Robust K-means algorithm with weighted window for seismic facies analysis

Chengyun Song1,*, Lin Li1, Lingxuan Li1 and Kunhong Li2

1 School of Computer Science and Engineering, Chongqing University of Technology, 40054, Chongqing, China
2 School of Resources and Environment, University of Electronic Science and Technology of China, 611731, Chengdu, China

*Corresponding author: Chengyun Song. E-mail: scyer123@163.com

Received 26 April 2021, revised 11 July 2021
Accepted for publication 20 July 2021

Abstract
Seismic facies analysis can generate a map to describe the spatial distribution characteristics of reservoirs, and therefore plays a critical role in seismic interpretation. To analyse the characteristics of the horizon of interest, it is usually necessary to extract seismic waveforms along the target horizon using a selected time window. The inaccuracy of horizon interpretation often produces some inconsistent phases and leads to inaccurate classification. Therefore, the developed adaptive phase K-means algorithm proposed a sliding time window to extract seismic waveforms. However, setting the maximum offset of the sliding window is difficult in a real data application. A value that is too large may cause the cross-layer problem, whereas a value that is too small reduces the flexibility of the algorithm. To address this disadvantage, this paper proposes a robust K-means (R-K-means) algorithm with a Gaussian-weighted sliding window for seismic waveform classification. The used weights punish those windows distant from the interpretation horizon in the objective function, consequently producing a smaller range of horizon adjustments even when using relatively large maximum offsets and benefitting the generation of stable and reliable seismic facies maps. The application of real seismic data from the F3 block proves the effectiveness of the proposed algorithm.

Keywords: seismic facies analysis, horizon interpretation, K-means, noise

1. Introduction
Seismic facies analysis aims to describe and interpret extracted seismic reflection parameters, including geometry, continuity, amplitude, frequency, velocity, and coherence de Matos et al. (2007). Due to 3D seismic data covering large areas, the manual interpretation of seismic facies consumes a considerable amount of time and effort, even with the aid of user-friendly software, and can generate subjective results which rely on the experience of the interpreters Saggaf et al. (2003). Consequently, automatic seismic facies analysis technology has been under constant development in recent decades (Coléou et al. 2003). This technique utilises pattern recognition methods to group similar seismic reflections into one class, which represents a sedimentary feature and can also be defined as seismic facies (Bloch et al. 2003).

In general, the machine-learning methods used in seismic facies analysis can be divided into supervised and unsupervised approaches (Roy et al. 2013). Supervised techniques require labeled data to train a classification model for identifying unlabeled seismic data (Wrona et al. 2018). However, well data or seismic interpretation are limited in new exploration areas. The problem of incomplete and unbalanced labels has emerged, which brings challenges for supervised learning. Instead, unsupervised seismic facies analysis uses a clustering algorithm to generate natural clusters presented in...
seismic data space (Marroquín et al. 2009). The data-driven method can perform well in predicting properties without a professional interpreter. The commonly used unsupervised methods include the K-means algorithm (Barnes & Laughlin 2002), self-organising mapping (SOM) (Strecker 2012), generated topographic maps (GTMs) (Roy et al. 2014), neural networks (West et al. 2002; Coléou et al. 2003; Banchs & Rondon 2007; Ross & Cole 2017), deep learning (Qian et al. 2018; Mousavi et al. 2019), and so on.

With respect to the input of clustering methods, there are two main research branches in unsupervised seismic facies analysis: seismic waveform clustering (Andersen & Boyd 2004; Arshin et al. 2014) and multi-attribute clustering (Linari et al. 2003; Roden et al. 2015). The extracted seismic attributes from the seismic data are usually robust to noise and can reveal hidden reservoir information. However, selecting appropriate representations from numerous and redundant attributes is the main challenge in multi-attribute analysis (Zhao et al. 2017). Meanwhile, the optimised attributes using some multi-feature fusion techniques (such as principal component analysis and deep learning) may lose the physical meaning of the original seismic attributes and are difficult to interpret. Hence, waveform classification is also quite often used in seismic facies analysis because of its simplicity and effectiveness. Moreover, seismic reflection configurations, such as strong and weak amplitude reflections and continuous and discontinuous reflections, are easily identified from the model trace of the seismic facies map. Nevertheless, there are some problems, such as time-window selection and signal-to-noise ratio, needing to be resolved in order to improve the analysis efficiency. Xie et al. (2017) extracted Mel-frequency cepstral coefficients (MFCCs) based on the bottom and top horizons, thus avoiding the use of a fixed-size time window. Song et al. (2019) proposed a new similarity measure for calculating the distance of unequal length waveforms directly from the target interval. Saraswat & Sen (2012) combined an artificial immune system with the self-organising map to form a new approach called AI-SOM for unsupervised seismic facies analysis, which is robust to noise in seismic data. In addition, stratigraphic continuity is considered in Song et al. (2017a) and Song et al. (2017b) to overcome the noise influence.

Another important issue in the waveform classification is horizon interpretation noise (de Matos et al. 2007). This is shown in figure 1, where the green lines represent the interpreted horizons. Because of the uncertainty of interpretation, a smooth horizon is expected in the middle positions existing with obvious changes of seismic characteristics. Thus the interpreted horizon does not follow the peaks, turning points, or other points with the consistent phase, and thus it is conducive to seismic facies analysis. Gao (2008, 2011) introduced texture model regression (TMR) as a way to perform seismic facies analysis, which uses a variable phase model to distinguish seismic phase characteristics. However, model selection and training are difficult when confronted with sparse and inaccurately labeled data. For this reason, Song et al. (2016) developed the adaptive phase K-means algorithm (Ap-K-means) for unsupervised waveform classification to alleviate the horizon interpretation noise. The Ap-K-means algorithm proposes a sliding window to extract waveforms along the target horizon. The waveform that best matches the model traces (that is, centroids of clustering) is used to represent the feature of the reflection point. With this approach, the interpretation horizon is slightly adjusted with the iterative optimisation process of the algorithm. For example, the green line in figure 1 is the original horizon, for which there is some interpretation noise. Ap-K-means can modify the given horizon to that shown by the red line in figure 1a, thus producing satisfactory seismic facies maps.

However, the difficulty of horizon interpretation varies with geological structure, which will cause inconsistent horizon noise at different reflection times with larger coverage of seismic data. For instance, the noise in the zone of the black ellipse in figure 1b is larger than that in the other locations, and this mainly contributes to the strong discontinuity. As a result, when using a large maximum offset for the sliding window to tolerate different horizon noise, the cross-layer problem may arise and thus generate unreliable seismic facies maps. To alleviate this problem, instead of using the same weights for different sliding windows, we adapt a Gaussian weight to punish more distant sliding windows to ensure that the adjustment is close to the original horizon. At the same time, the algorithm also has a certain ability to adjust the horizon in a distant time range if the waveform near the horizon cannot match model traces very well. Based on this idea, we propose a weighted adaptive phase distance to calculate the resemblance between the waveform extracted around
the horizon of interest and the model trace and then construct the objective function of the proposed method. Due to its robustness in the face of horizon interpretation noise, we name the proposed algorithm as robust K-means (R-K-means) for short. Finally, an alternative approach is proposed to solve the optimisation problems of R-K-means.

The rest of this paper proceeds as follows. In Section 2, we first give descriptions of the traditional K-means and the improved Ap-K-means algorithms. Then, the details of our proposed R-K-means algorithm and its corresponding method of optimization are described. After that, we present an experiment on the real data from the F3 block in Section 3. Finally, we conclude this work.

2. Method

2.1. K-means clustering

K-means (Macqueen 1967) is a classic and the most commonly used clustering algorithm aiming to classify data into $K$ clusters. When the input data are large, K-means can maintain certain stability and computational efficiency and thus is widely used in seismic facies analysis. Given $N$ samples, $\{x_1, \ldots, x_N\}$, the traditional K-means can be expressed as follows (Guo et al. 2020):

$$E = \min \sum_{i=1}^{K} \sum_{j=1}^{N} z_{ij} \| x_j - \mu_i \|^2,$$  \hspace{1cm} (1)

where $E$ is the sum of squared errors of all samples, $K$ denotes the number of clusters, and $z_{ij}$ is the binary indicator variable $z_{ij} \in \{0, 1\}$; if the input sample $x_j$ belongs to cluster $i$, then $z_{ij} = 1$, otherwise, $z_{ij} = 0$; $\mu_i$ denotes the centroid of the $i$th cluster. The main processing procedures are as follows:

(i) Set the cluster number $K$ and select $K$ samples from all data to represent the centers $\mu_i$.
(ii) Calculate the distance between samples and each center $\mu_i$, and then divide each sample into its nearest center.
(iii) Update $\mu_i$, as the mean of samples belonging to the $i$th cluster.
(iv) Repeat steps (ii)–(iii) until convergence.

The choice for $K$ (the number of seismic facies) is crucial in unsupervised seismic facies analysis. A small value will offer a very rough classification, and some facies of interest can not be identified. By contrast, a large $K$ will produce lots of redundant facies, bringing difficulties to interpretation (Du et al. 2015). In practice, we use more seismic facies than the number of expectant geofacies because the additional facies may represent data noise (Roy et al. 2012). However, the data complexity of research zone may be unknown. It is difficult to evaluate the number of expectant geofacies and noise. Therefore, the ideal number of classification is usually obtained using iterative testing and prior knowledge. In addition, the positions of the initial $K$ points are selected arbitrarily, without a standard rule. To obtain an optimal position, one may need to repeat the experiments with different selections of the initial $K$ points, and compare the final results. Wang (2012) also suggested using the hierarchical clustering to overcome the problems with the K-means method.

2.2. Adaptive phase K-means algorithm

To perform unsupervised seismic analysis, an important step is extracting waveforms from seismic data along the horizon of interest. However, interpretation uncertainty causes some noise to the interpreted horizon, which furthermore brings about the phase changes of an extracted waveform. The current methods with fixed windows cannot deal with this problem particularly well. Hence, Song et al. (2016) improved the traditional K-means algorithm and introduced the adaptive phase distance $Apd(A, B)$ to calculate the similarity between model $A$ (one center of K-means) and waveform $B$ (the input samples), defined as follows:

$$Apd(A, B) = \min_p \| A - Win(B, p) \|^2,$$  \hspace{1cm} (2)

where $Win(B, p)$ represents the waveform extracted from $B$ with a sliding window, as shown in figure 2, $p$ controls the offset of the window from the interpretation horizon point and can be regarded as the phase adjustment of the extraced waveform, and $Apd(A, B)$ is the Euclidean distance between model $A$ and waveform $Win(B, p^*)$ with the best-matched phase. Based on this distance, the objective function is defined as

$$E_{Ap} = \min \sum_{i=1}^{K} \sum_{j=1}^{N} z_{ij} Apd(\mu_i, x_j).$$  \hspace{1cm} (3)
It is noted that the lengths of model $\mu_j$ and sample $x_i$ are different, which can be seen in figure 2. Similar to K-means, an iterative alternative optimisation method is used to optimise equation (3). Each input vector $x_i$ is assigned to its closest cluster based on equation (2), and the phase $p^*$ can be obtained. Then $\mu_i$ is updated as the means of the waveforms $Win(x_j, p^*)$ in the $i$th cluster.

2.3. Proposed robust K-means algorithm

2.3.1. Objective function. In the adaptive phase K-means algorithm, it is necessary to set the maximum phase offset $p_{\text{max}}$ to control the horizon adjustment. Because the difficulty of horizon interpretation varies with geological structure, a large $p_{\text{max}}$ is selected to ensure the ability of horizon adjustment and thus tolerate different horizon noise generally, but doing so may lead to serious and incorrect adjustment.

For this reason, the proposed R-K-means algorithm adapts a Gaussian weight to punish distant sliding windows and thus ensure that the adjustment is close to the original horizon. The objective function can be defined as follows:

$$E_{\text{robust}} = \min \sum_{j=1}^{K} \sum_{i=1}^{N} z_{ij} Rd(\mu_j, x_i),$$  \hspace{1cm} (4)

where

$$Rd(\mu_j, x_i) = \min_p \left\{ W(p) \cdot \| \mu_i - Win(x_j, p) \|^2 \right\},$$  \hspace{1cm} (5)

$$W(p) = 1 - \frac{1}{\sigma \sqrt{2\pi}} \exp \left( - \frac{p^2}{2\sigma^2} \right).$$  \hspace{1cm} (6)

Parameter $\sigma$ is the standard deviation of the Gaussian function and controls the magnitude of the weights. A very large $\sigma$ implies the assignment of nearly equal weights to different sliding windows, thus generating similar cluster results to Ap-K-means. However, a very small $\sigma$ will degrade the ability of the horizon adjustment, and the performance of R-K-means is very close to that of the traditional K-means algorithm. Consequently, the proposed R-K-means is an extension of the Ap-K-means and traditional K-means algorithms. In addition, since horizon noise is within a certain range, and a smaller offset is expected, we also use $p_{\text{max}}$ to prevent excessive offset of the sliding window and improve computing efficiency.

2.3.2. Optimisation. The objective function of R-K-means contains three variables that need to be solved, and it is difficult to find the optimal solution of those variables directly at the same time. Therefore, an alternative approach is proposed to address the optimisation problems of R-K-means. Specifically, three steps are iteratively performed until $E_{\text{robust}}$ is less than the set threshold, or until the maximum number of iterations is reached: (1) Update $z$ with the fixed $p$ and $\mu$; (2) Update $\mu$ with the fixed $p$ and $z$; and (3) Update $p$ with the fixed $z$ and $\mu$.

- Update $z$ with the fixed $p$ and $\mu$. When $p$ and $\mu$ are fixed in equation (4), the optimal solution for $z$ can be obtained in closed form:

$$z_{kj} = \begin{cases} 1 & \text{if } k = \text{argmin}_i (\mu_j, x_i) \\ 0 & \text{else} \end{cases}$$  \hspace{1cm} (7)

The process is identical to the nearest-neighbor classification and is used in the K-means algorithm.

- Update $\mu$ with the fixed $p$ and $z$. When $p$ and $z$ are fixed, equation (4) is simplified to

$$E_{\text{robust}} = \min_\mu J, \hspace{1cm} (8)$$

where

$$J = \sum_{i=1}^{K} \sum_{j=1}^{N} z_{ij} W(p_j) \cdot \| \mu_i - Win(x_j, p_j) \|^2. \hspace{1cm} (9)$$

We get the derivative of equation (9) as follows:

$$\frac{\partial J}{\partial \mu_i} = 2 \sum_{j=1}^{N} z_{ij} W(p_j) \{ \mu_i - Win(x_j, p_j) \}. \hspace{1cm} (10)$$

Setting equation (10) to zero, we can get

$$\mu_i = \frac{\sum_{j=1}^{N} z_{ij} W(p_j) Win(x_j, p_j)}{\sum_{j=1}^{N} z_{ij} W(p_j)}, \hspace{1cm} (11)$$

which is consistent with the center $\mu_i$ update in the K-means algorithm with a weighted approach.

- Update $p$ with the fixed $z$ and $\mu$. When $z$ and $\mu$ are fixed, equation (4) has the same solution as equation (5):

$$p_j = \text{arg} \min_p \left\{ W(p) \cdot \| \mu_i - Win(x_j, p) \|^2 \right\}. \hspace{1cm} (12)$$

Since $p$ represents the window offset, and its precision is related to the sampling ratio, the solution space is relatively small under the limitation of $p_{\text{max}}$. Therefore, we can directly solve $p$ using an exhaustive search method.

3. Application

To prove the effectiveness of our method, we apply the proposed method to free and actual data from the F3 block in the Netherlands. Figure 3 shows a seismic section, and we select the MFS4 horizon, marked as the blue line, to analyse seismic facies. Figure 4 displays the seismic amplitude slice along the MFS4 horizon. Several seismic facies, such as faults, fractures and channels, need to be delineated along the horizon of interest. Since the sampling interval is 4 ms, fine interpolation is necessary to enhance the accuracy of horizon adjustment before extracting the waveform. According to the main frequency of the seismic data along the horizon, we use a time
Figure 3. A vertical section from the F3 data. The blue line represents the horizon of interest, called MFS4.

Figure 4. A horizontal seismic section along MFS4.

Figure 5. The initial centroids used in all compared methods.

gate of $[-8 \text{ ms}, +32 \text{ ms}]$ to extract seismic waveforms. The parameters $p_{\text{max}} = 10 \text{ ms}$, $\sigma = 15$, and $K = 6$ are selected to perform seismic facies analysis using the traditional K-means, Ap-K-means and proposed R-K-means algorithms. Because of the sensitivity of those methods to the initial centroids, we adopt the same initial seismic waveforms shown in figure 5. The average objective function (the objective function value divided by the number of samples) in the iteration procedure is provided in figure 6, which indicates the good convergence of R-K-means.

Figure 6. The mean objective function in the iteration.

Figure 7 shows the final facies map based on the proposed R-K-means algorithm. The main seismic facies, such as channels and faults, can be easily identified. By contrast, the facies maps generated by the Ap-K-means and K-means algorithms (figures 8a and c) present some discontinuous and noisy classifications, especially in the zone marked by white ellipses. Meanwhile, two channels are not easy to distinguish in figure 8. Moreover, the first peaks of the model traces in figure 7b lie at approximately 12 ms, which is similar to figure 8d. Therefore, we can say that the horizon adjustment is slightly attributed to the Gaussian weight of the proposed algorithm. Serious adjustment in Ap-K-means affects the generation of clustering centers. In addition, we show the section along line AB and the seismic facies analysis results of different methods (figure 9). The adjusted horizon (green line) is closer to the original horizon (blue line) than that in
Figure 8. The facies maps and model traces generated by the Ap-K-means and traditional K-means algorithms.

Ap-K-means (red line), thus demonstrating the effectiveness of the weighted window. Meanwhile, the classification result from R-K-means is more consistent with the reflection characteristics of the seismic section than the other methods. The two channels located at roughly lines 665 and 700 are visible using the R-K-means algorithm.

We further calculate the statistics of the final adjusted phase in both the Ap-K-means and R-K-means algorithms shown in figure 10. Despite using the maximum offset of 10 ms, the most adjusted phase lies between −5 and 5 ms. This also confirms that small adjustments can be generated based on R-K-means compared with the Ap-K-means method. However, the proper $p_{\text{max}}$ may be 5 ms, as seen in figure 10. $p_{\text{max}}$ with 10 ms may be large and thus work poorly for Ap-K-means. Hence, we generate the seismic facies using the Ap-K-means algorithm with $p_{\text{max}}$ of 5 ms (figure 11). The main facies is very similar to figure 7, but the channel has less continuity in the zone of the white ellipse. Figure 12 shows the statistics of the final adjusted phase for both Ap-K-means with $p_{\text{max}} = 5$ ms and for...
Figure 10. Statistics of the final adjusted phase.

Figure 11. The facies maps generated by Ap-K-means with the best parameter (\(p_{\text{max}} = 5\) ms).

Figure 12. Statistics of the final adjusted phase in the Ap-K-means with the best parameter \(p_{\text{max}}\) and R-K-means algorithm.

Figure 13. The facies map and model traces generated by the proposed R-K-means algorithm using the original noisy data from the F3 block.

4. Discussion

We test the proposed algorithm using the filtered data from Opendtect for generating clear seismic facies maps in the previous application. However, the original data from the F3 block is noisy (Saraswat & Sen 2012). To confirm the adaptability of the R-K-means algorithm to seismic noise, we perform seismic facies analysis based on the original noisy data. The same parameters, including the \(K\) value, \(p_{\text{max}}\), \(\delta\) and locations of initial waveforms, are introduced to the experiment. Figure 13 shows the resulting facies map. The main facies remains consistent with figure 7 based on filtered data. Only a few noises are presented in the boundaries of the seismic facies. Therefore, we can say that our proposed method can be directly applied to the seismic data with poor quality.

Although the R-K-means can obtain satisfactory seismic facies maps in the real application, it takes a longer running time than the K-means algorithm. Suppose \(n\) represents the number of seismic traces; the K-means algorithm computes the distances between each trace and cluster centroids per iteration, and requires a time of \(O(Kn)\). Ap-K-means and R-K-means algorithms need to calculate the distances with a series of sliding windows in order to find the matching phase. Let \(w\) denote the number of sliding windows, which is determined by \(p_{\text{max}}\) and the sampling interval of seismic data. The time complexity of both Ap-K-means and R-K-means algorithms is \(O(Knw)\) per iteration. We use a personal computer with a CPU of 3.6 GHz Intel Core i9 and 32 GB of RAM to perform seismic facies analysis in this paper. The running times of K-means, Ap-K-means and R-K-means algorithms with 20 iterations are 138 s, 2506 s and 2555 s, respectively. An accelerated version can be implemented with parallel computing. Also, sampling can reduce the number of training data and thus helps to save time. In brief, the R-K-means algorithm is a worthy method in unsupervised seismic facies analysis.
5. Conclusion

In this paper, we propose a robust K-means algorithm for seismic facies analysis, which improves the adaptive phase K-means algorithm using Gaussian weights to reduce the sensitivity of the maximum offset. As a result, a smaller range of horizon adjustment is preferred in the proposed algorithm and thus enhances the stability of seismic facies analysis. The satisfactory result in the application of real seismic data proves the horizon adjustability of the proposed algorithm. The final adjusted horizon is closer to the original horizon in this method than in the comparison method, confirming the validity of the weight function used. Additionally, the implementation of the proposed R-K-means is simple and efficient and thus can be used for seismic data with horizons of poor interpretative quality.

Acknowledgements

This work is supported by the Natural Science Foundation of China (grant no. 41804112), the Youth Project of Science and Technology Research Program of Chongqing Education Commission of China (grant no. KJQN202001143) and the Scientific Research Foundation of Chongqing University of Technology. We thank DBA Earth Science for providing free and public data sets from block F3.

Conflict of interest statement. None declared.

References

Andersen, E. & Boyd, J., 2004. Seismic waveform classification: techniques and benefits. CSEG Recorder, 29.

Arshin, B.M., Ghazali, A.R., Amin, Y.K. & Barnes, A.E., 2014. Hybrid waveform classification applied to delineate compartments in a complex reservoir in the Malay Basin, in International Petroleum Technology Conference, Kuala Lumpur, 2014, Society of Petroleum Engineers, Richardson, TX, USA, p. 3030.

Banchs, R. & Rondon, O., 2007. Seismic pattern recognition based upon hopfield neural networks, Exploration Geophysics, 38, 220–224.

Barnes, A.E. & Laughlin, K.J., 2002. Investigation of methods for unsupervised classification of seismic data, SEG Technical Program Expanded Abstracts, 2221–2222.

Bloch, G., El Deeb, M., Badaam, H., Cailly, F., Lecante, G. & Meunier, A., 2003. Seismic facies analysis for fracture detection: a powerful technique, in Middle East Oil Show, Bahrain 2003, Society of Petroleum Engineers, Richardson, TX, USA, doi: 10.2118/81526-MS.

Coléou, T., Poupon, M. & Azbel, K., 2003. Unsupervised seismic facies classification: A review and comparison of techniques and implementation, The Leading Edge, 22, 942–953.

de Matos, M.C., Osorio, P.L. & Johann, P.R., 2007. Unsupervised seismic facies analysis using wavelet transform and self-organizing maps, Geophysics, 72, P9–P21.

Du, H.-K., Cao, J., Xue, Y.-J. & Wang, X., 2015. Seismic facies analysis based on self-organizing map and empirical mode decomposition, Journal of Applied Geophysics, 112, 52–61.

Gao, D., 2008. Application of seismic texture model regression to seismic facies characterization and interpretation, The Leading Edge, 27, 394–397.

Gao, D., 2011. Latest developments in seismic texture analysis for subsurface structure, facies, and reservoir characterization: A review, Geophysics, 76, W1–W13.

Guo, X., Liu, X., Zhu, E., Zhu, X., Li, M., Xu, X. & Yin, J., 2020. Adaptive self-paced deep clustering with data augmentation, IEEE Transactions on Knowledge and Data Engineering, 32, 1680–1693.

Linari, V., Santiago, M., Pastore, C., Azbel, K. & Poupon, M., 2003. Seismic facies analysis based on 3D multiattribute volume classification, La Palma field, Maracaibo, Venezuela, The Leading Edge, 22, 32–36.

Macqueen, J.B., 1967. Some methods for classification and analysis of multivariate observations, in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 1965, University of California Press, Berkeley, CA, p. 1.

Marroquin, I., Brault, J.J. & Hart, B.S., 2009. A visual data-mining methodology for seismic facies analysis: Part 1– testing and comparison with other unsupervised clustering methods, Geophysics, 74, P1–P11.

Mousavi, S.M., Zhu, W., Ellsworth, W. & Beroza, G., 2019. Unsupervised clustering of seismic signals using deep convolutional autoencoders, IEEE Geoscience and Remote Sensing Letters, 16, 1693–1697.

Qian, F., Yin, M., Liu, X.Y., Wang, Y.J., Lu, C. & Hu, G.M., 2018. Unsupervised seismic facies analysis via deep convolutional autoencoders, Geophysics, 83, A39–A43.

Roden, R., Smith, T. & Sacrey, D., 2015. Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps, Interpretation, 3, SAES59–SAE83.

Ross, C.P. & Cole, D.M., 2017. A comparison of popular neural network facies-classification schemes, The Leading Edge, 36, 340–349.

Roy, A., Dowell, B.L. & Marfurt, K., 2012. Characterizing a Mississippian Tripolitic Chert reservoir using 3D unsupervised seismic facies analysis and well logs: an example from Osage County, Oklahoma, SEG Technical Program Expanded Abstracts, 1–5.

Roy, A., Jayaram, V. & Marfurt, K.J., 2013. Active learning algorithms in seismic facies classification, in SEG Technical Program Expanded, 1467–1471.

Roy, A., Romero-Pelaez, A., Kwiatkowski, T.J. & Marfurt, K.J., 2014. Generative topographic mapping for seismic facies estimation of a carbonate wash, Veracruz Basin, southern Mexico, Interpretation, 2(1), SA31–SA47.

Saggaf, M.M., Toksoz, M.N. & Marhoon, M.L., 2003. Seismic facies classification and identification by competitive neural networks, Geophysics, 68, 1984–1999.

Saraswat, P. & Sen, M.K., 2012. Artificial immune based self organizing maps for seismic facies analysis, Geophysics, 77, 45–53.

Song, C., Liu, Z., Wang, Y., Xu, F., Li, X. & Hu, G., 2018. Adaptive phase k-means algorithm for waveform classification, Exploration Geophysics, 49, 213–219.

Song, C., Liu, Z., Cai, H., Wang, Y., Li, X. & Hu, G., 2017a. Unsupervised seismic facies analysis with spatial constraints using regularized fuzzy c-means, Journal of Geophysics and Engineering, 14, 1535–1543.

Song, C., Liu, Z., Wang, Y., Li, X. & Hu, G., 2017b. Multi-waveform classification for seismic facies analysis, Computers & Geosciences, 101, 1–9.

Song, C., Liu, Z., She, B., Li, K. & Hu, G., 2019. Seismic waveform classification via a new similarity measure, SEG Technical Program Expanded Abstracts, 2438–2442.

Strecker, U., 2012. Data mining of 3D poststack seismic attribute volumes using Kohonen self-organizing maps, The Leading Edge, 21(10), 1032.

Wang, Y., 2012. Reservoir characterization based on seismic spectral variations, Geophysics, 77, M89–M95.
West, B.P., May, S.R., Eastwood, J.E. & Rossen, C., 2002. Interactive seismic facies classification using textural attributes and neural networks, *The Leading Edge*, 21, 1042–1049.

Wrona, T., Pan, I., Gawthorpe, R.L. & Fossen, H., 2018. Seismic facies analysis using machine learning, *Geophysics*, 83, 1–34.

Xie, T., Zheng, X. & Zhang, Y., 2017. Seismic facies analysis based on speech recognition feature parameters, *Geophysics*, 82(3), O23–O35.

Zhao, T., Li, F. & Marfurt, K.J., 2017. Seismic attribute selection for unsupervised seismic facies analysis using user guided data-adaptive weights, *Geophysics*, 83, O31–O44.