JOURNAL OF PETROLEUM GEOMECHANICS
(JPG)

Original article

Optimizing Deep Learning Models for Shear Wave Velocity Estimation Utilizing Petrophysical Logs: A Case Study on an Oil Reservoir in Southern Iran

Fatemeh Eil Saadatmand [1], Andisheh Alimoradi [1*], Mirhassan Moosavi [2], Mohammad Mehrad [3], Mohammad Ali Davari [1], Parisa Rezakhani [1]

1- Department of Mining Engineering, Imam Khomeini International University, Qazvin, Iran
2- Department of petroleum Engineering, Masjed-Soleiman Branch, Islamic Abad University, Khuzestan, Iran
3- Department of Mining, Petroleum and Geophysics, Shahrood University of Technology, Semnan, Iran

Received: 17 January 2024; Accepted: 18 February 2024

DOI: 10.22107/JPG.2024.436270.1225

Keywords

Shear wave velocity,
Petrophysical logs,
Deep learning,
Multi-layer perceptron,
Particle swarm optimization,
Social ski drive,
Sarvak formation

Abstract

Full identification and understanding of hydrocarbon reservoirs depends on knowing the mechanical properties. One of the main parameters that indicates mechanical properties is shear wave velocity. Bipolar sound recorder is among thee best tools for measuring shear wave velocity. This tool is not very popular due to the high costs of driving in the well despite the high accuracy. Shear wave velocity estimation methods include three main branches of experimental methods, regression and the use of machine learning algorithms or in other words artificial neural networks. The studied formation in this research is Sarvak in one of the oil fields in the south of Iran. The input data of the estimator model is the usual petrophysical logs that are driven and measured in many wells, and the output data is obtained from the DSI tool. In this research, data are pre-processed by removing noise effects. Then, to improve the estimation effectiveness, data with a high correlation coefficient are selected as input data. After that, shear wave velocity is estimated from petrophysical data with three types of multi-layer perceptron (MLP), multi-layer perceptron optimized by particle swarm optimization (MLP-PSO), and the introduction of a relatively new method of multi-layer perceptron-social ski drive (MLP-SSD). To compare the efficiency of the neural network method, two traditional experimental and regression methods used. The validation results show the better performance of the MLP-SSD method.

1. Introduction

The mechanical and geophysical properties of rock and formations are very important in predicting strategies and engineering planning in the oil industry. These properties play an essential role in the selection of casing pipe, the stability of the well, determining the lithology of the reservoir, detecting the fluid in the holes, and measuring the pore pressure. Direct and indirect methods are used to determine these properties, and recently indirect methods have been more widely used due to their lower cost and the use of well petrophysical information. The most important parameters for determining the mechanical properties are the compressional and shear wave velocity which by using petrophysical data and estimating these parameters, it is possible to obtain a more reliable answer for the needs of the oil industry.

Neural networks, by simulating human learning and the training and testing system, can calculate linear and non-linear equations between the user's desired inputs and the desired output and have provided acceptable results in the field of petroleum engineering and related sciences [1].

Russell proposed the multi-layer neural

*** Corresponding Author:** alimoradi@eng.ikiu.ac.ir

network in geophysics and geological sciences to evaluate reservoir properties so that other researchers can use this method in earth sciences [2].

Using the genetic algorithm, Moatazedian and colleagues predicted the compressional and shear wave velocity in the carbonate reservoir located in Abuzar and Handijan fields and compared it with the available real values [3].

Alimoradi et al research focuses on improving the prediction of shear wave velocity in subsurface strata. This study uses geotechnical investigations, specifically the spectral analysis of surface waves (SASW) approach, and compares them to down hole test (DHT) data acquired from wells in the same location. They use a trained artificial neural network (ANN) to find nonlinear correlations between SASW results and DHT values. The results show that a well-trained neural network can accurately predict shear wave velocity between wells [4].

In 2013, Hamza estimated the average shear wave velocity for seismographic measurements of the California basin by using a combination of genetic algorithm and neural network and concluded that the hybrid correlation coefficient of neural network and genetic algorithm is better than multiple regression [5].

Also, in the same year, Zoveidavianpoor and colleagues predicted the shear wave velocity from petrophysical charts along a well in a carbonate field using the adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression [6].

Maleki has used the back-propagation (BP) method and support vector regression (SVR) in an article. Regarding this research, he has used the parameters related to resistivity tools, density, gamma-ray, caliper, and dipole sonic imager and proposed the limited effectiveness of the BP compared to the SVR method [7].

Singh has used the BP neural network to compare the function of shear wave estimation in horizontal and vertical wells and concluded that this network had a more appropriate estimation in vertical wells [8].

Asoodeh has used the alternative condition expectation (ACE) method in a gaseous carbonate reservoir to find the best relationship between input and output and to transform the data space into a larger space. In this research, the efficiency of converting the data space into a larger space has been remarkable [9].

Nourafkan and Kadkhodai Ilkhchi used the ant colony method which is based on finding the best food path for ants. They also used the fuzzy logic method for neural network input in the form of a neural network with an observer. GR, Depth, NPHI, PHOB, and PEF were the neural network's inputs from the Cheshmeh-Khosh oil field [10].

Mehrgini calculated the shear wave in the Mansouri formation using four types of neural networks. These four networks are ELMAN, ELMAN-PSO, MLP, and MLP-PSO. In the comparison between these methods, Mehrghini concluded that the ELMAN method had the best function in estimating the shear wave velocity [11].

Salaheddin Elkatatny and colleagues estimated shear wave velocity with three neural network techniques, ANFIS algorithm, and SVR, and among the three mentioned methods, the artificial network techniques provided more acceptable results [12].

Anemangely research about shear wave slowness estimation suggested that to achieve more efficient results, first, it is necessary to pre-process the data, then, a non-dominated sorting algorithm (NSGA-II) should be used along with MLP neural network to select the best input parameters, and then, by using the selected input parameters and two algorithms: ANFIS-GA and ANFIS-PSO estimated shear wave slowness. The results of using these two algorithms showed that the ANFIS-PSO model has high accuracy in estimation [13].

Learning based on the particle swarm optimization method was used in the article by Yang et al. This learning increases the improvement of the target parameter estimation. In this research, the combination of this method with the multi-layer perceptron neural network has been used [14].

In relatively new research, the social-ski drive algorithm has been used. Tharwat and Gabel has mentioned that this algorithm has acceptable stability, because it is associated with modeling the movement of skiers in a certain interval and the limited fluctuation of values towards the goal. The combination of this algorithm with the multi-layer perceptron neural network has been used in this research [15].

Zhang et al proposed a 1D convolutional neural network (1D-CNN) with different well reports to predict S-wave velocity. By combining reports from diverse wells, 1D-CNN outperforms

traditional machine learning methods, thereby increasing exploration accuracy and increasing oil and gas production in other fields [16].

Rajabi et al investigated the application of hybrid machine learning (HML) and deep learning (DL) algorithms for VS prediction using data from the Marun oil field in Iran. Comparison with HML algorithms and empirical equations shows that the DL model is slightly better than HML and significantly better than empirical equations. The findings showed potential improvements in the understanding of well instability, casing collapse problems, and increasing the stability of oil reservoirs through accurate VS predictions [17].

In this article, first, all available common log data were reviewed and the data that had the highest correlation with the shear wave velocity data were selected. Corrections and preparation conditions were made to these data. Then, the value of the shear wave velocity was first obtained from the empirical equations. Then it was calculated using the regression method and finally using three types of multi-layer perceptron (MLP), multi-layer perceptron-particle swarm optimization (MLP-PSO), and introducing a relatively new method of multi-layer perceptron-social ski drive (MLP-SSD) to estimate shear wave velocity from petrophysical data.

2. Theory and Methodology

2.1. Methodology of multi-layer perceptron

An artificial neural network consists of neurons connected in different layers that send information to each other [18–20]. In general, the role of neurons in this network is information processing, which is done by a mathematical processor called the activation function. The feedback multi-layer neural network is one of the most widely used neural network architectures, also known as multi-layer perceptron networks [21,22]. This network has the following specifications:

1- Network processors consist of several different layers.
2- As its name suggests, there must be at least 2 layers in the network.
3- The neurons of each layer are only allowed to receive signals from the processors of their previous layer, and the output signal is also applied to the next neurons.
4- The inputs of the network are the influencing parameters of the output, so the number of input and output layer nodes is actually known from the beginning of using the network.

The number of hidden layer nodes and the number of hidden layers is determined through trial and error. In fact, the number of nodes and hidden layers is determined when the network gives the best response. There are two major destructive and constructive methods in this type of network [23]. In the first method, the network starts with a large number of nodes and layers and reaches the optimal structure by removing additional units and related connections. The second method starts with a simple network with few hidden nodes, and by increasing the number of nodes, it moves forward until the minimum error. In most cases, one or two hidden layers are enough for proper information processing. The number of nodes in the hidden layer is usually half to two or three times the number of nodes in the input layer of the network. The initial weights are often considered random, although it is possible to give more weight to more effective inputs by knowing the input parameters from the beginning [24].

2.2. Methodology of Particle Swarm Optimization (PSO)

The particle swarm optimization (PSO) method is a type of optimization method that can be used to solve problems whose response is a point or a surface in n-dimensional space [14]. The PSO algorithm based on a type of social behavior was first introduced by Kennedy and Eberhart.The creators of this algorithm believed that information sharing between members of a society plays a fundamental role in the evolution of that society [25].

A simple PSO algorithm starts by considering a population that contains possible responses. Then, each particle (response) moves continuously and by repeating the process in the search space, and in this interval, it is attracted to the best fitting location found by itself (local optimum) or particles located in its neighborhood (global optimum). In fact, in a PSO algorithm, instead of determining the way a particle moves according to the common rules in the genetic algorithm, based on the values of a function, the type of movement is based on the experience of

the particle's previous movements and the experience of the movements of neighboring particles. Applying chaos theory can also help to adjust the basic constants of the algorithm as well as its related random variables [26].

In the primary PSO algorithm, the velocity of each particle was determined according to an equation and based on the previous velocity of the particle, the direction in which each particle reached the best and most optimal location, and also the direction of the best population in reaching the optimal location. After several experiments and mathematical simulations with this algorithm, a researcher suggested adding an inertia parameter to the velocity calculation equation so that trade-offs between global exploration and local extraction of each particle can be controlled [22,26].

The evolution of the PSO velocity equation is as follows:

If we consider g^* as the global optimum, and x_i^* as the local optimum, the speed of each particle is obtained from the following equation:

$$v_i^{t+1} = v_i^t + \alpha\varepsilon_1[g^* - x_i^t] + \beta\varepsilon_2[x_i^* - x_i^t] \quad (1)$$

In this equation, ε_1 and ε_2 are two random vectors whose input values are all between 0 and 1.

It should be noted that the Hadamard product of two matrices u and v is used in this equation, not the normal product of two matrices, that is:

$$[u \odot v]_{ij} = u_{ij} v_{ij} \quad (2)$$

The α and β parameters are the learning parameters or acceleration constants of the algorithm, which usually have a value close to 2:

$$\alpha \approx \beta \approx 2 \quad (3)$$

The initial location of the particles should be considered according to a continuous uniform distribution so that the particles can cover all the environment or at least most of it. This is a very important thing, especially in multi-mode problems. The initial velocity of the particles is considered to be zero; then their next location is obtained according to the following equation:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (4)$$

Finally, as it was mentioned, the object of the velocity equation underwent a functional change by adding the inertia index. This equation is:

$$v_i^{t+1} = \theta v_i^t + \alpha\varepsilon_1 \odot [g^* - x_i^t] + \beta\varepsilon_2 \odot [x_i^* - x_i^t] \quad (5)$$

In this equation, θ has a value between 0 and 1. In the simplest form of the network, the inertia function can be considered as a numerical constant and approximately $\theta \approx 0.5 - 0.9$. In this case, the algorithm performs the convergence process at a much higher speed.

The main setting parameters of this algorithm are the number of particles and the number of rounds of repetition of the algorithm, which are determined according to the number of input data. Other parameters, such as the degree of excitation and the impact of particles on each other, can also be changed, and it is better to use them in the default mode [27]

2.3. Methodology of Social Ski-drive (SSD)

The social ski-drive (SSD) algorithm is a new optimization method that is based on the movement of skiers on the snow, or in other words, the track of their skis on the snow and downhill. This method was used in an article in 2020 by [15] in an article combining the SVR method and compared with the SVR-PSO hybrid method. Since skiers use a spiral path for the best path to reach the goal, this optimization method is introduced. This method is a type of optimization method that can be used to solve problems whose response is a point or a surface in the n-dimensional space.

Therefore, as mentioned, the movement path of the particles is like sinusoidal equations and the particles do not go to infinity like the PSO algorithm and they move between certain values in search of the best response. Fig. 1 describes how two particles move in the search space. The particle A in the initial position in a non-linear path to reach the value of M first in the position A' and then in the position A" and finally approaches one of the three best responses. This algorithm gives the ability of better exploration and diverse directions to particles in a directed state [15].

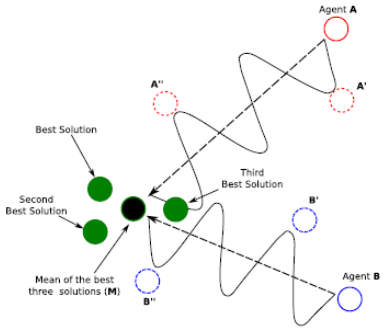

Fig. 1. General schematic of the particle movements of the SSD algorithm towards the top three responses

In this algorithm, the initial position of the particles and their initial velocity are randomly selected, and the number of particles and the maximum number of repetitions are determined by the user.

An SSD algorithm starts by considering a population that contains possible responses. First, it determines intervals for the data and divides the particles into parts. Then, each particle (response) moves continuously and by repeating the process in the search space and is attracted to the best fitting location found by itself (local optimum) or particles located in its neighborhood (global optimum). In fact, in an SSD algorithm, the way a particle moves according to the common rules in the genetic algorithm is not determined in terms of the values of a function, but the type of movement is based on the experience of the previous movements of the particle and the experience of the movements of neighboring particles. Then, each group of particles reaches the most optimal response and finally, it is averaged among these three responses and published as the best response [15].

In the primary SSD algorithm, the velocity of each particle was determined according to the equation (8) and based on the previous velocity of the particle, the direction in which each particle reached the best and most optimal location, and also the direction of the best population in reaching the optimal location. Various factors are effective in this algorithm, they will be explained in more detail later [15].

- Position of particles in n-dimensional space ($X_i \in R^n$): This value is useful for calculating the movement function of particles in n-dimensional space to estimate subsequent movements according to the initial position.
- Best previous position (P_i): The best values for all particles use optimization functions. Also, the best values are compared with the current values of the particles, and finally, the best position is saved. This technique of the SSD algorithm is something similar to the PSO algorithm.
- Average of the best response (M_i): In this algorithm, according to equation (6), the particles move to the three best responses and it is average between these three responses.

$$M = \frac{X_\alpha + X_\beta + X_\gamma}{3} \quad (6)$$

In the mentioned equation, the values of X_α, X_β, and X_γ are the three best responses.

- Particle velocity (V_i): The position of each particle is updated with the sum of the previous position and this parameter, according to equation (7).

$$X_i^{t+1} = X_i^t + V_i^t \quad (7)$$

In the above equation, the velocity of the particles is expressed by the following equation:

$$V_i^{t+1} = \begin{cases} c\sin(r_1)\,(P_i^t - X_i^t) + \sin(r_1)\,(M_i^t - X_i^t) \\ c\cos(r_1)\,(P_i^t - X_i^t) + \cos(r_1)\,(M_i^t - X_i^t) \end{cases} \quad (8)$$

In the above equation, the values of r_1 and r_2 are randomly selected from the interval [0 1]. The value of P_i is the best response of particle i and M_i is the average response among all particles.

The c value plays a balancing role between the search and exploration space and is calculated from the following equation:

$$c^{t+1} = \alpha c^t \quad (9)$$

The value of t is the current repetition number of the particle and the value of α is between zero and one to reduce the value of c. When the value of c tends to zero, in fact, the value of t tends to the value of t_{max}. The maximum value of t is chosen by the user.

Therefore, as seen in equation (8), the movement path of the particles does not go to infinity like the PSO algorithm, and they move between certain values in search of the best response. As a result, this procedure makes the stability of this algorithm more reliable than PSO [15]. In this algorithm, the initial position of the particles and their initial velocity are randomly selected, and the number of particles and the maximum number of repetitions are determined by the user. Fig. 2 briefly describes the general procedure of this algorithm [15].

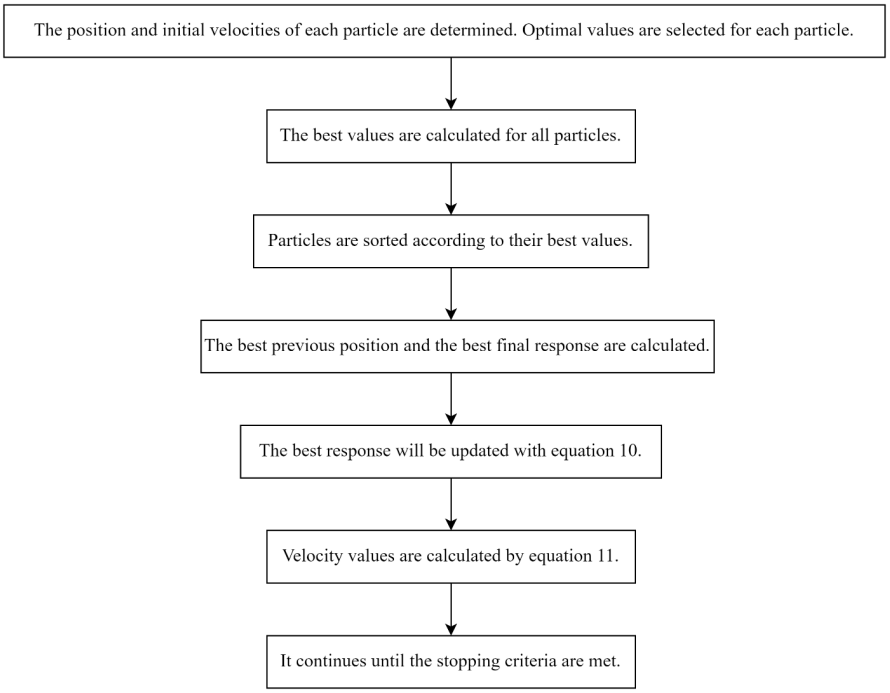

Fig. 2. The general trend of the SSD algorithm

2.4. MLP-PSO

The two feedback multilayer networks and particle motion swarm were explained in the previous section. In this section, the procedure of one of the main methods of this research is described.

In other words, when a hybrid network is used, a combination of two types of algorithms is used for optimization. Optimization algorithms adjust the weights and biases in such a way that the amount of estimation error is minimized. That is, the objective function in optimization algorithms is the mean square of the error. So first, the number of neurons and hidden layers of the multilayer perceptron neural network is determined, and then the particle swarm learning algorithm is used to determine the optimal values of weights and bias with the number of particles used by the user and the maximum repetition values in the particle training algorithm. The best response is selected to estimate the shear wave velocity.

2.5. MLP-SDD

The working procedure of this hybrid network is the same as the MLP-PSO hybrid network that was explained in the previous section. In this algorithm, by optimizing the weights and biases, the values of neurons and intermediate layers are selected by the user for the objective function, i.e. the mean square of the errors, and then the SSD algorithm is used to learn the network, and the SSD algorithm procedure is used to learn the network and by determining the values of the number of particles and the maximum number of repetitions to achieve the best response, the error values and correlation coefficients between the estimated and measured shear wave velocity are obtained.

To estimate the shear wave velocity, the data is entered into three main networks, namely MLP, MLP-PSO, and MLP-SSD. The important point is that for these three networks, the values of intermediate layers, the number of neurons, and the percentage of training and testing data must be the same. Fig. 3 explains the steps schematically in a completely summarized way.

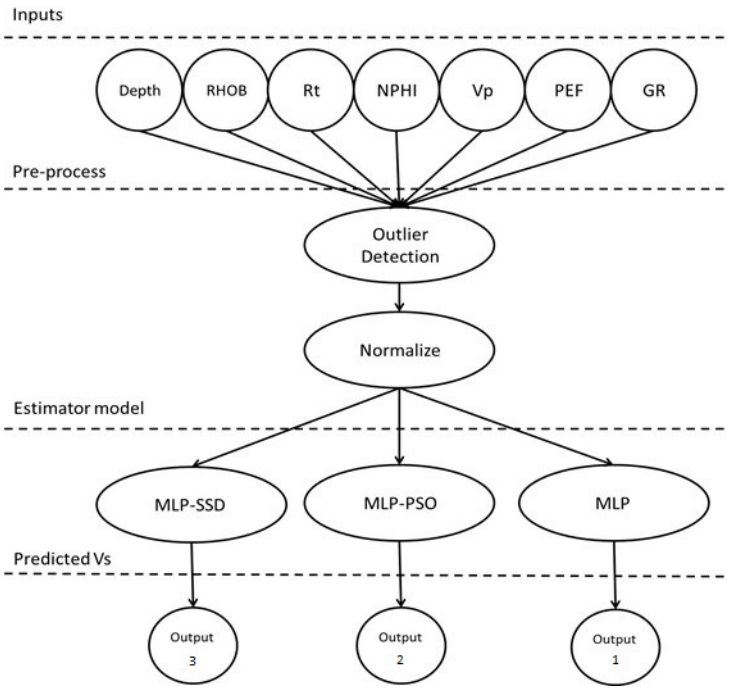

Fig. 3. General procedure

3. Data processing and analysis and geological background

Abteymour oil field is located 25 kilometers from the southwest of the Ahvaz field, between the Sosangard and Mansouri fields. This field, along with Susangerd and Mansouri oil fields, is located on a single structural height that has a northwest-southeast direction and is located in the west of Khuzestan Fig. 4 [28].

Abteymour oil field has no outcrops on the surface of the earth and its structural form has been determined by geophysical operations. The structure of the Abteymour oil field is a gentle and symmetrical anticline with a length of 23 kilometers and an average width of 5.6 kilometers on the horizon of Ilam. (The contact levels of water and initial oil are considered as the last parallel closed curve). The maximum slope of the northern and southern ridges on the horizon of Ilam is 5.5 and 6 degrees, respectively, and in the northwestern and southeastern capes, it is 3 and 2.2 degrees[3].

On the horizon of Sarvak, the maximum slope of the northern ridges is 6 degrees and the southern ridges are 8 degrees and in the northwest and southeast capes, it is 3.2 and 2.2 degrees, respectively.

The studied formation in this field is the Sarvak Formation, which contains 5.821 meters of limestone in the sample section, and its petrology includes the following three parts [3]:

- 5.254 meters at the base: dark gray limestones with nodular, fine-grained, clay layering and containing traces of small ammonites with thin layers of dark gray marl interlayered
- 534 meters in the middle: It consists of massive, rough, chalky light brown limestones along with pieces of Rudists, which 7.109 meters from the base of this part has brown to red siliceous nodules and in the middle part of this part cross classification has been observed.
- 43 meters at the top: It consists of very thick to thick-layered limestones, with irregular weathering and stained with iron oxides. The lower limit of this formation in the sample section is gradual and concordant with the Kazhdomi formation, and the upper limit of the Sarvak limestone formation is mixed with the marls and shales of the Gurpi formation.

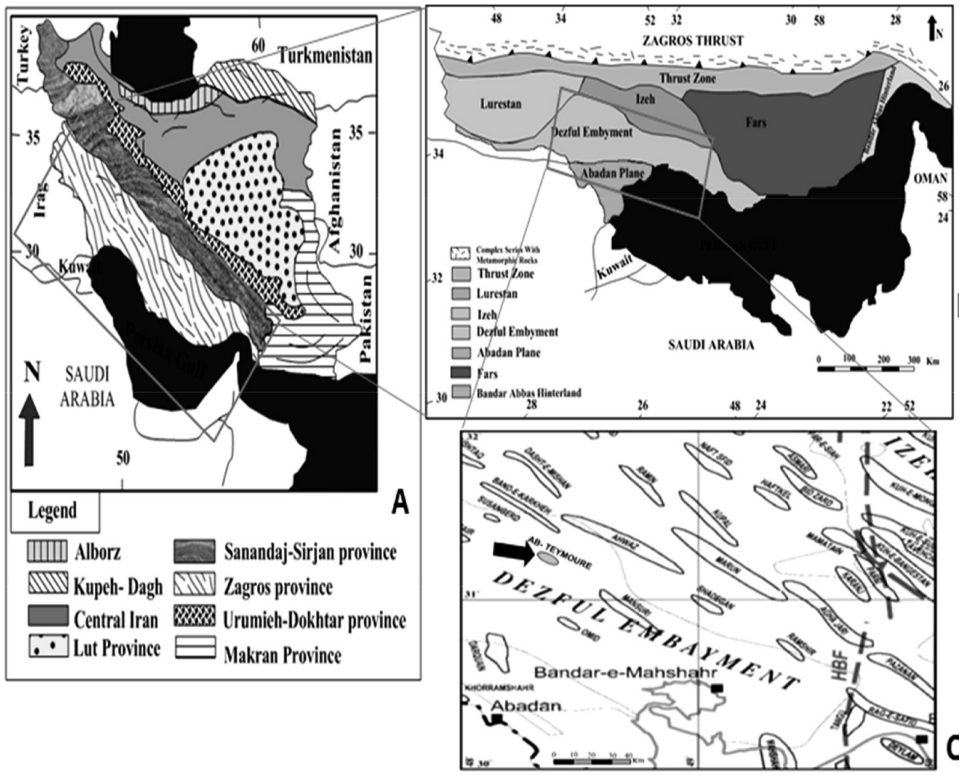

Fig. 4. Geographical location of Abteymour oil field in southwest Iran[29]

3.1. Data preparation

There are 1042 available points that are located from the depth of 3315 to 3473.65 meters of well A and the data of depth, caliper, photoelectric factor (PEF), gamma ray radiation rate, resistivity, neutron porosity (NPHI), bulk density (RHOB), compressional wave velocity and shear wave velocity (which is also considered as an output parameter) are used by the DSI probe that is driven separately in the well and the collected data are placed in the equations related to the calculations of this tool and the values have been taken as compressional wave transit time (DTCO) and shear wave transit time (DTSM). A section of the data of well A are showed in Table 1.

Table 1. A section of the data bank extracted from well A

Row	Depth	Caliper	PEF	GR	RT	NPHI	RHOB	Vp	Vs
1	3379.318	7.1989	4.7199	21.4273	38.5558	0.3248	2.5663	5.233732	2.771526
2	3379.47	7.1989	4.8095	20.3052	120.4304	0.3248	2.5957	5.303517	2.745132
3	3379.622	7.1929	4.8789	21.3181	155.5504	0.3248	2.6085	5.136519	2.72085
4	3379.775	7.1567	4.9484	22.42	150.3951	0.3248	2.6215	5.081414	2.604754
5	3379.927	7.1506	4.9565	26.659	72.2965	0.3248	2.6116	5.019928	2.59952
6	3380.08	7.1567	4.9773	28.111	100.4453	0.3248	2.591	4.919589	2.566322
7	3380.232	7.1397	4.9334	29.2803	99.7479	0.3248	2.5728	5.100658	2.527824
8	3380.385	7.1325	4.8741	27.81	100.6385	0.3248	2.5663	5.277365	2.556026
9	3380.537	7.1265	4.8192	25.4117	89.8469	0.3248	2.5786	5.291621	2.591991
10	3380.689	7.1365	4.7784	25.0227	77.8917	0.3248	2.6019	5.214838	2.701062
11	3380.842	7.1144	4.7894	24.18	56.7269	0.3248	2.6194	5.399009	2.757525
12	3380.994	7.1083	4.7754	25.3495	45.7972	0.3248	2.6365	5.25801	2.672866
13	3381.147	7.1083	4.7987	24.91	40.4876	0.3248	2.6414	5.247193	2.719399
14	3381.299	7.1023	4.8305	26.1886	39.9267	0.328	2.6401	5.174451	2.758518
15	3381.451	7.1144	4.8593	22.5894	39.5016	0.3312	2.6272	5.139932	2.753519

Fig. 5 shows the matrix of correlation coefficients of all petrophysical data and shear wave velocity. The slope of the lines that are close to one or minus one indicates a high correlation between the independent and dependent variables. Values close to one indicate a positive correlation. This means that as the independent variable increases, the dependent variable also increases. On the other hand, values close to minus one mean that as the independent variable increases, the dependent variable decreases. The efficiency of the correlation coefficient values is to select the parameters that have the greatest effect on the dependent variable. In this way, considering them as the input of the artificial intelligence algorithm, while achieving a model with high accuracy, the processing time and complexity of the model will be reduced.

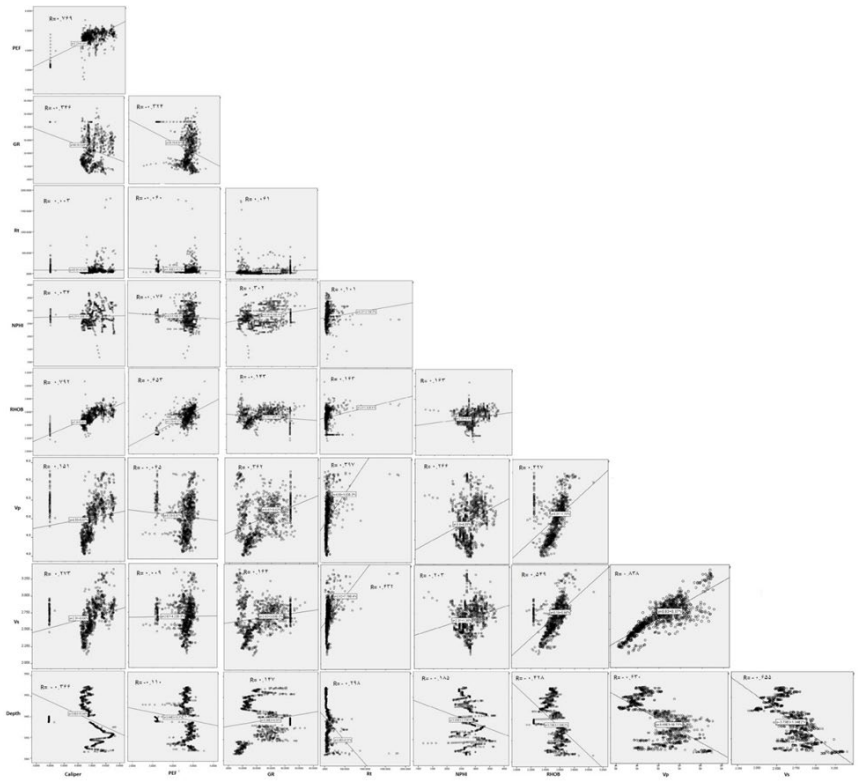

Fig. 5. Matrix of correlation coefficients between all input data and output data

To prepare the input data, pre-processing has been done on the data, which includes quality control, environmental corrections, depth mismatching, bad hole, cycle skipping, ski pies, tool pick-ups, etc.

Based on the Tukey method, in each category of input parameters, such as depth, caliper, PEF, etc., the first and third quartiles are determined using MATLAB software. Then, according to the following equations, the value of the lower inner fence (LIF) and the upper inner fence (UIF) are determined using the interquartile range (H). In the feature, according to LIF and UIF values, only the values that are smaller than LIF and larger than UIF are identified as outliers and the points corresponding to them are removed [30].

$$H=1.5\times (Q_3 - Q_1) \quad (10)$$

$$LIF=Q_1 - H \quad (11)$$

$$UIF=Q_3 + H \quad (12)$$

As a result of applying this step to the input parameters of this research, out of a total of 1042 available points, 194 points were identified as outliers and discarded. Depth, neutron, density, and shear wave velocity factors have the highest

correlation coefficients among the 8 petrophysical factors, and these factors are selected as the input data of the estimator network.

The input data have different scales, so they have been normalized so that by making their scales equal, a suitable weight can be assigned according to the importance of that parameter in the target estimation in the model training stage. Different linear and non-linear methods have been presented to normalize the data, and in this research, the linear method will be used. This stage is such that to map the data scale between zero (not zero itself) and one, the largest numerical value of each parameter is determined and all parameter values are divided by that value. To review and analyze the methods used in this research, among all the previous data, which includes 848 points, 10% of them are randomly selected and discarded. All the methods that will be mentioned estimate the shear wave velocity with the remaining 90%, and after that, ten percent of the data that was discarded becomes the input of the estimator model. The importance of this step is how much stability the estimation methods have in dealing with the data that has been left out.

4. Data processing

4.1. Experimental and regression methods

To investigate the methods of determining the shear wave velocity, first, the experimental methods to determine the shear wave speed were discussed. To check the accuracy and efficiency of these methods, the shear wave velocity obtained from DSI diagrams and the shear wave velocity obtained from experimental equations were compared.

Pickett's method calculated the linear relationship between shear wave velocity and compressional velocity for limestone formations [31].

$$v_s = \frac{v_p}{1/9} \quad (13)$$

Also, for the data related to limestone formations, Castagna and Grinberg presented non-linear relationships with dependence on compressional wave velocity in single-grained pure and water-saturated lithologies [32].

$$v_s = av_p^2 + bv_p + c \quad (14)$$

Brocher by examining a huge range of different lithologies, reached the following nonlinear equation of compressional and shear wave velocity in limestone formations [33]:

$$v_s = 0/7858 - 1/2344v_p + 0/794v_p^2 - 0/1238v_p^3 + 0/0064v_p^4 \quad (15)$$

In the regression method, using statistical software such as SPSS that has the regression estimator module, the shear wave velocity was estimated using the compressional wave velocity that is available from the petrophysical data of the studied field. The resulting equation of this estimate is equation (16).

$$V_s = -0/1395V_p^2 + 1/766V_p - 2/636 \quad (16)$$

Table 3 shows all the results of the experimental and regression methods of shear wave velocity estimation. As it is clear in this table, among these methods, the regression method is more accurate. The reasons for this conclusion are first the lower values of RMSE and then the higher values of the coefficient of determination.

4.2. MLP network

The neural network was embedded and all the values were recorded during 20 repetitions. In each part of the test and training, the highest and the lowest amount of error were separated and subtracted from each other. This interval is very important to evaluate the stability of the network. The lower the variability of this interval, the more stable the network is, and as a result, the network will be more favorable. Fig. 6 shows the correlation coefficient between the actual and predicted shear wave velocity in the two stages of training and testing. Fig. 7 and Fig. 8 Comparison diagram between the estimated shear wave velocity of the MLP method and the actual shear wave velocity along the well in the training phase - model development phase. Comparison of test and training error values of this network are shown in Table 2.

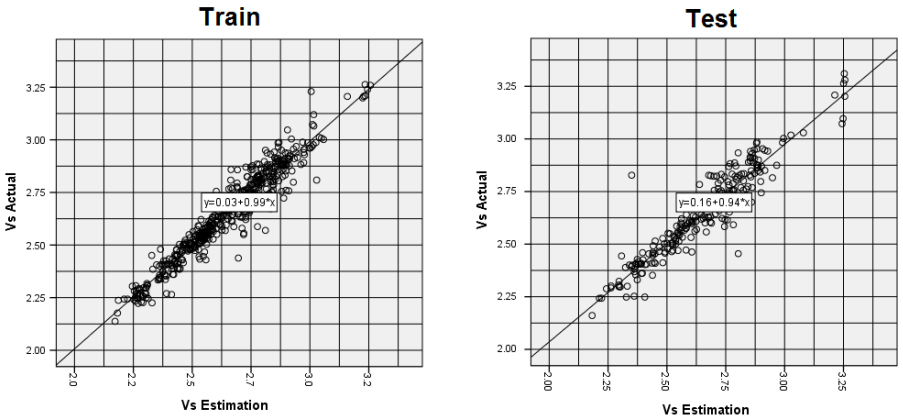

Fig. 6. Correlation coefficient between real and estimated shear wave speed in two parts of training (left image) and test (right image) of MLP network in the model development phase

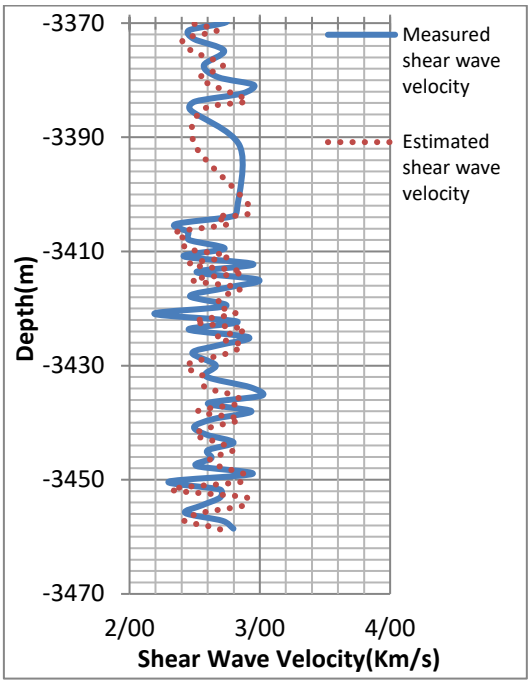

Fig. 7. Comparison diagram between the estimated shear wave velocity of the MLP method and the actual shear wave velocity along the well in the training phase - model development phase

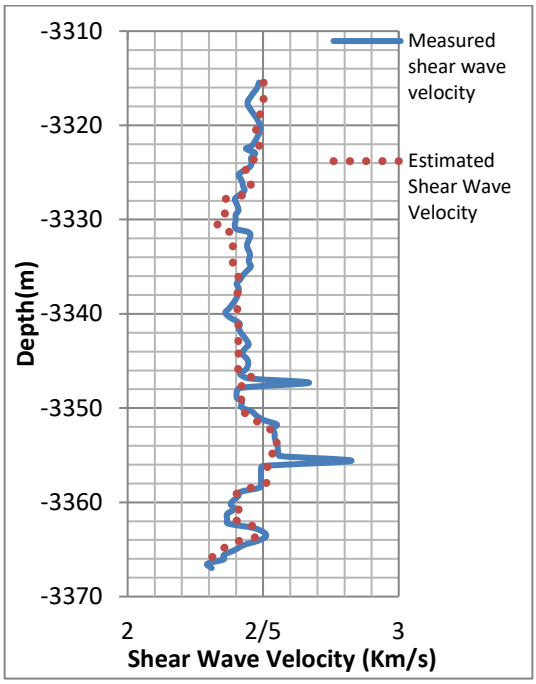

Fig. 8. Comparison diagram between the estimated shear wave velocity of the MLP method and the actual shear wave velocity along the well in the test phase - model development phase

4.3. hybrid network MLP-PSO

The architecture of this network is like the multi-layer perceptron network, the three middle layers were designed with 9, 11, and 12 neurons respectively. This hybrid network has two other effective parameters: the number of particles and maximum repetition. The number of particles was assigned to 50 particles and the maximum number of repetitions of particle swarm training was assigned to 120 repetitions. In this part, like the previous stage, the hybrid network was performed 20 times, and the maximum and minimum values of the mean squared errors of the test and training stages were extracted from the tables related to all

twenty stages and subtracted from each other. Fig. 9 shows the correlation coefficient between the actual and predicted shear wave velocity in the two stages of training and testing. Fig. 10 and Fig. 11 Comparison diagram between the estimated shear wave velocity of the MLP-PSO method and the actual shear wave velocity along the well in the training phase - model development phase. Comparison of test and training error values of this network are shown in Table 2.

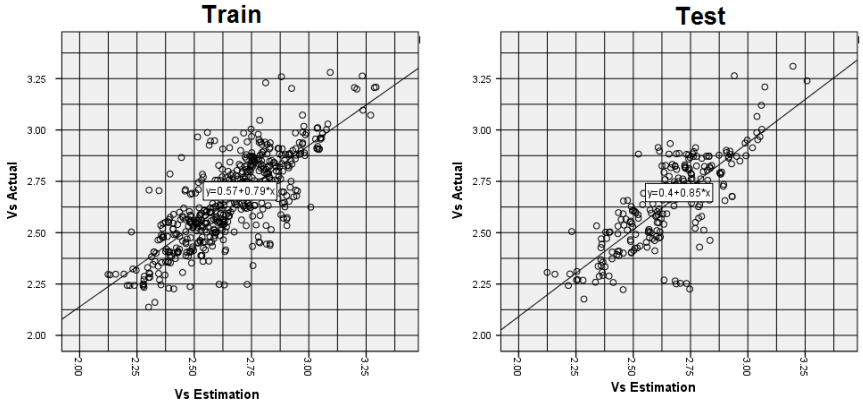

Fig. 9. Correlation coefficient between real and estimated shear wave velocity in two parts of MLP-PSO network training and test in the model development phase

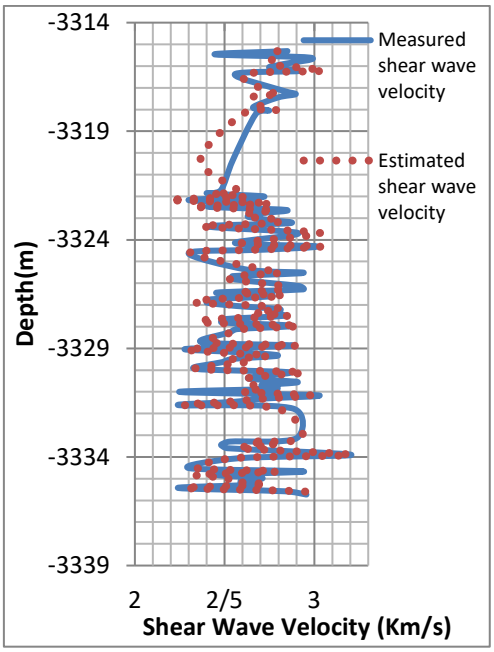

Fig. 10. Comparison diagram between the estimated shear wave velocity of the MLP-PSO method and the actual shear wave velocity along the well in the training phase - development phase

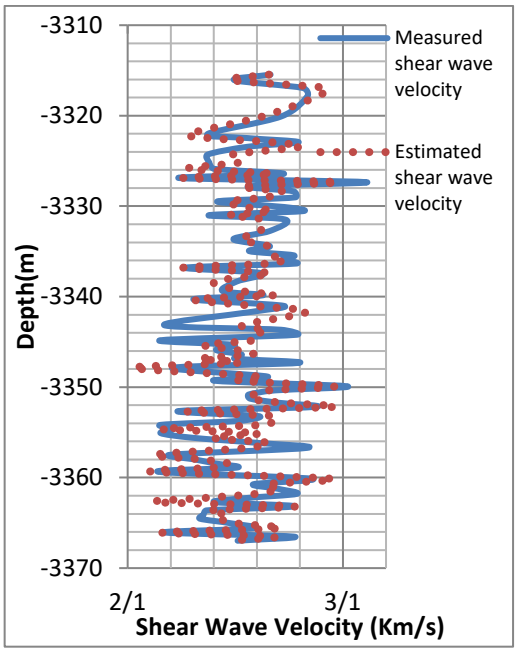

Fig. 11. Comparison diagram between the estimated shear wave velocity of the MLP-PSO method and the actual shear wave velocity along the well in the test phase - model development phase

4.4. Hybrid network MLP-SSD

In this network, the values of neurons and intermediate layers are selected by the user, and then the SSD algorithm is used to learn the network, by determining the values of the number of particles and the maximum number of repetitions, the network is performed and towards the best response and the error values and

correlation coefficients between the estimated and actual shear wave velocity are obtained. In this way, three intermediate layers with the number of neurons 9, 11, and 12 and the maximum repetition in the SSD training algorithm 120 times and 50 particles were presented for this integrated network by trial-and-error method and comparing different situations. Fig. 12 shows the correlation coefficient between the actual and predicted shear wave velocity in the two stages of training and testing. Fig. 13 and Fig. 14 Comparison diagram between the estimated shear wave velocity of the MLP-SSD method and the actual shear wave velocity along the well in the training phase - model development phase. Since for each of the methods, the optimal variable values and the maximum number of repetitions of the network were implemented, the results reached the best answers and as a result, the error values obtained are very small and minor. Comparison of test and training error values of this network are shown in Table 2.

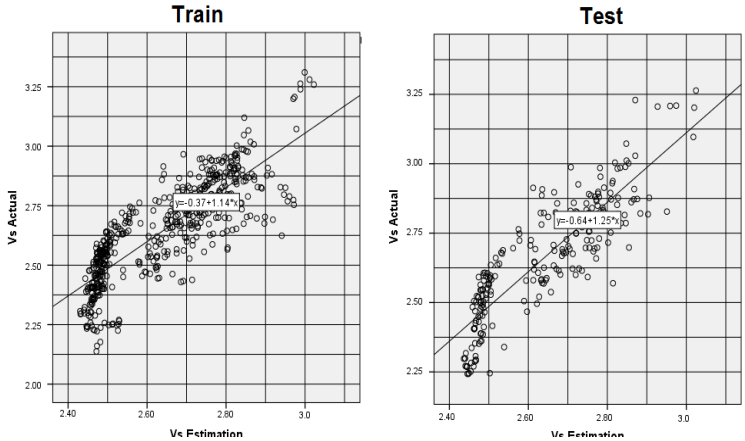

Fig. 12. The correlation coefficient between actual and estimated shear wave velocity in two parts of training (left image) and test (right image) of MLP-SSD network - model development phase

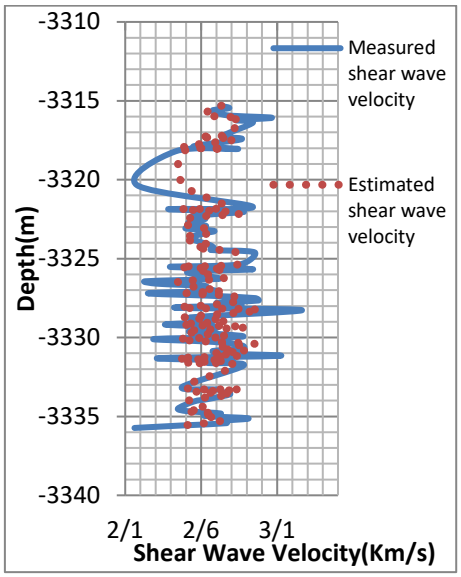

Fig. 13. Comparison diagram between the estimated shear wave velocity of the MLP-SSD method and the actual shear wave velocity along the well in the training phase - model development phase

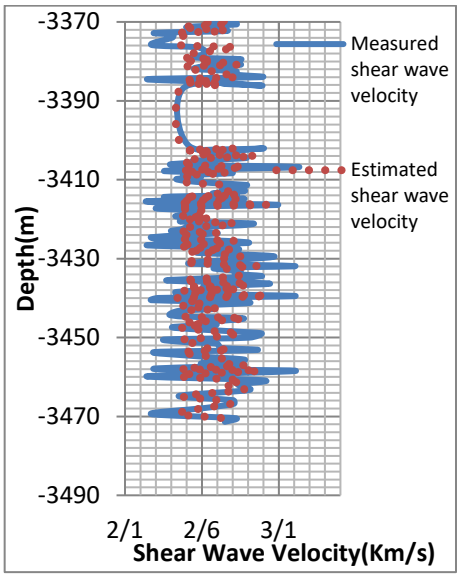

Fig. 14. Comparison diagram between the estimated shear wave velocity of the MLP-SSD method and the actual shear wave velocity along the well in the test phase - model development phase

Table 2. Comparison of test and training error values of three neural networks: MLP, MLP-PSO, and MLP-SSD

Method	Min RMSE tr	Max RMSE tr	Max RMSE tr-Min RMSE tr	Min RMSE ts	Max RMSE ts	Max RMSE tr-Min RMSE ts
MLP	0.0524	0.0755	0.0221	0.0682	0.0821	0.0139
MLP-PSO	0.0959	0.1854	0.0895	0.0927	0.2066	0.1139
MLP-SSD	0.1026	0.1523	0.0497	0.0997	0.1614	0.0617

5. Conclusion and Suggestions

The results of the research in all stages related to the estimation of the shear wave speed and the validation of the methods are shown in Table 3. The most common way of evaluation is based on the lowest number of errors. So, if the basis of evaluation of the methods is based on the lowest amount of evaluation error, according to Table 3, the regression method is the best method among experimental methods with an error of about 8% in the estimation stage and 9% error in the validation stage. Based on this, the MLP method also has the lowest error value in the model development phase among the machine learning methods with an error of 6% in the estimation phase, but contrary to expectations, it is not able to estimate the shear wave velocity well in the validation phase and the growth It has 7% error and the error rate is 12%. On the other hand, the MLP-PSO method, which had an error of about 14% in the model development phase, has been reduced to about 5% in the validation phase with a significant increase in the estimation power. Among the new methods of artificial intelligence for estimation, if only one estimation or validation step is considered, the best estimation methods are MLP and MLP-PSO, respectively.

On the other hand, according to Table 3, if the measurement is based on the correlation coefficient, the regression method is the best among the traditional methods with slight changes in both estimation and validation stages, respectively, with correlation coefficient values of 0.8697 and 0.8464. Among the machine learning methods, the MLP method has the highest correlation coefficient value of 0.943 in the estimation phase, but this network has a significant drop in the correlation coefficient in the validation phase. This value is 0.535. Also, if the evaluation stages are changed, the MLP-PSO network has a correlation coefficient of 0.923 in the validation stage, which has the best correlation coefficient compared to all traditional and new methods, but the major flaw of this method is in evaluation based on the correlation coefficient, the value of the correlation coefficient of 0.788 in the main estimation stage and the huge difference of these values.

The evaluation criteria of this research are based on the stability of estimation methods. In such a way that it is an accepted method that does not suffer gross errors in the training and testing phase when dealing with new data, small data, or massive data. In other words, the uniformity of network behavior is of great importance in estimating and solving various problems. In simple words, you should choose a method that has a reasonable error but with a small difference in the two steps of the main estimation and validation, and the correlation coefficients are close to each other. As can be seen in Table 3, the MLP-SSD method has an error of 12% in the estimation stage, and the determination coefficient of this stage is 0.746. In the validation stage, the error rate of this method is 11% and the coefficient of determination is 0.671. It is evident that the difference between these values is within a short interval. The stable behavior of this method refers to the behavior that originates from sinusoidal relationships. In this way, the particles move at a certain interval to reach convergence.

Although it should be mentioned that the regression method also has the same features among the traditional methods, due to the mechanism of this method that was obtained for the data of this research, it also has good stability. However, since the importance of the results of this research does not only belong to this formation, and perhaps the need to use the MLP-SSD method will be felt in other sciences as well, the chosen method is the MLP-SSD. Due to limitations in access to regional data and information preservation, it was not possible to check estimation methods in other fields and wells. In this way, for the better application of the chosen final method in other wells, the ability to generalize, and reliability in other wells, the degree of stability is the focus of this research.

Table 3. Summarizing the results and comparing methods in two stages of estimation and validation

		Estimation Stage			Validation Stage		
Row	**Method**	**R**	R^2	**RMSE**	**R**	R^2	**RMSE**
1	Picket	0.8539	0.7292	0.1092	0.821	0.674	0.1302
2	Castagna-Greenberg	0.1572	0.0247	0.2069	0.8315	0.6914	0.1070
3	Broocher	0.8698	0.7566	0.1150	0.1431	0.0205	0.2945
4	Regression	0.8697	0.7565	0.0894	0.8464	0.7164	0.0918
5	MLP	0.934	0.889	0.06445	0.535	0.286	0.1204
6	MLP-PSO	0.788	0.620	0.14065	0.923	0.851	0.0509
7	MLP-SSD	0.864	0.746	0.12745	0.819	0.671	0.11605

Suggestions

Since the available data of this research are very limited and are located at a certain depth of a well in Abteymour formation and the nature of this formation is limestone, it is suggested:

1- By using the selected MLP-SSD method and obtaining the optimal number of particles, repetitions, the number of intermediate layers and neurons, if the information of petrophysical logs is available in other wells, from this method and in limestone, dolomite, and sand formations, and the presence of at least one well with information related to the DSI profile, can estimate the shear wave velocity in other wells of that field to obtain complete information about the mechanical properties of the formations.

2- With the optimal use of this method, it is possible to estimate the required quantities for other projects in the oil industry and related sciences and reduce the possible exorbitant costs.

3- The use of optimization algorithms to choose the appropriate structure, i.e. the number of hidden layers and the number of neurons in each layer, can lead to a better performance of the multilayer perceptron neural network in its simple or hybrid form. Therefore, it is suggested to determine the structure with an optimization algorithm in future studies.

4- The use of deep neural networks in other fields of engineering sciences is increasing. Therefore, it is suggested to use this tool to estimate shear wave velocity from petrophysical logs.

5- It is suggested to use other feature selection methods that are based on fit and cost functions (i.e. wrapper methods) which have shown better performance than the method used in this research.

6- It is advisable to employ advanced machine learning techniques, such as the Kyoshin network (K-NET), or a hybrid approach combining various methods for achieving a more accurate determination of Shear Wave Velocity. This can result in a higher correlation coefficient and a reduced error value.

References

[1] Mohaghegh, S. D., (2005), "Recent Developments in Application of Artificial Intelligence in Petroleum Engineering," Journal of Petroleum Technology, 57, 04, pp. 86–91.

[2] Russell, B. H., (2004), ″The Application of Multivariate Statistics and Neural Networks to the Prediction of Reservoir Parameters Using Seismic Attributes″, PhD Thesis, Department of Geology and Geophysics, Calgary, Alberta.

[3] Moatazedian, I., Rahimpour-Bonab, H., Kadkhodaie-Ilkhchi, A., and Rajoli, M., (2011), "Prediction of Shear and Compressional Wave Velocities from Petrophysical Data Utilizing Genetic Algorithms Technique: A Case Study in Hendijan and Abuzar Fields Located in Persian Gulf," Geopersia, 1, 1, pp. 1–17.

[4] Alimoradi, A., Shahsavani, H., and Rouhani, A. K., (2011), "Prediction of Shear Wave Velocity in Underground Layers Using SASW and Artificial Neural Networks."

[5] Güllü, H., (2013), "On the Prediction of Shear Wave Velocity at Local Site of Strong Ground Motion Stations: An Application Using Artificial Intelligence," Bulletin of Earthquake Engineering, 11, 4, pp. 969–997.

[6] Zoveidavianpoor, M., Samsuri, A., and Shadizadeh, S. R., (2013), "Adaptive Neuro Fuzzy Inference System for Compressional Wave Velocity Prediction in a Carbonate Reservoir," J Appl Geophy, 89, pp. 96–107.

[7] Maleki, S., Moradzadeh, A., Riabi, R. G., Gholami, R., and Sadeghzadeh, F., (2014), "Prediction of Shear Wave Velocity Using Empirical Correlations and Artificial Intelligence Methods," NRIAG Journal of Astronomy and Geophysics, 3, 1, pp. 70–81.

[8] Singh, S., and Kanli, A. I., (2016), "Estimating Shear Wave Velocities in Oil Fields: A Neural Network Approach," Geosciences Journal, 20, pp. 221–228.

[9] Asoodeh, M., and Bagheripour, P., (2014), "ACE Stimulated Neural Network for Shear Wave Velocity Determination from Well Logs," J Appl Geophy, 107, pp. 102–107.

[10] Nourafkan, A., and Kadkhodaie-Ilkhchi, A., (2015), "Shear Wave Velocity Estimation from Conventional Well Log Data by Using a Hybrid Ant Colony–Fuzzy Inference System: A Case Study from Cheshmeh–Khosh Oilfield," J Pet Sci Eng, 127, pp. 459–468.

[11] Mehrgini, B., Izadi, H., and Memarian, H., (2019), "Shear Wave Velocity Prediction Using Elman Artificial Neural Network," Carbonates Evaporites, 34, pp. 1281–1291.

[12] Elkatatny, S., Tariq, Z., Mahmoud, M., Mohamed, I., and Abdulraheem, A., (2018), "Development of New Mathematical Model for Compressional and Shear Sonic Times from Wireline Log Data Using Artificial Intelligence Neural Networks (White Box)," Arab J Sci Eng, 43, pp. 6375–6389.

[13] Anemangely, M., Ramezanzadeh, A., and Tokhmechi, B., (2017), "Shear Wave Travel Time Estimation from Petrophysical Logs Using ANFIS-PSO Algorithm: A Case Study from Ab-Teymour Oilfield," J Nat Gas Sci Eng, 38, pp. 373–387.

[14] Yang, H., Xu, Y., Peng, G., Yu, G., Chen, M., Duan, W., Zhu, Y., Cui, Y., and Wang, X., (2017), "Particle Swarm Optimization and Its Application to Seismic Inversion of Igneous Rocks," Int J Min Sci Technol, 27, 2, pp. 349–357.

[15] Tharwat, A., and Gabel, T., (2020), "Parameters Optimization of Support Vector Machines for Imbalanced Data Using Social Ski Driver Algorithm," Neural Comput Appl, 32, pp. 6925–6938.

[16] Zhang, Y., Zhang, C., Ma, Q., Zhang, X., and Zhou, H., (2022), "Automatic Prediction of Shear Wave Velocity Using Convolutional Neural Networks for Different Reservoirs in Ordos Basin," J Pet Sci Eng, 208, p. 109252.

[17] Rajabi, M., Hazbeh, O., Davoodi, S., Wood, D. A., Tehrani, P. S., Ghorbani, H., Mehrad, M., Mohamadian, N., Rukavishnikov, V. S., and Radwan, A. E., (2023), "Predicting Shear Wave Velocity from Conventional Well Logs with Deep and Hybrid Machine Learning Algorithms," J Pet Explor Prod Technol, 13, 1, pp. 19–42.

[18] Davari, M. A., Senemari, S., Alimoradi, A., and Safavi, S. J., (2024), "Permeability Prediction from Log Data Using Machine Learning Methods," Journal of Petroleum Geomechanics, 7, 3.

[19] Sarkheil, S., Hassani, H., Alinya, F., Enayati, A. A., and Motamedi, H., (2009), "A Forecasting System of Reservoir Fractures Based on Artificial Neural Network and Borehole Images Information-Exemplified by Reservoir Fractures in Tabnak Feild, Fars, Iran," International Multidisciplinary Scientific GeoConference: SGEM, 1, p. 563.

[20] Sarkheil, H., Hassani, H., and Alinia, F., (2013), "Fractures Distribution Modeling Using Fractal and Multi-Fractal–Neural Network Analysis in Tabnak Hydrocarbon Field, Fars, Iran," Arabian Journal of Geosciences, 6, pp. 945–956.

[21] Alimoradi, A., Hajkarimian, H., Hemati Ahooi, H., and Salsabili, M., (2022), "Comparison between the Performance of Four Metaheuristic Algorithms in Training a Multilayer Perceptron Machine for Gold Grade Estimation," International Journal of Mining and Geo-Engineering, 56, 2, pp. 97–105.

[22] Azimi, Y., Talaeian, M., Sarkheil, H., Hashemi, R., and Shirdam, R., (2022), "Developing an Evolving Multi-Layer Perceptron Network by Genetic Algorithm to Predict Full-Scale Municipal Wastewater Treatment Plant Effluent," J Environ Chem Eng, 10, 5, p. 108398.

[23] Bahri, E., Alimoradi, A., and Yousefi, M., (2023), "Investigating the Performance of Continuous Weighting Functions in the Integration of Exploration Data for Mineral Potential Modeling Using Artificial Neural Networks, Geometric Average and Fuzzy Gamma Operators," International Journal of Mining and Geo-Engineering, 57, 4, pp. 405–412.

[24] Ebrahimi, A., Izadpanahi, A., Ebrahimi, P., and Ranjbar, A., (2022), "Estimation of Shear Wave Velocity in an Iranian Oil Reservoir Using Machine Learning Methods," J Pet Sci Eng, 209, p. 109841.

[25] Kennedy, J., and Eberhart, R., (1995), "Particle Swarm Optimization," *Proceedings of ICNN'95-International Conference on Neural Networks*, IEEE, pp. 1942–1948.

[26] Hemati Ahooi, H. R., (2021), "Optimizing Extreme Learning Machine Algorithm Using Particle Swarm Optimization to Estimate Iron Ore Grade," Journal of Mining and Environment, 12, 2, pp. 397–411.

[27] Fathi, M., Alimoradi, A., and Hemati Ahooi, H. R., (2021), "Optimizing Extreme Learning Machine Algorithm Using Particle Swarm Optimization to Estimate Iron Ore Grade," Journal of Mining and Environment, 12, 2, pp. 397–411.

[28] Omidi, R., Sadeghi, A., Hosseini-Barzi, M., and Akbar Bas Kelayeh, N., (2021), "New Findings in Biostratigraphy of the Sarvak and Ilam Formations of Abteymour Oil Field (Dezful Embayment)," Journal of Stratigraphy and Sedimentology Researches, 37, 1, pp. 23–44.

[29] Heydari, E., Hassanzadeh, J., Wade, W. J., and Ghazi, A. M., (2003), "Permian–Triassic Boundary Interval in the Abadeh Section of Iran with Implications for Mass Extinction: Part 1–Sedimentology," Palaeogeogr Palaeoclimatol Palaeoecol, 193, 3–4, pp. 405–423.

[30] Rousseeuw, P. J., and Leroy, A. M., (2005), ″Robust Regression and Outlier Detection″, John wiley & sons.

[31] Pickett, G. R., (1963), "Acoustic Character Logs and Their Applications in Formation Evaluation," Journal of Petroleum Technology, 15, 06, pp. 659–667.

[32] Castagna, J. P., Batzle, M. L., and Kan, T. K., (1993), "The Link between Rock Properties and AVO Response."

[33] Brocher, T. M., (2005), "Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust," Bulletin of the seismological Society of America, 95, 6, pp. 2081–2092.