УДК 551: 004.89 DOI: https://doi.org/10.24028/gj.v47i3.320290

Applying spectral decomposition to seismic facies clustering with unsupervised machine learning

R. Malikov, G. Babayev, 2025

Institute of Geology and Geophysics of the Ministry of Science and Education of the Republic of Azerbaijan, Baku, Azerbaijan
Received 5 January 2025

Seismic facies analysis, essential for subsurface geological exploration, has traditionally challenged the ability to capture subtle variations in complex stratigraphic environments. This study uses spectral decomposition and unsupervised machine learning, specifically the Kohonen Self-Organizing Map, to improve the identification of detailed seismic facies. Spectral decomposition enables frequency-based seismic data analysis, capturing intricate geological features often missed by traditional methods. The Continuous Wavelet Transform was applied to decompose seismic signals, and the resulting frequency components were clustered using a Self-Organizing Map to classify seismic facies. This paper validated this approach using seismic data from the South Caspian Basin. The results successfully identified channel systems and facies boundaries, enhancing their delineation and enabling a more accurate interpretation of channel systems and their internal variability. This automated methodology offers valuable insights for reservoir characterization and hydrocarbon exploration, potentially reducing exploration risks and enhancing resource estimation.

Key words: seismic facies, spectral decomposition, self-organizing map, continuous wavelet transform, unsupervised learning, machine learning.

Introduction. One of the techniques utilized to explore subsurface geology is seismic facies analysis, which uses seismic data to classify seismic wave behavior and features to extract subsurface stratigraphic characteristics [Brown, 2011; Kourki, Ali Riahi, 2014; Song et al., 2017; Wrona et al., 2018]. While effective in identifying major geological features, traditional seismic interpretation methods often fall short in capturing subtle variations in complex stratigraphic environments [Partyka et al., 1999; Chopra, Marfurt, 2008]. The recent rise in machine learning technologies, particularly unsupervised learning algorithms such as the Kohonen Self-Organizing Map (SOM), provides new opportunities for automating the classification of seismic facies, thereby improving the accuracy and efficiency of subsurface interpretations [Roden et al., 2015; Meyer et al., 2022].

The frequency content of seismic data contains more information than is immediately apparent. Spectral decomposition, a technique that decomposes seismic signals into their component frequencies, is known for its ability to highlight subtle geological features such as thin beds, fluid content, and stratigraphic variations [Partyka et al., 1999; Shan et al., 2019; Castagna et al., 2003]. Focusing on the frequency domain, this approach provides a more detailed representation of subsurface structures than traditional amplitude-based methods, which often fail to capture the full complexity of the geological signal [Sinha et al., 2005]. The power of spectral decomposition lies in its ability to reveal geological

Citation: Malikov, R., & Babayev, G. (2025). Applying spectral decomposition to seismic facies clustering with unsupervised machine learning. *Geofizychnyi Zhurnal*, *47*(3), 43—53. https://doi.org/10.24028/gj.v47i3.320290.

heterogeneities at multiple scales, making it a valuable tool in seismic facies analysis.

Extracting the full potential of frequency content through manual interpretation alone is challenging, especially in the context of seismic attribute analysis, where the human factor can contribute to missing important geological features [Chopra, Marfurt, 2008; Naseer, 2021]. Therefore, automated facies classification algorithms are employed that incorporate various combinations of frequency components, blending techniques, and both conventional and unconventional attributes [Guo et al., 2009; Kourki,Ali Riahi, 2014; Zhao et al., 2016; Wrona et al., 2018; Ray et al., 2022]. These algorithms enable the generation of highly detailed, multi-dimensional facies distributions that increase interpretation accuracy, improve well planning, and refine reservoir evaluation.

In recent years, unsupervised machine learning techniques, particularly the SOM, have emerged as effective tools for seismic facies classification [Kohonen, 1982]. The SOM uses unsupervised neural networks to represent high-dimensional data in an accessible two-dimensional format, effectively clustering similar data points [Kohonen, 1982; Zhou, Fu, 2005; Roden et al., 2015]. This capability is particularly significant in geosciences, where the intricate relationships embedded in seismic data can obscure important geological insights, highlighting the growing need for advanced analytical techniques in an era of ever-increasing data collection [Kourki, Ali Riahi, 2014]. The application of SOM provides a better understanding of subsurface conditions, allowing geoscientists to more accurately interpret geological structures such as splays, levees, and subtle geological elements [de Matos et al., 2006; Roden et al., 2015; Zhao et al., 2016; Song et al., 2017;Meyer et al., 2022]. SOM's multi-attribute analysis combines various seismic attributes such as amplitude, frequency, and dip into coherent visualizations that help identify subtle geologic features often missed by traditional seismic interpretation methods [Kourki, Ali Riahi, 2014; Roden et al., 2015; Zhu et al., 2022]. Successful implementations of SOM in areas such as the Eagle Ford Shale demonstrate its effectiveness in revealing geological complexities critical to resource exploration and reservoir characterization [Roden et al., 2015].

Despite the advantages of SOM, its application in seismic analysis faces several challenges. The massive amounts of data generated from modern seismic surveys can overwhelm interpreters, complicating the extraction of meaningful interpretations. The presence of outliers and the inherent complexity of geological features can impact the accuracy of SOM classification, raising concerns about the reliability of the results [Meyer et al., 2022]. Addressing these challenges is crucial for refining the utility of SOM in seismic interpretation and enhancing geological understanding derived from complex datasets [Kourki, Ali Riahi, 2014].

This study aims to evaluate the effectiveness of integrating the unsupervised SOM technique with multiple spectral decompositions of seismic data to analyze seismic facies. To assess the proposed methodology's potential, it was applied it to a real seismic dataset from the Caspian Basin. This workflow offers an alternative approach to seismic facies analysis by automating the interpretation process while providing higher resolution and detail. This empowers geoscientists to make more informed decisions, reducing exploration risk and improving the accuracy of subsurface characterization.

Methodology. The workflow consists of interpretating seismic horizons followed by multiple frequency decompositions around the surface. The approach allows using spectral decomposition for a wide frequency range, starting from low frequencies up to the Nyquist limit. The choice of the frequency boundaries depends on the dominant frequency range and the signal-to-noise level of the target seismic data.

The whole process starts with a surgical horizon interpretation. This step requires precise matching to the target horizon, as the algorithm is intended to work at high frequencies as well, which narrows the frequency search window. Multiple spectral decompositions are applied to the seismic data at the

horizon level to extract different frequency ranges and obtain all possible geological information.

Spectral decomposition methods are used to obtain spectral amplitudes in specified frequency ranges. For seismic facies classification, various methods can be used to convert seismic reflection data into sets of spectral characteristics. Common approaches to spectral decomposition include the Short-Time Fourier Transform, wavelet transform, S-transform, and their combinations [Chakraborty, Okaya, 1995; Castagna et al., 2003; Sinha et al., 2005; Kazemeini et al., 2008; von Hartmann et al., 2012; Wu, Castagna, 2017; Huang et al., 2018].

In this study, the Continuous Wavelet Transform (CWT) method was applied. This method offers advantages over the Short-Time Fourier Transform, particularly when high time-frequency resolution is required for spectral information, which is typical in analysing seismic reflected signals [Castagna et al., 2003; Qodri et al., 2019].

A key aspect of the CWT is choosing the appropriate wavelet for signal transformation. With numerous wavelet functions available, the selection should be based on the characteristics of the signal and the features of interest. In CWT, the two main operations performed on the selected mother wavelet: scaling (ξ) and translation (ϑ). The scaling parameter ξ is employed to stretch and compress the wavelet, while the parameter ϑ controls the wavelet's position along the time axis (Fig. 1).

The CWT of a signal $s(t)$ at scale ξ and translation (shift) ϑ is expressed by the integral (Eq. 1):

$$Z(\xi,\vartheta)=\frac{1}{|\xi|^{1/2}}\int_{-\infty}^{\infty} s(t)\psi\left(\frac{t-\vartheta}{\xi}\right)dt\,, \qquad (1)$$

where $Z(\xi,\vartheta)$ is the CWT coefficient at (ξ,ϑ), $\psi(t)$ is the mother wavelet [Addison, 2017]. The mother wavelet, $\psi(t)$, is the base function from which all scaled and shifted versions are derived in the CWT. Its selection is important as it determines the time-frequency characteristics of the transform. Various mother wavelets have been proposed, each tailored to specific applications [Ngui et al., 2013].

The CWT transforms a one-dimensional signal into a two-dimensional matrix, where the two axes represent translation and scale. When the input signal is defined in the time domain, the translation dimension corresponds to time. In the context of signal processing, however, the concept of scale is more precisely interpreted as pseudo-frequency.

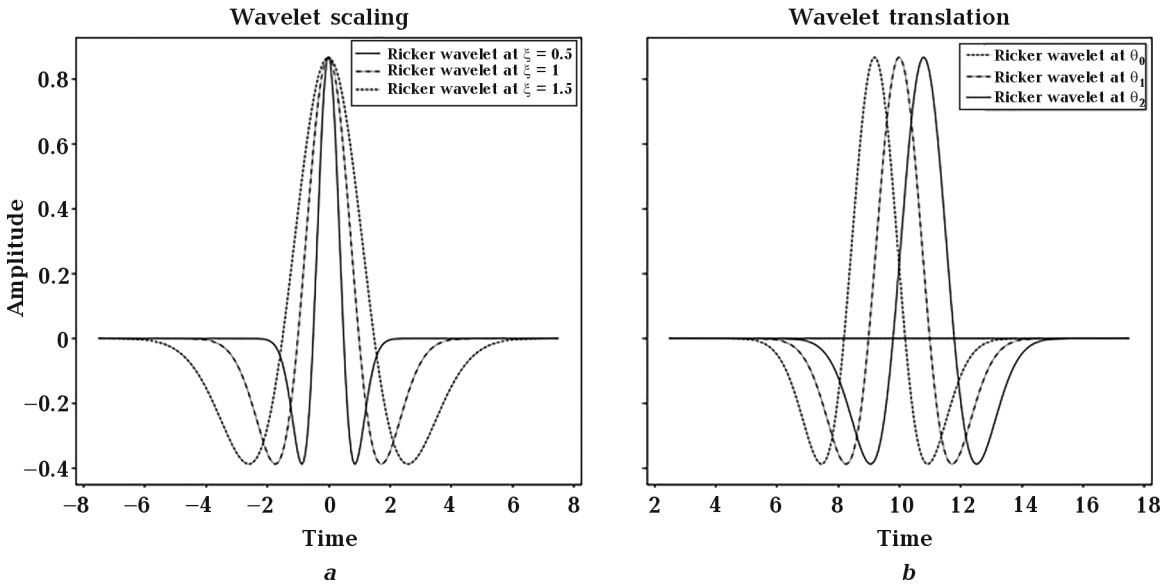

Fig. 1. Two processes applied to the mother Ricker wavelet: scaling (*a*) and translation (*b*).

This pseudo-frequency establishes a relationship between the wavelet scale and a specific frequency band, thus enabling an analysis of the signal's frequency components across different time intervals. The formal relationship between scale and pseudo-frequency is expressed in (Eq. 2):

$$f_{\mathrm{ps}} = \frac{f_{\mathrm{c}}}{\xi}, \qquad (2)$$

where f_{c} is the center frequency and f_{ps} is the pseudo-frequency of the mother wavelet.

In this study, a Ricker wavelet (Mexican hat wavelet) [Addison, 2017] was chosen due to its similarity to wavelets derived from the seismic cube and because the regional seismic data is assumed to be zero-phase. The Ricker wavelet is the second derivative of the Gaussian function and a zero-phase wavelet. It has a central frequency of 0.251 Hz and is commonly used for creating synthetic seismic data. Its time-domain expression is provided in (Eq. 3):

$$\psi(t) = \frac{2}{\sqrt{3}\pi^{1/4}}\left(1-t^2\right)e^{-\frac{t^2}{2}}. \qquad (3)$$

Spectral decomposition maps corresponding to the predefined frequency range were calculated by applying the CWT to the target horizon for each trace using a given time window centered at the picked time sample. These maps are input to the SOM algorithm for the subsequent training phase. In this phase, SOM organizes the high-dimensional spectral data into a two-dimensional grid of neurons while preserving the spatial relationships between the data points. This topological preservation is essential for identifying latent patterns and structures within the spectral data. As a result, the SOM map provides an intuitive and effective visualization that facilitates the exploration and interpretation of underlying geological features and potential anomalies [de Matos et al., 2006; Song et al., 2017].

The training process of a SOM follows a workflow that begins by initializing a grid of neurons; each represented as a vector of random values with the same dimensionality as the input data. These vectors are commonly called the neurons' weights (w_{ij}). At each iteration of the training process, a random data point from the input dataset is selected. The algorithm then identifies the Best Matching Unit (BMU), the neuron whose weight vector is the most similar to the selected input data. The similarity between the input and the neuron's weight is measured using the Euclidean distance, as defined by the following equation:

$$d_j = \sqrt{\sum_{i=1}^{n}\left(x_i - w_{ij}\right)^2}, \qquad (4)$$

where x_i and n represent the components of the input data vector and the number of data samples in the training data, respectively. Once the BMU is identified, its weight vector is updated to become more similar to the input data. Simultaneously, the weights associated with neighboring neurons in the grid undergo a lesser adjustment. This degree of adjustment is controlled by the neighborhood function, which decreases as the distance from the BMU increases. The weight update at each time step t is governed by the following equation:

$$w_i(t+1) = w_i(t) + \\ +\lambda(t)\,h_{\mathrm{nei}}(t, w_i(t))\left[x - w_i(t)\right], \qquad (5)$$

where $\lambda(t)$ is the learning rate, which decreases over time, and $h_{\mathrm{nei}}(t,w_i(t))$ is the neighborhood function. The learning rate is defined by the following formula of exponential decay:

$$\lambda(t) = \lambda(0)\exp\left(-\frac{t}{n}\right), \qquad (6)$$

where t is the iteration number and n is the total number of iterations. The neighborhood function $h_{\mathrm{nei}}(t,w_i(t))$, which dictates how neighboring neurons adjust their weights, is defined as:

$$h_{\mathrm{nei}}(t, w_i(t)) = \exp\left(-\frac{d^2_{\mathrm{BMU}}(w_i(t))}{2\sigma^2(t)}\right). \qquad (7)$$

In Eq. 7, $d^2_{\mathrm{BMU}}(w_i(t))$ represents the squared distance between the BMU and another neuron in the grid, and $\sigma(t)$ is the neighborhood radius at iteration t. The neighborhood radius decreases over time, following the equation:

$$\sigma(t) = \sigma(0)\exp\left(-\frac{\sigma(0)t}{n}\right). \qquad (8)$$

As the training progresses, the neighborhood function decays, reducing the BMU's influence on distant neurons. This gradual reduction ensures that the SOM converges toward a stable state, where the topological relationships of the input data are preserved in the two-dimensional neuron grid. This process allows the SOM to capture and represent the structure of high-dimensional data effectively.

As a result of the learning phase, the SOM returns generalized patterns characterized by their frequency distributions. The input data with similar frequency variations are grouped together, forming clusters reflecting these shared characteristics. These generalized patterns provide a coherent representation of the data, with closely related inputs organized based on their frequency pattern similarities (Figs 2 and 3).

During the application phase, each data pattern vector is presented to the trained SOM, where the corresponding winning neuron is identified based on the minimal distance between the input sample and the generalized patterns learned during training. The position of the winning neuron on the SOM

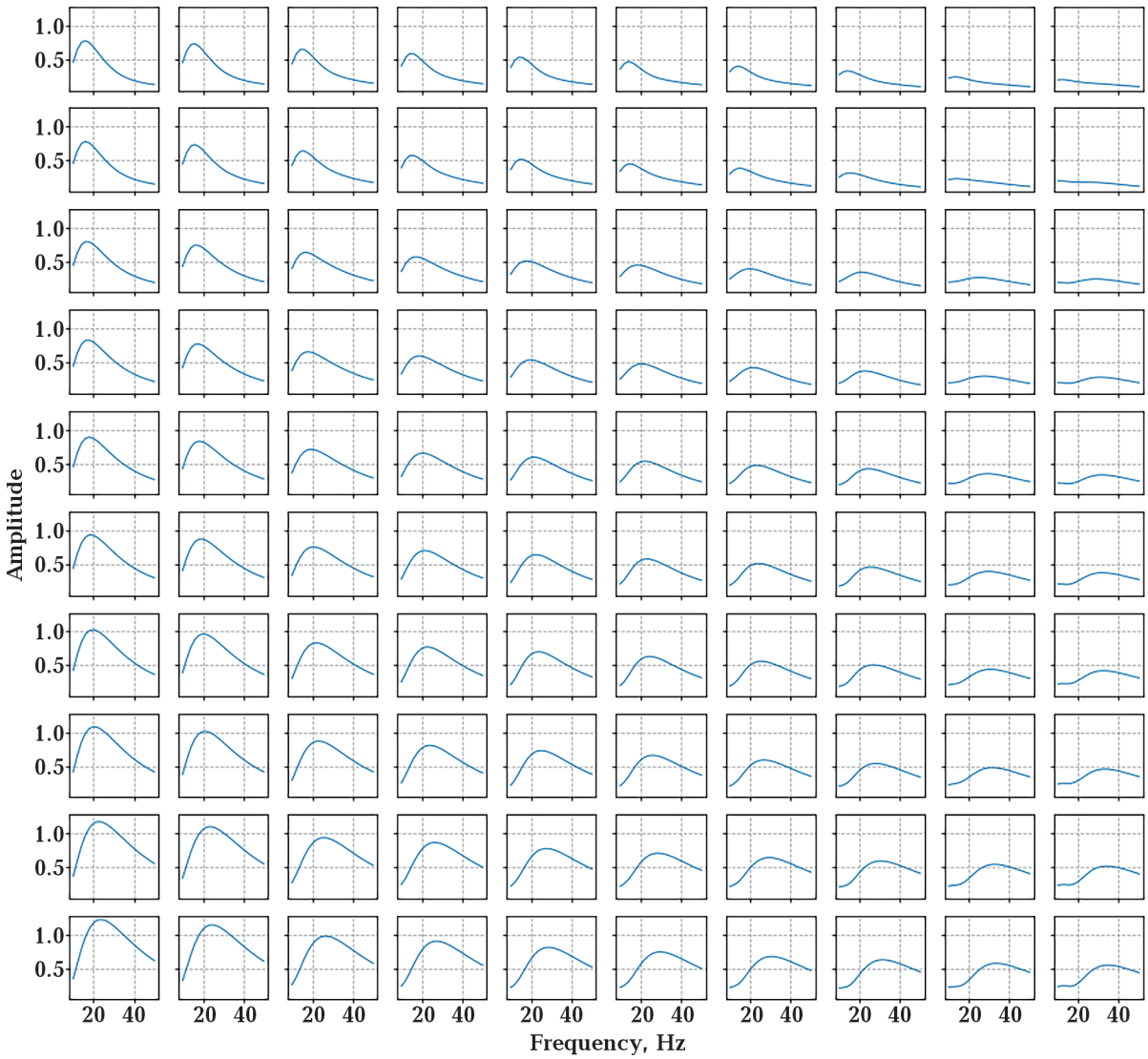

Fig. 2. A map of nodes with 10×10 topology of generalized frequency patterns generated after SOM application. All plots share a common *x*-axis and *y*-axis.

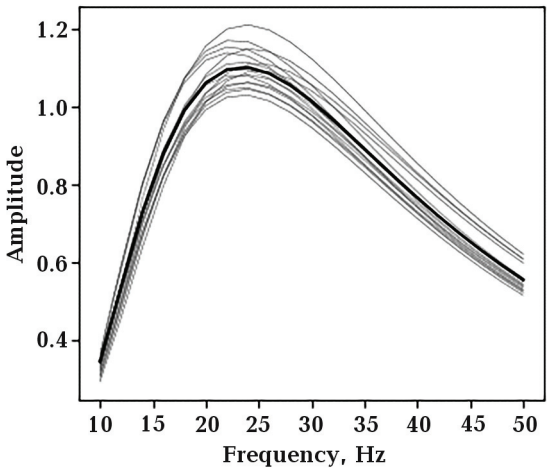

Fig. 3. An example of the corresponding patterns for one generalized frequency pattern. The bold line is the generalized pattern obtained through the SOM algorithm for the group of similar patterns in grey color at the background.

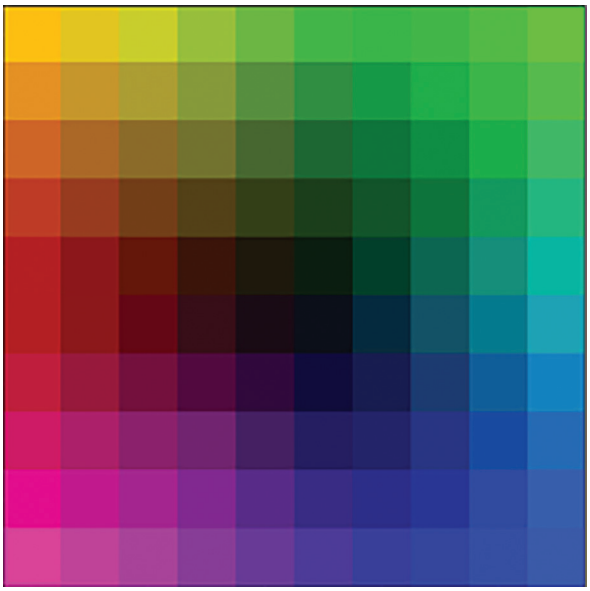

Fig. 4. A 2D colorbar with 10×10 grid used to visualize the obtained SOM grid after the training process.

grid defines the seismic facies type, which is further associated with a specific color code for visualization purposes.

Results and discussion. The above-described workflow was appliedto analyse the horizon picked within the entire post-stack seismic data from the South Caspian Sea. The South Caspian Sea, located east of Azerbaijan, west of Turkmenistan, and north of Iran, serves as the central point of the South Caspian Basin (SCB). The SCB, a Tertiary basin, is notable for its high sedimentation rates, sediment thickness reaching up to 20 km, low compaction, low geothermal gradient, and exceptionally high-pressure gradients [Smith, 2006]. This area's structures were shaped by large amounts of deltaic sediments, sea-surface subsidence, and intense regional tectonic activity. The SCB is a remnant of the back-arc basin from the Tethys paleo-ocean, dating back to the Pliocene [Kaz'min, Verzhbitskii, 2011]. During the Miocene to Recent period, the Paleo-Amu Darya and Paleo-Volga deltas were the main sediment contributors, with lesser input from the Paleo-Kura and Paleo-Sefid Rud systems. Though the SCB is currently a deepwater basin, it experienced predominantly shallow conditions throughout most of the Pliocene and early Pleistocene. A regional clinoform horizon is chronostratigraphically linked to the Absheron Stage of the Lower Pleistocene, characterized by thick shale sequences interbedded with marl and sand. This stage represents a shale-dominated lacustrine environment with minor riverine input. During the Pleistocene, the contribution of Amu Darya was significant among the various paleo-river systems.

As a result of the SOM application, a map of nodes with 10×10 topology is generated (Fig. 2). Each node represents a generalized pattern of amplitude distribution in the frequency domain. An example of a single node, with patterns derived from the original data assigned to that cluster, is shown in Fig. 3. To visualize a SOM grid with a large number of nodes, a 2D colorbaris utilized (Fig. 4). Each cluster is assigned a distinct color that corresponds to the generalized patterns represented on the 2D colorbar. Subsequently, each sample point in the predicted map is depicted using the corresponding clustercolor.

Fig. 5, *a* presentsthe clustered map of multiple frequency components obtained using SOM. The frequency components range from a minimum of 10 Hz to a maximum of 50 Hz, with a sampling rate of 2 Hz, resulting in a total of 21 frequency components used for clustering. This frequency range was selected based on the –10 dB bandwidth of the aver-

age frequency spectrum of the target interval (Fig. 6).The final map consists of 100 clusters, effectively highlighting channels, channel boundaries, and lateral heterogeneity. Fig. 5, *b* displays an RGB blend composed of three frequency components: 10 Hz (red), 30 Hz (green), and 50 Hz (blue), which further illustrates the channel system.

The meandering channels, characterized by a relatively high degree of sinuosity and oriented from North-West to South-East, are clearly depicted within the two primary channel systems on both the spectral decomposition and clustered maps. A comparison between Fig. 5 reveals enhanced visibility of these channel features (green arrows), attributed to the inclusion of higher-frequency components, particularly in the Eastern region, where smaller tributary channels become evident (yellow arrows). An additional strength of the clustering method is its ability to delineate sharp facies boundaries, allowing for precise interpretation of channel margins. Furthermore, the method preserves intra-channel heterogeneity (white arrows), essential for accurate volumetric estimations.

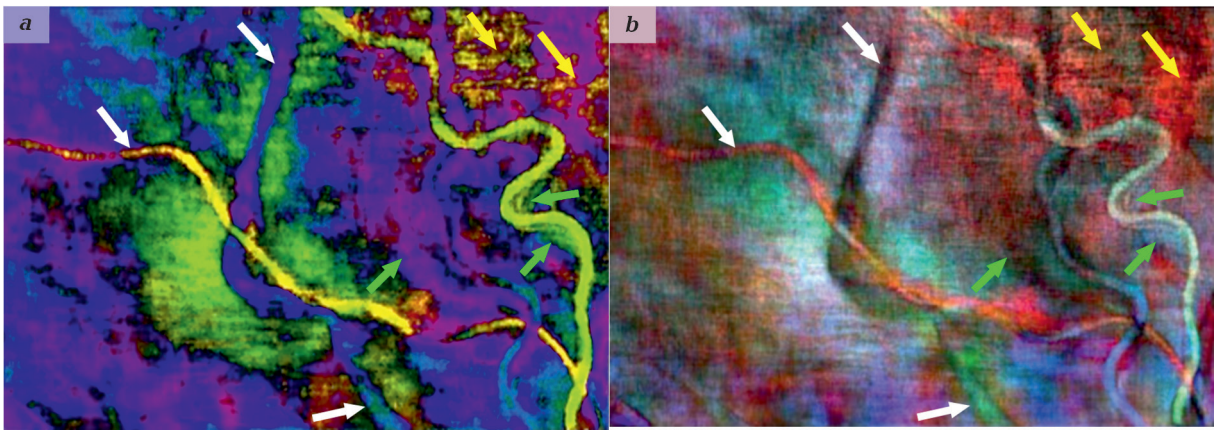

Fig. 5. The clustered map for the target horizonis generated using a trained SOM based on multiple frequency components and visualized with an applied 2D colorbar (*a*). An RGB blend was generated from three seismic frequency decompositions: red corresponds to 10 Hz, green to 30 Hz, and blue to 50 Hz (*b*).Colored arrows highlight key geological features: green arrows indicate enhanced channel visibility, yellow arrows mark the smaller tributary channels, and white arrows emphasize intra-channel heterogeneity.

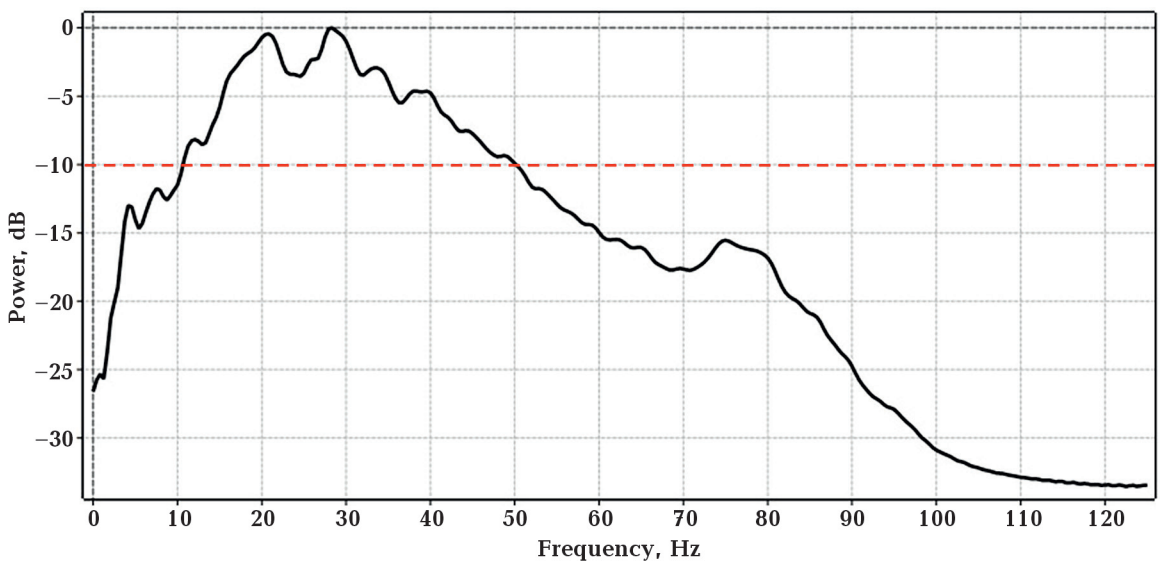

Fig. 6. The average frequency spectrum of the target interval from the seismic data, with the red dashed line indicating the bandwidth at the –10 dB level.

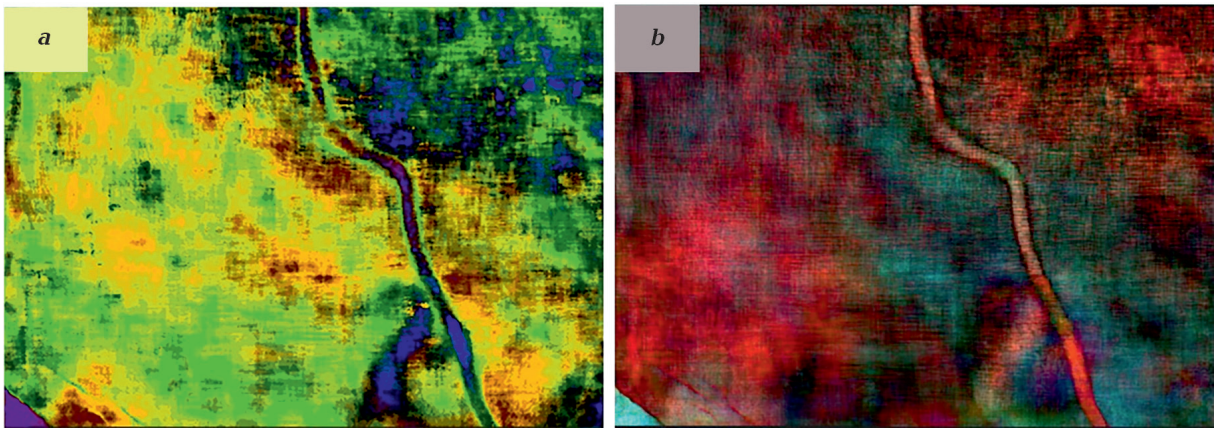

Fig. 7. The clustered map for the horizon 50 ms deeper than the target horizon is generated using a trained SOM based on multiple frequency components and visualized with an applied 2D colorbar (*a*). An RGB blend was generated from three seismic frequency decompositions: red corresponds to 10 Hz, green to 30 Hz, and blue to 50 Hz (*b*).

These detailed insights facilitate a more comprehensive understanding of the channel systems and their associated facies, making the approach valuable for geological analysis.

Fig. 7 provides an additional example of applying the trained SOM model to a different horizon 50 ms deeper than the reference surface. The results demonstrate that the model, trained on the initial horizon, could effectively adapt by allocating the channel to a distinct part of the SOM map, allowing to separate it from the surrounding clusters corresponding to the floodplain. This outcome can be achieved using seismic properties, which have remained relatively consistent over the specified interval. Consequently, these suggest that the methodology has the potential for application to a 3D seismic volume, enabling the extraction of 3D geobodies.

It is important to note that certain geological features are not separated into distinct clusters due to the non-uniqueness of the seismic response. In other words, similar seismic trace behaviors result in similar frequency characteristics, which are, therefore, assigned to the closest clusters on the SOM map. In addition, the algorithm shows a high sensitivity to seismic amplitudes, which is evident from the imprint of seismic acquisition footprints on the obtained maps. All of this requires further study to find a more robust feature extraction algorithm and modification of the SOM algorithm to consider both the amplitudes and patterns of the provided features.

Conclusion. Integrating the SOM technique with spectral decomposition has proven to be a powerful tool for automated seismic facies classification, offering significant improvements in identifying complex subsurface geological structures. By exploiting the frequency content of seismic data through the CWT application, the method effectively enhances the detection of subtle geological features, such as meandering channels and their tributaries. The use of multiple frequency components has improved the visibility and delineation of facies boundaries, allowing a precise interpretation of channel systems and their internal heterogeneity that could be used for reserve estimation. However, the results show a high sensitivity of the SOM algorithm to seismic amplitude variations and the challenges posed by non-unique seismic responses, which can result in ambiguous clustering of certain geological features. These issues point to the need for further refinement of feature extraction methods and potential enhancements to the SOM algorithm, such as incorporating additional seismic attributes or modification of the algorithm for better facies discrimination.

Application of the trained SOM model to a deeper horizon demonstrated the adaptability and consistency of the model in identifying similar geological features over intervals of

relatively stable seismic properties. This suggests the potential for extending the methodology to the entire 3D seismic data for volumetric clustering. Analyzing the entire seismic dataset for facies distributions could provide a more comprehensive understanding of the reservoir, map spatial variations, and reveal heterogeneities. Additionally, this volumetric clustering approach offers an opportunity to create 3D geo-bodies that facilitate improved reservoir characterization and provide deeper insights into reservoir connectivity, ultimately contributing to more accurate and reliable subsurface models.

References

Addison, P.S. (2017). *The Illustrated Wavelet Transform Handbook*. CRC Press, 464 p. https://doi.org/10.1201/9781315372556.

Brown, A.R. (2011). *Interpretation of Three-Dimensional Seismic Data*. Society of Exploration Geophysicists and American Association of Petroleum Geologists, 665 p. https://doi.org/10.1190/1.9781560802884.

Castagna, J.P., Sun, S., & Siegfried, R.W. (2003). Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons. *The Leading Edge, 22*(2), 120—127. https://doi.org/10.1190/1.1559038.

Chakraborty, A., & Okaya, D. (1995). Frequency-time decomposition of seismic data using wavelet-based methods. *Geophysics, 60*(6), 1906—1916. https://doi.org/10.1190/1.1443922.

Chopra, S., & Marfurt, K.J. (2008). Emerging and future trends in seismic attributes. *The Leading Edge, 27*(3), 298—318. https://doi.org/10.1190/1.2896620.

de Matos, M.C., Osorio, P.L., & Johann, P.R. (2006). Unsupervised seismic facies analysis using wavelet transform and self-organizing maps. *Geophysics, 72*(1), P9—P21. https://doi.org/10.1190/1.2392789.

Guo, H., Marfurt, K.J., & Liu, J. (2009). Principal component spectral analysis. *Geophysics, 74*(4), P35—P43. https://doi.org/10.1190/1.3119264.

Huang, Y., Zheng, X., Duan, Y., & Luan, Y. (2018). Robust time-frequency analysis of seismic data using general linear chirplet transform. *Geophysics, 83*(3), V197—V214. https://doi.org/10.1190/geo2017-0145.1.

Kazemeini, S.H., Juhlin, C., Zinck-Jørgensen, K., & Norden, B. (2008). Application of the continuous wavelet transform on seismic data for mapping of channel deposits and gas detection at the CO2SINK site, Ketzin, Germany. *Geophysical Prospecting, 57*(10), 111—123. https://doi.org/10.1111/j.1365-2478.2008.00723.x.

Kaz'min, V.G., & Verzhbitskii, E.V. (2011). Age and origin of the South Caspian Basin. *Oceanology, 51*(1), 131—140. https://doi.org/10.1134/s0001437011010073.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics, 43*(1), 59—69. https://doi.org/10.1007/bf00337288.

Kourki, M., & Ali Riahi, M. (2014). Seismic facies analysis from pre-stack data using self-organizing maps. *Journal of Geophysics and Engineering, 11*(6), 065005. https://doi.org/10.1088/1742-2132/11/6/065005.

Meyer, S.G., Reading, A.M., & Bassom, A.P. (2022). The use of weighted self-organizing maps to interrogate large seismic data sets. *Geophysical Journal International, 231*(3), 2156—2172. https://doi.org/10.1093/gji/ggac322.

Naseer, M.T. (2021). Seismic attributes and quantitative stratigraphic simulation' application for imaging the thin-bedded incised valley stratigraphic traps of Cretaceous sedimentary fairway, Pakistan. *Marine and Petroleum Geology, 134*, 105336. https://doi.org/10.1016/j.marpetgeo.2021.105336.

Ngui, W.K., Leong, M.S., Hee, L.M., & Abdelrhman, A.M. (2013). Wavelet Analysis: Mother Wavelet Selection Methods. *Applied Mechanics and Materials, 393*, 953—958. https://doi.org/10.4028/www.scientific.net/amm.393.953.

Qodri, M.N., Mulyani, M.C., Kaisagara, A.W., Sukmono, S., & Ambarsari, D.S. (2019). Evaluation of Continuous Wavelet Transform (CWT) Attribute in Analysis of Gas Reservoir Distribution on Talang Akar Reservoir in «QDR» Field of Northwest Java Basin. *IOP Conference Series:*

Earth and Environmental Science (Vol. 318, Is. 1, 012043). https://doi.org/10.1088/1755-1315/318/1/012043.

Partyka, G., Gridley, J., & Lopez, J. (1999). Interpretational applications of spectral decomposition in reservoir characterization. *The Leading Edge, 18*(3), 353—360. https://doi.org/10.1190/1.1438295.

Ray, A.K., Khoudaiberdiev, R., Bennett, C., Bhatnagar, P., Boruah, A., Dandapani, R., Maiti, S., & Verma, S. (2022). Attribute-assisted interpretation of deltaic channel system using enhanced 3D seismic data, offshore Nova Scotia. *Journal of Natural Gas Science and Engineering, 99*, 104428. https://doi.org/10.1016/j.jngse.2022.104428.

Roden, R., Smith, T., & Sacrey, D. (2015). Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps. *Interpretation, 3*(4), SAE59—SAE83. https://doi.org/10.1190/int-2015-0037.1.

Shan, X., Tian, F., Cheng, F., Yang, C., & Xin, W. (2019). Spectral Decomposition and a Waveform Cluster to Characterize Strongly Heterogeneous Paleokarst Reservoirs in the Tarim Basin, China. *Water, 11*(2), 256. https://doi.org/10.3390/w11020256.

Sinha, S., Routh, P.S., Anno, P.D., & Castagna, J.P. (2005). Spectral decomposition of seismic data with continuous-wavelet transform. *Geophysics, 70*(6), P19—P25. https://doi.org/10.1190/1.2127113.

Smith, L.S.(2006).*Oligocene-Miocene Maykop/Diatom total petroleum system of the South Caspian Basin province, Azerbaijan, Iran, and Turkmenistan*. U.S. Geological Survey, Bulletin, 2201-I, 27 p.https://doi.org/10.3133/b2201I.

Song, C., Liu, Z., Wang, Y., Li, X., & Hu, G. (2017). Multi-waveform classification for seismic facies analysis. *Computers & Geosciences, 101*, 1—9. https://doi.org/10.1016/j.cageo.2016.12.014.

von Hartmann, H., Buness, H., Krawczyk, C.M., & Schulz, R. (2012). 3-D seismic analysis of a carbonate platform in the Molasse Basin — reef distribution and internal separation with seismic attributes. *Tectonophysics, 572-573*, 16—25. https://doi.org/10.1016/j.tecto.2012.06.033.

Wrona, T., Pan, I., Gawthorpe, R.L., & Fossen, H. (2018). Seismic facies analysis using machine learning. *Geophysics, 83*(5), O83—O95. https://doi.org/10.1190/geo2017-0595.1.

Wu, L., & Castagna, J. (2017). S-transform and Fourier transform frequency spectra of broadband seismic signals. *Geophysics, 82*(5), O71—O81). https://doi.org/10.1190/geo2016-0679.1.

Zhao, T., Zhang, J., Li, F., & Marfurt, K.J. (2016). Characterizing a turbidite system in Canterbury Basin, New Zealand, using seismic attributes and distance-preserving self-organizing maps. *Interpretation, 4*(1), SB79—SB89. https://doi.org/10.1190/int-2015-0094.1.

Zhou, J., & Fu, Y. (2005). Clustering High-Dimensional Data Using Growing SOM. In *Lecture Notes in Computer Science* (pp. 63—68). Springer Berlin Heidelberg. https://doi.org/10.1007/11427445_11.

Zhu, Z., Chen, X., Ren, H., Tao, L., Jiang, J., Wang, T., Cheng, M., Ding, S., & Du, R. (2022). Seismic Facies Analysis Using the Multiattribute SOM-K-Means Clustering. In D. Zhang (Ed.), *Computational Intelligence and Neuroscience* (Vol. 2022, pp. 1—11). HindawiLimited. https://doi.org/10.1155/2022/1688233.

Застосування спектральної декомпозиції до кластеризації сейсмічних фацій за допомогою неконтрольованого машинного навчання

Р. Меліков, Г. Бабаєв, 2025

Інститут геології та геофізики Міністерства науки та освіти Азербайджанської Республіки, Баку, Азербайджан

Аналіз сейсмічних фацій, необхідний для геологічних досліджень надр, традиційно стикається з труднощами у виявленні дрібних варіацій у складних стратиграфічних умовах. У цьому дослідженні для поліпшення ідентифікації деталізованих сейсмічних фацій використовуються спектральна декомпозиція та методи неконтрольованого машинного навчання, зокрема карти Кохонена, що самоорганізується. Спектральна декомпозиція дає змогу проводити аналіз сейсмічних даних на основі частот, виявляючи складні геологічні особливості, які часто залишаються непоміченими при використанні традиційних методів. Для розкладання сейсмічних сигналів застосовували безперервне вейвлет-перетворення, а отримані частотні компоненти класифікували за допомогою карти Кохонена, що самоорганізується, для визначення сейсмічних фацій. Цей підхід було перевірено на сейсмічних даних Південнокаспійського басейну. За результатами успішно виявлено системи каналів та межі фацій, покращено їхню деталізацію, що дало змогу точніше інтерпретувати системи каналів та їх внутрішню варіативність. Цей автоматизований метод надає цінні дані для характеристики колектора та розвідки вуглеводнів, що може знизити ризики розвідки та підвищити точність оцінювання ресурсів.

Ключові слова: сейсмічні фації, спектральна декомпозиція, карта, що самоорганізується, безперервне вейвлет-перетворення, неконтрольоване машинне навчання, машинне навчання.