock zones, respectively

CGR logs. Finally, output data, including TOC and HI, were predicted.

Geochemical evaluation of the Pabdeh and Gurpi source rocks

After completing the steps mentioned above, the geochemical indices of the Pabdeh and Gurpi Formations were assessed using the MLPNN approach utilizing wireline data from five wells. For these Formations in all wells, TOC, HI, and kerogen type were estimated every 15 cm (Figs. 12, 13, 14, 15, 16).

With certain restrictions, TOC may be used to assess the potential for the formation of hydrocarbons. TOC is one of the standard criteria used to evaluate the abundance of OM (Peters 1986; Hakimi et al. 2016; Diab et al. 2023). With a TOC estimate, it is easier to evaluate shale gas or oil source rocks (Chalmers and Bustin 2007; Gasparik et al. 2014; Hakimi et al. 2023). The average TOC value for the Pabdeh Formation using the MLPNN is 1.19, 0.75, 0.7, 0.61, and 1.06 wt% for the five studied wells, NS-13, KL-36, KL-38, and PL-2, respectively. The Rock–Eval TOC approach yields a TOC value for the Pabdeh Formation in the NS-13 well that varies from 0.3 to 2.59 wt%, with a mean value of 1.07 wt%. The five available wells, NS-13, KL-36, KL-38, KL-48, and PL-2, provided the MLPNN with average TOC readings for the Gurpi Formation that are 0.96, 0.52, 0.46, 0.5, and 0.68 wt%, respectively. The laboratory approach yields a TOC value for the Gurpi Formation in the NS-13 well that varies from 0.95 to 1.48 wt%, with a mean value of 1.2 wt%.

Organic material that is scattered across sediments is used to make petroleum. According to Peters (1986), Peters and Cassa (1994), and Valdon et al. (2023), the quantity of generated petroleum is direct to the amount and quality (i.e., kerogen type) of organic matter, thickness, and thermal maturity of potential source rock intervals. Peters and Cassa (1994) state that a source rock must have at least 1.0 wt% of total organic carbon to have a "good" hydrocarbon potential. A rock with a high TOC is often considered a superior hydrocarbon source, even if it is not the only factor used to assess its ability to generate hydrocarbons (Dembicki 2009). Rock production and storage volume are substantially influenced by OM properties (enrichment, type, and maturity). Type I/II kerogen has a higher potential generation capacity

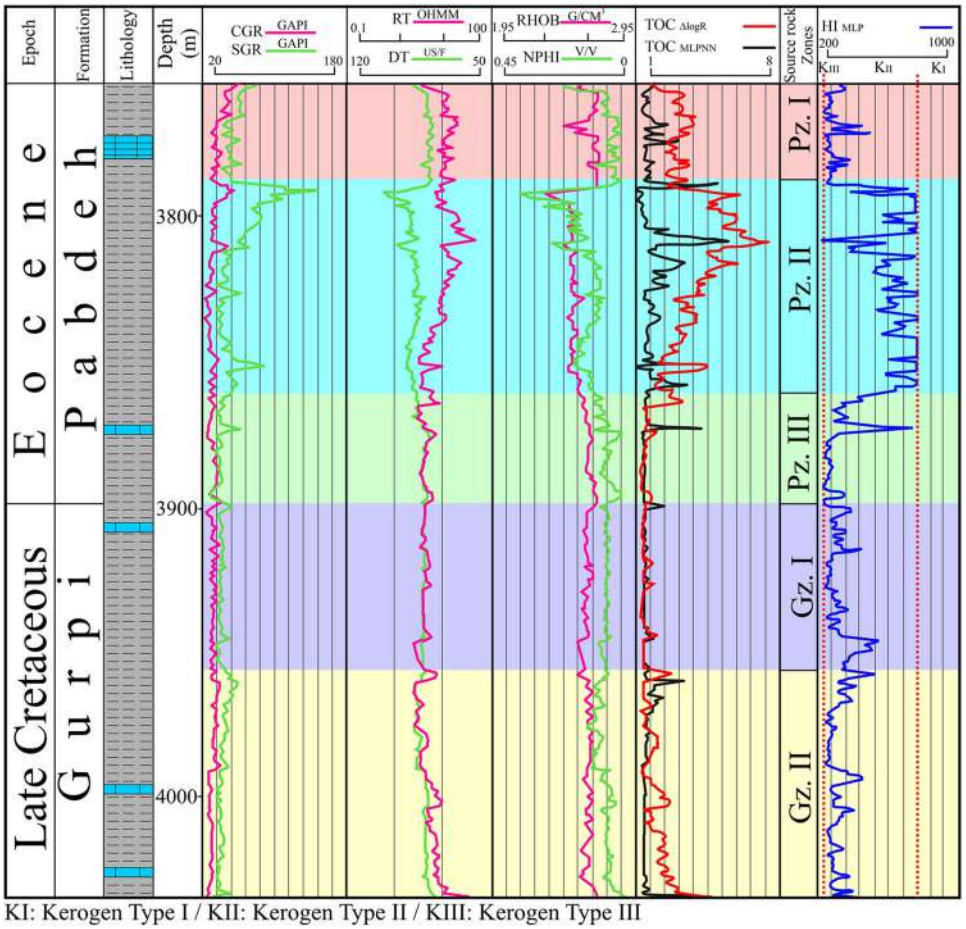

Fig. 15 Geochemical and wireline log for Pabdeh and Gurpi Formations in well PL-2. Pabdeh and Gurpi Formations have 3 and 2 source rock zones, respectively

than Type III and Type IV kerogens (Curtis 2002; Mahdi et al. 2022; Zheng et al. 2023; Fathy et al. 2023).

These findings show that the Pabdeh and Gurpi have higher OM concentrations in certain zones of their parent rocks. The acceptable to excellent levels of TOC in this study may point to a sea level rise during the deposition of these brown to sporadically dark-colored shales. Sea-level transgression encourages the upwelling of nutrients from the ocean's depths, which may lead to a high level of biological production, according to Awan et al. (2021). Additionally, the amount of organic carbon in sediments may rise due to anoxia development and rapid burial (organic preservation) (Miller et al. 2005; Müller et al. 2008; Valdon et al. 2023; Li et al. 2023). The assessment of source rocks may sometimes be impacted by hydrocarbons that move and collect in fine-grained sediments (Peters 1986).

The kind of organic materials in a shale source rock is essential. Type I sapropelic, Type II sapropelic-humic combination, and Type III humic have been used to categorize organic materials, notably kerogen (Tissot and Welte 1984; Sanders et al. 2022; Wu et al. 2023). These different types of kerogen have different capabilities to produce hydrocarbons and other products depending on the structure and chemical makeup of the OM. The composition and maceral variants are significant in the hydrocarbon generation potential and kerogen-type variation. Both type II and type I kerogen can generate oil and gas, while type III kerogen is generally limited to gas generation only. In this study, kerogen type was identified via pyrolysis analysis. The quantity of organic hydrogen in the kerogen is indicated by the hydrogen index (HI). It reveals if a chemical can create gas or hydrocarbons (Xu et al. 2022). With a mean value of 224 and 99 mg HC/g TOC, the HI in the Pabdeh and Gurpi Formations ranges from 99 to 388 and 56 to 125 mg HC/g TOC. The average HI value for the Pabdeh Formation utilizing the MLPNN is 221, 292, 367, 430, and 363 mg HC/g TOC for the five available wells (NS-13, KL-36, KL-38, KL-48, and PL-2). The average HI value for the Gurpi Formation utilizing the MLPNN is 95, 104, 209, 173, and 298 mg HC/g TOC for the five available wells (NS-13, KL-36, KL-38, KL-48, and PL-2).

II and III kerogens may be oil and gas-prone, according to the hydrogen index data of the Pabdeh and Gurpi Formations (Figs. 11, 17). Humic (Type I) kerogen enters the gas window sooner than sapropelic (Type II) kerogen. According to Shuangfang et al. (2012) and Al-Yaseri et al. (2023), the mass gas potential of the sapropel kerogen group is later

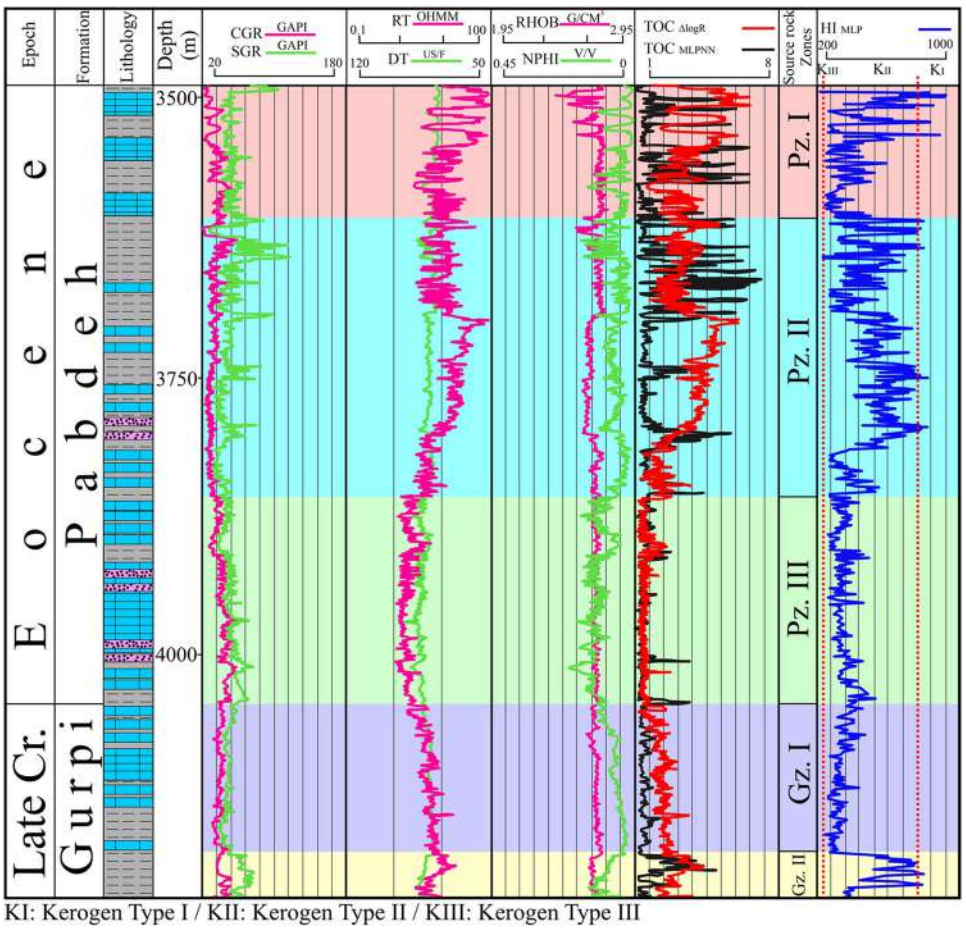

Fig. 16 Cross plots of kerogen type for Pabdeh and Gurpi Formations in well NS-13. **a** The cross plot of Tmax versus HI displays the kerogen type and maturity (Modified after Yandoka et al. 2016). **b** Modified Van Krevelen diagram indicating the kerogen type

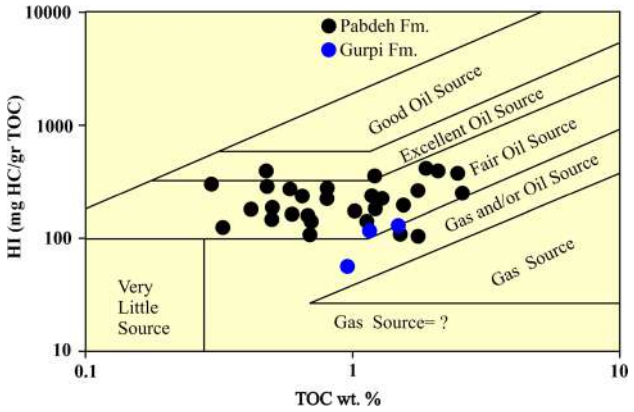

Fig. 17 The distribution of the Pabdeh and Gurpi Formation samples on a plot of TOC versus HI (after Jackson et al. 1985)

than that of the other kerogen kinds. The ability of Type III kerogen (low HI) to generate oil is much lower than kerogen Types I and II (higher HI).

Source rocks zonation

Pabdeh and Gurpi Formations were zoned in the studied wells based on the source rock potential. TOC classification was based on Peters (1986) and Peters and Cassa (1994). Kerogen classification was based on HI (Yandoka et al. 2016). Pabdeh Formation has three zones, and Gurpi Formation has two zones (Fig. 18):

- *Upper zone of Pabdeh Formation (Pz. I)*: This zone consists of gray to calcareous and marly dark shales and dark-colored limestones. The thickness of this zone is between 25 and 122 m in the studied wells. TOC is, on average, 0.86 wt% by weight. The highest amount of TOC is in KL-48 (1.3) and PL-2 (1.27) wells. HI index in this zone is between 88 and 390 mg HC/g TOC and type II kerogens.
- *Middle zone of Pabdeh Formation (Pz. II)*: This zone consists of brown to dark-colored limy shales, sometimes with interlayers of limestone and marl. The thickness of this zone is between 69 and 256 m in the studied wells. TOC is, on average, 2.6 wt%. The decrease in the RHOB log and the difference between the two gamma-

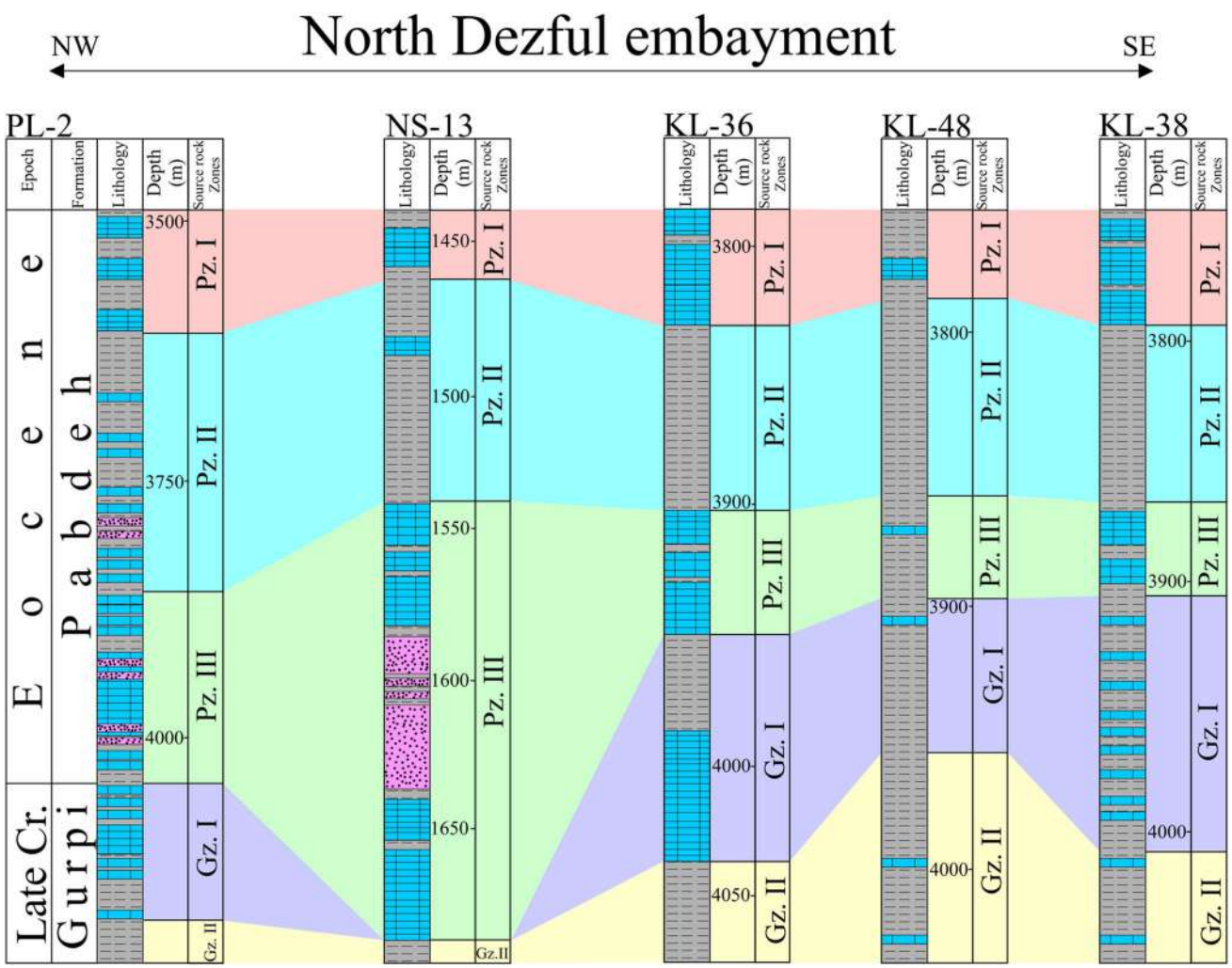

Fig. 18 Correlation of Pabdeh and Gurpi source rock zonation from NW to SE of north Dezful embayment in studied wells

ray logs (SGR and CGR), DT and the RT logs, are signs of organic materials in this zone. This zone was formed before the Pyrenean orogenic processes, and deep facies indicate the sedimentary basin's deepening and suitable conditions for forming source rock rich in organic materials (Alizadeh et al. 2018, 2020). The HI index in this zone is between 95 and 750 mg HC/g TOC and mainly type II kerogens. This zone has the highest amount of TOC and Highest potential for hydrocarbon production compared to other zones of the Pabdeh Formation. Its kerogen is mainly sapropelic, and its oil production ratio is more than gas.

- *Lower zone of Pabdeh Formation (Pz. III)*: Same as Pz. I; it consists of gray to sometimes dark green calcareous, marly shales, and dark limestones. The thickness of this zone is between 43 and 175 m in the studied wells. TOC is, on average, 0.72 wt% by weight. The amount of TOC increases in the wells (NS-13 and PL-2) where this zone has marl. The HI index in this zone is between 7 and 410 mg HC/g TOC and type II kerogens.
- *Upper zone of Gurpi Formation (Gz. I)*: This zone includes brown to dark limestones with gray shales. The thickness of this zone is between 58 and 135 m. TOC is, on average, 0.34 wt%. HI index in this zone is between 16 and 370 mg HC/g TOC and type II and III kerogens.
- *Lower zone of Gurpi Formation (Gz. II)*: This zone includes gray shales and sometimes brown limestone interlayer. The thickness of this zone is between 6 and 78 m. TOC is, on average, 0.98 wt%. HI index in this zone is between 79 and 623 mg HC/g TOC and type II kerogens. In a limited number of horizons in this zone, the TOC level reaches three wt%, which is suitable for hydrocarbon production.

During the geological period, the Zagros region had three major sedimentary basins, including the epicontinental platform (from the middle of the Permian to the Jurassic),

the platform to the southwest, and the shelf to the northeast (from the Jurassic to the end of the Cretaceous), and the proforland depositional system (from the middle of the Cretaceous to now) (Alavi 2004). Pabdeh and Gurpi formations were formed early in this proforland system. These two formations' tectonic settings have been located in two foredeeps according to the formation age. Many researchers have reported deep sedimentary environments in foredeep (e.g., Alavi 2004; Heydari 2008). On the other hand, according to Heydari's (2008) studies, the Pabdeh Formation is in the Sassan supersequence, and the Gurpi Formation was formed in the Ardavan supersequence. At the beginning of these supersequences, with the increase of the water level and reaching the maximum flooding surface, a deep marine environment had formed, and in these environments, suitable conditions such as anoxia level, abundance of Organic materials, and fine-grain sediment to form the source rock of hydrocarbon (Heydari 2008; Alizadeh et al. 2017, 2020). The brown shales of the Pabdeh Formation (Pz. II) were formed at the maximum flooding surface and have the most significant source rock potential. The Gurpi Formation has also formed in a deep environment. However, based on reasons such as low thickness of layers and reduction of organic matter, compared to Pabdeh Formation, it has less potential for oil production. Nevertheless, some of the layers of this formation show relatively good potential (Alavi 2004; Heydari 2008).

Conclusions

A robust multi-layer perceptron neural network (MLPNN) model was built to accurately estimate the total organic carbon (TOC) and hydrogen index (HI) for Pabdeh and Gurpi Formations. The following conclusions could be drawn from the results obtained in this study:

- In this study, the log*R* technique proved inappropriate for determining TOC. This is the result of a combination of several reasons, such as (a) only taking into account the total resistivity log (RT) and the sonic transit time log (DT) logs; (b) not making use of any additional logs; and (c) failing to verify the relationships between the logs. Only the areas of the NS-13 well that are particularly abundant in organic matter (OM) will be seen using this technique.
- MLPNN, with its structure based on the multi-layered perceptron, can be useful for both TOC and HI values. In this technique, due to the small number of datasets of a lbfgs solving function and the adaptive learning rate, the network could be trained with high accuracy (TOC R^2: 0.95; HI R^2: 0.93). This model's accuracy was significantly improved by several aspects, including the high precision of the log data and their strong convergence with each other. During the validation process, this model likewise demonstrated high accuracy (TOC R^2: 0.91; HI R^2: 0.89). The overfitting data hypothesis was debunked in the last phase by applying validation on log data from other wells. As a result, a large quantity of anticipated TOC was projected and calculated on organic matter (OM) prone horizons. This method has an accuracy of 0.93 for predicting TOC, and it has an accuracy of 0.9 for predicting HI in the well that was tested.
- The amount of laboratory-measured TOC in the Pabdeh and Gurpi Formations was 1.07 and 1.2 wt% on average. The highest amount of laboratory-measured TOC (> 1.5 wt%) has been recorded in the dark gray to brown shales of the Pabdeh Formation. These formations can produce hydrocarbons in a suitable amount and are of kerogen type II (sapropelic-humic combination, oil–gas prone).
- The Pabdeh Formation consists of three zones. The middle zone of the Pabdeh Formation (Pz. II), compared to the other two zones, has more organic matter (The average TOC is 2.6 wt%) and can produce more hydrocarbons. These are due to some factors such as (a) the decrease in density log, (b) the difference between the two gamma-ray logs, and values of DT and RT data for this formation. The kerogen of this formation is mainly type II, and according to the HI index, it can produce oil and gas.
- Gurpi Formation has two zones. This formation has a weak source rock potential due to the amount of TOC (> 1 wt%), and only in a few limited horizons of gray shales in the lower zone (Gz. II) can play a role in hydrocarbon production. The kerogen of this formation is of type II and III, and due to the relatively low HI index, it can produce oil and gas (mainly gas).

Appendices

Tables 1, 2, 3 and 4.

Table 1 Rock–Eval pyrolysis derived parameters in NS-13

Formation	Sample	Depth	S1	S2	S3	HI	OI	PI	T max	TOC
		m	mg HC/g	mg HC/g	mg HC/g	mg HC/g	mg HC/g		°C	wt%
Pabdeh	1	1440.3	0.93	1.86	1.16	388	241	0.33	425	0.48
Pabdeh	2	1454.4	0.44	1.39	1.10	290	230	0.24	423	0.48
Pabdeh	3	1459.8	0.38	0.87	1.07	290	357	0.30	422	0.3
Pabdeh	4	1469.3	1.17	4.08	1.30	340	108	0.22	428	1.2
Pabdeh	5	1476.3	1.41	8.75	1.57	350	63	0.14	438	2.5
Pabdeh	6	1484.8	1.35	7.33	1.49	386	78	0.16	435	1.9
Pabdeh	7	1493.2	0.56	2.28	1.16	190	97	0.20	425	1.2
Pabdeh	8	1503.4	1.68	6.37	1.45	246	56	0.21	434	2.59
Pabdeh	9	1508.5	0.41	2.75	1.18	220	94	0.13	425	1.25
Pabdeh	10	1515.8	1.36	7.70	1.51	372	73	0.15	436	2.07
Pabdeh	11	1520.8	0.56	2.76	1.19	230	99	0.17	425	1.2
Pabdeh	12	1529.5	0.78	4.50	1.30	253	73	0.15	429	1.78
Pabdeh	13	1536.3	0.70	1.62	1.13	108	75	0.30	424	1.5
Pabdeh	14	1544.4	0.66	1.74	1.14	171	111	0.27	424	1.02
Pabdeh	15	1554.9	0.43	0.99	1.08	141	154	0.30	423	0.7
Pabdeh	16	1559.4	0.27	0.72	1.06	167	245	0.27	422	0.43
Pabdeh	17	1568.0	0.25	0.74	1.06	105	151	0.26	422	0.7
Pabdeh	18	1574.3	0.35	0.74	1.06	148	212	0.32	422	0.5
Pabdeh	19	1589.2	0.53	1.91	1.14	239	142	0.22	424	0.8
Pabdeh	20	1604.9	0.53	2.96	1.20	190	77	0.15	426	1.56
Pabdeh	21	1612.5	0.42	2.01	1.14	251	142	0.17	424	0.8
Pabdeh	22	1619.1	0.34	1.59	1.11	269	188	0.18	423	0.59
Pabdeh	23	1631.7	0.30	1.78	1.12	99	62	0.14	424	1.8
Pabdeh	24	1641.5	0.48	0.95	1.08	159	180	0.33	423	0.6
Pabdeh	25	1647.9	0.37	1.45	1.10	227	172	0.20	423	0.64
Pabdeh	26	1655.7	0.47	0.90	1.08	173	207	0.34	423	0.52
Pabdeh	27	1663.8	0.26	0.39	1.04	118	314	0.40	422	0.33
Pabdeh	28	1678.1	0.49	1.03	1.09	149	157	0.32	423	0.69
Gurpi	29	1686.2	0.62	1.37	1.11	117	95	0.31	423	1.17
Gurpi	30	1691.2	0.43	1.85	1.13	125	76	0.19	424	1.48
Gurpi	31	1692.1	0.27	0.53	1.05	56	110	0.33	422	0.95

Table 2 Petrophysical training and testing dataset in NS-13 (N: not detected)

Formation	Sample	Dataset type	DEPTH m	DT us/f	RHOB gr/cm^3	RT ohmm	CGR gapi	SGR gapi	NPHI v/v
Pabdeh	1	Train	1440.3	N	N	N	N	52.70	N
Pabdeh	2	Train	1454.4	N	N	3.27	41.52	87.49	0.33
Pabdeh	3	Test	1459.8	N	N	3.99	46.02	103.45	0.25
Pabdeh	4	Train	1469.3	82.28	2.62	3.42	33.56	107.00	0.26
Pabdeh	5	Train	1476.3	83.75	2.57	2.83	30.29	91.14	0.33
Pabdeh	6	Train	1484.8	85.39	2.57	2.93	29.32	92.16	0.29
Pabdeh	7	Train	1493.2	73.15	2.63	3.74	30.46	79.60	0.26
Pabdeh	8	Test	1503.4	73.05	2.62	5.92	23.67	69.78	0.17
Pabdeh	9	Train	1508.5	73.24	2.65	4.84	29.87	84.05	0.20
Pabdeh	10	Train	1515.8	76.59	2.62	4.44	31.62	89.69	0.27
Pabdeh	11	Train	1520.8	73.93	2.64	6.18	29.50	86.85	0.21
Pabdeh	12	Train	1529.5	71.12	2.64	6.04	26.58	73.81	0.18
Pabdeh	13	Test	1536.3	67.01	2.65	7.11	22.47	62.25	0.14
Pabdeh	14	Train	1544.4	66.74	2.63	6.45	19.04	47.03	0.13
Pabdeh	15	Train	1554.9	66.87	2.62	6.65	18.25	44.37	0.12
Pabdeh	16	Train	1559.4	68.93	2.58	9.85	14.43	34.92	0.13
Pabdeh	17	Test	1568.0	69.41	2.61	5.86	18.05	44.87	0.14
Pabdeh	18	Train	1574.3	87.90	2.47	5.39	31.13	56.25	0.30
Pabdeh	19	Train	1589.2	76.92	2.54	16.49	16.36	67.63	0.19
Pabdeh	20	Train	1604.9	74.44	2.46	29.63	8.48	33.86	0.14
Pabdeh	21	Test	1612.5	71.15	2.51	20.55	8.50	30.72	0.14
Pabdeh	22	Train	1619.1	80.69	2.51	20.01	13.09	66.70	0.18
Pabdeh	23	Train	1631.7	56.87	2.56	42.24	9.73	47.43	0.08
Pabdeh	24	Test	1641.5	64.22	2.62	9.35	17.43	40.00	0.13
Pabdeh	25	Train	1647.9	62.55	2.63	10.89	17.01	40.12	0.11
Pabdeh	26	Train	1655.7	66.64	2.60	7.04	16.47	33.01	0.14
Pabdeh	27	Train	1663.8	69.07	2.61	4.57	18.15	40.92	0.15
Pabdeh	28	Train	1678.1	76.62	2.51	3.23	14.80	26.16	0.24
Gurpi	29	Test	1686.2	69.89	2.60	8.60	17.85	49.80	0.15
Gurpi	30	Train	1691.2	66.34	2.64	8.62	20.99	57.93	0.15
Gurpi	31	Train	1692.1	63.25	2.64	13.24	23.86	68.16	0.11

Table 3 Training and testing TOC and HI values from Rock–Eval pyrolysis, Δlog*R*, and MLPNN in NS-13 (N: not detected)

Formation	Sample	Dataset type	DEPTH m	$TOC_{Rock–Eval}$ wt%	$HI_{Rock–Eval}$ mg HC/g	$TOC_{\Delta logR}$ wt%	TOC_{MLPNN} wt%	HI_{MLPNN} mg HC/g
Pabdeh	1	Train	1440.3	0.5	388	N	0.50	450
Pabdeh	2	Train	1454.4	0.5	290	1.09	0.43	321
Pabdeh	3	Test	1459.8	0.3	290	0.85	0.20	263
Pabdeh	4	Train	1469.3	1.2	340	0.50	1.38	339
Pabdeh	5	Train	1476.3	2.5	350	0.35	2.30	385
Pabdeh	6	Train	1484.8	1.9	386	0.48	2.26	441
Pabdeh	7	Train	1493.2	1.2	190	0.10	1.00	194
Pabdeh	8	Test	1503.4	2.6	246	0.64	2.80	216
Pabdeh	9	Train	1508.5	1.3	220	0.41	1.37	232
Pabdeh	10	Train	1515.8	2.1	372	0.49	1.93	341
Pabdeh	11	Train	1520.8	1.2	230	0.74	1.38	255
Pabdeh	12	Train	1529.5	1.8	253	0.56	1.95	234
Pabdeh	13	Test	1536.3	1.5	108	0.53	1.94	86
Pabdeh	14	Train	1544.4	1.0	171	0.40	0.66	135
Pabdeh	15	Train	1554.9	0.7	141	0.44	0.57	120
Pabdeh	16	Train	1559.4	0.4	167	1.03	0.45	202
Pabdeh	17	Test	1568.0	0.7	105	0.43	0.63	110
Pabdeh	18	Train	1574.3	0.5	148	1.35	0.31	113
Pabdeh	19	Train	1589.2	0.8	239	2.08	0.65	185
Pabdeh	20	Train	1604.9	1.6	190	2.65	1.47	247
Pabdeh	21	Test	1612.5	0.8	251	2.03	0.89	257
Pabdeh	22	Train	1619.1	0.6	269	2.52	0.40	223
Pabdeh	23	Train	1631.7	1.8	99	2.10	2.00	97
Pabdeh	24	Test	1641.5	0.6	159	0.70	0.46	166
Pabdeh	25	Train	1647.9	0.6	227	0.79	0.46	184
Pabdeh	26	Train	1655.7	0.5	173	0.50	0.45	192
Pabdeh	27	Train	1663.8	0.3	118	0.11	0.52	123
Pabdeh	28	Train	1678.1	0.7	149	0.12	0.43	160
Gurpi	29	Test	1686.2	1.2	117	0.92	0.98	93
Gurpi	30	Train	1691.2	1.5	125	0.72	1.32	78
Gurpi	31	Train	1692.1	1.0	56	1.07	0.77	57

Table 4 Comparison between this study and other studies that used neural network, based on the input data and R^2 factor

This study/other references	Technique	Input data	R^2
This study	MLPNN	DT, RT, CGR, SGR, NPHI, RHOB	0.93
Tan et al. (2013)	RBF	AC, GR, DEN, CNL, RT	0.85
Alizadeh et al. (2012)	ANN	GR, CGR, SGR, THOR, POTA, DT, NPHI	0.94
Wang et al. (2019)	CNN	RT, DN, CNL, AC	0.79
Wang et al. (2019)	BPANN	RT, DN, CNL, AC	0.75
Zheng et al. (2021)	MLPNN	GR, AC, DNL, RT, DEN	0.9
Chan et al. (2022)	ANN	GR, CGR, SGR, THOR, POTA, DT, NPHI, RT	0.83

Acknowledgements This research is part of the senior author's Ph.D. thesis and was supported by the Department of Geology, Ferdowsi University of Mashhad (Research Code: 3/58338), Iran. The authors would like to acknowledge the Ferdowsi University of Mashhad, Iran, and the National Iranian South Oil Company, Iran, for financial and technical support.

Funding This study received no definite grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of interest The authors declare that they have no known competing commercial interests or personal relationships that could have performed to affect the work reported in this paper.

References

Abubakar R, Adomako-Ansah K, Marfo SA, Fenyi C, Owusu JA (2023) Integrated geochemical and statistical evaluation of the source rock potential in the deep-water, Western Basin of Ghana. J Pet Sci Eng 220:111164

Aghanabati A (2004) Geology of Iran. GSI Publication, Tehran, p 586

Ala MA, Kinghorn RRF, Rahman M (1980) Organic geochemistry and source rock characteristics of the Zagros petroleum province, southwest Iran. J Pet Geol 3(1):61–89

Alavi M (2004) Regional stratigraphy of the Zagros fold-thrust belt of Iran and its pro-foreland evolution. Am J Sci 304(1):1–20

Alavi M (2007) Structures of the Zagros fold-thrust belt in Iran. Am J Sci 307(9):1064–1095

Al-Husseini MI (2000) Origin of the Arabian plate structures: Amar Collision and Najd Rift. GeoArabia 5:527–542

Alizadeh B, Najjari S, Kadkhodaie-Ilkhchi A (2012) Artificial neural network modeling and cluster analysis for organic facies and burial history estimation using well log data: a case study of the South Pars Gas Field, Persian Gulf, Iran. Comput Geosci 45:261–269

Alizadeh B, Alipour M, Chehrazi A, Mirzaie S (2017) Chemometric classification and geochemistry of oils in the Iranian sector of the southern Persian Gulf Basin. Org Geochem 111:67–81

Alizadeh B, Maroufi K, Heidarifard MH (2018) Estimating source rock parameters using wireline data: an example from Dezful Embayment, South West of Iran. J Pet Sci Eng 167:857–868

Alizadeh B, Opera A, Kalani M, Alipour M (2020) Source rock and shale oil potential of the Pabdeh Formation (Middle–Late Eocene) in the Dezful Embayment, southwest Iran. Geol Acta 18:1–22

Al-Yaseri A, Abu-Mahfouz IS, Yekeen N, Wolff-Boenisch D (2023) Organic-rich source rock/H2/brine interactions: implications for underground hydrogen storage and methane production. J Energy Storage 63:106986

Awan RS, Liu C, Aadil N, Yasin Q, Salaam A, Hussain A, Yang S, Jadoon AK, Wu Y, Gul MA (2021) Organic geochemical evaluation of Cretaceous Talhar Shale for shale oil and gas potential from Lower Indus Basin, Pakistan. J Pet Sci Eng 200:108404

Behar F, Beaumont V, De Penteado BHL (2001) Rock–Eval 6 technology: performances and developments. Oil Gas Sci Technol 56(2):111–134

Bolandi V, Kadkhodaie A, Farzi R (2017) Analyzing organic richness of source rocks from well log data by using SVM and ANN classifiers: a case study from the Kazhdumi formation, the Persian Gulf basin, offshore Iran. J Pet Sci Eng 151:224–234

Bordenave ML (2002) The middle cretaceous to early Miocene petroleum system in the Zagros domain of Iran, and its prospect evaluation. AAPG Annu Meet Am Assoc Pet Geol Houst 6:1–9

Bordenave ML, Burwood R (1994) The Albian Kazhdumi Formation of the Dezful Embayment, Iran: one of the most efficient petroleum generating systems. In: Petroleum source rocks. Springer, Berlin, p 227

Bordenave ML, Hegre JA (2010) Current distribution of oil and gas fields in the Zagros Fold Belt of Iran and contiguous offshore as the result of the petroleum systems. In: Tectonic and stratigraphic evolution of Zagros and Makran during the Mesozoic-Cenozoic, vol 330. Geological Society, London, Special Publications, pp 291–353

Chalmers GRL, Bustin RM (2007) The organic matter distribution and methane capacity of the Lower Cretaceous strata of Northeastern British Columbia, Canada. Int J Coal Geol 70(1–3):223–239

Chan SA, Hassan AM, Usman M, Humphrey JD, Alzayer Y, Duque F (2022) Total organic carbon (TOC) quantification using artificial neural networks: improved prediction by leveraging XRF data. J Pet Sci Eng 208:109302

Curtis JB (2002) Fractured shale-gas systems. Am Assoc Pet Geol Bull 86(11):1921–1938

Dembicki H Jr (2009) Three common source rock evaluation errors made by geologists during prospect or play appraisals. Am Assoc Pet Geol Bull 93(3):341–356

Diab AI, Sanuade O, Radwan AE (2023) An integrated source rock potential, sequence stratigraphy, and petroleum geology of (Agbada–Akata) sediment succession, Niger delta: application of well logs aided by 3D seismic and basin modeling. J Pet Explor Prod Technol 13(1):237–257

Espitalie J, Deroo G, Marquis F (1985) Rock–Eval pyrolysis and its applications. Part One. Rock–Eval Pyrolysis and its applications (Part One) Oil & Gas Science and Technology - Rev. IFP, vol. 40, No. 5, pp 563–579

Fathy D, El-Balkiemy AF, Makled WA, Hosny AM (2023) Organic geochemical signals of Paleozoic rocks in the southern Tethys, Siwa basin, Egypt: implications for source rock characterization and petroleum system. Phys Chem Earth 130:103393

Gasparik M, Bertier P, Gensterblum Y, Ghanizadeh A, Krooss BM, Littke R (2014) Geological controls on the methane storage capacity in organic-rich shales. Int J Coal Geol 123:34–51

Hakimi MH, Ahmed AF, Abdullah WH (2016) Organic geochemical and petrographic characteristics of the Miocene Salif organic-rich shales in the Tihama Basin, Red Sea of Yemen: implications for paleoenvironmental conditions and oil-generation potential. Int J Coal Geol 154:193–204

Hakimi MH, Lashin A, Varfolomeev MA, Rahim A, Sen S, Naseem W, Saeed SA, Al-Muntaser AA, Qadri SMT, Mustapha KA (2023) Oil generation and expulsion modeling of the syn-rift Salif oil-source rock in the Tihamah Basin, Yemeni Red Sea: implications for shale oil exploration. J Afr Earth Sci 202:104924

Harland WB (1990) A geologic time scale 1989. Cambridge University Press, Cambridge, p 256

Hassan A, Mohammed E, Oshaish A, Badhafere D, Ayranci K, Dong T, Waheed U, El-Husseiny A, Mahmoud M (2023) Prediction of total organic carbon in organic-rich shale rocks using thermal neutron parameters. ACS Omega 8(5):4790–4801

He J, Ding W, Zhang J, Li A, Zhao W, Dai P (2016) Logging identification and characteristic analysis of marine—continental transitional organic-rich shale in the Carboniferous-Permian strata, Bohai Bay Basin. Mar Pet Geol 70:273–293

Hessami K, Koyi HA, Talbot CJ, Tabasi H, Shabanian E (2001) Progressive unconformities within an evolving foreland fold–thrust belt, Zagros Mountains. J Geol Soc 158(6):969–981

Heydari E (2008) Tectonics versus eustatic control on supersequences of the Zagros Mountains of Iran. Tectonophysics 451(1–4):56–70

Hosseini Asgarabadi Z, Khodabakhsh S, Mohseni H, Halverson G, Hao Bui T, Abbassi N, Moghaddasi A (2019) Microfacies, geochemical characters and possible mechanism of rhythmicdeposition of the Pabdeh Formation in SE Ilam (SW Iran). Geopersia 9(1):89–109

Jackson KS, Hawkings PJ, Bennett AJR (1985) Regional facies and geochemica evaluation of Southern Denison Trough. Aust Pet Explor Assoc J J 20:143–158

James GA, Wynd JG (1965) Stratigraphic nomenclature of Iranian Oil Consortium agreement area. Am Assoc Pet Geol Bull 49:2182–2245

Kadkhodaie-Ilkhchi A, Rahimpour-Bonab H, Rezaee MR (2009) A committee machine with intelligent systems for estimation of total organic carbon content from petrophysical data: an example from Kangan and Dalan reservoirs in South Pars Gas Field, Iran. Comput Geosci 35:457–474

Kamali MR, Mirshady AA (2004) Total organic carbon content determined from well logs using ΔlogR and neuro-fuzzy techniques. J Pet Sci Eng 45:141–148

Karimi AR, Rabbani AR, Kamali MR (2016) A bulk kinetic, burial history and thermal modeling study of the Albian Kazhdumi and the Eocene-Oligocene Pabdeh formations in the Ahvaz anticline, Dezful Embayment, Iran. J Pet Sci Eng 146:61–70

Khalil Khan H, Ehsan M, Ali A, Amer MA, Aziz H, Khan A, Bashir Y, Abu-Alam T, Abioui M (2022) Source rock geochemical assessment and estimation of TOC using well logs and geochemical data of Talhar Shale, Southern Indus Basin, Pakistan. Front Earth Sci 10:969936

Lai H, Li M, Mao F, Liu J, Xiao H, Tang Y, Shi S (2020) Source rock types, distribution, and their hydrocarbon generative potential within the Paleogene Sokor-1 and LV formations in Termit Basin, Niger. Energy Explor Exploit 38(6):2143–2168

Li Y, Lu J, Liu X, Wang J, Chen S, He Q (2023) Geochemical characteristics of source rocks and gas exploration direction in Shawan Sag, Junggar Basin, China. Nat Gas Geosci 8(2):95–107

Liu Q, Jin Z, Liu W, Lu L, Meng Q, Tao Y, Han P (2013) Presence of carboxylate salts in marine carbonate strata of the Ordos Basin and their impact on hydrocarbon generation evaluation of low TOC, high maturity source rocks. Sci China Earth Sci 56:2141–2149

Liu Y, Zeng J, Qiao J, Yang G, Cao W (2023) An advanced prediction model of shale oil production profile based on source-reservoir assemblages and artificial neural networks. Appl Energy 333:120604

Liu WH (2016) Study on the evaluation model of shale reservoirs and well logging interpretation of Zhen Jing Region. Doctoral thesis, China University of Geosciences, Wuhan, Hubei, China, p 327

Mahdi AQ, Abdel-Fattah MI, Hamdan HA (2022) An integrated geochemical analysis, basin modeling, and palynofacies analysis for characterizing mixed organic-rich carbonate and shale rocks in Mesopotamian Basin, Iraq: insights for multisource rocks evaluation. J Pet Sci Eng 216:110832

McQuarrie N, Van Hinsbergen DJJ (2013) Retrodeforming the Arabia–Eurasia collision zone: age of collision versus magnitude of continental subduction. Geology 41:315–318

Miller KG, Kominz MA, Browning JV, Wright JD, Mountain GS, Katz ME, Sugarman PJ, Cramer BS, Christie-Blick N, Pekar SF (2005) The Phanerozoic record of global sea-level change. Science 310(5752):1293–1298

Müller RD, Sdrolias M, Gaina C, Steinberger B, Heine C (2008) Long-term sea-level fluctuations driven by ocean basin dynamics. Science 319(5868):1357–1362

Ouadfeul SA, Aliouane L (2014) Shale gas reservoirs characterization using neural network. Energy Procedia 59:16–21

Passey QR, Creaney S, Kulla JB, Moretti FJ, Stroud JD (1990) A practical modal for organic richness from porosity and resistivity logs. Am Assoc Pet Geol Bull 74:1777–1794

Peters KE (1986) Guidelines for evaluating petroleum source rock using programmed pyrolysis. Am Assoc Pet Geol Bul 70(3):318–329

Peters KE, Cassa MR (1994) Applied source rock geochemistry: Chapter 5: Part II. Essential elements. In: Magoon LB, Dow WG (eds) The petroleum system—from source to Trap. AAPG Mem 60, New York, pp 93–120

Pirouz M, Avouac JPh, Hassanzadeh J, Kirschvink JL, Bahroudi A (2017) Early Neogene foreland of the Zagros, implications for the initial closure of the Neo-Tethys and kinematics of crustal shortening. EPSL 477:168–182

Rabbani AR, Kamali MR (2006) Source rock evaluation and petroleum geochemistry, offshore SW Iran. J Pet Geol 28(4):413–428

Sanders MM, Jubb AM, Hackley PC, Peters KE (2022) Molecular mechanisms of solid bitumen and vitrinite reflectance suppression explored using hydrous pyrolysis of artificial source rock. Org Geochem 165:104371

Schmoker JW (1981) Determination of organic-matter content of Appalachian Devonian shales from gamma-ray logs. Am Assoc Pet Geol Bul 65(7):1285–1298

Schmoker JW, Hester TC (1983) Organic carbon in Bakken Formation, United States portion of Williston Basin. Am Assoc Pet Geol Bul 67(12):2165–2174

Setudehnia A (1978) The Mesozoic sequence in south-west Iran and adjacent areas. J Pet Geol 1(1):3–42

Shi X, Wang J, Liu G, Yang L, Ge X, Jiang S (2016) Application of extreme learning machine and neural networks in total organic carbon content prediction in organic shale with wireline logs. J Nat Gas Sci Eng 33:687–702

Shuangfang L, Huang W, Fangwen C, Jijun L, Min W, Haitao X, Weiming W, Xiyuan C (2012) Classification and evaluation criteria of shale oil and gas resources: discussion and application. Pet Explor Dev 39(2):268–276

Stow DAV, Huc AY, Bertrand P (2001) Depositional processes of black shales in deep water. Mar Pet Geol 18(4):491–498

Tabatabaei SMH, Kadkhodaie-Ilkhchi A, Hosseini Z, Asghari Moghaddam A (2015) A hybrid stochastic-gradient optimization to estimating total organic carbon from petrophysical data: a case study from the Ahwaz oilfield, SW Iran. J Pet Sci Eng 127:35–43

Tan M, Liu Q, Zhang S (2013) A dynamic adaptive radial basis function approach for total organic carbon content prediction in organic shale. Geophysics 78(6):445–459

Tissot BP, Welte DH (1984) Petroleum formation and occurrence. Springer, New York, p 966

Valdon YB, Oboh-Ikuenobe FE, Maigari SA, Zobaa MK (2023) Source-rock potential and paleoenvironment of the Coniacian Numanha Formation, northeastern Nigeria: an integrated palynofacies and geochemical approach. J Afr Earth Sci 198:104790

Vega-Ortiz C, Beti DR, Setoyama E, McLennan JD, Ring TA, Levey R, Martínez-Romero N (2020) Source rock evaluation in the central-western flank of the Tampico Misantla Basin, Mexico. J S Am Earth Sci 100:102552

Wang P, Peng S, He T (2018) A novel approach to total organic carbon content prediction in shale gas reservoirs with well logs data, Tonghua Basin, China. J Nat Gas Sci Eng 55:1–15

Wang H, Wu W, Chen T, Dong X, Wang G (2019) An improved neural network for TOC, S1, and S2 estimation based on conventional well logs. J Pet Sci Eng 176:664–678

Wu Y, Liu C, Jiang F, Hu T, Lv J, Zhang C, Guo X, Huang L, Hu M, Huang R, Awan RS, Zhao Y (2022) Geological characteristics and shale oil potential of alkaline lacustrine source rock in Fengcheng Formation of the Mahu Sag, Junggar Basin, Western China. J Pet Sci Eng 216:110823

Wu G, Kong F, Tian N, Ma T, Tao C (2023) Structural characteristics and deep-water hydrocarbon accumulation model of the Scotian Basin. Eastern Canada. Energy Geosci 4(3):100152

Xu K, Chen S, Liu C, Liu Z, Li C (2022) Geochemical characteristics of source rocks and natural gas in Fudong area, Junggar Basin: implications for the genesis of natural gas. Arab J Geosci 15(6):544

Yandoka BMS, Abdullah WH, Abubakar MB, Hakimi MH, Jauro A, Adegoke AK (2016) Organic geochemical characterization of shallow marine cretaceous formations from Yola sub-basin, northern Benue trough, NE Nigeria. J Afr Earth Sci 117:235–251

Zheng D, Wu S, Hou M (2021) Fully connected deep network: an improved method to predict TOC of shale reservoirs from well logs. Mar Pet Geol 132:105205

Zheng T, Grohmann S, Arysanto A, Baniasad A, Zhang Q, Littke R (2023) Petrographical and geochemical investigation on maturation and primary migration in intact source rock micro-plugs: Insight from hydrous pyrolysis on Woodford Shale. Int J Coal Geol 266:104170