E3S Web of Conferences **35**, 03008 (2018) https://doi.org/10.1051/e3sconf/20183503008
POL-VIET 2017

Applying of the Artificial Neural Networks (ANN) to Identify and Characterize S*weet Spots* in Shale Gas Formations

Edyta Puskarczyk[1, *]

[1] AGH University of Science and Technology, Faculty of Geology, Geophysics and Environmental Protection, A. Mickiewicza Av. 30, 30-059 Krakow, Poland

Abstract. The main goal of the study was to enhance and improve information about the Ordovician and Silurian gas-saturated shale formations. Author focused on: firstly, identification of the shale gas formations, especially the *sweet spots* horizons, secondly, classification and thirdly, the accurate characterization of divisional intervals. Data set comprised of standard well logs from the selected well. Shale formations are represented mainly by claystones, siltstones, and mudstones. The formations are also partially rich in organic matter. During the calculations, information about lithology of stratigraphy weren't taken into account. In the analysis, self-organizing neural network – Kohonen Algorithm (ANN) was used for *sweet spots* identification. Different networks and different software were tested and the best network was used for application and interpretation. As a results of Kohonen networks, groups corresponding to the gas-bearing intervals were found. The analysis showed diversification between gas-bearing formations and surrounding beds. It is also shown that internal diversification in sweet spots is present. Kohonen algorithm was also used for geological interpretation of well log data and electrofacies prediction. Reliable characteristic into groups shows that Ja Mb and Sa Fm which are usually treated as potential sweet spots only partially have good reservoir conditions. It is concluded that ANN appears to be useful and quick tool for preliminary classification of members and sweet spots identification. **Key words** – shale gas, sweet spots, Artificial Neural Networks, classification.

1 Introduction

Advanced statistical methods and artificial neural networks can provide useful information for identification of productive horizons in oil and gas reservoirs [1]. Commercial software, used by oil and gas companies usually contain modules based on them. One of the basic types of self-organizing neural networks are Kohonen's networks. This type of neural networks can be used for data classification. As a result electrofacies were created. The term electrofacies was originally defined by Serra and Abbot [2]. An electrofacies represents a unique set of log responses, which characterizes physical properties of the rocks and fluids contained in the volume investigated by the logging tools.

* Corresponding author: puskar@agh.edu.pl

For determining electrofacies, authors like Serra [3], Doveton [4], Moss [5] suggest to use cluster analysis in *n* dimensions with each wireline log as a dimension. In this paper there are shown the results of using Kohonen algorithms for identification beds with different properties recorded on the well logs. After an automatic clustering and classification, each group was described for sweet spot identification. Several neural networks were tested using two different software, one dedicated for well logs analysis and the other one dedicated for statistical analysis.

2 Materials and methods

A single well located at the Polish Baltic Basin was chosen for the ANN tests. Data set contains data from the Silurian and Ordovician gas-saturated shale formations from Pa Fm to Sl Fm (Fig. 1). The depth intervals of each formations/members range as follow: Pa Fm (3736.5–3915 m), Pe Fm (3915–3938 m), Ja Mb (3938–3951 m), Pr Fm (3951– 957.5 m), Sa Fm (3957.5–3975 m), Ko Fm (3975–3980 m) and Sl Fm (3980–3983 m). An analysed depth interval is almost 250 m with sample rate equal to 0.1 m. That means that about 2500 objects were taken into account for the ANN calculations. Shale formations are represented mainly by claystones, siltstones, and mudstones. One of the formations, Ko Fm is represented mainly by limestones. Usually, bituminous shales have significant content of bentonite and pyrite. It was also observed the high content of marly limestones. The following standard well logs were chosen as an input data for ANN:

- CALI (mm) – calliper log, used for well QC, primary lithology indicator, in shales usually shows caves,
- GR (API) – natural gamma ray log, total radioactivity of the formation, reach high value in the shale rich formation, is treated as a primary shale indicator,
- POTA (%), THOR (ppm) and URAN (ppm) – individual concentration of potassium, thorium and uranium respectively, from spectral gamma ray log. Potassium and thorium tend to be concentrated in clays while uranium and its derivatives tend to be concentrated in source rocks because of adsorption by organic matter,
- DT (μs/ft) – compressional slowness log, in shales lowers sonic velocities are observed because of presence of organic matter,
- MSFL (omm), LLS (omm) and LLD (omm) – micro, shallow and deep resistivity logs respectively,
- NPHI (%) – neutron porosity, high values in shales, caused by high hydrogen index in organic matter and presence in clays free water, adsorbed water (clay bound water) as well as lattice-water, which is a part of clay mineral structure,
- RHOB (g/cm^3) – bulk density log, thick series of shales demonstrates progressive increase in density due to the compaction, when formation is overpressure there is a break in normal shale compaction trend manifested by a drop of bulk density, organic shales are distinguished by clear low anomalies on RHOB log because organic matter has much lower density in comparison with non-source shales,
- PE (b/e) – photoelectric factor log, is a function of chemical (mineral) composition and thus is indirectly related to lithology, but is independent from porosity.

Artificial neural networks proposed by Teuvo Kohonen [6] are the most popular model of self-organizing networks. Generalized learning algorithm for Kohonen networks is the following:

- the *n*-dimensional weight vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, …,$\boldsymbol{v}_i$ of the j computing units are randomly selected. Then an initial radius r, a learning constant a and a neighbourhood function α are selected,
- an input vector $\boldsymbol{\varepsilon}$ is selected by using the probability distribution over the input space,

- the unit k with the maximum excitation and for which the distance between $\boldsymbol{v}_i$ and ε is minimal is selected,
- using the neighbourhood function, the weight vectors are updated

$$\boldsymbol{v}_i \leftarrow \boldsymbol{v}_i + a\alpha(i,k)(\varepsilon - \boldsymbol{v}_i) \ \ for \ \ i = 1,\ldots,j \tag{1}$$

- modify a and/or α and continue from the beginning or stop after N iterations. By repeating the process several times, arriving at a uniform distribution of weight vectors in input space is expected. During the learning process both the size of the neighbourhood and the value of neighbourhood function fall gradually, so that the influence of each unit upon its neighbours is reduced. From an initial distribution of random weights, and over N iterations, the networks settles into a map of stable zones. Each zone is effectively a feature classifier.

Two different and independent software were used for electrofacies identification. The first one, IPSOM module at Techlog Schlumberger software was designed for geological interpretation of well log data and facies prediction. This module provides classification solutions based on 'classical' clustering and the Kohonen networks. IPSOM provides also QC tool for the results evaluation as a probability of occurrence of predicted facies at each depth. In this study hierarchical cluster analysis was performed based on the Euclidean metrics. The second, ANN module at Statistica software developed by StatSoft is an universal statistical tool used in many different fields, e.g. economics, electronics. It has the ability to select the type of network and its parameters by the user.

3 Statistical characterization of shales

Descriptive statistics were calculated for each lithostratigraphic formation and member. The statistics were calculated for both the well logs and the results of well logs comprehensive interpretation. The average value of each petrophysical parameter and standard deviation (values are in brackets, arithmetic average and standard deviation respectively) observed and interpreted in two selected formations (which can be treated as potential *sweet spots*) are presented below:

- **Ja Mb (a part of Pa Fm)** – 130 samples; CALI (221 and 0.15 mm), GR (134 and 27 API), POTA (3 and 0.5%), THOR (10 and 1 ppm), URAN (5.5 and 1.6 ppm), DT (80 and 5 us/ft), MSFL (94 and 70 omm), LLS (52 and 33 omm), LLD (60 and 41 omm), NPHI (18 and 5%), RHOB (2.6 and 0.05 g/cm^3); data from comprehensive interpretation: PHI (4 and 1%), VCL (58 and 14%), VSAND (22 and 11%), VLIME (10 and 10%), VDOLO (3 and 3%), VMIN Fe (0.04 and 0.1%), VKEROGEN (2.2 and 0.8%) and TOC (1.06 and 0.3%wt),
- **Sa Fm** – 165 samples; CALI (217 and 4.5 mm), GR (142 and 23 API), POTA (2.9 and 0.4%), THOR (11 and 1 ppm), URAN (6.3 and 2 ppm), DT (81 and 8 us/ft), MSFL (333 and 2203 omm, but median equal 35 omm), LLS (40 and 59 omm), LLD (50 and 75 omm), NPHI (19 and 4%), RHOB (2.47 and 0.2 g/cm^3); data from comprehensive interpretation: PHI (6 and 3%), VCL (52 and 11%), VSAND (21 and 16%), VLIME (15 and 16%), VDOLO (2 and 6%), VMIN Fe (0.08 and 0.3%), VKEROGEN (3.4 and 1%) and TOC (1.7 and 0.7%wt).

Similar mean GR values were obtained for Ja Mb, Sa Fm and other shale formations. These values were much lower (average 60 API) only for Ko Fm and much higher for Sl Fm. Similar relationships were obtained for POTA, THOR and URAN logs. DT average values ranged from 69 to 81 us/ft for all formations except for Ko Fm, where they reached lower (average 55 us/ft) values. Lower densities were observed for Ja Mb and Sa Fm. On the Pe curve only Ja Mb had lowered values and Ko Fm increased, the remaining formations reached an average about 3.8 b/e. At each shale formation, the NPHI values were high, while the decrease value was recorded for Ko Fm. Resistivity curves recorded large variation within each formation, but mean values did not show differences between the formations.

The statistical analysis of the petrophysical parameters of comprehensive well logs interpretation in individual formations shows that at Ja Mb and Sa Fm the highest porosities (average 5%) is observed. In sweet spots intervals also the highest content of sandstones (about 22%) is observed. Most limestone content (average 70%) has been observed for Ko Fm, and most dolomite content (average 12%) for Pr Fm. Highest values of kerogen content and total organic carbon were observed for Sa Fm and Ja Fm. The remaining shale formations (Pa Fm, Pe Fm, Pr Fm) and Ko Fm registered the TOC content below 1%wt.

The analysis of descriptive statistics shows that individual lithostratigraphic intervals differ from others, but also often show large internal variations.

4 Classification and *sweet spots* identification

Several dozen Kohonen networks were tested to determine the electrofacies. The tests started with a 10 × 10 grid. Finally, the best network was chosen, with a 2 × 3 grid. The networks created with exactly the same parameters and input in the two tested software gave very similar results. This article presents the results of two networks. The results were compared with the results of classical cluster analysis (Fig. 1). Network called IPSOM (Techlog) was created based on 12 input logs (described in the paragraph 2). Network called ANN 14_6 (Statistica) as an input included additionally TOC and PHI logs. We have decided to add an extra logs because when exactly the same input logs were applied for exactly the same type of networks, the results were almost the same. We have decided add two logs, total porosity (PHI) and total organic carbon content (TOC), obtained from comprehensive well log interpretation as a parameters which have high influence on the possible gas presents. In both cases, there were obtained 6 groups, but obtained division was no identical. By analysing the last two tracks shown in the Figure 1, it can be noted that the IPSOM network has allocated thinner layers within the individual lithostratigraphic units (shaded alternately on the Fig. 1). However, within the potential *sweet spots*, both networks showed a very similar distribution. By comparing the results of Kohonen networks to the results of ‘classical’ clustering analysis it can be concluded that: on the basis of cluster analysis, the division is approximately proportional to the lithostratigraphic division, while the Kohonen networks have shown the variation within these units.

Figures 2 and 3 show the distribution of mean values (square inside the box) and the range of variability (box frame sets minimum and maximum) of the standardized petrophysical parameters of comprehensive interpretation in created groups based on the IPSOM Techlog (Fig. 2) and ANN 14-6 Statistica (Fig. 3) networks.

Results of analysis:

- *sweet spots* can be identified by using ANN. For the IPSOM network within the Ja Mb and Sa Fm, groups 3 and 5 were identified. For the ANN 14_6 network within Ja Mb groups 1 and 5 were identified and within Sa Fm 1 and 4 were identified. Group 3 on IPSOM correspond to group 1 on ANN 14_6. In this groups high values of porosity and TOC are observed. Group 5 in IPSOM correspond to group 5 in ANN 14_6. Average values of each log are similar.
- it is therefore possible to conclude that in both *sweet spots* intervals there are levels with great potential for hydrocarbon presence in the top of the member/formation and in both *sweet spots* the lower part of the intervals has worse properties,

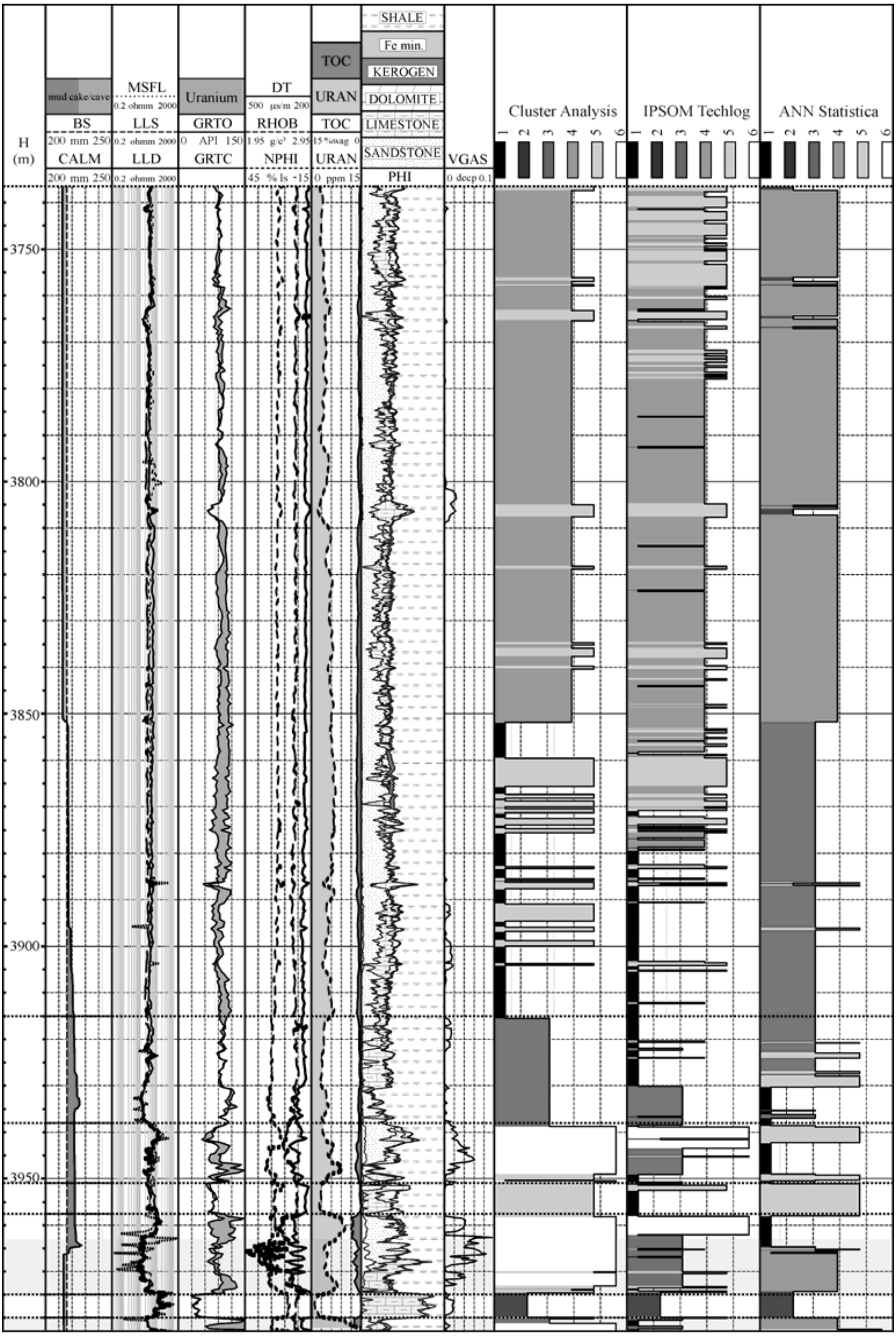

Fig. 1. Well logging data from measurements (first 5 tracks), interpretation (lithology track, TOC and gas content) and results of 'classical' cluster analysis and artificial neural networks application

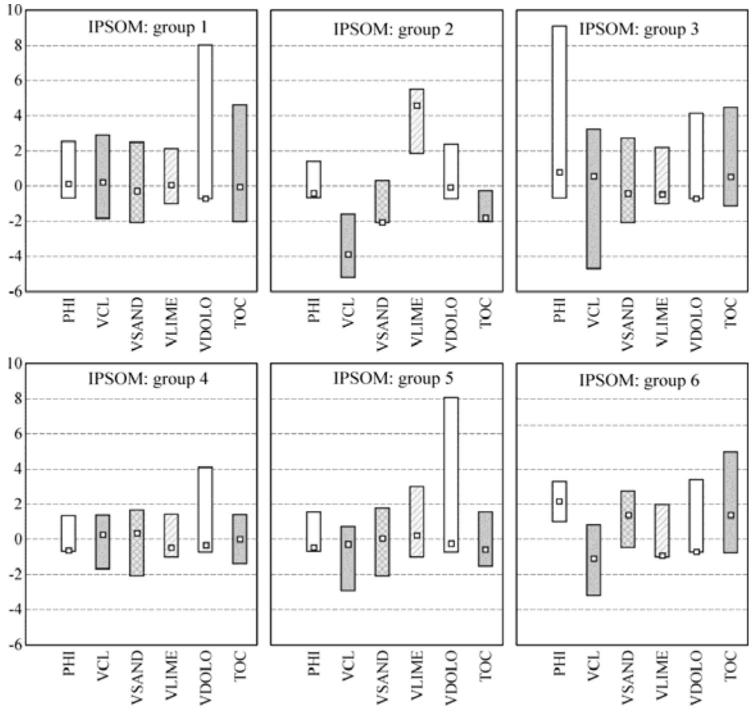

Fig. 2. Average values (square inside the box) and minimum – maximum (box), calculated for standardized results of comprehensive interpretation in each group based on IPSOM module

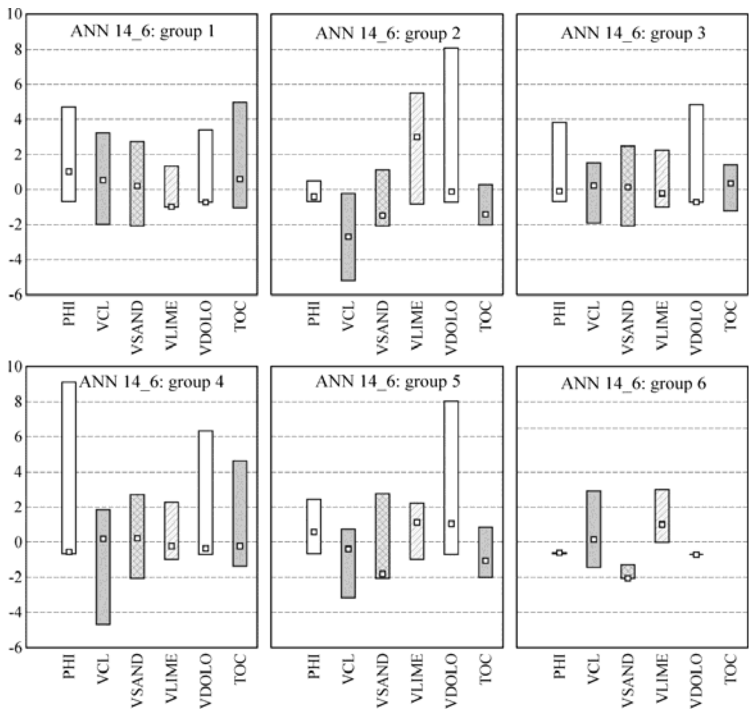

Fig. 3. Average values (square inside the box) and minimum – maximum (box), calculated for standardized results of comprehensive interpretation in each group based on ANN Statistica module

- it is worth to check the interval at which IPSOM shows electrofacies 1, because it has a high TOC value. This electrofacies was found at Pa Fm, Pe Fm and Pr Fm,
- in both networks electrofacies number 2 correspond to the Ko Fm interval. In this group limestone content is extremely high compared to other electrofacies,
- a comparison of 'classical' cluster analysis results to neural networks shows a similarity in the general distribution of clusters, but also differences in the detailed analysis of individual small layers,
- using two different software for Kohonen networks didn't show significant difference in terms of using identical inputs,
- adding two additional variables (PHI and TOC) to the ANN14_6 network resulted in large variations in the grouping compared to IPSOM. The selection of variables for analysis is one of the key elements of correct analysis.

4 Conclusions

The ANN were used for the classification and grouping of data accordingly to natural petrophysical features of rocks. Three were found groups corresponding to the potentially gas-saturated zones. The Kohonen networks confirm and refine the cluster analysis results.

- Complex analysis showed diversification between shale gas formations and adjacent beds.
- *Sweet spots* intervals were separated from the shale formation. It shows that internal diversification in each gas formation is present.
- Electrofacies determined on the basis of standard logs turned out to be useful in discrimination of reservoir and sealing parts of formation.
- All statistical methods that were applied to the rock classification gave satisfactory results.

It is concluded that artificial neural networks and especially self-organizing maps appear to be a useful and quick tool for preliminary classification of members and hydrocarbon-saturated zones identification. Statistical methods offer petrophysicists an additional tool for improving standard well log analysis. This offers possibility to define zones of interests based on the criteria petrophysical parameters defined from well logs interpretation.

Acknowledgements

Data was allowed by POGC Warsaw co, Poland for the MWSSSG Polskie Technologie dla Gazu Łupkowego project (2013-2017). Techlog 13 software was used under the AGH UST grant from Schlumberger. Statistica 12 software was used under the AGH UST grant from StatSoft.

References

1. E. Puskarczyk, J. Jarzyna, Sz. J. Porębski Geol. Q **59**, 1 (2015)
2. O. Serra and H. Abbot, SPE. 55th Annual Fall Technical Conference and Exhibition Dallas, Texas, paper No. 9270 (1980)
3. O. Serra, Edu. Serv., **272**, Buenos Aires (1987)
4. J. Doveton, AAPG Computer Applications in Geology, No. 2, 169, Tulsa (1994)
5. B. Moss, Geological Society Special Publication **122** (1997)
6. T. Kohonen, Biol. Cybern., **43**, 1 (1982)