Click to see slides of presentation

EA Applying Machine Learning Methods to Study Compartmentalization in Complex Reservoirs Based on Static Pressure Information*

Jose Victor Contreras[1]

Search and Discovery Article #42473 (2019)**
Posted December 11, 2019

*Adapted from extended abstract based on oral presentation given at AAPG 2019 International Conference and Exhibition, Buenos Aires, Argentina, August 27-30, 2019
 DOI:10.1306/42473Contreras2019

[1]Baker Hughes, a GE company (jose.contreras2@bhge.com)

Abstract

Static pressure history data provide very important information from reservoirs, not only associated with the energy behavior over time during its production life, but also the degree of compartmentalization of the reservoir from heterogeneities or structural boundaries. However, engineers must handle huge numbers of static pressure measurements that have been collected in hundreds of wells during large time spans, resulting in thousands of pressure points to analyze for extracting reservoir compartmentalization information. Analyzing these amounts of information manually is impractical and inefficient. In these cases, it is more efficient to develop a pattern recognition workflow, based on unsupervised learning, to identify and classify wells with pressure measurements that share similar characteristics and associate these wells to a specific compartment.

A workflow for pattern recognition based on unsupervised learning was developed to classify wells with similar pressure and spatial coordinates and assign those results to the corresponding reservoir compartment. An unsupervised learning algorithm based on the Self-Organizing Map (SOM) technique was trained to identify pressure patterns and classify wells. This workflow automatically assigns wells to each compartment according to the similarities in characteristics previously described. The first step in this workflow was to process the pressure information for each well for use in the SOM. Then, an unsupervised learning algorithm was applied to the resulting data to classify wells with similar pressures and spatial coordinates. Finally, each classified well was assigned to the corresponding compartment.

Experience shows that unsupervised learning is a powerful tool to identify different compartments in complex reservoirs based on static pressure information, because of the powerful capacity to identify hidden pressure patterns (compartments). This workflow considerably

reduces the time to obtain compartmentalization studies. Moreover, this workflow reduces the error introduced by human intervention during classification processes and unifies criteria when the job is done by different engineers.

Introduction

Static pressure information is crucial to estimate the reservoir energy level and comprehend the degree of depletion in reservoirs due to fluid production. Additionally, when static pressure data for several wells is analyzed as a function of time, it is possible to obtain pressure trends with dissimilar behaviors which can be associated to different compartments resulting from heterogeneities or structural boundaries (layers, lobes or channels). Once these pressure trends are identified, it is relatively easy to classify wells that share similar depletion characteristics.

Pressure trend analysis is performed by human experts with simple tools, whose use their knowledge and expertise to categorize wells with similar characteristics. However, experts must deal with thousands of static pressure measurements collected in hundreds of wells over time, which result in a time-consuming process to digest this vast amount of information and very difficult task to capture pressure trend similarities in a single plot.

Machine learning methods can be used to overcome this kind of limitation, specifically unsupervised learning algorithms which are well-known methods utilized to implement clustering analysis or pattern recognition processes. These sorts of techniques allow to classify large amounts of observations (static pressures) in certain numbers of classes, where each class is composed of observations that share similar attributes, and different compared with observations from other classes. In recent years, many Oil and Gas disciplines like geophysics, petrophysic, production engineering, etc., have applied effectively unsupervised learning algorithms to facilitate classification or pattern recognition process. Successful utilizations have been described in seismic facies classification (Roy et al., 2013), image category detection (Seebock et al., 2017), hydraulic flow units identification (Contreras, 2018) and production-data classification (Subrahmanya et al., 2014). One of the most widely used algorithm to carry out unsupervised learning tasks is the Self-Organizing Map (Kohonen, 1982). This algorithm is a topographic organization in which nearby locations in the map represent observations with similar properties. This work shows the benefits of using unsupervised learning algorithms based on static pressure information to identify compartments in complex reservoirs. The theory of SOM and the workflow construction will be discussed in detail. Results for an application in a complex reservoir will be presented as well as comments regarding the reduction of time to obtain compartmentalization studies. This result demonstrates the potential of unsupervised learning algorithm to identify compartments in complex reservoirs and classify wells belonging to those compartments.

Machine Learning

Machine learning is defined as computational methods that use experience to improve performance or to make predictions (Mohri et al., 2012). In computer sciences, experience refers to the past information available to the learner, which takes the form of data collected in electronic format. This data could be training sets digitized and labeled by human or information obtained from measurements from the environment (Mohri et al., 2012).

Machine learning explores the construction and study of algorithms that can learn from and make predictions on data. Such algorithms operate by building a model from example inputs in order to make data-driven predictions or decisions, rather than following strictly static program instructions. The success of a learning algorithm relays on the data used, that's why it is related to data analysis and statistics. Learning techniques are data-driven methods that combine fundamental concepts in computer science with theories from statistics, probability and optimization (Mohri et al., 2012).

The most common machine learning types of training are supervised learning, unsupervised learning and semi-supervised learning. These techniques contrast in the types of training data available to the learner, the method by which training data is received and the type of test data used to evaluate the learning algorithm (Mohri et al., 2012). In supervised learning the learner receives a set of labeled information as training data and makes predictions for all unseen points (classification, regression, and ranking problems). In unsupervised learning the learner receives only unlabeled training data, and makes predictions for all unseen points. Since no labeled example is available, it could be challenging to quantitatively evaluate the performance of the learner. This procedure is associated with clustering and dimensionality reduction. Finally, in semi-supervised learning the learner receives a training sample consisting of both labeled and unlabeled data, and makes predictions for all unseen points. This technique could be used for applications such as classification, regression, or ranking tasks.

Unsupervised Learning

Unsupervised learning involves a process of auto association of information from the inputs in a set of classes with similar characteristics. Unsupervised learning algorithms are used to find patterns or features from data sets consisting of input data without labeled responses (Figure 1). The main objective is to find subclasses by exploring the structure of the data sets to extract meaningful information without the guidance of a known target output, discovering hidden structures based on similarities (Contreras, 2018). Unsupervised learning needs a criterion to terminate the process, otherwise the learning process continues even when a pattern has been found.

There are several kinds of unsupervised learning techniques that can accomplish the mission of discovering hidden structures in data where the right answer is not known in advance. One of the most useful procedure employed is clustering analysis, with which it is possible to yield a data description in terms of groups of data points that possess strong internal similarities. Clustering procedures use a criterion function to finalize the learning process, such as the sum of the squared distances from the cluster centers, and seek the grouping that maximizes the criterion function.

Clustering procedures rely mainly on two measures that govern the learning process, first measure is the similarity between two samples, which is the distance between them. The second measure is the partition of a set of samples into clusters, which is the optimum number of groups that maximize the distance between clusters. Clustering analysis involve the use of any distance or similarity function that can be evaluated for example with the following equation:

$$d(X,X') = \left(\sum_{k=1}^{d} |x_k - x'_k|^q\right)^{1/q}$$

Where *d* is the distance between vectors *X* and *X'*, *k* is the number of pair of samples and *q* is a selectable parameter greater or equal than 1. Setting q = 2 gives the familiar Euclidean metric while setting q = 1 provides the Manhattan or city block metric (Duda et al., 2000).

The other important task performing during the clustering analysis comprises the use of a criterion function to be optimized. The criterion function can be defined as a function that measures the clustering quality of any partition of the data. The most widely used criterion function for clustering is the sum-of-squared-error criterion, which can be represented by (Duda et al., 2000):

$$J_e = \sum_{i=1}^{c} \sum_{x \in \mathcal{D}_i} \|x - m_i\|^2$$

Where *c* and *Di* is the number of subsets or clusters, *D* is the set of samples, *x* is the vector of samples and *mi* is the mean of samples, given by:

$$m_i = \frac{1}{n_i} \sum_{x \in \mathcal{D}_i} x$$

Where *ni* is the number of samples *x1*, . . . , *xn*.

In other words, the criterion function minimizes the sum of the squared lengths of the vectors x and mi in a given cluster *Di*, where the mean vector mi is the best representative of the samples in *Di*. In consequence, *Je* measures the total squared error incurred in representing the *n* samples *x1*, . . . , *xn* by the c cluster centers *m1*, . . . , *mc*.

This work will be focused on unsupervised algorithm based on competitive learning, in which the output neurons of a neural network compete between themselves for the right to respond, resulting in only one neuron being activated at any one time. This activated neuron is called the winning neuron. The competition is induced by lateral inhibition connections between the neurons, forcing neurons to organize themselves. This network is called a Self-Organizing Map (Kohonen, 1982) (Figure 2).

Self-Organizing Map

Kohonen SOM is a topographic organization in which nearby locations in the map represent inputs with similar properties. SOM is a group of neurons organized in a low dimension mesh where each neuron is represented by a weight vector of m dimensions in which m is equal to the input vector dimension. Neurons are connected to adjacent neurons by a vicinity relationship that dictates topographic organization to the map.

The Self-Organizing Map represents a set of high-dimensional data items as a quantized two-dimensional image in an organized manner. Every data item is mapped into one point (node) in the map, and the distances of the items in the map reflect similarities between the items. (Kohonen, 2014) The goal of Self-Organizing Maps is to represent all points in the source space by Kohonen points in a target space, such that distance and proximity relationships are preserved in SOM maps as much as possible (Duda et al., 2000).

The process that must be followed to construct a self-organizing map involved a series of steps that take into account the neuron competition, neuron organization and a criterion to terminate the process. The algorithm comprises the following tasks (Contreras, 2018):

<u>Initialization:</u> In this step, all connection weights (w_i) are initialized with small random values. Weights are represented by a sequence of real n-dimensional Euclidean vectors $w_i(t)$, where t is an integer that denotes a step in the sequence. The sub-index i is the spatial index of the node with which w_i is associated.

<u>Competition:</u> In each pattern there are input units defined as $\mathbf{x} = \{x_i : i = 1, \ldots, D\}$ and the connection weights between the input units i and the neurons j, written as $\mathbf{w_j} = \{w_{ji} : j = 1, \ldots, N; i = 1, \ldots, D\}$ where N is the total number of neurons. Neurons compute their respective values of a discriminant function which provides the basis for competition. The discriminant function can be defined as the Euclidean distance between the input vector x and the weight vector w_j for each neuron j:

$$d_{ij}(\mathrm{X}) = \sum_{i=1}^{D} \left\| x_i - w_{ji} \right\|$$

The particular neuron with the smallest value of the discriminant function is declared the winner. In other words, the neuron whose weight vector comes closest to the input vector is declared the winner.

<u>Cooperation:</u> The winning neuron determines the spatial location of a topological neighborhood of excited neurons thereby providing the basis for cooperation among neighboring neurons. When one neuron is stimulated, its closest neighbor tends to get excited more than those neurons located further away. The Gaussian topological neighborhood for the neurons is defined as:

$$T_{i,I(X)} = \exp\left(-S_{j,I(x)}^{2} / 2\sigma^{2}\right)$$

Where $S_{i,I(x)}$ is the lateral distance between neurons j on the grid of neurons and the winning neuron, and $\sigma = \sigma(t)$ is a suitable decreasing time dependent function that contains information about the neighborhood radius.

Adaptation: In this process, excited neurons decrease their individual values of the discriminant function in relation to the input units through suitable adjustment of the associated connection weights. It is very important to the formation of ordered maps that not only the winning neuron gets its weights updated, but as topologically related subsets, its neighbors will have their weights updated as well, with a similar kind of correction imposed, although not as much as the winner itself. The appropriate weight update equation is:

$$w_i(t+1) = \mathrm{w}_i(t) + \alpha(t)\, T_{i,I(X)}(t)[x_i - \mathrm{w}_i(t)]$$

Where $\alpha(t)$ is the time dependent decreasing scalar function and could be represented by a suitable decreasing time dependent function for example, a decreasing exponential function (Figure 3), hyperbolic or piecewise linear functions. The learning rate decreases as the learning process proceeds.

The weight vector $\boldsymbol{w_j}$ associated with the winner neuron is updated in such a way that it moves towards the input vector $\boldsymbol{x}$ (Figure 4).

The factor $\alpha(t)T_{I,i(x)}$ in the above equation is called the neighborhood function. This function has the most central role in self organization process. This function resembles the kernel that is applied in usual smoothing processes. (Kohonen, 2014) The original SOM algorithm assumes that the above process converges and produces the wanted ordered values for the models. (Kohonen, 2014)

Application of SOM to Compartmentalization Study in Complex Reservoirs

Before the application of the workflow to study the reservoir compartmentalization in this complex reservoir, it was necessary to estimate the level of pressure available for each well at certain period of time. This plan started with a validation procedure applied to all static pressure tests recorded during the reservoir productivity life. Subsequently, all pressure measurements were referred to a single depth (datum), in order to compare pressure values in the same plane of reference. Finally, pressure values at a specific period of time were selected from the existing data set for each well, and in those wells with very few pressure tests available another procedure was followed to estimate pressure values at that specific period of time.

Data Preparation

The static pressure information used in this study belongs to a complex sandstone reservoir in a turbidite formation. The data set comprises coordinate X, coordinate Y and one static pressure information per well. There is static pressure information for 100 wells.

The data set was organized in an input matrix in order to be used in the unsupervised learning algorithm. In the input matrix, each ith column have elements that denote the characteristic that represent each well (coordinates and pressure) and each ith row represents the number of observations (wells). In this specific case, the matrix has a size of 3 columns (characteristics) and 100 rows (number of wells with static pressure information).

Applying the Unsupervised Learning Algorithm

Creating the Self-Organizing Maps

The SOM will be created with a neural network that will learn how to classify wells and assign those results to the corresponding compartments. First step is the initialization process of the SOM, where the map size, type of cell and initial weights with random values are specified. In this case, the SOM will have an arrangement of 2-dimension layer of 25 neurons (5x5) and hexagonal topology. During this process values for learning rate (α) = 0.1 and neighborhood radius (r) = 2 were specified.

Model Selection

This process consists of the selection of the optimum SOM size, which involves the selection of the map that minimizes the average distance between each data vector and its best matching unit (BMU), and the topographic error. During the selection process, a sensitivity analysis with different sizes of SOM was performed, in order to study and analyze the impact of different map sizes in the pattern recognition task. The following map sizes were used 5x5, 8x8, 10x10, 12x12, 14x14 and 16x16. Table 1 contains different SOM characteristics and its performance during the training process.

Table 1 shows that in a square map with random initialization vectors, the distance between data vector and its BMU decreases as the map size increases. However, the minimum value for topographic error was accomplished by the map size of 12x12 cells. In conclusion, the SOM that presents the best performance during the pattern recognition process is the map size 12x12.

The SOM selection process explained above corresponds to 100 pressure observation. Because the selection of the optimum SOM size is highly dependent on the number of observations, the selection process must be carried out when the number of observations change greatly. Other studies have demonstrated that the SOM size is proportional to the number of observations (Contreras, 2018). Additionally, some studies also state that there is no other method to select the optimum SOM size than a trial-and-error method. Once the optimum size of the SOM was identified, the trained map could be used to perform the compartmentalization study.

Case Study (Complex Reservoir)

This reservoir belongs to a turbidite formation, which are characterized for being highly complex deposits composed of several layers, lobes or channels. Pressure values in this reservoir ranges from 105 psi to 1615 psi, showing the degree of compartmentalization presents in the reservoir. The ample range of pressure depletion is the result of fluid production in different compartments with diverse sizes and characteristics. As stated before, one pressure value is available per well with a total of 100 wells. Figure 5 shows the well distribution in the reservoir with static pressure information.

The goal of the SOM is to identify automatically the number of compartments present in this reservoir by using the static pressure information to better understand the degree of complexity in the reservoir. The unsupervised algorithm was applied to the available information and the result is shown in Figure 6. The bar code in the SOM from Figure 6 denotes similarity between well characteristics (coordinates and static pressure values), ranging from yellow (very different) to dark blue (very similar). In other words, connection neurons with yellow, orange, green and cyan colors indicate wells with very different or different pressure trend and spatial coordinates, while connection neurons with blue and dark blue indicate wells with very similar characteristics.

Results

The SOM analysis shows that there is approximate 23 group of wells with different pressure trend (Figure 7), revealing that the reservoir is highly compartmentalized. Each group of wells are separated by neurons with yellow or cyan colors that represent different compartments. The more connecting neurons have the yellow or cyan color between the neurons with information, the more different the pressure trend of the well will be. Additionally, there are certain group of wells with small differences in coordinates and pressure values, separated by 1 row of neurons with a light blue color which can be interpreted as different compartments.

The SOM evaluation identified a total of 29 compartments which contain, in most of the cases, very different pressure trends and location characteristics (Figure 8). In Figure 8 is possible to observe the high degree of complexity in the reservoir by interpreting the pressure variability. It can be noted pressure differences up to 1190 psi in neighbor compartments separated by less than 500 meters. Figure 9 reveals that static pressure values for wells in the same compartment are very similar, noting that in most of the cases the pressure range in the same compartment is about 150 psi or less.

The compartments identification process by utilizing SOM honors the selection of well with similar coordinates and similar pressure information. Only six wells were observed to be misclassified, which represent only 6% of the total number of available wells. This indicates that there is a high probability to obtain a good compartmentalization study using this technique. This SOM incorporates the X and Y coordinates and the static pressure values at a specific period of time, however, the rate of pressure declination could be used to improve the classification process.

Discussion

The number of observations significantly influences the optimum SOM size to be used. Results in Table 1 corroborates this statement. Additionally, SOM features affect the classification performance. The compartments identification process by utilizing SOM honors the selection of well with similar coordinates and similar pressure information. Only six wells were observed to be misclassified, which represent only 6% of the total number of available wells. The rate of pressure declination could be used to improve the classification process.

The unsupervised learning is a powerful tool to identify different compartments in complex reservoirs, and additionally, reduces the time and the error introduced by human to obtain compartmentalization studies.

Conclusions

Unsupervised learning by using SOM is a great tool to perform compartmentalization studies in complex reservoirs based on static pressure information. The classification result indicates that SOM was able to identify compartments with large or small differences in its attributes (coordinates and pressure).

The SOM size is proportional to the number of observations, reason for which it is very important to execute a sensitivity analysis by varying map size in order to select the optimum SOM size with the best performance.

Results indicate that the compartmentalization study highly depends on well's coordinates and static pressure information. The number of misclassified wells was about 6%, which indicates that there is a high probability to obtain a good compartmentalization study using this procedure along with these attributes.

References Cited

Bar-Joseph, Z. et al., 2001, Fast optimal leaf ordering for hierarchical clustering: Proceeding of the Ninth International Conference on Intelligent Systems for Molecular Biology.

Contreras, Jose, 2018, Unsupervised Learning Applied to Flow Unit Identification Based on Wireline Formation Pressure Data: Search and Discovery Article #42260 (2018).

Duda, Richard, Peter Hart, and David Stork, 2000, Pattern Classification, second edition: Wiley-Interscience, New York.

Kohonen, T., 1982, Clustering, Taxonomy, and Topological Maps of Patterns: International Conference on Pattern Recognition (ICPR), p. 114-128.

Kohonen, T., 2014, MATLAB Implementations and Applications of the Self-Organizing Map: Unigrafia Oy, Helsinki, Finland.

Kohonen, T., 1996, Self-Organizing Map: 2nd ed., Springer-Verlag, Berlin, p. 78-82.

Kiviluoto, K., 1996, Topology Preservation in Self-Organizing Maps: Proceeding of International Conference on Neural Networks (ICNN), p. 294-299.

Mohri M., et al., 2012, Foundations of Machine Learning: Massachusetts Institute of Technology, the MIT Press, Cambridge, Massachusetts.

Roy et al., 2013, Active learning algorithms in seismic facies classification: SEGAM 2013-0769, presented at the SEG Houston 2013 Annual Meeting, Houston, Texas.

Seebock et al., 2017, Unsupervised Learning for Image Category Detection: 22nd Computer Vision Winter Workshop, Retz, Austria, February 6-8.

Shalev-Shwartz, S., and S. Ben-David, 2014, Understanding Machine Learning, from theory to algorithms: Cambridge University Press, New York, p. 309-310.

Subrahmanya et al., 2014, Advanced Machine Learning Methods for Production Data Pattern Recognition: presented at the SPE Intelligent Energy Conference and Exhibition, Utrecht, The Netherlands.

Zhu, X., and Z. Ghahramani, 2002, Learning from labeled and unlabeled data with label propagation: Technical Report CMU-CALD-02-107, Carnegie Mellon University.

Click to see slides of presentation

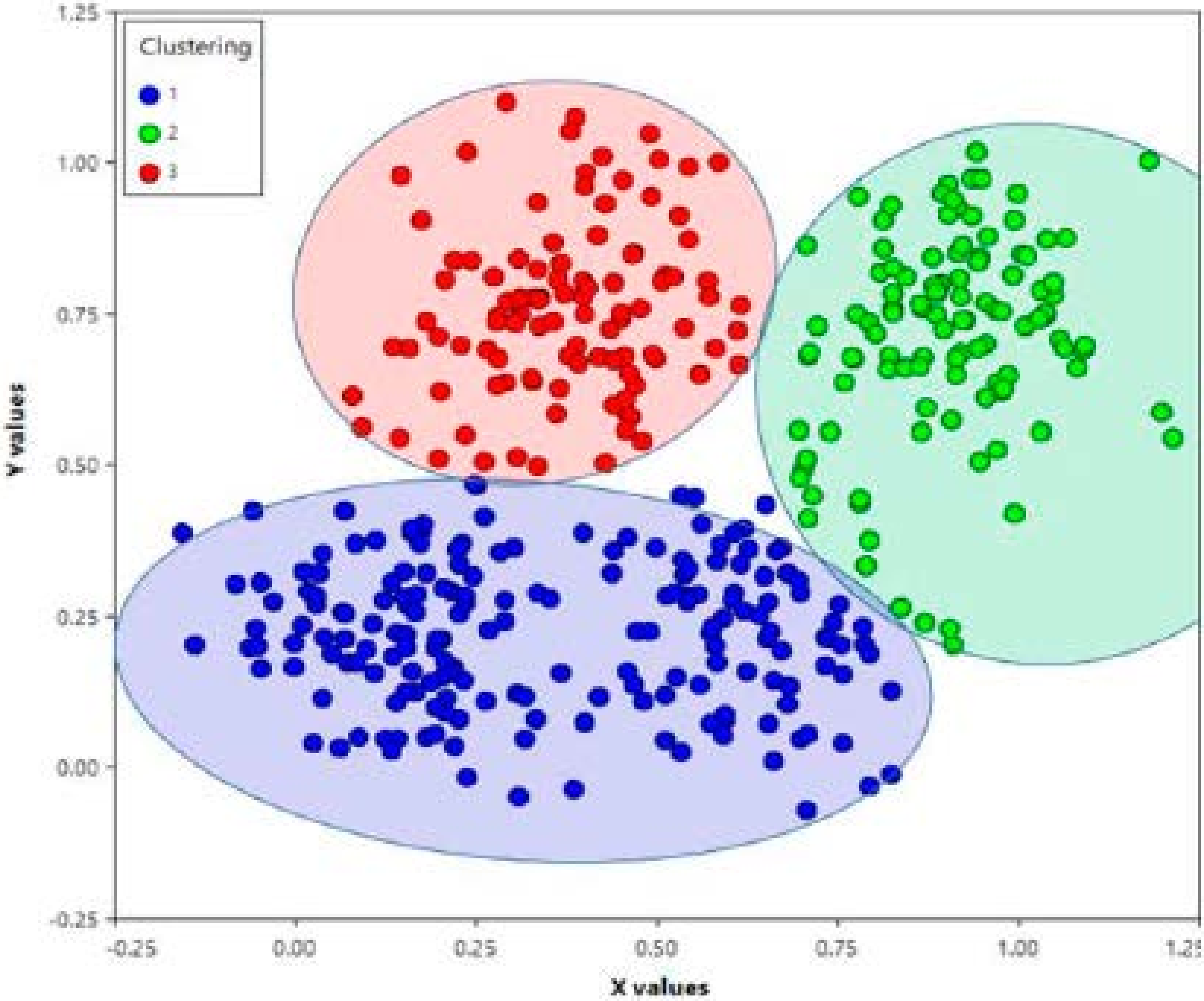

Figure 1. Pattern recognition of unlabeled data.

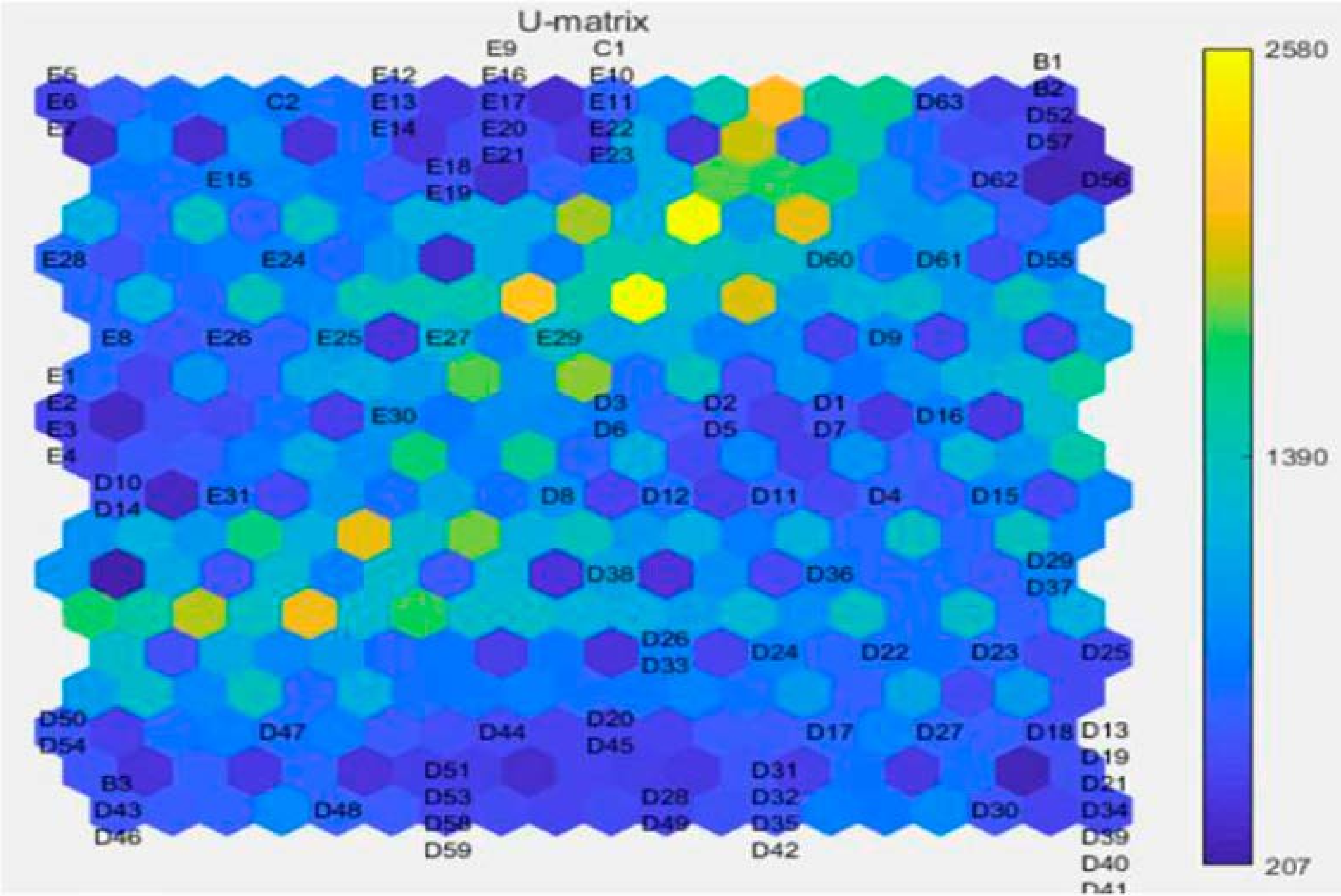

Figure 2. Classification with Self-Organizing Map.

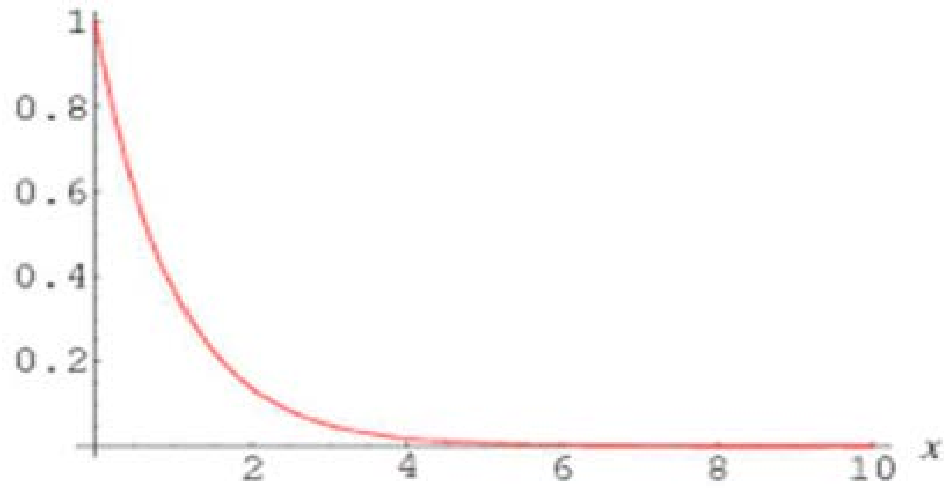

Figure 3. Learning rate with decreasing function.

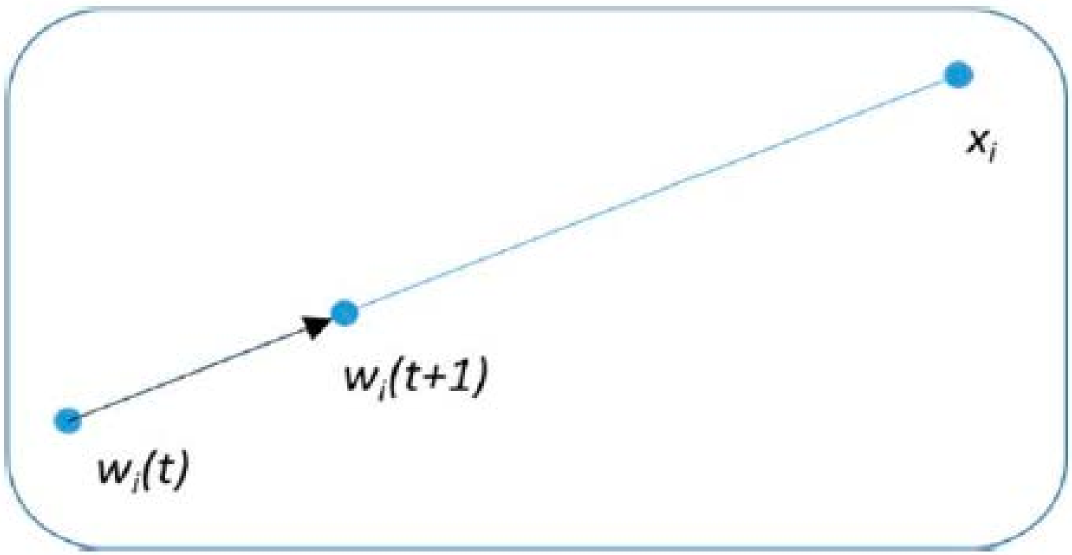

Figure 4. Adaptation process (Contreras, 2018).

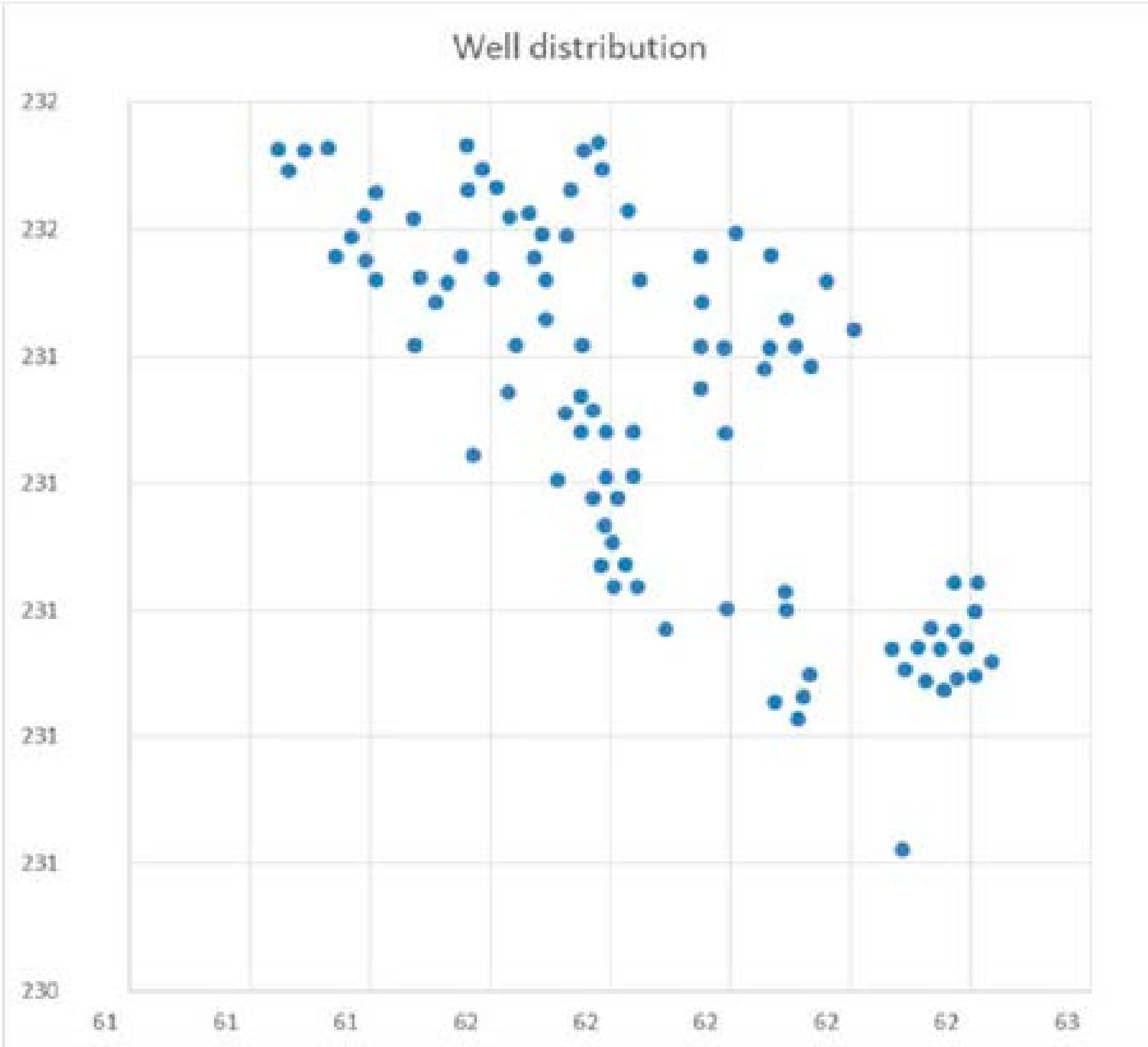

Figure 5. Well distribution with static pressure information in the complex reservoir.

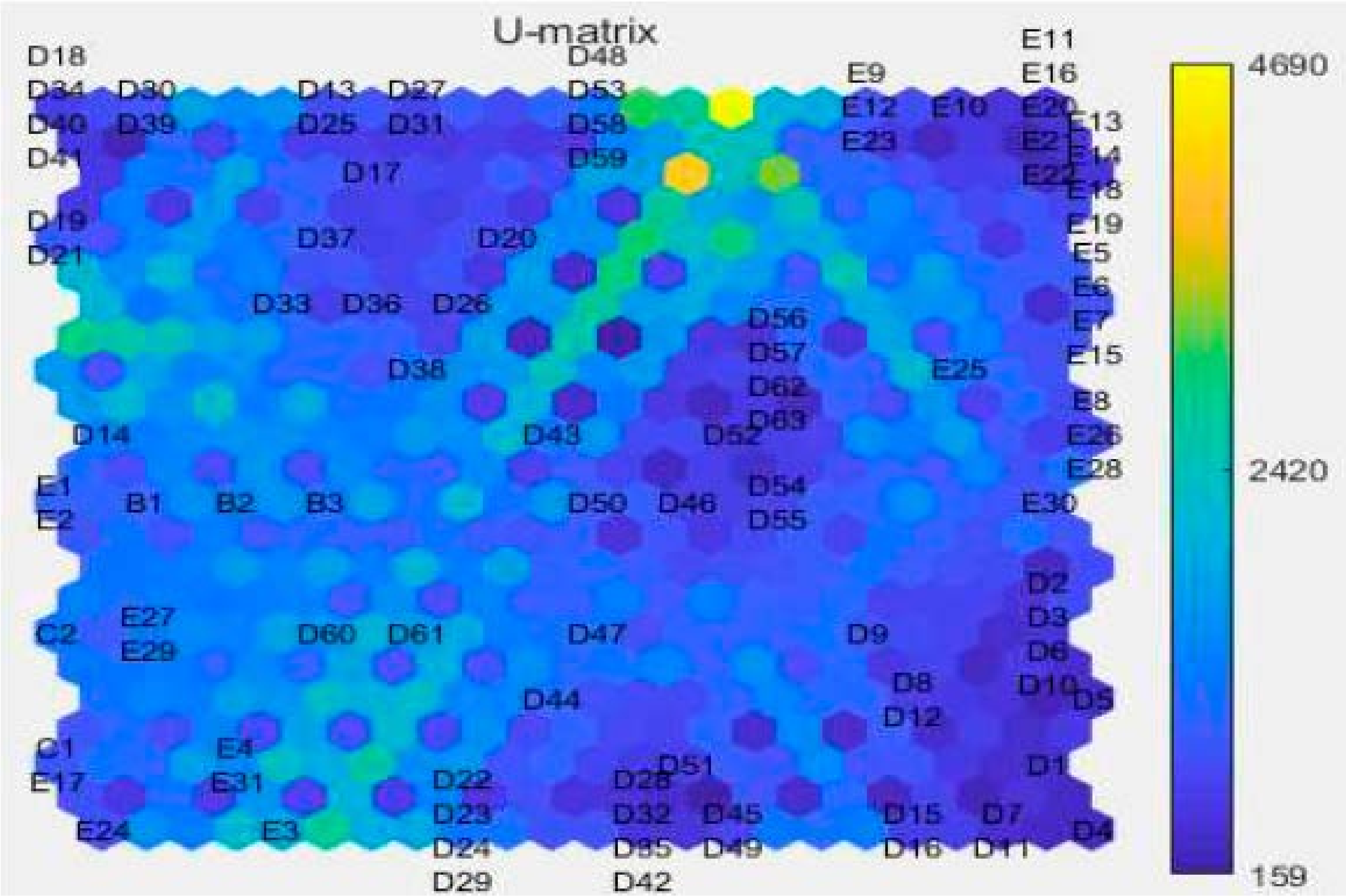

Figure 6. Self-Organizing Map identifying wells that share similar characteristics.

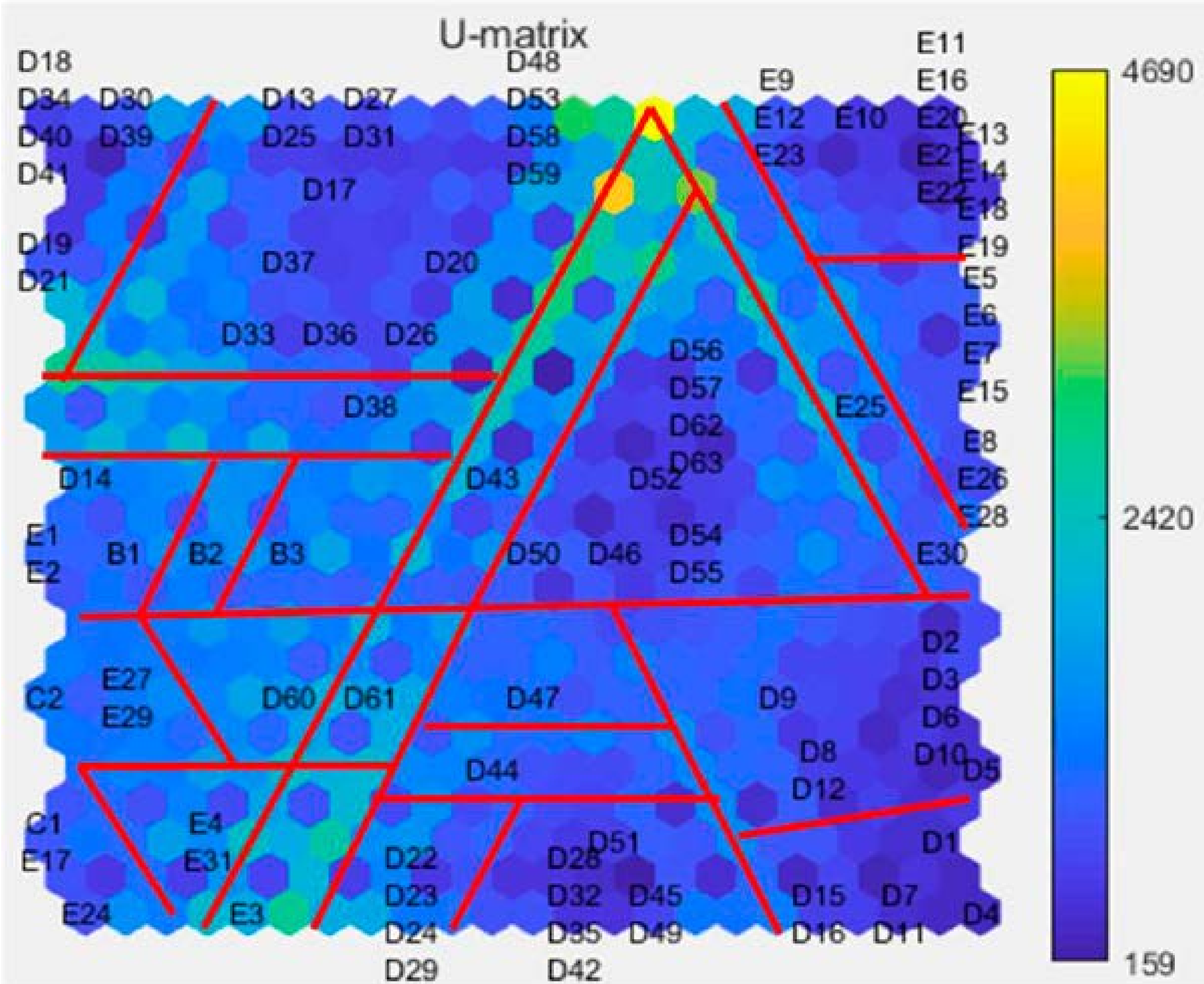

Figure 7. Identification of group of wells that follow the same pressure trend.

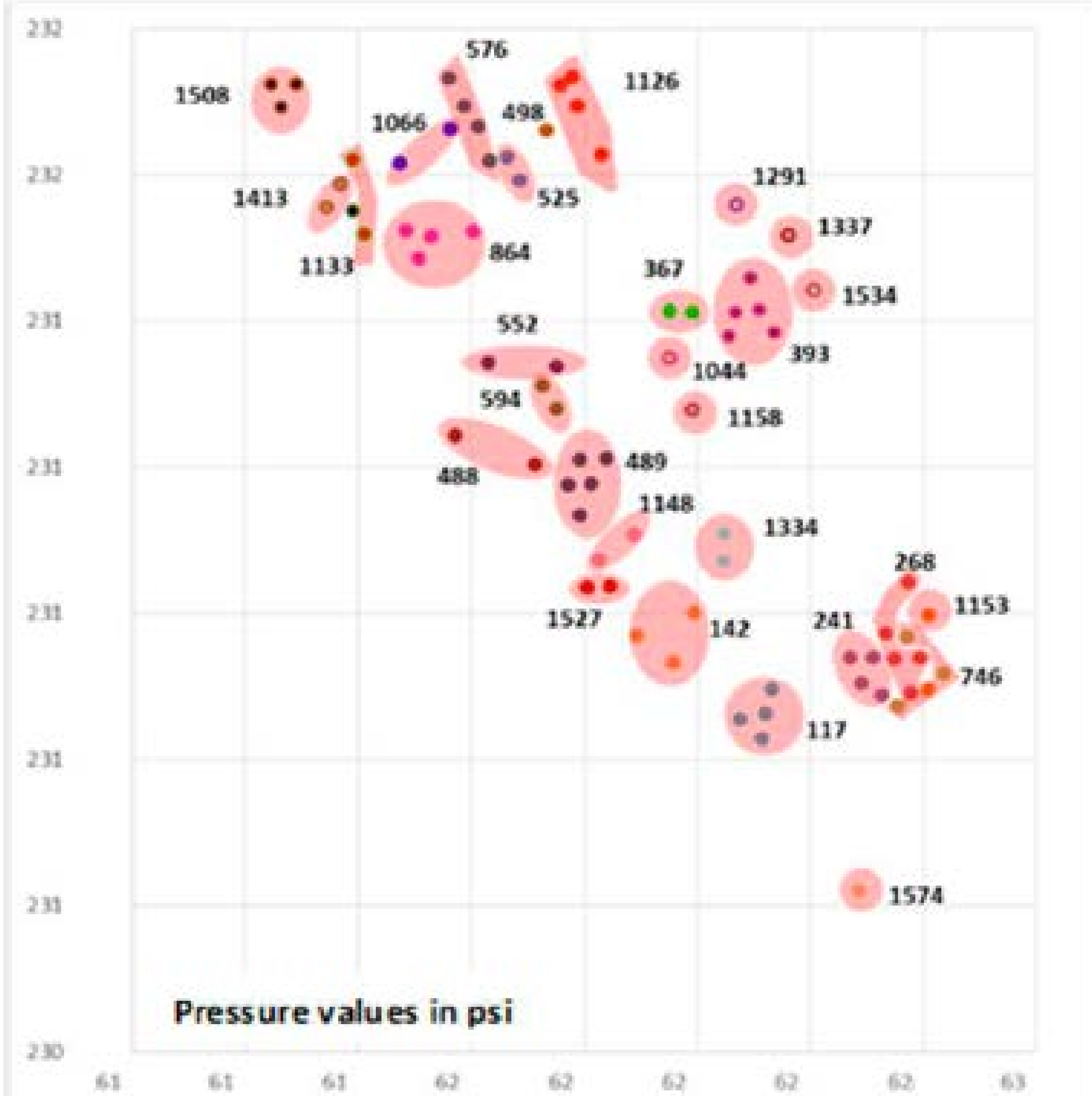

Figure 8. Well distribution with compartments identification.

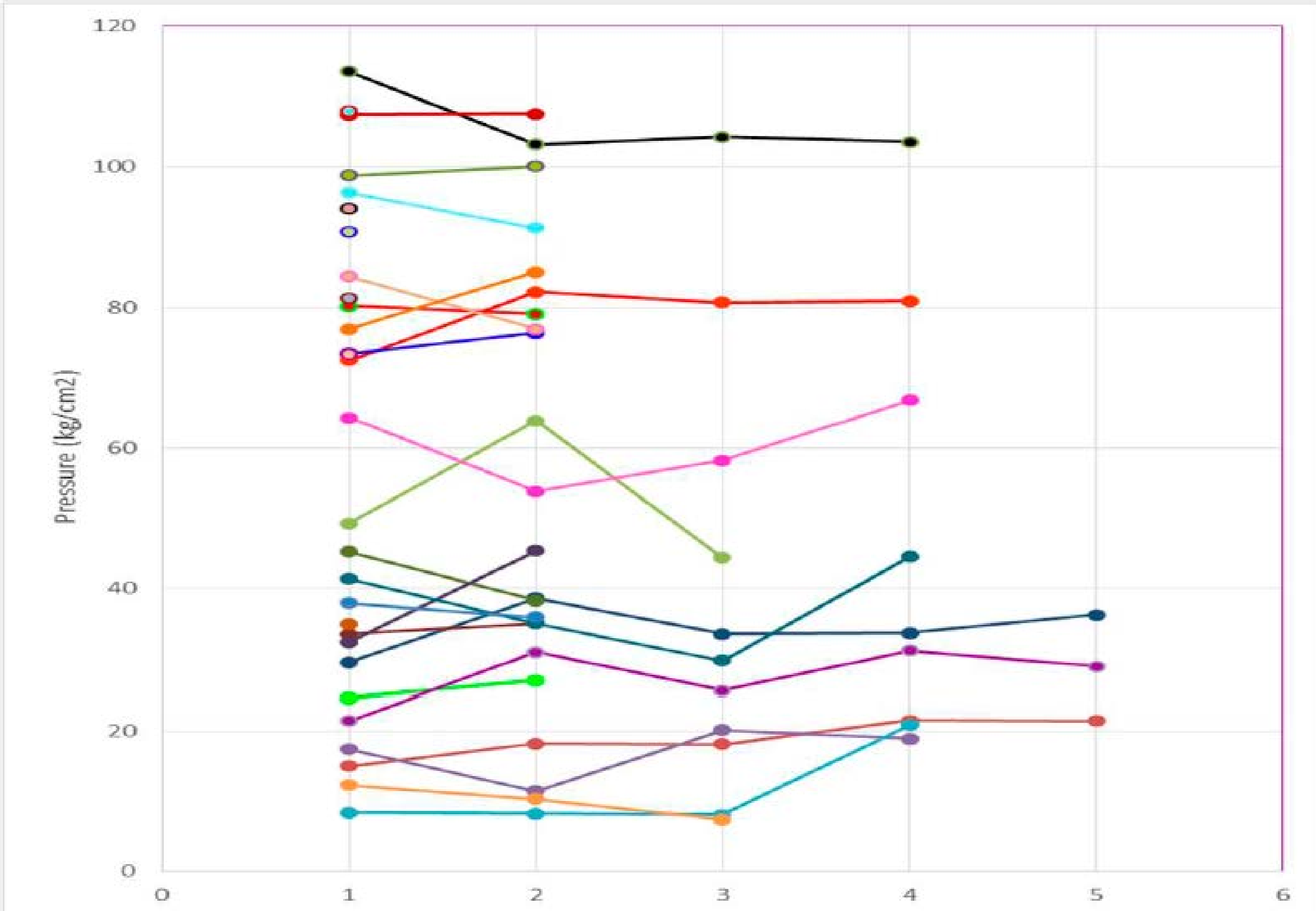

Figure 9. Static pressure values in each compartment.

SOM	α	r	Size	Initialization of SOM	Topology	Average distance to BMU	Topographic error
1	0.1	2	5x5	Random weight	Hexagonal	957.7260	0.0505
2	0.1	2	8x8	Random weight	Hexagonal	586.1954	0.1313
3	0.1	2	10x10	Random weight	Hexagonal	514.5578	0.1313
4	0.1	2	12x12	Random weight	Hexagonal	426.4821	0.0808
5	0.1	2	14x14	Random weight	Hexagonal	391.3766	0.1313
6	0.1	2	16x16	Random weight	Hexagonal	327.6043	0.1313
7	0.1	1	12x12	Random weight	Hexagonal	418.3134	0.2727
8	0.1	3	12x12	Random weight	Hexagonal	527.8874	0.1616
9	0.2	2	12x12	Random weight	Hexagonal	355.2159	0.1111

Table 1. Comparison results for different Self-Organizing Map characteristics.