Application of Artificial Neural Networks and Support Vector Machines for carbonate pores size estimation from 3D seismic data

A. Alimoradi[1*], A. Moradzadeh[1], M. R. Bakhtiari[2]

1. Faculty of Mining, Petroleum and Geophysics, ShahroodUniversity of Technology; Shahrood, Iran
2. Department of Petroleum Engineering, Amir Kabir University of Technology (Tehran Polytechnic), Tehran, Iran

Abstract

This paper proposes a method for the prediction of pore size values in hydrocarbon reservoirs using 3D seismic data. To this end, an actual carbonate oil field in the south-western part of Iran was selected. Taking real geological conditions into account, different models of reservoir were constructed for a range of viable pore size values. Seismic surveying was performed next on these models. From seismic response of the models, a large number of seismic attributes were identified as candidates for pore size estimation. Classes of attributes such as energy, instantaneous, and frequency attributes were included amongst others. Applying sensitivity analysis, we determined Instantaneous Amplitude and asymmetry as the two most significant attributes. These were subsequently used in our machine learning algorithms. In particular, we used feed-forward artificial neural networks (FNN) and support vector regression machines (SVR) to develop relationships between the known attributes and pore size values in a given setting. The FNN consists of twenty one neurons in a single hidden layer and the SVR method uses a Gaussian radial basis function. Compared with real values from the well data, we observed that SVM performs better than FNN due to its better handling of noise and model complexity.

Keywords: *Seismic Inversion, Seismic Attributes, Synthetic Data, Feed Forward Neural Network.*

1. Introduction

A challenging problem in quantitative reservoir modeling is the characterization of the carbonate reservoirs. These reservoirs, as one of the major hydrocarbon settings, include heterogeneous pore spaces with unknown and irregular distributions (from microscopic pore spaces of less than 1 mm in size to macroscopic pores of about 1 cm). Without proper determination of the distribution of pore spaces, it is difficult to perform reliable characterization of the carbonate reservoirs. Many researchers have worked on the problem of pore space detection and carbonate reservoir characterization in the past and the summary of their findings is briefly presented here (see for example [1], [2], [3], [4], [5]). Some of the developments have been an attempt in correlating the pores size with parameters such as water saturation, permeability, and porosity ([6], [7]).

Others have studied the detection of faults in a carbonate reservoir using sharp contrasts between acoustic impedances [8]. Siripitayananon et al [9] developed a method for facies detection using back-propagating artificial neural networks. Other noteworthy contributions have been intelligent inversion of seismic attributes to determine carbonate facies ([10], [11]), using multivariable statistical procedures to determine lateral changes of porosity in a carbonate field [12], and development of relationships between porosity and seismic attributes of amplitude, phase, and frequency [13]. Zhou et al [14] utilized amplitude variation with offset (AVO) and prestacked seismic data to obtain information about liquids in a carbonate formation. Most of the works done illustrate that the results cannot be completely dependable due to the distribution of pore spaces

and their effects on the values of reservoir properties of interest. Ellis et al [15] believe that pores have an effect on cementation factor in reservoirs. Lucia [16] showed that fluid saturation is an important property of hydrocarbon reservoirs that depends significantly on the pores size [17] investigated the effects of pore structure in carbonate reservoir on velocity using a dual porosity DEM theory [18] studied the effects of pores system on electrical conductivity of carbonates and concluded that the change in pores size can change the values of electrical conductivity and hence the values of water saturation in the reservoirs.

Although reservoir pores size can be derived reliably from core samples or well-log measurements, this property varies laterally from one well to another. Seismic data, particularly 3-D surveys, contain valuable information about the lateral variation of reservoir properties. When there are wells inside the seismic coverage, it is natural to infer the reservoir property between the wells by interpreting the seismic data and using the reservoir property at well locations as spatial control points. Assuming that there exists a functional or statistical relationship between the seismic data and the reservoir property, intelligent methods can probably be applied to establish a model of the relationship using the training sample set. This model can then be used to predict the reservoir properties away from the wells ([19],[20]).

This paper suggests an intelligent technique for reservoir characterization using artificial neural network and support vector machine to determine reservoir pores size from seismic attributes. We subsequently use a carbonate reservoir in southern Iran for which the values of pores size are readily availableas test bed for our proposed methodology.

2. Methodology
2.1. Site geology
One of the Iranian carbonate oil field which is located in the south western part of Iran was selected. This field consists of all of the necessary data for this study including 3D seismic data and well data (cores and logs). There are also two wells drilled in this field. Both wells contain hydrocarbon in Sarvak level (one of the famous hydrocarbon zones in Iranian carbonate oil fields) at the depth of 2850 meters. The thickness of the reservoir is about 200 meters. Since the data of well 1 are so noisy and incomplete, we decided to implement well 2 in this study. Geological investigations illustrate that the reservoir through this well (well 2) consists of pure limestone.

2.2. Seismic data acquisition
Using a realistic example, the proposed methodology will be explained. 3D seismic survey has been performed over this field. Figure 1 illustrates the seismic line which passes both wells. Since OpendTect is one of the most powerful packages in seismic data interpretation, the application of this software was considered for seismic attribute extraction. As previously mentioned, the data of well 1 are not suitable for analysis; therefore, it is inevitable to work on the data of well 2 only. According to limited resolution of seismic survey which leads to the limited number of data points in discrete well analysis, it is necessary to generate adequate synthetic data.

Forward modeling was done to simulate a reservoir level in Sarvak zone for well 2 using modified velocity form of the Gassmann rock physics equation ([21]):

$$\rho_{sat} V_{Psat}{}^2 = \left[\frac{\left(1 - \dfrac{\alpha K_{Gdry}}{K_0}\right)^2}{\dfrac{\varphi}{K_{fl}} + \dfrac{1-\varphi}{K_0} - \dfrac{\alpha K_{Gdry}}{K_0{}^2}} + \alpha K_{Gdry} + \frac{4}{3}\mu \right] \quad (1)$$

where,

ρ_{sat} = density of the saturated rock,

V_{Psat} = P-wave saturated rock velocity,

μ = rock shear modulus.

K_{Gdry} = dry rock effective bulk modulus from Geertsma equation,

K_0 = bulk modulus of the mineral material making up the rock,

K_{fl} = effective bulk modulus of the pore fluid,

φ = porosity, and

α = coefficient of pores sizes.

Saturated rock P-wave velocity (V_{Psat}) data were generated by changing the values of φ and α. Other parameters in Gassmann equation were considered to be constant (according to their real values in the reservoir). These parameters are: K_0 = 63 GPa, $\mu = 26$ GPA, $\rho = 2479$ kg/m^3 and S_W = 0.3279, respectively. 81 data were generated in this way (Table 1) each called from model 101 to model 909. The first value in each model name refers to the values of porosity (0.1 to 0.9) and the third one refers to values of α (0.1 to 0.9). For example, model 207 refers to the synthetic model which has the porosity of 0.2 and α of 0.7.

Preparing suitable codes in Seismic Unix forward modeling package, it is possible to construct the synthetic geological model. Figure 2 shows the

geological model for synthetic data 101. In this figure, "Zone L" illustrates the reservoir level. To make the model similar to the real reservoir, all levels above reservoir level were exactly considered with regard to their thickness and velocity. The objective is to perform seismic survey on the model and determine the seismic response of the model.

After constructing geological model for each synthetic data, a pre-defined seismic survey (by writing suitable codes in the Seismic Unix package) was performed over constructed models in order to extract the seismic response of each model. The output of this step were then processed using ray tracing technique and were stacked thoroughly to obtain the seismic section of the studied model. Therefore, 81 seismic sections that each one points to the specific pore size situation in the reservoir were extracted. Figure 3 illustrates the stacked seismic section of the model 101. These models can be used to extract attributes, make attribute analysis and study the effect of the changes in pore size on different attributes. Hence, it is possible to find related attributes with pore size parameter and model the relationship between those attributes and the values of pore size.

2.3. Attribute extraction and analysis

To investigate the effect of pore size changes on attribute values, the synthetic models were classified into specific groups. Different α values in each φ state were considered as a group (Table 2). Seismic attributes should be extracted in each group and attribute analysis should be performed over them. According to the proper capabilities of OpendTect software, in seismic attribute analysis and interpretation, this package was considered to extract attributes in our study. OpendTect is an open source system for seismic data interpretation that interprets huge volume of seismic data using attributes and new techniques of imaging. In this study, 43 different seismic attributes were extracted for all models in Table 2. Table 3 illustrates the values of these attributes for model 101.

In the next step, attributes were analyzed to form a correlation matrix. This matrix for all groups indicated that two attributes of *Instantaneous Amplitude* and *Asymmetry* have the highest correlation values with the values of pore size. Table 4 shows the values of correlation coefficient for these two attribute in the first three groups of pore size.
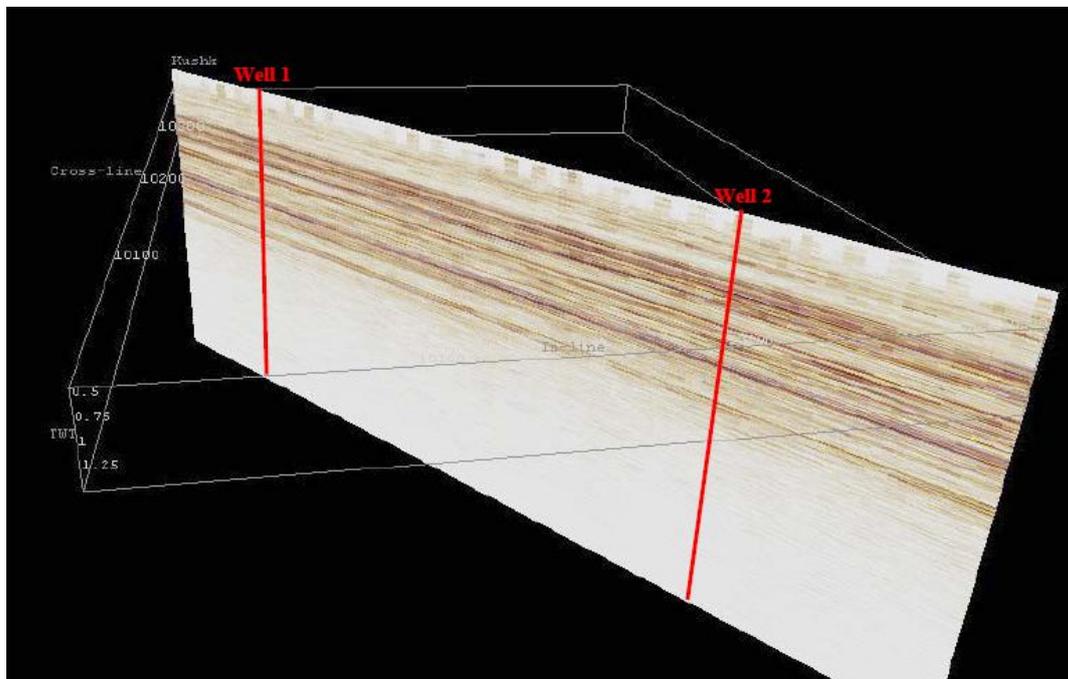
Figure 1. Seismic line over wells 1 and 2.

Table 1. 81 synthetic models generated using modified Gassmann equation.

φ	K_{Gdry}	α	K_{dryNew}	V_P	Model	φ	K_{Gdry}	α	K_{dryNew}	V_P	Model
		0.1	1.05	4573	101			0.1	0.573	4243	201
		0.2	2.1	4598	102			0.2	1.146	4263	202
		0.3	3.15	4624	103			0.3	1.719	4283	203
		0.4	4.2	4649	104			0.4	2.292	4302	204
0.1	10.5	0.5	5.25	4675	105	0.2	5.73	0.5	2.865	4322	205
		0.6	6.3	4700	106			0.6	3.438	4341	206
		0.7	7.35	4726	107			0.7	4.011	4361	207
		0.8	8.4	4752	108			0.8	4.584	4380	208
		0.9	9.45	4778	109			0.9	5.157	4399	209
		0.1	0.394	4101	301			0.1	0.3	4021	401
		0.2	0.788	4116	302			0.2	0.6	4034	402
		0.3	1.182	4132	303			0.3	0.9	4047	403
		0.4	1.576	4147	304			0.4	1.2	4059	404
0.3	3.94	0.5	1.970	4163	305	0.4	3	0.5	1.5	4072	405
		0.6	2.364	4178	306			0.6	1.8	4084	406
		0.7	2.758	4193	307			0.7	2.1	4097	407
		0.8	3.152	4208	308			0.8	2.4	4109	408
		0.9	3.546	4224	309			0.9	2.7	4122	409
		0.1	0.242	3971	501			0.1	0.203	3935	601
		0.2	0.484	3981	502			0.2	0.406	3945	602
		0.3	0.726	3992	503			0.3	0.609	3954	603
		0.4	0.968	4003	504			0.4	0.812	3963	604
0.5	2.42	0.5	1.210	4013	505	0.6	2.03	0.5	1.015	3972	605
		0.6	1.452	4024	506			0.6	1.218	3981	606
		0.7	1.694	4034	507			0.7	1.421	3991	607
		0.8	1.936	4045	508			0.8	1.624	4000	608
		0.9	2.178	4056	509			0.9	1.827	4009	609
		0.1	0.175	3909	701			0.1	0.154	3890	801
		0.2	0.350	3918	702			0.2	0.308	3897	802
		0.3	0.525	3926	703			0.3	0.462	3904	803
		0.4	0.700	3934	704			0.4	0.616	3911	804
0.7	1.75	0.5	0.875	3942	705	0.8	1.54	0.5	0.770	3919	805
		0.6	1.050	3950	706			0.6	0.924	3926	806
		0.7	1.225	3958	707			0.7	1.078	3933	807
		0.8	1.400	3966	708			0.8	1.232	3940	808
		0.9	1.575	3974	709			0.9	1.386	3948	809
		0.1	0.137	3874	901						
		0.2	0.274	3880	902						
		0.3	0.411	3887	903						
		0.4	0.548	3894	904						
0.9	1.37	0.5	0.685	3900	905						
		0.6	0.822	3907	906						
		0.7	0.959	3913	907						
		0.8	1.096	3920	908						
		0.9	1.233	3926	909						
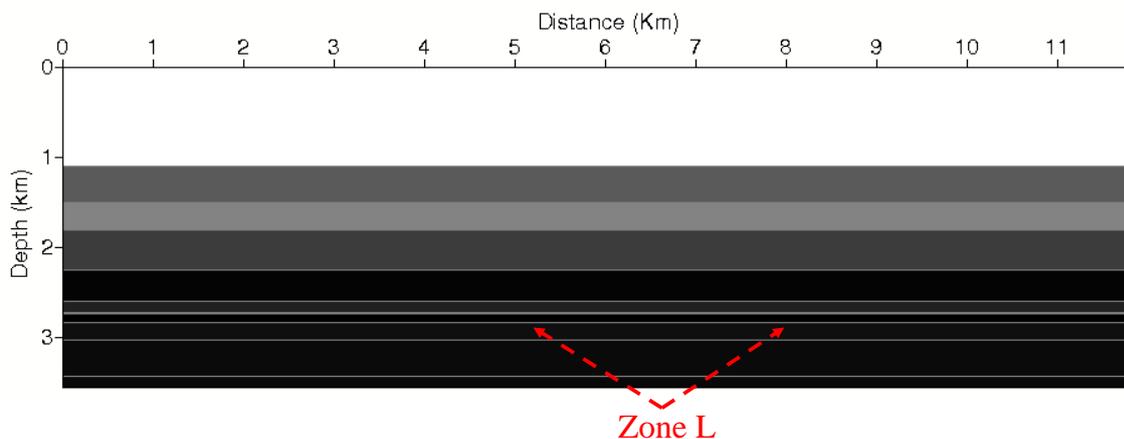
Figure 2. Geological model of the reservoir constructed by Seismic Unix.

Table 2. Different groups of pore size according to constant values of φ.

				Pore size Groups				
= 0.1 φ	= 0.2 φ	= 0.3 φ	= 0.4 φ	= 0.5 φ	= 0.6 φ	= 0.7 φ	= 0.8 φ	= 0.9 φ
101	201	301	401	501	601	701	801	901
102	202	302	402	502	602	702	802	902
103	203	303	403	503	603	703	803	903
104	204	304	404	504	604	704	804	904
105	205	305	405	505	605	705	805	905
106	206	306	406	506	606	706	806	906
107	207	307	407	507	607	707	807	907
108	208	308	408	508	608	708	808	908
109	209	309	409	509	609	709	809	909

Table 3. Values of 43 seismic attributes for model 101.

Energy/Energy	0.001516	Velocity Fan Filter	0.000213274
Energy/Sqrt	0.038938	Frequency/Dominent Frequency	13.989712
Energy/Ln	-6.491552	Frequency/Average Frequency	45.645275
Instantaneous/Amplitude	0.001748	Frequency/Median Frequency	38.888885
Instantaneous/Phase	0.250555	Frequency/Average Frequency Squared	3006.6557
Instantaneous/Frequency	101.56539	Frequency/Maximum Spectral Amplitude	0.03902
Instantaneous/Hilbert	0.000433462	Frequency/Spectral Area Beyond Dominent Frequency	0.258528
Instantaneous/Amplitude/1st Derivative	0.205685	Frequency/Frequency Slope Fall	0.202982
Instantaneous/Amplitude/2nd Derivative	59.849281	Frequency/Absorption Quality Factor	14.671811
Instantaneous/Cosine Phase	0.968775	Spectral Decomp	3.73E-09
Instantaneous/Envelope Weighted Phase	0.224808	Event/Peakedness	0.000132157
Instantaneous/Envelope Weighted Frequency	109.31568	Event/Steepness	0.004993
Instantaneous/Phase acceleration	-3067.375	Event/Assymetry	0.597961
Instantaneous/Thin bed indicator	-7.75029	Volume Statistics/Average	-0.013966
Instantaneous/Bandwidth	18.724953	Volume Statistics/Median	0.001047
Instantaneous/Q factor	-2.712033	Volume Statistics/Variance	0.001308
Convolve/Lowpass	0.001834	Volume Statistics/Min	-0.119581
Convolve/Laplacian	-0.000703483	Volume Statistics/Max	0.004804
Convolve/Prewitt	0.002109	Volume Statistics/Sum	-0.628459
Frequency Filter/ LowPass	0.003497	Volume Statistics/Norm Variance	4.014766
Frequency Filter/HighPass	-0.009834	Volume Statistics/RMS	0.038396
Frequency Filter/BandPass	-0.007492		

Table 4. Values of correlation coefficient for Instantaneous Amplitude and Asymmetry.

Group	R (Instantaneous Amplitude)	R (Asymmetry)
1 (Models 101 – 109)	0.925	-0.956
2 (Models 201 – 209)	0.884	-0.911
3 (Models 301 – 309)	0.875	-0.882
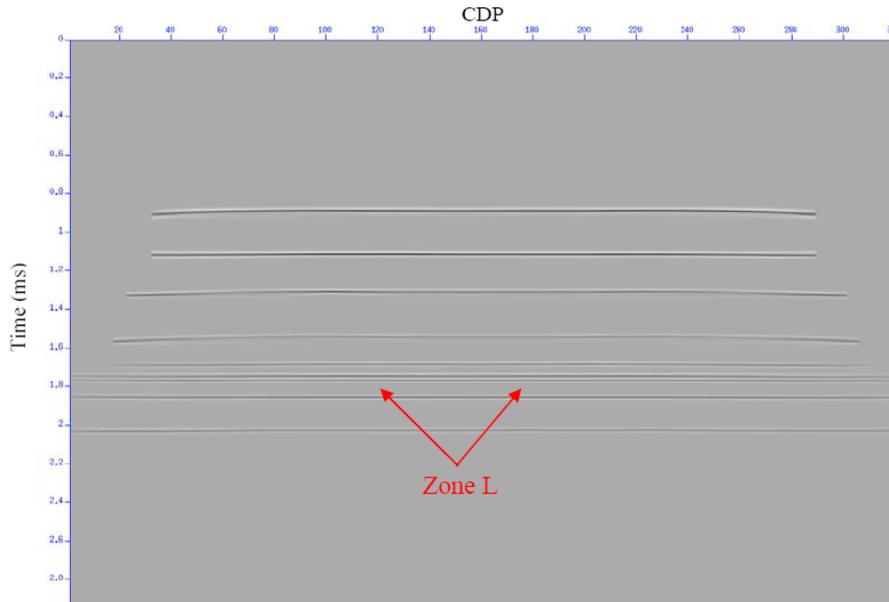
Figure 3. The stacked seismic section of the model 101.

2.4. Back-propagating Artificial Neural Networks (BANN)

Artificial neural networks (ANNs) are computational models based on human's understanding of cortical structure of the brain and cognition. Algorithmically, ANNs are parallel adaptive systems and therefore require training. Back-propagation is a powerful method of supervised learning that is developed after the seminal work by Werbos and Rumelhart in seventies and eighties [22]. Details of various methods of ANN design and training are beyond the scope of this paper and as such they are explained elsewhere (see [23] for example); nevertheless, a brief description of the terminology is provided here.

The structure of a neural network, in general, consists of an interconnected group of artificial neurons (simple processors that are connected to many other neurons). These processing units receive the information, apply some simple processing on them and pass them to other neurons. The flow of information creates a computational model for information processing. Each neuron is assigned a weight that is changed adaptively to improve the performance of the network based on pairs of external and internal signals (training information, input-output mapping). Practically, neural networks may be used in nonlinear statistical data modeling, system identification, extraction of complex relationships between inputs and outputs of a system, and for pattern recognition.

In addition to weight, each node (neuron) in the network is equipped with an activation function (or transfer function) that is part of the information processing unit of the neuron. The flow of information could be imagined from left to right, such that each neuron performs the processing on the data in parallel with other neurons in the layer. The response of the network is compared at the terminating layer with a set of desired outputs and the weights of the neurons are thus corrected following a training algorithm to minimize the output error. Issues with regards to the number of nodes per layer, number of layers, and the type of activation function that could be used are dealt with in the design of the architecture of the network. This is explained later on in this paper.

There are numerous methods of training of a neural network. Categorically, these methods are grouped into three main classes: supervised learning, unsupervised learning, and reinforcement learning. In a supervised learning scheme, the network is provided with a set of examples in the input-output space: $(x, y), x \in X, y \in Y$ and the goal of the training process is to find function f in a set of valid functions that could match the input/output pairs reliably. By doing so, the network becomes capable of making inferences in mapping that is implied by the training data. This procedure involves minimizing a cost function. The cost function is often defined as the mismatch between the network's mapping and the actual data.

A commonly used cost function is the mean-squared error between the average of network's output, *f(x)*, and the target value *y* over all

example pairs presented to the network. Minimizing this cost function in a gradient descent algorithm for a class of neural networks called Multi-Layer Perceptrons constitutes the basis of back-propagation algorithm [22] .In this study, we successfully developed and implemented a network with one hidden layers of 6 nodes.

Dataset of 81 data points to train and test the neural network was used (Table 5). From this, 60 points (80% of the total data) were selected randomly for the network training and the remaining 20% of the data was used for testing the network. Each data point is a vector of three input values, namely, *Porosity*, *Instantaneous Amplitude* and *Asymmetry* as described earlier. The desired network output is *pore size* value. The input layer of the network receives input data at three nodes and the network generates an output at the final layer. We used the *Scaled Conjugate Gradient* (*SCG*) method for training because it generally results in faster and more reliable convergence for our application.

The best and the worst results of 20 iterations for training of the network are presented in Table 6. In Table 6, RMS_{train} is the root-mean-square of the training error, and RMS_{test} is the root-mean-square of error during testing of the network. Considering the limited amount of data available for network training, the results shown in Table 6 appear to be reasonable for practical applications. Other training algorithms such as Levenberg-Marquardt, One-Step Secant, and Fletcher-Powell Conjugate Gradient were also used but were discarded due to higher tolerance for the test errors and lower reliability in our application [22]. The results of the training are presented in Figure 4.

In Figure 4, *R* is the correlation coefficient between the real and the predicted *pore size* values; *A* and *T* are the predicted and real values respectively. The correlation coefficient is close to 1.0, implying a good network performance. The gap between values is caused by simulating data in special pore sizes (0.1, 0.2, … , 0.5).

We used the abovementioned neural network to classify the test data. The results are shown in Figure 5.

During testing, a correlation coefficient of greater than 0.85 was generally obtained (as exhibited in Figure 5). This shows that the pore size values in the test data were practically well-correlated with the network predictions; but the other parameter that should be mentioned is the error value. Table 6 shows that the root mean square error value is about 0.17 which means the high value of error.

Table 5. Dataset used for ANN.

Input Data			Output Data
Porosity	Instantaneous Amplitude	Assymetry	α
0.1	0.001748	0.597961	0.1
0.1	0.002019	0.597663	0.2
0.1	0.001525	0.60772	0.3
0.1	0.002473	0.588113	0.4
0.1	0.002912	0.56398	0.5
0.1	0.003173	0.560266	0.6
0.1	0.003345	0.542739	0.7
0.1	0.003691	0.532852	0.8
0.1	0.003396	0.512257	0.9
0.2	0.000931859	0.685173	0.1
0.2	0.000791402	0.688684	0.2
0.2	0.00090889	0.686031	0.3
0.2	0.000899879	0.683106	0.4
0.2	0.001065	0.673531	0.5
0.2	0.000999837	0.680311	0.6
0.2	0.00122	0.664729	0.7
0.2	0.00113	0.651461	0.8
0.2	0.00127	0.647098	0.9
0.3	0.000837592	0.71386	0.1
0.3	0.000885393	0.708296	0.2
0.3	0.000883955	0.702708	0.3
0.3	0.000892952	0.693833	0.4
0.3	0.000939133	0.692525	0.5
0.3	0.000900433	0.689285	0.6
0.3	0.000913071	0.68192	0.7
0.3	0.000843878	0.684671	0.8
0.3	0.000808572	0.691255	0.9
0.4	0.000648431	0.730728	0.1
0.4	0.000709612	0.73075	0.2
0.4	0.000656809	0.731734	0.3
0.4	0.000651827	0.749313	0.4
0.4	0.000736101	0.727364	0.5
0.4	0.000821421	0.717924	0.6
0.4	0.000811904	0.709992	0.7
0.4	0.000916465	0.712966	0.8
0.4	0.000783791	0.70152	0.9
0.5	0.000628134	0.73861	0.1
0.5	0.000627863	0.734825	0.2
0.5	0.000695581	0.738426	0.3
0.5	0.000629897	0.746579	0.4
0.5	0.000658908	0.740022	0.5
0.5	0.000681385	0.733785	0.6
0.5	0.000685294	0.72986	0.7
0.5	0.000770238	0.72278	0.8
0.5	0.000736661	0.725888	0.9
0.6	0.000612668	0.732602	0.1
0.6	0.000550543	0.738258	0.2
0.6	0.000587072	0.735914	0.3
0.6	0.000736215	0.730293	0.4
0.6	0.000688032	0.735972	0.5
0.6	0.000653682	0.736834	0.6
0.6	0.00064337	0.738835	0.7
0.6	0.000652874	0.735078	0.8
0.6	0.000691799	0.731495	0.9
0.7	0.000509322	0.734671	0.1
0.7	0.000734904	0.738778	0.2
0.7	0.000676518	0.73741	0.3
0.7	0.00084061	0.737717	0.4
0.7	0.000944629	0.735134	0.5
0.7	0.000826004	0.738131	0.6
0.7	0.000875064	0.73847	0.7
0.7	0.000691957	0.734863	0.8
0.7	0.000630033	0.73149	0.9
0.8	0.000569378	0.741718	0.1
0.8	0.000645451	0.744131	0.2
0.8	0.00093433	0.739249	0.3
0.8	0.000839511	0.738052	0.4
0.8	0.000999948	0.738124	0.5
0.8	0.000850102	0.738184	0.6
0.8	0.000765388	0.737913	0.7
0.8	0.000723143	0.73148	0.8
0.8	0.000650587	0.737991	0.9
0.9	0.000567052	0.739472	0.1
0.9	0.000656579	0.74714	0.2
0.9	0.001046	0.741409	0.3
0.9	0.000745751	0.746428	0.4
0.9	0.000541761	0.739048	0.5
0.9	0.001028	0.738884	0.6
0.9	0.00062945	0.739424	0.7
0.9	0.000601762	0.738741	0.8
0.9	0.000554786	0.743716	0.9

This is evident in Figure 6 with the medium performance of the trained network remarkably demonstrated for a set of synthetic test data. The real values of pore size, shown by small circles in Figure 6 could not be easily predicted by the back-propagating neural network, shown by small inverted triangles.

Table 6. Error values for the best and the worst results.

RMS_{train}	RMS_{test}
0.08	0.11
0.09	0.28
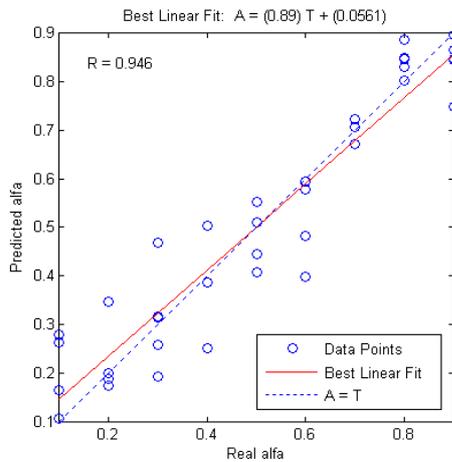
Figure 4. Correlation coefficient for train data.
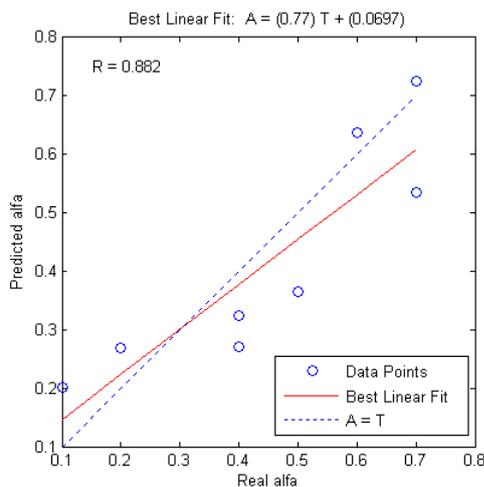
Figure 5. Correlation coefficient for the test data

As shown in Table 6, the reduction in the network error will increase the reliability of network's predictions. To do this, either additional training data should be available or another method should be used. In some cases, ANNs detect the relative optimum point instead of global optimum point as a solution for the problem which is the main weak point for ANNs [24]. Recognizing the computational power of support vector machines

in rule generation and function approximation and their robustness particularly in the area of data classification, we embarked on developing and training of a support vector regression machine (SVR) for the purpose of classification of pores size in this study.
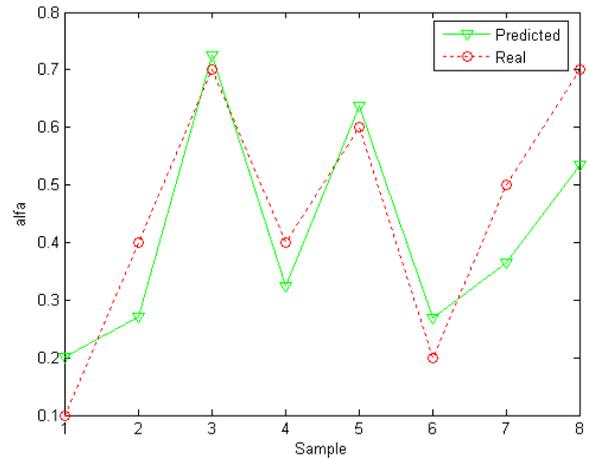
Figure 6. Predicted results for a set of test data.

2.5. Support vector machines (SVM) and their application for this Study

In pattern recognition, the SVM algorithm constructs nonlinear decision functions by training a classifier to perform a linear separation in some high dimensional space which is nonlinearly related to input space. To generalize the SVM algorithm for regression analysis, an analogue of the margin is constructed in the space of the target values (y) using Vapnik's ε-insensitive loss function (Figure 7) [25].

$$|y - f(x)|_\varepsilon := \max\{0, |y - f(x) - \varepsilon|\} \quad (2)$$

To estimate a linear regression

$$f(x) = (w.x) + b \quad (3)$$

where, w is the weighted matrix, x is the input vector and b is the bias term. With precision, one minimizes

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}|y - f(x)|_\varepsilon \quad (4)$$

where, C is a trade-off parameter to ensure that the margin ε is maximized and error of the classification ξ is minimized. Considering a set of constraints, one may write the following equations as a constrained optimization problem:

$$L(w, \xi, \xi') = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i') \qquad (5)$$

$$Subject\ to \begin{cases} y_i - w^T.x - b \le \xi_i + \varepsilon & (6) \\ w^T.x + b - y_i \le \xi_i' + \varepsilon & (7) \\ \xi_i, \xi_i', x_i \ge 0 & (8) \end{cases}$$

According to Equations (6) and (7), any error smaller than ε does not require a nonzero ξi or ξ_i', and does not enter the objective function 2 [27].

By introducing Lagrange multipliers (α and α') and allowing for C > 0, ε > 0 chosen a priori, the equation of an optimum hyper plane is achieved by maximizing of the following equations:

$$L(\alpha, \alpha') = \frac{1}{2} \sum_{i=1}^{N} \begin{matrix} \sum_{i=1}^{N} (\alpha_i - \alpha_i') x_i' x_i (\alpha_i - \alpha_i') + \\ \sum_{i=1}^{N} ((\alpha_i - \alpha_i') y_i - (\alpha_i + \alpha_i')\varepsilon) \end{matrix} \qquad (9)$$

$$Subject\ to\ 0 \le (\alpha_i - \alpha_i') \le C \qquad (10)$$

where, x_i appears only inside an inner product. To get a potentially better representation of the data in non-linear case, the data points can be mapped into an alternative space, generally called feature space (a pre-Hilbert or inner product space) through a replacement:

$$x_i.x_j \rightarrow \varphi(x_i).\varphi(x_j) \qquad (11)$$

The functional form of the mapping $\varphi(x_i)$ does not need to be known since it is implicitly defined by the choice of kernel: $k(x_i, x_j) = \varphi(x_i).\varphi(x_j)$ or inner product in Hilbert space. With a suitable choice of kernel the data can become separable in feature space while the original input space is still non-linear. Thus, whereas data for n-parity or the two spirals problem is non-separable by a hyper plane in input space, it can be separated in the feature space by the proper kernels. Table 7 gives some of the common kernels.

Then, the nonlinear regression estimate takes the following form:

$$y_i = \sum_{i=1}^{N} \sum_{j=1}^{N} \begin{matrix} (\alpha_i - \alpha_i')\varphi(x_i)^T \varphi(x_j) + b = \\ \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i - \alpha_i') K(x_i, x_j) + b \end{matrix} \qquad (12)$$

where, b is computed using the fact that Equation (6) becomes an equality with $\xi i = 0$ if $0<\alpha_i<C$, and Equation (7) becomes an equality with $\xi_i' = 0$ if $0 < \alpha_i' < C$ [28].
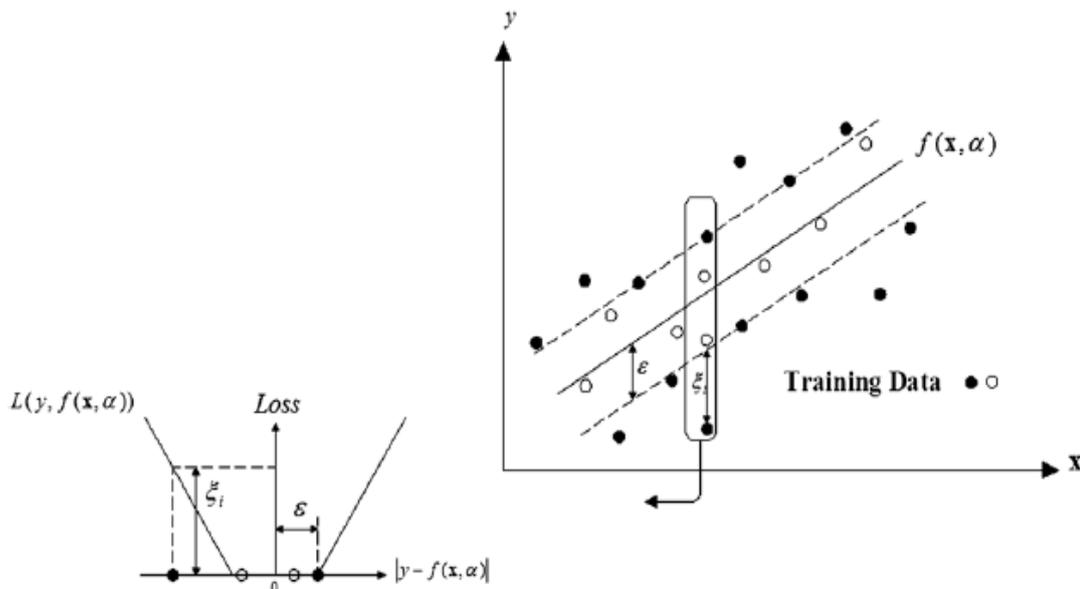
Figure 7. Concept of ε-insensitivity. Only the samples out of the $\pm$ ε margin will have a non-zero slack variable, so they will be the only ones that will be part of the solution [26]

Table 7.Polynomial, normalized polynomial and Radial Basis Function (Gaussian) Kernels [29].

Kernel Function	Type of Classifier
$K(x_i, x_j) = (x_i^T x_j + 1)^\rho$	Complete polynomial of degree ρ
$K(x_i, x_j) = \dfrac{(x_i^T x_j + 1)^\rho}{\sqrt{(x_i^T x_j) - (y_i^T y_j)}}$	Normalized polynomial kernel of degree ρ
$K(x_i, x_j) = \exp\left[\left\| x_i - x_j \right\|^2 / 2\sigma^2 \right]$	Gaussian (RBF) with parameters σ which control the half-width of the curve fitting peak

Similar to other multivariate statistical models, the performances of SVM for regression depend on the combination of several parameters. They are capacity parameter C, ε of ε-insensitive loss function, the kernel type K and its corresponding parameters. C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the training error. In order to make the learning process stable, a large value should be set up for C (e.g., C = 100). The optimal value for ε depends on the type of noise present in the data, which is usually unknown. Even if enough knowledge of the noise is available to select an optimal value for ε, there is the practical consideration of the number of resulting support vectors. ε-insensitivity prevents the entire training set meeting boundary conditions, and so allows for the possibility of sparsity in the dual formulations solution. Therefore, choosing the appropriate value of ε is critical in theory.

Since in this study the nonlinear SVM is applied, it would be necessary to select a suitable kernel function. The obtained results of previously published researches indicate the Gaussian radial basis function has superior efficiency than other kernel functions [30]. As shown in Table 7, the form of the Gaussian kernel is as follow:

$$K(x_i, x_j) = e^{-\left| x_i - x_j \right|^2 / 2\sigma^2} \qquad (13)$$

In addition, as σ is a constant parameter of the kernel and can either control the amplitude of the Gaussian function and the generalization ability of SVM, we have to optimize σ and find the optimal one. In order to find the optimum values of two parameters (σ and ε) and prohibit the over-fitting of the model, the data set was separated into a training set (80% of available data for each borehole), a test set of 20% and the leave-one-out cross-validation of the whole training set was performed. The leave-one-out (LOO) procedure consists of removing one example from the training set, constructing the decision function on the basis only of the remaining training data and

then testing on the removed example [31]. In this fashion, one tests all examples of the training data and measures the fraction of errors over the total number of training examples. The root mean square error (RMS) was used as an error function to evaluate the quality of model.

To obtain the optimal value of σ, the SVM with different σ were trained, the σ varying from 0.01 to 0.9, every 0.01. According to the generalization ability of the model based on the LOO cross-validation for the training set, we calculated the RMS on different σ, in order to determine the optimal one. The optimal σ was found as 0.6. In order to find an optimal ε, the RMS on different ε was calculated. The optimal ε was found as 0.11. From the above discussion, the σ, ε and C were fixed to 0.6, 0.11 and 100, respectively. Figure 8 is a schematic diagram showing the construction of the SVM.

The best and the worst results of 6 iterations for training of the network are presented in Table 8.

Table 8. Error values for the best and the worst results.

R_{train}	R_{test}	RMS_{train}	RMS_{test}
0.95	0.89	0.05	0.07
0.98	0.96	0.08	0.16

The results of the training are presented in Figure 9.As shown in Figure 9, the correlation coefficient of training data is 0.96, implying the proper performance of SVM. The abovementioned support vector machine was used to classify the test data. The results are shown in Figure 10. During testing, a correlation coefficient of greater than 0.95 was generally obtained. This implies that the pore size values in the test data were practically well predicted using SVM. This is evident in Figure 11 with the superior performance of the trained SVM remarkably demonstrated for a set of synthetic test data.
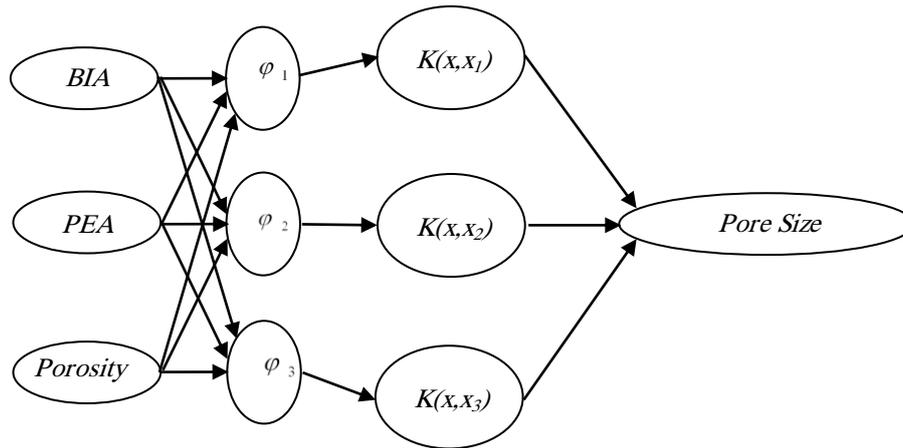
Figure 8. schematic diagram of construction of the SVM.

3. Application of the proposed SVM method for data of Well 2

The results of previous chapters indicate that SVR method could predict the values of pores size reliably for synthetic data. One of the major preferences of SVMs to ANNs is their superior performance facing little amounts of training data [32]. Because of the little number of synthetic data in training process (60 data) and the differences between synthetic and real data (complicacy and existence of noise in real data), the authors decided to use trained SVR for real porosity values prediction in this well.

To test the capability of the proposed methodology, data acquired from logs and 3D seismic measurements in well 2 were used. The measurements were Porosity, Instantaneous Amplitude andAsymmetry (see Table 9).

Figure 12 shows the pore size values from the well log measurements (4th column in Table 9) as real values and those obtained from support vector regression machine. For this test, the already trained SVR of synthetic data was used to make the prediction. The values of correlation coefficient and root mean square error of the prediction are 0.97 and 0.04 respectively. Figure 12 exhibits a considerable coincidence between the results of support vector machine approach and that of well measurements.

5. Conclusions

Measurement noise and nonlinear relationship between seismic data and pore size quantities cause difficulties in performing seismic data interpretation reliably. Consequently, other viable methods of prediction, such as the one proposed in this paper, may be deemed necessary in realistic cases. We successfully implemented and tested an artificially intelligent computational agent (a back-propagating neural network) and support vector regression machine (SVR) to consider the unknown nonlinear relationships between system variables in our prediction problem (foreseeing the pore size values) for synthetic data. Our approach uses instantaneous amplitude and asymmetry as input system variables. The ANN and SVR seek the relationship between these input variables adaptively and strive to a desirable output which is, in our case, the values of pore size.

Synthetic data showed that only SVR could train itself very well with practically complete correlation between real pore size values and the predicted ones (correlation coefficient R of almost one). This method also exhibited a remarkable capability in estimating the unknown zones (test data).

Since the number of synthetic data in training process is limited and due to the differences between synthetic and real data (complicacy and existence of noise in real data), it seems that SVR can be used in predicting values of pore size for real data of well 2. Applying the machine to well 2 case, while showing acceptable precision in prediction pore size, proved the performance of the machine. The SVR did predict the pore size values in well 2 reliably.

In this study, there was an access only to the data of well 2. To generalize the results of the abovementioned procedure, it is suggested to obtain sufficient reliable data samples from more wells in specific oil field and augmenting the training of the support vector machine with the new data. We speculate this would enhance the capability of the machine to be used for the similar reservoirs.
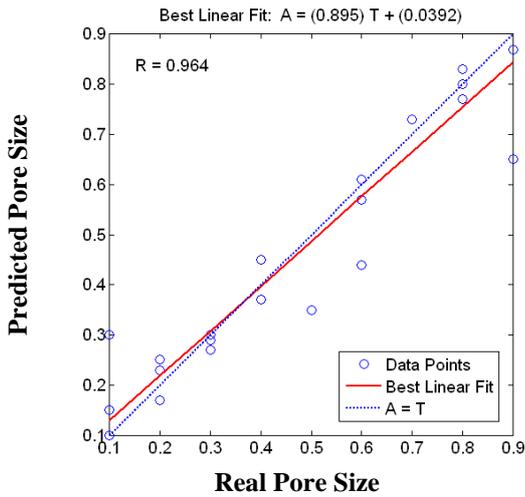
Figure 9. Correlation coefficient for train data.
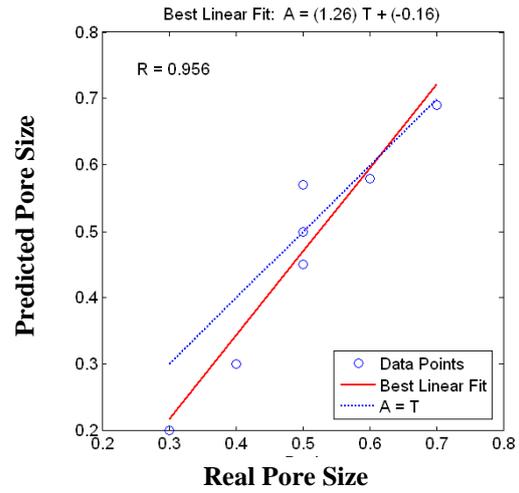
Figure 10. Correlation coefficient for test data.
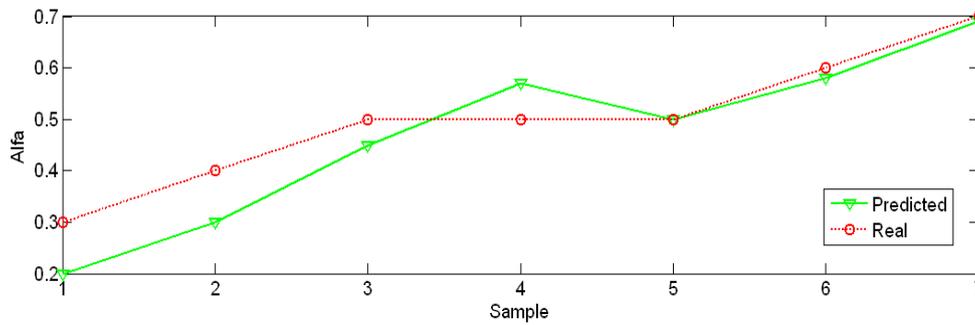
Figure 11. Predicted results for a set of test data.

Table 9. Data of well 2 used in trained SVR

Porosity	BIA	PEA	alfa
0.97206	0.42171	0.27902	0.915
0.47079	0.3639	1	0.855
0.65635	0.64461	1	0.86
0.64566	0.94775	-0.07698	0.855
0.91746	0.81664	-0.07698	0.89
0.92458	0.60243	-4.97	0.9
0.85231	0.6088	0.015595	0.915
0.87557	0.84257	0.015595	0.915
1	1	0.015595	0.915
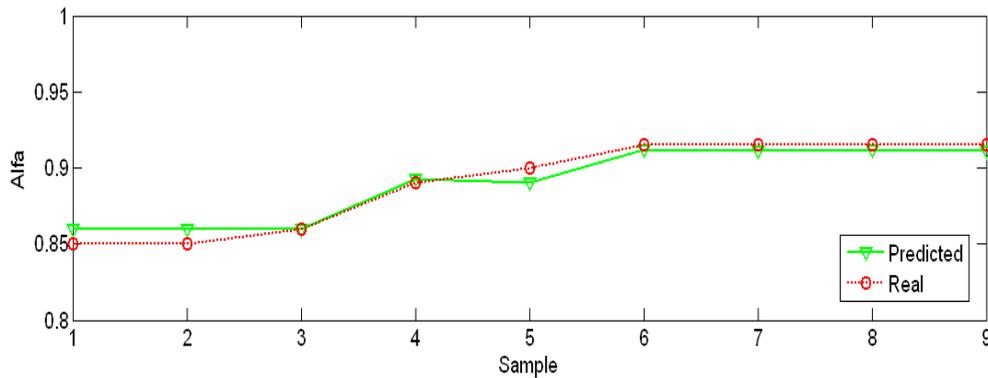
Figure 12. Predicted results for a set of data in well 2.

12

References

[1]. Okabe, H. and Blunt, M. J. (2005). Pore Space Reconstruction Using Multiple-Point Statistics, Journal of Petroleum Science and Engineering, 46: 121-137.

[2]. Al-Kharusi, A. S. and Blunt, M. J. (2007). Network Extraction from Sandstone and Carbonate Pore Space Images, Journal of Petroleum Science and Engineering, 56: 219-231.

[3]. Padhy, G. S.,Lemaire, C., Amirtharaj, E. S. and Ioannidis, M. A. (2007). Pore Size Distribution in Multiscale Porous Media as Revealed by DDIF-NMR, Mercury Porosimetry and Statistical Image Analysis, Colloids and Surfaces, 300: 222-234.

[4]. Schoenfelder, W., Glase, H. R., Mitreiter, I. and Stallmach, F. (2008). Two-Dimensional NMR Relaxometry Study of Pore Space Characteristics of Carbonate Rocks from a Permian Aquifer, Journal of Applied Geophysics, 65: 21-29.

[5]. Hollis, C., Vahrenkamp, V., Tull, S., Mookerjee, A., Taberner, C. and Huang, Y. (2010). Pore System Characterisation in Heterogeneous Carbonates: An Alternative Approach to Widely-Used Rock-Typing Methodologies, Marine and Petroleum Geology, 27: 772-793.

[6]. Alger, R. P., Luffel, D. L. and Truman, R. B. (1989).New Unified Method of Integrating Core Capillary Pressure Data with Well Logs, SPE, 16793.

[7]. Obeida, T. A., Al-Mehairi, Y. S., and Suryanarayana, K. (2005).Calculations of Fluid Saturations from Log-Derived J-Functions in Giant Complex Middle-East Carbonate Reservoir, E-Journal of Petrophysics, 1(1).

[8]. Lawrence, P. (1998). Seismic Attributes in the Characterization of Small-Scale Reservoir Faults in Abqaiq Field, The Leading Edge, April.

[9]. Siripitayananon, P., Chen, H. C., and Hart B. S. (2001). A New Technique for Lithofacies Prediction: Back-Propagation Neural Network, Association for Computing Machinery, Inc.

[10]. Reddy, G. V., Kumar, P., Naik, S. B. R., Murthy, G. N., Vandana, Kaul, S. K., and Bhowmick, P. K. (2004). Reservoir Characterization Using Multi Seismic Attributes in B-172 Area of Heera-Panna- Bassein Block of Bombay Offshore Basin, India, 5th Conference & Exposition on Petroleum Geophysics, Hyderabad-2004, India, 716-721.

[11]. Raeisi, M. (2008).Application of Artificial Neural Networks in Facies Change Prediction for Hydrocarbon Reservoirs, M.Sc. Thesis, Shahrood University of Technology.

[12]. Edalat, A., and Siyahkoohi, H (2007). Using Seismic Facies in Characterizing One of the Iranian Hydrocarbon Reservoirs, Iranian Geophysics Journal, 1(1): 37-49.

[13]. Malvic, T., and Prskalo, S. (2007). Some Benefits of the Neural Approach in Porosity Prediction (Case Study from Benicanci Field), NAFTA 58 (9): 455-461.

[14]. Zhou, Y. J., Tao, J. Q., Guo, Y. B., Zhang, X. H., and Qiang, M. (2009). Rock Physics Based Prestack Seismic Reservoir Characterization - Application to Thin Bed, 71th EAGE Conference & Exhibition – Amsterdam, TheNetherlands.

[15]. Ellis, D. V., and Singer, J. M. (2007). Well Logging for Earth Scientists, Chapter 4, Springer.

[16]. Lucia, F. J. (2007). Carbonate Reservoir Characterization, An Integrated Approach, Second Edition, Springer.

[17]. Baechle, G. T., Colpaert, A., Eberli, G. P. and Weger, R. J. (2007). Modeling Carbonate Pore Structure Effects on Velocity Using a Dual Porosity DEM Theory, SEG 76th Annual Meeting, San Antonio, Texas.

[18]. Van Golf-Racht, T. D. (1982).Fundamentals of fractured Reservoir Engineering, Elsevier.

[19]. An, P. and Moon, W. M. (1993). Reservoir Characterization Using Feedforward Neural Networks, 63rd Ann. Internat.Mtg., Soc. Expl.Geophys., Expanded Abstracts, 258-262.

[20]. An, P. (1994). The Effect of Random Noise in Lateral Reservoir Characterization Using Feedforward Neural Networks, 64th Ann. Internat.Mtg., Soc. Expl.Geophys., Expanded Abstracts, 787-790.

[21]. Alimoradi, A., Moradzadeh, A. and Bakhtiari, M. R. (2011). Modifying Gassmann Equation to Determine the Effect of Pores Sizes in Carbonate Reservoirs Characterization, Petroleum Science (Springer), Under Review.

[22]. Demuth, H. and Beale, M. (2002).Neural Network Toolbox for Use with MATLAB, Version 3.0, 742p.

[23]. Hagan, M. T., Demuth, H. B., and Beale M. (1996).Neural Network Design, PWS Publishing Company, Boston, MA.

[24]. Yu Hen, H., Jenq-Neng, H. (2002).Handbook of Neural Network Signal Processing, CRC PRESS.

[25]. Quang-Anh, T., Xing, L. and Haixin, D. (2005). Efficient Performance Estimate for One-Class Support Vector Mmachine, Pattern Recognition Letters, 26: 1174–1182.

[26]. Liu, H., Wen, S., Li, W., Xu, C. and Hu, C. (2009). Study on Identification of Oil/Gas and Water Zones in Geological Logging Base on Support-Vector Machine, Fuzzy Information and Engineering, 2, AISC 62: 849–857.

[27]. Lia, Q., Licheng, J. and Yingjuan, H. (2007). Adaptive Simplification of Solution for Support Vector Machine, Pattern Recognition, 40: 972 – 980.

[28]. Chih-Hung, W., Gwo-Hshiung, T. andRong-Ho, L. (2009). A Novel Hybrid Genetic Algorithm for Kernel Function and Parameter Optimization in Iupport Vector Regression, Expert Systems with Applications, 36: 4725–4735.

[29]. Scholkopf, B., Smola, A. J. and Muller, K. R. (1998). Nonlinear Component Analysis as a Kernel Eigenvalues Problem, Neural Computer, 10, 1299–1319.

[30]. Wang, W.J., Xu, Z. B. Lu, W. Z. and Zhang, X. Y. (2003). Determination of the Spread Parameter in the Gaussian Kernel for Classification and Regression, Neurocomputing, 55, 643-663.

[31]. Liu, H., Yao, X., Zhang, R., Liu, M., Hu, Z. and Fan, B. (2006). The Accurate QSPR Models to Predict the Bioconcentration Factors of Nonionic Organic Compounds Based on the Heuristic Method and Support Vector Machine, Chemosphere, 63: 722–733.

[32]. Duda, R. A., Hart, P. E., and Stork, D. G. (2002). Pattern Classification, Springer.