

USING STACKED GENERALIZATION ENSEMBLE METHOD TO ESTIMATE SHEAR WAVE VELOCITY BASED ON DOWNHOLE SEISMIC DATA: A CASE STUDY OF SARAB-E-ZAHAB, IRAN

SAIRAN ALIZADEH¹, RASHED POORMIRZAE^{2*}, RAMIN NIKROUZ¹ and SIAMAK SARMADY³

¹*Faculty of Science, Geophysics group, Urmia University, Urmia, Iran.
sairan.alizadeh@gmail.com*

²*Department of Mining Engineering, Faculty of Environment, Urmia University of technology, Urmia, Iran. Rashed.poormirzaee@gmail.com*

³*Department of Information Technology and Computer Engineering, Faculty of Industrial Technologies, Urmia University of Technology, Urmia, Iran.
siamak.sarmady@gmail.com*

(Received May 26, 2020; revised version accepted January 6, 2021)

ABSTRACT

Alizadeh, S., Poormirzaee, R., Nikrouz, R. and Sarmady, S., 2021. Using stacked generalization ensemble method to estimate shear wave velocity based on downhole seismic data: a case study of Sarab-e-Zahab, Iran. *Journal of Seismic Exploration*, 30: 281-301.

The proper estimation of shear wave velocity (V_s), because of its direct relation to the soil dynamic properties, for the study of Site effects is an important task in the engineering geophysics. Because of the direct travel of waves from sources to receivers, the downhole seismic method, among others, is suitable for accurate estimation of shear wave velocity. However, the main challenge is the high cost of borehole surveys, which limits the amount of downhole seismic data when studying a large area. In order to tackle this problem, an ensemble system is proposed that estimates the shear wave velocity using a limited amount of data. For this purpose, the downhole seismic data at 4 points were collected in Sarab-e-Zahab area, Iran. Then, the data were processed and the shear wave velocity profile was obtained for each borehole. Finally, using an ensemble of neural networks, a 3- and 2-dimensional model of V_s was constructed for the study area. Feed-forward neural networks were used as the base classifiers in an ensemble system and two methods, namely averaging and stacked generalization were employed to combine the results of base classifiers. The performances of the two methods were compared and the shear wave velocity was estimated as a function of depth. The results of the ensemble neural networks method in the study area were compared with Kriging geostatistical method. The results show the ensemble neural networks in the V_s modeling in comparison to the Kriging method has better performance. Also, the findings showed that the stacked generalization method outperformed the averaging method in the estimation of shear wave velocity.

KEY WORDS: shear waves, downhole seismic data, ensemble systems, Kriging method.

INTRODUCTION

Earthquake is a natural phenomenon that occasionally shakes part of the earth, causing devastation, injury and death. Seismic microzonation is the first step to mitigate the devastating effects of earthquakes. In the seismic microzonation, the study area is classified according to the degree of importance and also the extent of the potential earthquake vulnerability of the area to varying degrees of risk. The results can be used for urban planning, disaster preparedness and risk reduction (Ansal et al., 2010). The observations of different earthquakes indicate that the earthquake damage in sedimentary basins is much greater than that in the areas located on the bedrock (Zaharia et al., 2008). The studies of seismic microzonation and site reaction include several steps and one of the most important and fundamental steps is to investigate the characteristics of subsurface layers in the study area. The subsurface layers of the area are explored using the geotechnical and geophysical studies. The commonly studied properties are: soil layering, average shear wave velocity, bedrock location, maximum shear modulus, layer thickness, main period of soil profiles, etc. (ISSMGE, 1993). In the geotechnical experiments, after the borehole drilling and sampling, the experiments are usually performed in the laboratory, since the results of these tests are limited to the area around the borehole, and since the soil is heterogeneous and anisotropic, the results of the experiments cannot be generalized to the whole area under investigation. Also, when complex geological conditions exist, the use of geophysical methods could increase the quality of borehole data (Earley and Rudenako, 1988). The studies of seismic microzonation and site reaction require the identification of subsurface soil properties at the site. The knowledge of shear wave velocity is a key parameter for evaluating the dynamic behavior of soil in the shallow subsurface (Kanli et al., 2006). Information about the shear wave velocity in depths is needed to predict the ground response to earthquakes (Hunter et al., 2002). Burcherdt (1994) recommended an average shear wave velocity of 30 m (V_{S30}) for the soil classification and seismic design. The geophysical seismic methods are based on the fact that the wave propagation velocity in an elastic object is a function of the elastic modulus and material density (Hvorslev, 1949). The downhole seismic method is one of the most accurate geophysical techniques to investigate the characteristics of subsurface layers in the study area and also the site effects or seismic microzonation. Despite the usefulness of the downhole method in examining the characteristics of subsurface layers, if it is being used for exploring the characteristics of regional subsurface layers in a large area, a large number of downhole boreholes are required. This is sometimes not practical due to the costs. This study attempts to use a neural network model to predict the shear wave velocity in areas where it is not possible to drill boreholes. In order to increase the accuracy and performance of the neural network model, an ensemble of neural networks is used. Combining multiple networks and creating an ensemble neural network can reduce the risk of

incorrect results and potentially improve the accuracy and generalization capability. One of the earliest ensemble systems was created by Dasarathy and Sheela (1979). They partitioned the attribute space to two or more classifiers. Hansen and Salamon (1990) showed that the performance of neural networks can be improved using a set of neural networks with similar configurations. There are various approaches to the integration of multiple neural networks which all attempt to minimize the error of estimation, including the following. Zhou et al. (2002) used genetic algorithm to combine the results of multiple neural networks using the GASEN method. Wolpert (1992) proposed a stacked generalization method for combining the base classifiers. Several studies have utilized the artificial neural networks (ANN) in geophysics, some of which are briefly listed here. Shimshoni and Intrator (1998) used artificial neural networks to differentiate natural earthquake waves from artificial explosion waves. Carderon et al. (2000) estimated the thickness and velocity of seismic layers using artificial neural networks and geoelectric data. In a study, Van der Baan and Jutten (2000) examined the use of neural networks for approximating geophysical problems. Liu et al. (2004) used the radial basis function (RBF) neural networks as base networks and combined their results with majority voting technique to predict earthquakes. The results of this study showed that the use of ensemble neural networks increases the earthquake prediction accuracy. Abdideh (2012) estimated the permeability in an oil field in Iran using artificial neural networks. Polo et al. (2017) introduced a new approach based on deep neural network (DNN) on seismic images to detect faults. Nyein and Hamada (2018) applied ANN in the prediction of water saturation and porosity of shaly sandstone reservoirs using well logging data. In current study, we introduce a new method using machine learning techniques, i.e., ensemble systems, for the construction of 3- and 2-dimensional shear wave velocity models in Sarab-e-Zahab (Iran) located in a region with a high level of seismic hazard.

METHODS

Artificial Neural Network

Artificial neural networks are modeled on the neurons of brain. They make inference through simple mathematical calculations. The artificial neurons are put together in layers to form a feed-forward network. The input weights of neurons are adjusted so that feeding the training dataset to the network could produce the expected outputs with a good accuracy. It is then expected that the network can calculate appropriate outputs for new unseen inputs. Since the relation between inputs and outputs could be nonlinear, nonlinear activation functions like sigmoid are used for hidden neurons of the network.

Ensemble Neural Network

Creating a neural network with the highest possible performance and decreasing the risk and level of error is a major goal when building a neural network. A lot of effort is usually spent on fine tuning the network to make sure it provides the best results. However, a single classification or regression model can hardly produce the best results for every sub-section of a complex problem space. Creating multiple neural networks (or any other machine learning model) and combining the results of those networks could potentially improve the performance and decrease the risk of error (Zhou et al., 2002). An ensemble neural network has two main components, namely an ensemble of base models and a combiner. Several base neural networks are trained with either different settings or different subsets of the training data. The predictions of the individual networks are then combined. The success of an ensemble system, relies on the diversity of classification or regression models that make it. If all the models make the same mistakes, no improvement can be made. If each model makes different mistakes, then by combining them in an efficient way, total error could potentially be reduced. Diversity can be achieved in many ways. The most popular method, is to use different subsets of training datasets for each base model. Such subsets can be obtained using various methods such as bootstrapping, bagging and n-fold subset selection (Polikar, 2006). Other approaches for creating diversity include usage of different Meta parameters (e.g., number of neurons and layers in a neural network), or using different classes of models (e.g., decision trees and kNN).

Combination of results in ensemble systems

A combiner takes the outputs of the base models and combines them into a single result. Combination methods are generally of two types: trainable and non-trainable. The non-trainable methods combine the results using simple rules like majority voting (for classification) and average calculation (for regression models). The trainable methods learn to map outputs of base models into the final result by looking into the outputs generated by individual models (for the training data) and the expected correct result. The techniques such as the mixture of experts (Jacobs, 1991) and stacked generalization (Wolpert, 1992) fall within this category of methods.

In the stacked generalization method, a neural network is used for combining the results of the base models. The network is trained by providing the output of base models and expected correct result for a given dataset (Wolpert, 1992).

Geology of the study area

The study area, Sarab-e-Zahab, is located in the central part of the Sarpol-e-Zahab county in Kermanshah province, Iran. The area is geologically located on the zone of Ghasr-e-Shirin and conforms to the characteristics of this zone. The study area is shown in Fig. 1. This area has a high level of estimated hazard. The last powerful earthquake (7.3 Mw) in the study area happened on November 12, 2017 at 18:18 pm (World Time) about 37 km northwest of Sarpol-e-Zahab. In this earthquake, the study area was entirely destroyed. Besides the earthquake magnitude, the geological formations of the area and the conditions of subsurface layers appear to play an important role in this disaster. In terms of geological structure, Kermanshah province consists of two structural units, namely Sanandaj-Sirjan and Zagros structures. The Zagros Mountains are bounded on the north by the Iranian plateau and on the south by the active basins of Mesopotamia and the Persian Gulf. The province can also be geologically divided into eastern and western parts. The eastern part, which is higher and mainly mountainous, comprises a series of igneous and metamorphic rocks (Sanandaj-Sirjan unit), calcareous and radiolarian rocks (Zagros drifting) and calcareous dolomitic high folds (folded Zagros).

The geological formations of Ghasr-e-Shirin region during the Neogene period include the sediments of Fars Group (Gachsaran, Mishan and Aghajari formations). These formations generally include sandstone, red clay, marl, dolomitic limestones and gypsum (Agha Nabati, 1995). Due to lack of local geological map and absence of geotechnical data in the study region, 3-dimensional distribution of subsurface soil not available. But the observations during the drilling boreholes up to 30 m depth, for downhole seismic test, show drilling cores consisted of soft clay, sand, clay, marl and, to some extent, gypsum and lime. Table 1 shows the range of shear wave velocities for the mentioned formations.

Table 1. Shear wave velocity range for some types of sediments (Mavko, 2005).

Type of formation	S wave velocity (m/s)
Dry sands	100-500
Wet sands	400-600
Clay	175-375
Marls	750-1500
Chalk	1100-1300
Limestone	2000-3300

Data acquisition

The study area is located in Sarab-e-Zahab village, 15 km from Sarpol-e-Zahab, Kermanshah province, Iran. The study area covers an area of 74000 square meters. To gather downhole seismic data, four boreholes were drilled in the study area. The downhole seismic test can directly measure the P- and S-wave velocities in a borehole by penetrating the seismic cone into the soil or rock. The seismic waves are generated by a seismic source on the surface near the borehole. The downhole receivers are located at a specific distance from the borehole to detect incoming seismic wave. Then, the travel time of the seismic wave between the source and the receiver is measured. The P- and S-wave velocities are determined after measuring the travel time for the P- and S-waves produced. The position of the study area and boreholes are shown on Google Earth map in Fig. 2. In each of the boreholes, the downhole tests were performed using the shear and longitudinal waves. The specifications of the boreholes are set out in Table 2.

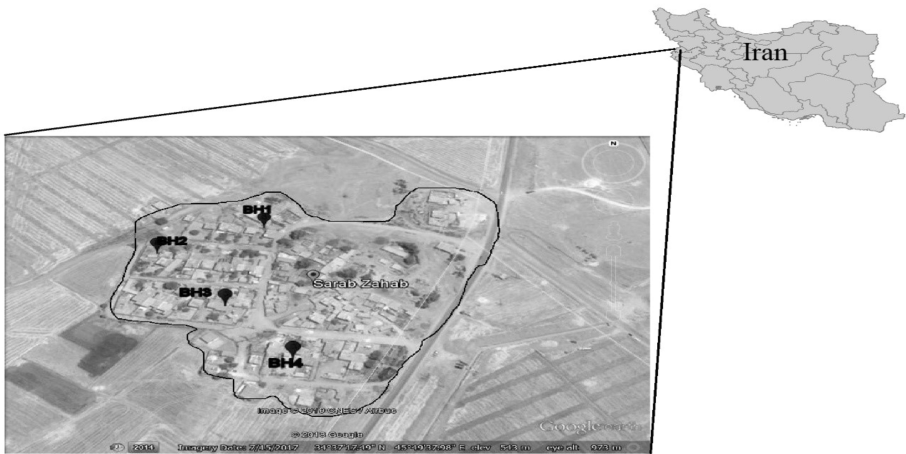


Fig. 2. Location of the boreholes in the study area (Google Earth imagery).

Table 2. The specification of boreholes in the study area.

No Downhole	Distance metal plate (m)	Distance shear beam (m)	No.records SH waves	No.records SV&P waves	Distance between records (m)	Depth (m)	long	lat
BH1	3	3	20	20	1.5	30	34°37'20"	45°49'35"
BH2	3	3	20	20	1.5	30	34°37'19"	45°49'32"
BH3	3	3	20	20	1.5	30	34°37'16"	45°49'34"
BH4	3	3	20	20	1.5	30	34°37'14"	45°49'37"

After gathering the data, namely travel times and shear wave velocities, the processing was performed by Downhole 2016 software. For this purpose, the noises are filtered and then, the first arrival times of P- and S-waves were determined at different depths. Moreover, for calculating the corrected arrival time, eq. (1) is used. After calculating the first arrival times of P- and S-wave velocities, it is time to calculate the P- and S-wave velocities for each of the boreholes:

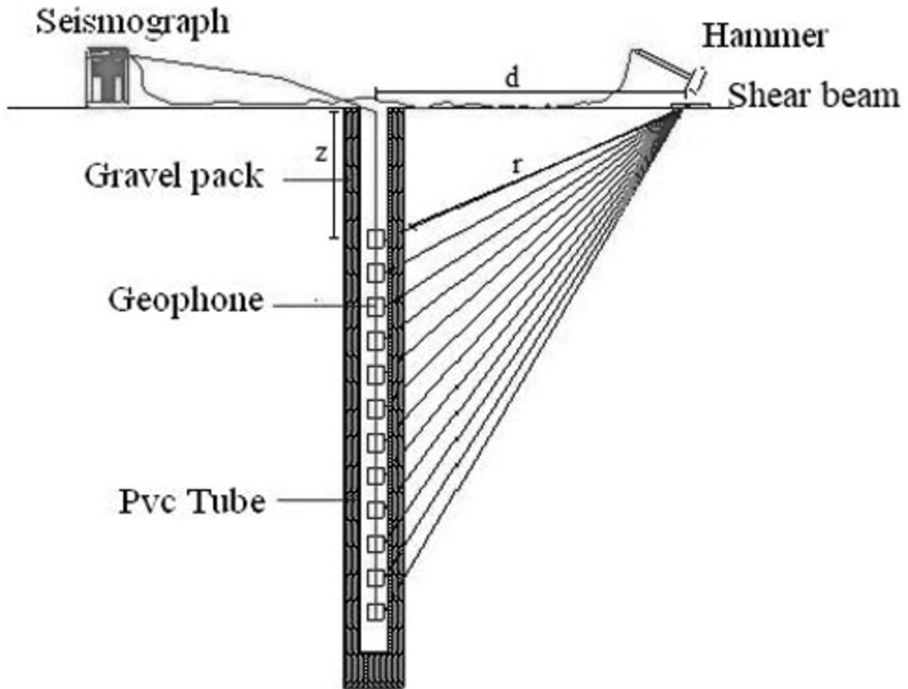


Fig. 3. Schematic representation of downhole seismic method.

$$T_{\text{cor}} = \frac{z}{r} t \quad . \quad (1)$$

In Fig. 3, d is the distance from the borehole center to the energy source, z is the depth of investigation where the receiver is located, r is the direct route of the source to the receiver, and T_{cor} is the corrected time. The obtained corrected times and wave velocities for the borehole data processing in all boreholes are presented in Table 3.

Table 3. The obtained corrected times and wave velocities for the boreholes in the study area.

Depth (m)	Corrected time(S) wave Tscor [msec]	(S) wave Velocity (m/s)	Corrected time(S) wave Tscor [msec]	(S) wave Velocity (m/s)	Corrected time(S) wave Tscor [msec]	(S) wave Velocity (m/s)	Corrected time(S) wave Tscor [msec]	(S) wave Velocity (m/s)
	BH1	BH1	BH2	BH2	BH3	BH3	BH4	BH4
	1.5	10.20	147.11	13.06	114.87	9.21	162.82	7.56
3	18.74	175.61	21.78	172.01	18.38	163.54	13.51	252.19
4.5	25.29	228.80	30.70	168.09	30.29	126.03	22.05	175.57
6	31.39	245.9	39.62	168.15	36.40	245.23	30.23	183.32
7.5	38.25	218.7	44.38	315.25	45.77	160.07	33.24	498.69
9	44.68	233.29	49.24	308.93	50.28	332.87	35.67	617.03
10.5	47.50	532.61	56.54	205.45	54.90	324.48	38.75	487.19
12	55.49	187.67	64.13	197.66	61.12	241.31	47.25	176.45
13.5	58.67	472.2	71.07	216.14	64.89	405.43	57.00	153.64
15	65.70	213.37	75.01	379.94	70.21	278.25	62.07	296.35
16.5	69.17	432.63	81.17	243.71	76.25	248.33	70.54	177.04
18	72.30	478.23	90.16	166.91	81.87	266.87	73.19	566.70
19.5	77.29	300.72	97.26	211.29	85.00	479.34	78.18	300.61
21	81.87	327.65	104.34	211.72	89.49	333.97	82.56	342.34
22.5	86.83	302.25	108.04	405.07	93.37	386.37	86.14	419.47
24	90.10	459.11	116.00	188.60	97.44	368.73	96.75	141.38
25.5	104.68	102.89	120.87	308.07	106.46	166.22	101.80	296.98
27	110.22	270.58	124.93	369.02	113.00	229.40	108.04	240.49
28.5	113.37	475.9	127.99	489.92	122.93	151.27	113.27	286.31
30	118.21	310.12	134.23	240.48	129.45	229.59	118.41	292.10

Estimation of Vs using ensemble neural network technique

In order to process the data using an ensemble system, the map of the study area was first extracted from Google Earth and then transferred to ArcGIS software. Afterwards, the area was gridded and 19 points as hypothetical boreholes were designated for the estimation of shear wave velocity. Also, an origin point was considered outside the designed grid. The arrangement of points, drilled boreholes in the study area, and origin point are depicted in Fig. 4. The coordinates of the boreholes and 19 points were defined as (X, Y, Z), where X and Y are the horizontal and vertical distance from the origin point, respectively. Also, Z is the depth of the points divided

into the 1.5 m intervals. Table 4 shows the coordinates of boreholes and points in the study area. Then, an ensemble of neural networks and the available downhole seismic data were used to estimate The Vs profiles for the whole study area, namely 19 designated points. In this study, different preprocessing methods were used prior to training the neural networks. In order to normalize inputs of neural networks, logarithm of the basis 10 was found to work better than other methods.

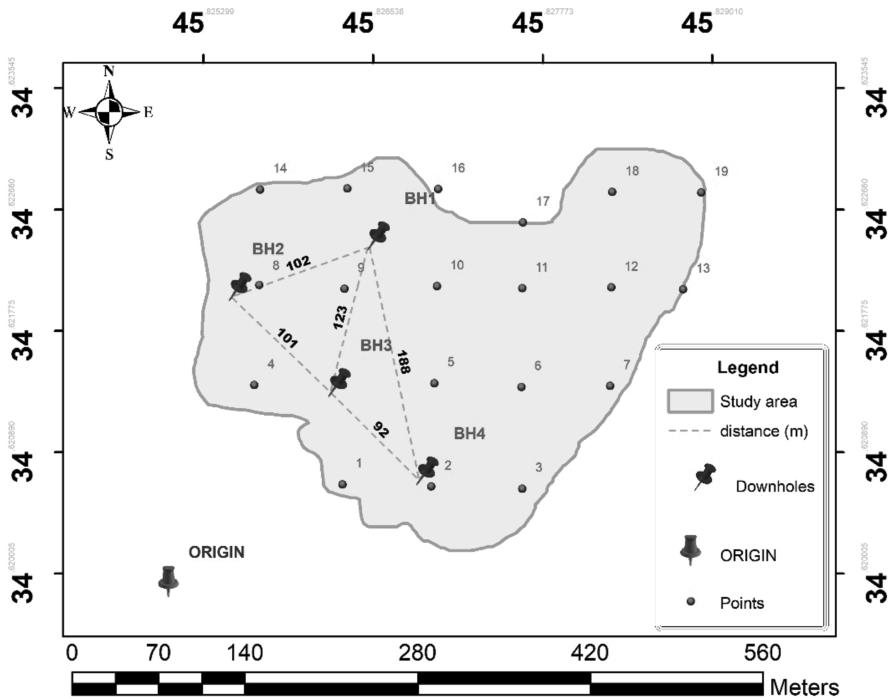


Fig. 4. The arrangement of points, boreholes and origin in the study area.

In order to train the base neural network of the ensemble system, which estimate the depth profile of shear wave velocity at 19 single points, 80 data items of 4 boreholes (20 data for each borehole) were prepared and randomly arranged in quadruple form (X , Y , Z , V_s). The (X , Y , Z) are the neural network inputs, and the output, shear wave velocity (V_s), is going to be predicted by the networks. Furthermore, to improve the performance of predictions, the data items were divided into two depth ranges, namely (1.5 to 15) meters and (16.5 to 30) meters. For each range of data, a separate ensemble system was used. For each ensemble system, 10 base models (i.e., neural networks) were trained. Ten sets of training and test data were

prepared using 10-fold subset selection method. For each subset, 9/10 of the whole data items were put together to form the training data. The remaining 1/10 was used as the test data. Base networks have the same hyper-parameters and consist of three layers. There are 7 neurons in the first hidden layer and 3 neurons in the second hidden layer, all with sigmoid activation functions. The network performance was evaluated using mean square error (Table 5). After training the networks, the profiles of shear wave velocity in each of the 19 boreholes and points were estimated using the trained networks. Since 10 profiles (up to 30m deep) are estimated per borehole by 10 base networks, they need to be combined to obtain the final depth profile of shear wave velocity in boreholes and 19 points. In order to combine the estimated results of the base neural networks, two methods, namely simple averaging and stacked generalization, were used, and then, the results (i.e., the estimated profiles) of the two methods were compared.

Table 4. Boreholes and points coordinates in designed grid for the study area.

point	X [m]	Y [m]	Z [m]
BH1	136.9	281	-1.5,-3,-4.5,.....,-30
BH2	44.6	239.2	-1.5,-3,-4.5,.....,-30
BH3	111.6	162	-1.5,-3,-4.5,.....,-30
BH4	171.1	91.2	-1.5,-3,-4.5,.....,-30
Point1	115.4	79.8	-1.5,-3,-4.5,.....,-30
Point2	177.4	79.8	-1.5,-3,-4.5,.....,-30
Point3	234.4	79.8	-1.5,-3,-4.5,.....,-30
Point4	58.5	159.5	-1.5,-3,-4.5,.....,-30
Point5	177.4	159.5	-1.5,-3,-4.5,.....,-30
Point6	234.4	159.5	-1.5,-3,-4.5,.....,-30
Point7	293.9	159.5	-1.5,-3,-4.5,.....,-30
Point8	58.5	240.5	-1.5,-3,-4.5,.....,-30
Point9	115.4	240.5	-1.5,-3,-4.5,.....,-30
Point10	177.4	240.5	-1.5,-3,-4.5,.....,-30
Point11	234.4	240.5	-1.5,-3,-4.5,.....,-30
Point12	293.9	240.5	-1.5,-3,-4.5,.....,-30
Point13	353.3	240.5	-1.5,-3,-4.5,.....,-30
Point14	58.5	317.7	-1.5,-3,-4.5,.....,-30
Point15	115.4	317.7	-1.5,-3,-4.5,.....,-30
Point16	177.4	317.7	-1.5,-3,-4.5,.....,-30
Point17	234.4	292.4	-1.5,-3,-4.5,.....,-30
Point18	293.9	317.7	-1.5,-3,-4.5,.....,-30
Point19	353.3	317.7	-1.5,-3,-4.5,.....,-30

Table 5. Performance of 10 networks with different data sets and similar network parameters (first and second ranges of data).

Range 1 - depth (1.5 to 15) m		Range 2 - depth (16.5to 30) m	
MSE primary data	MSE data after training	MSE primary data	MSE data after training
0.0835	2.67×10^{-11}	0.0392	9.34×10^{-9}
0.0335	3.07×10^{-13}	0.0189	7.79×10^{-10}
0.101	4.71×10^{-11}	0.112	0.000177
0.520	1.22×10^{-9}	0.0366	0.00411
0.0513	0.000143	0.175	4.25×10^{-5}
0.283	0.000119	0.161	2.01×10^{-14}
0.281	1.63×10^{-10}	0.491	4.08×10^{-5}
0.140	4.15×10^{-10}	0.0659	1.92×10^{-11}
0.357	7.35×10^{-11}	0.586	0.000377
0.0799	1.76×10^{-5}	0.311	4.44×10^{-11}

Simple Averaging

In this method, the estimation of shear wave velocity profile provided by 10 base neural networks in 4 boreholes were combined using the averaging. In order to evaluate the method, the correlation chart was plotted for the estimated values and the observation data (obtained from downhole operations). The RMSE value and regression coefficient (R) were determined. The correlation is shown in Fig. 5

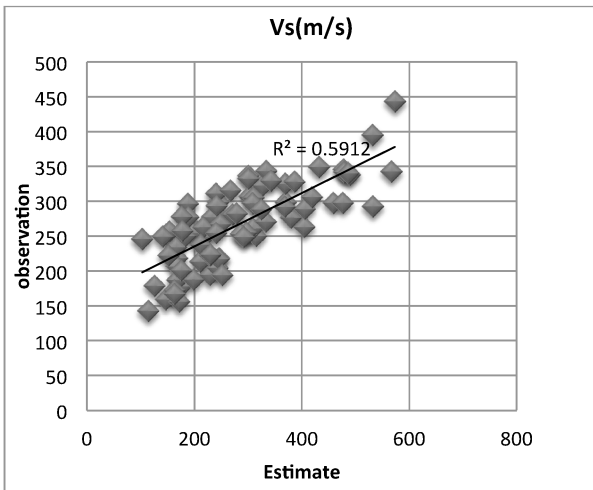


Fig. 5. Correlation graph for estimated data (average combiner) and observation (downhole data).

Stacked Generalization Method

In this method, the estimates of shear wave velocity profiles in 4 boreholes provided by the base neural networks were combined using stacked generalization method. A feed-forward neural network was used to combine the results of the base networks. The input data of the combiner network consists of 10 profiles of shear wave velocity in 4 boreholes and the target data are the results of the downhole boreholes.

A three-layer network with sigmoid activation function for the hidden layers was used for the combiner. There are 13 neurons in the first hidden layer and 3 neurons in the second hidden layer. A correlation diagram was prepared for the estimated data (using stacked generalization) and the observation data (obtained from downhole operation). Furthermore, the RMSE (root-mean-square error) and regression coefficient R were calculated. The correlation chart is shown in Fig. 6.

The results of the averaging and stacked generalization methods are compared in Table 6. The results show that the RMSE of stacked generalization method is lower than the averaging method, while the regression coefficient (R) is higher and better. In other words, the estimated results for the depth profile of shear wave velocity in 4 boreholes are much better in the stacked generalization method.

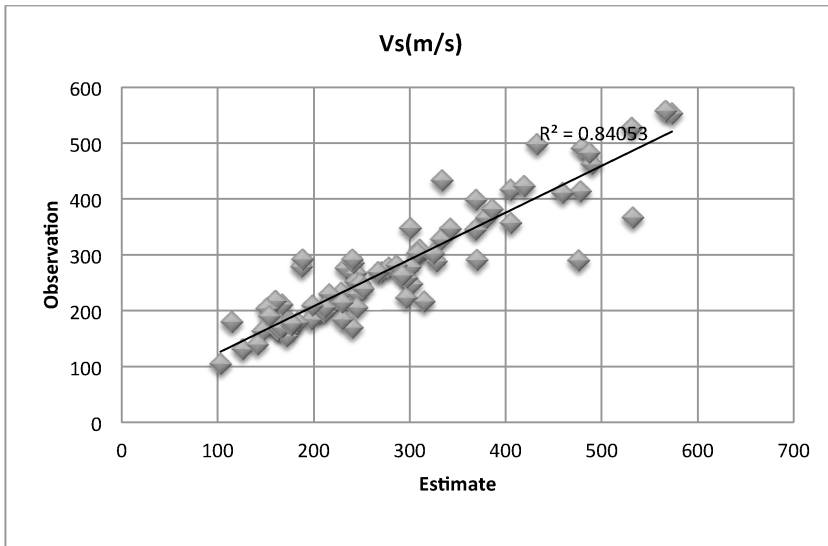


Fig. 6. Correlation graph for estimated data (stacked generalization combiner) and observation (downhole data).

Table 6. The results of averaging and stacked generalization methods.

Stacked Generalization method		Averaging method	
RMSE	R	RMSE	R
46.13067	0.8405	81.09671	0.5912

Soil classification in the study area

The earth layers are divided according to the shear wave velocity. In accordance with Eurocode 8 (Eurocode 8 , 2005), the average shear wave velocity in the upper 30 m is calculated using eq. (2), where d_i is the layer thickness, V_{si} is the shear wave velocity of layers, and n denotes the number of layers. In the study area, after estimating the shear wave velocity in 4 drilled boreholes and designated points, the average shear wave velocity of the study area was calculated. Results show the average shear wave velocities in top 30 m from surface, i.e., V_{S30} , in the study area varies between 187 m/s to 260 m/s. According to Eurocode 8 (Table 7) the study area is classified as category C.

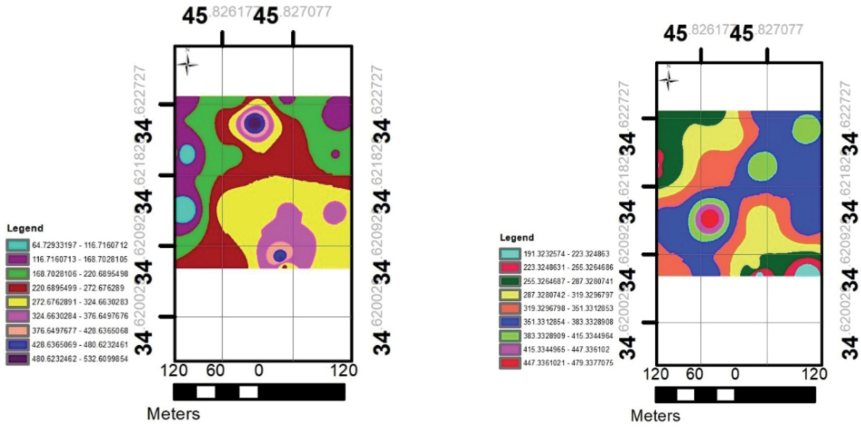
$$V_{S30} = \frac{30}{\sum_{i=1}^n \frac{d_i}{V_{si}}} \quad (2)$$

2 and 3-dimensional modeling of shear wave velocity in the study area

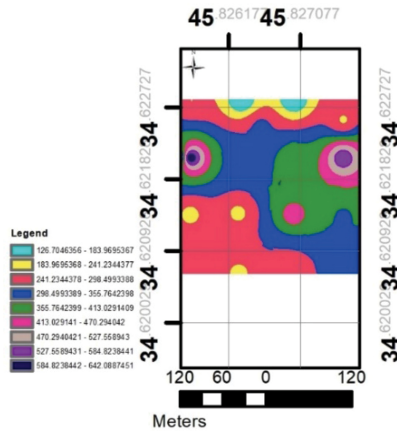
After estimating the depth of shear wave velocity (V_s), the two-dimensional maps containing the horizontal cross-sections at 10, 20 and 30m depths are plotted in Fig. 7(a, b, c). Also, Fig. 8 shows the V_{S30} in the study area. A 3-dimensional (3D) map of the shear wave velocity in the study area is also plotted in Fig. 9 (a, b).

Table 7. Soil classification according to Eurocode 8 (Eurocode 8., 2005).

Ground type	Description of stratigraphic profile	V_{s30} (m/s) Shear wave velocity	N_{spt} (standard penetration test)(blows/30cm)	C_u (kpa) (undrained shear strength)
A	Rock or other rock-like geological formation, including at most 5 m of weaker material at the surface.	>800	-	-
B	Deposits of very dense sand, gravel, or very stiff clay, at least several tens of meters in thickness, characterized by a gradual increase of mechanical properties with depth.	360-800	>50	>250
C	Deep deposits of dense or medium-dense sand, gravel or stiff clay with thicknesses from several tens to many hundreds of metres	180-360	15-50	70-250
D	Deposits of loose-to-medium cohesionless soil (with or without some soft cohesive layers), or of predominantly soft-to-firm cohesive soil.	<180	<15	<70
E	A soil profile consisting of a surface alluvium layer with v_s values of type C or D and thickness varying between about 5 m and 20 m, underlain by stiffer material with $v_s > 800$ m/s.			
S1	Deposits consisting of, or containing, a layer at least 10 m thick, of soft clays/silts with a high plasticity index ($PI > 40$) and a high water content.	<100	-	10-20
S2	Deposits of liquefiable soils, of sensitive clays, or any other soil profile not included in types A– E or S1.			



(a)



(c)

Fig. 7. The horizontal Cross section map of the study area at depth, (a): 10 m, (b): 20 m and (c): 30 m.

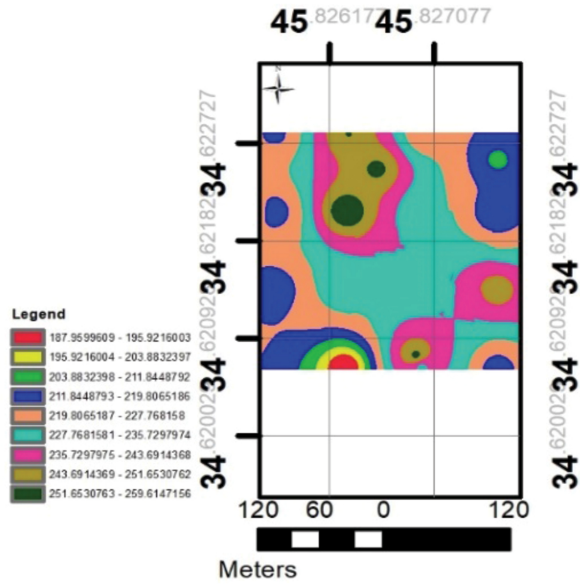


Fig. 8. The shear wave velocity in top 30 m from surface (V_{S30}) in the study area.

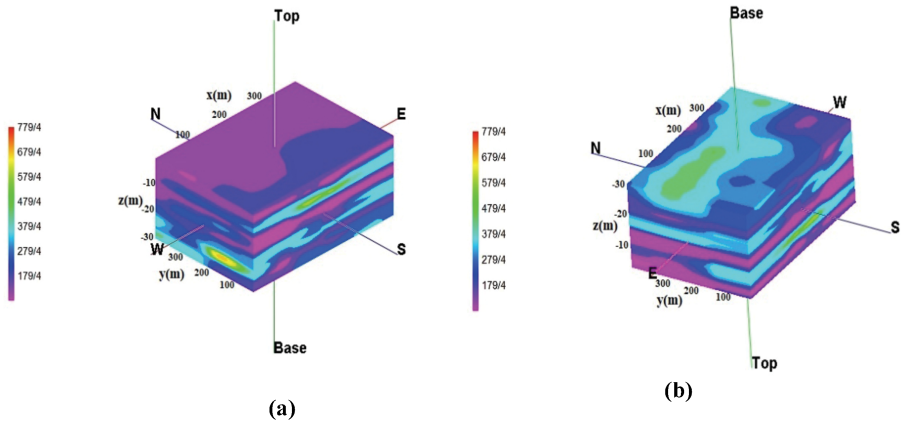


Fig. 9. 3D map of the estimated shear wave velocity in the study area, (a): top, (b): bottom.

Comparison of stacked generalization and Kriging geostatistical methods in Vs30 modelling

To explore the capability of the proposed method, i.e., stacked generalization method, the obtained results in the study area (Fig. 10) were compared with the Kriging geostatistical method. To this aim, the Kriging tool available on the GIS software package was used to estimate the S-wave velocities in target points (Fig. 11). However, the availability of a sufficient number of samples and proper spatial distribution of them to interpolate of variable values in the target points are the significant issues in the application of geostatistical methods. Based on the spread of target points and down holes locations, there is a poor horizontal data distribution. This issue impact Kriging and fails the results to converge a realistic estimation of Vs. In this situation, the use of neural networks because of its ability in the mapping complicated relationships is an appropriate choice, even in circumstances that there are few known data points. Based on geological information of the study area, estimation of Vs by the stacked generalization method has better performance than the Kriging geostatistical method.

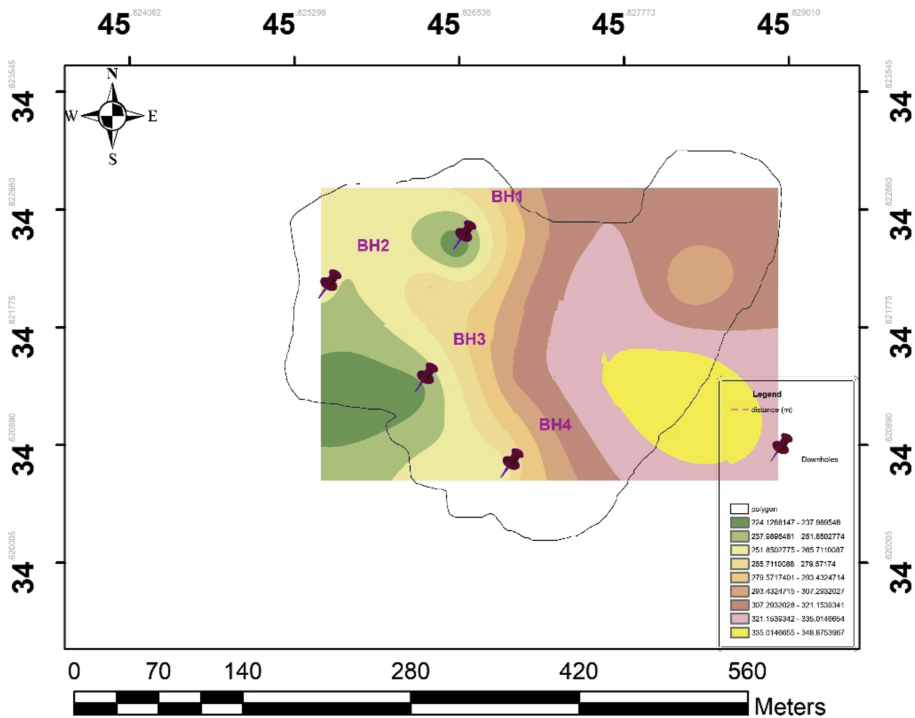


Fig. 10. Estimated Vs₃₀ in the study area by the stacked generalization method.

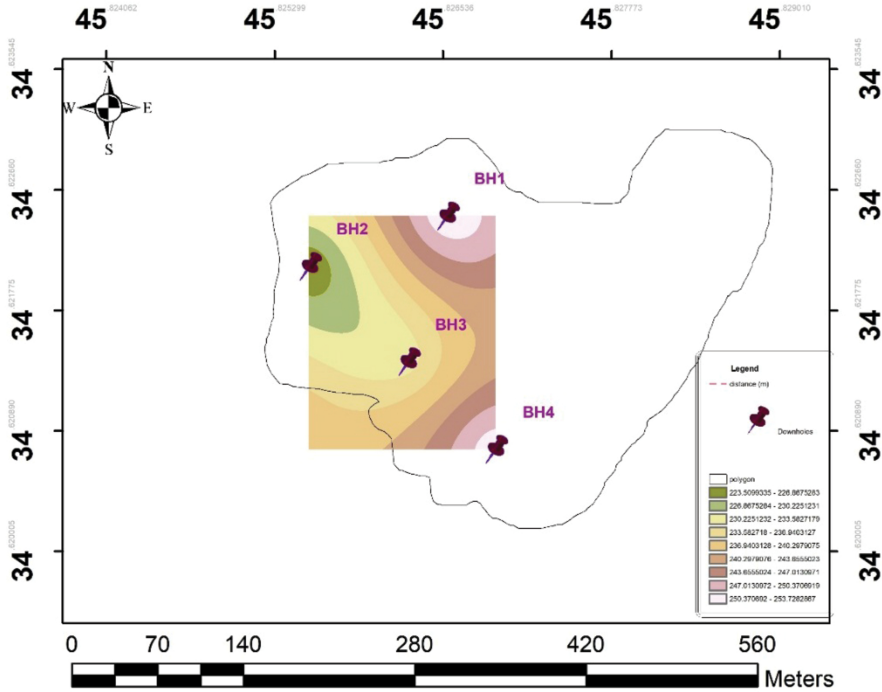


Fig. 11. Estimated V_{S30} in the study area by the Kriging geostatistical method.

CONCLUSIONS

The results show the success of the ensemble system in estimation of shear wave velocity in the study area. The ensemble system can be used to estimate the shear wave velocity, especially when there are few data items or when the data is very complex. The ensemble neural network method consists of two main components: (1) base neural networks, and (2) combiner, which combines the results of the base models. The results of the base neural networks are combined in two different ways, namely averaging and stacked generalization. Stacked generalization produced lower error and improved accuracy for the estimations.

Based on the results obtained for the shear wave velocity in the downhole boreholes and the results estimated by the ensemble system elsewhere in the region, we conclude that the approximate variations of the depth profile of shear wave velocity until the 30 m depth in the region range from 100 m/s to 800 m/s. As shown in Fig. 8, V_{S30} ranges from 187 m/s to 260 m/s. To probe the efficiency of the proposed method, the obtained

results in the study area were compared with the Kriging geostatistical method. The results show the ensemble neural networks in V_s modeling in comparison to the Kriging method has better performance, especially in circumstances that there are few datasets. In accordance with Eurocode 8, we conclude that the soil type is category C, which mainly comprises deep deposits of dense or medium-dense sand, gravel or stiff clay with thicknesses from several tens to many hundreds of meters. According to the geological data of the study area, the subsurface layers are mostly sand, clay and marl with approximate range of shear wave velocity from 100 m/s to 1500 m/s and possibly lime and gypsum with shear wave velocity range of 1100 m/s to 3300 m/s. Comparing the results of the depth profile of shear wave velocity in the region and the V_{s30} zoning in the region and its compliance with the geological information of the region, we conclude that the results of this study have a good agreement with the geological data of the area, and the probability of presence of gypsum and lime in this area is very low. Also, the results of this study confirm the high damage caused by the earthquake, because one of the major damages caused by the earthquake is the classification of the soil type at the site, as the lower the density of sediments, the more the earthquake destruction.

REFERENCES

- Abdideh, M., 2012. Estimation of permeability using artificial neural networks and regression analysis in an Iran oil field. *Internat. J. Physic. Sci.*, 7: 5308-5313.
- Agha Nabati, A., 1995. Description of the Geological Map of Kermansa. Geological Survey of Iran (GSI).
- Ansal, A., Kurtulus, A. and Tönük, G., 2010. Seismic microzonation and earthquake damage scenarios for urban areas. *Soil Dynam. Earthq. Engineer.*, 30: 1319-1328.
- Arava-Polo, M., Dahlke, T., Frogner, C., Zhang, C., Poggio, T. and Hohl, D., 2017. Automated fault detection without seismic processing. *The Leading Edge*, 36: 208-214.
- Borcherdt, R.D., 1994. Estimates of site-dependent response spectra for design (methodology and justification). *Earthq. Spect.*, 10: 617-617.
- Calderón-Macías, C., Sen, M.K. and Stoffa, P.L., 2000. Artificial neural networks for parameter estimation in geophysics. *Geophys. Prosp.*, 48: 21-47.
- Eurocode 8: Design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings, 2005. Brussels: European Committee for Standardization.
- Dasarathv, B.V. and Sheela, B.V., 1979. A composite classifier system design: concepts and methodology. *Proc. IEEE*, 67: 708-713.
- Hansen, L.K. and Salamon, P., 1990. Neural network ensembles. *IEEE Transact. Pattern Analvs. Mach. Intellig.*, 10: 993-1001.
- Havkin, S. and Havkin, S.S., 2009. *Neural networks and learning machines*. Vol. 3, Pearson, Upper Saddle River, NJ.
- Hunter, J.A., Beniumea, B., Harris, J.B., Miller, R.D., Pullan, S.E., Burns, R.A. and Good, R.L., 2002. Surface and downhole shear wave seismic methods for thick soil site investigations. *Soil Dynam. Earthq. Engineer.*, 22: 931-941.
- Hvorslev, M.J., 1949. Subsurface exploration and sampling of soils for civil engineering purposes. *Waterways Experiment Station*, 521.

- Internat. Soc. Soil Mechan. Foundat. Engineer. (ISSMGE). Technical Comm. Earthq. Geotechn. Engineer., 1993. Manual for Zonation on Seismic Geotechnical Hazards.
- Jacobs, R.A., Jordan, M.I., Nowlan, S.J. and Hinton, G.E., 1991. Adaptive mixtures of local experts. *Neural Computat.*, 3: 79-87.
- Kanli, A.I., Tildy, P., Prónav, Z., Pınar, A. and Hermann, L., 2006. VS30 mapping and soil classification for seismic site effect evaluation in Dinar region, SW Turkey. *Geophys. J. Internat.*, 165: 223-235.
- Liu, Y., Wang, Y., Li, Y., Zhang, B. and Wu, G., 2004. Earthquake prediction by RBF neural network ensemble. Abstracts, Internat. Symp. Neural Networks, ISNN.
- Mavko, G., 2005. Conceptual overview of rock and fluid factors that impact seismic velocity and impedance. Retrieved, 11: 2012.
- Nvein, C.Y., Hamada, G.M. and Elsakka, A., 2018. Artificial Neural Network (ANN) prediction of porosity and water saturation of shaly sandstone reservoirs. Abstracts, AAPG Asia Pacific Region, 4th AAPG/EAGE/MGS Oil and Gas Conf., Myanmar.
- Polikar, R., 2006. Ensemble based systems in decision making. *IEEE Circuits Systems Magaz.*, 6: 21-45.
- Earlev, K.H. and Rudenko, D., 1988. Use of geophysical methods in a geotechnical investigation. Abstr., 2nd Internat. Conf. Case Hist. Geotechn. Engineer., St.Louis, Missouri.
- Shimshoni, Y. and Intrator, N., 1998. Classification of seismic signals by integrating ensembles of neural networks. *IEEE Transact. Sign. Process.*, 46: 1194-1201.
- Van der Baan, M. and Jutten, C., 2000. Neural networks in geophysical applications. *Geophysics*, 65: 1032-1047.
- Wolpert, D.H., 1992. Stacked generalization. *Neural Netw.*, 5: 241-259.
- Zaharia, B., Radulian, M., Popa, M., Grecu, B., Bala, A. and Tataru, D., 2008. Estimation of the local response using the Nakamura method for the Bucharest area. *Roman. Rep. Phys.*, 60: 131-144.
- Zhou, Z.H., Wu, J. and Tang, W., 2002. Ensembling neural networks: many could be better than all. *Artif. Intellig.*, 137: 239-263.