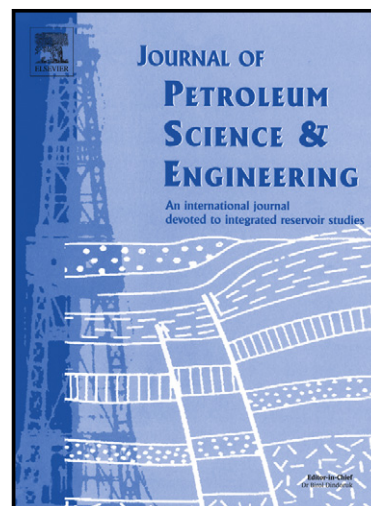


## Author's Accepted Manuscript

Support vector regression based determination of shear wave velocity

Parisa Bagheripour, Amin Gholami, Mojtaba Asoodeh



[www.elsevier.com/locate/petrol](http://www.elsevier.com/locate/petrol)

PII: S0920-4105(14)00395-7  
DOI: <http://dx.doi.org/10.1016/j.petrol.2014.11.025>  
Reference: PETROL2868

To appear in: *Journal of Petroleum Science and Engineering*

Received date: 19 April 2013  
Revised date: 12 October 2014  
Accepted date: 24 November 2014

Cite this article as: Parisa Bagheripour, Amin Gholami, Mojtaba Asoodeh, Support vector regression based determination of shear wave velocity, *Journal of Petroleum Science and Engineering*, <http://dx.doi.org/10.1016/j.petrol.2014.11.025>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Support Vector Regression Based Determination of Shear Wave Velocity

Parisa Bagheripour<sup>1</sup>; Amin Gholami<sup>2</sup>; Mojtaba Asoodeh<sup>3\*</sup>

<sup>1</sup> Department of Petroleum Engineering, Gachsaran Branch, Islamic Azad University, Gachsaran, Iran

<sup>2</sup> Abadan Faculty of Petroleum Engineering, Petroleum University of Technology, Abadan, Iran

<sup>3</sup> Islamic Azad University, Birjand Branch, Birjand, Iran

Address Correspondence to MojtabaAsoodeh, Islamic Azad University, Birjand Branch, Birjand, Iran, Email: Asoodeh.mojtaba@gmail.com , Tel.: +989395161149

### Abstract

Shear wave velocity in the company of compressional wave velocity add up to an invaluable source of information for geomechanical and geophysical studies. Although compressional wave velocity measurements exist in almost all wells, shear wave velocity is not recorded for most of elderly wells due to lack of technologic tools in those days and incapability of recent tools in cased holes. Furthermore, measurement of shear wave velocity is to some extent costly. This study proposes a novel methodology to remove aforementioned problems by use of support vector regression tool originally invented by Vapnik and his co-workers (1995). Support vector regression (SVR) is a supervised learning algorithm plant based on statistical learning (SLT) theory. It is used in this study to formulate conventional well log data into shear wave velocity in a quick, cheap, and accurate manner. SVR is preferred for model construction because it utilizes structural risk minimization (SRM) principle which is superior to empirical risk minimization (ERM) theory, used in traditional learning algorithms such as neural networks. A group of 2879 data points was used for model construction and 1176 data points were employed for assessment of SVR model. A comparison between measured and SVR predicted data showed SVR was

capable of accurately extract shear wave velocity, hidden in conventional well log data. Finally, a comparison among SVR, neural network, and four well-known empirical correlations demonstrated SVR model outperformed other methods. This strategy was successfully applied in one of carbonate reservoir rocks of Iran Gas-Fields.

**Keywords:** Shear Wave Velocity; Support Vector Regression (SVR); Structural Risk Minimization (SRM); Empirical Risk Minimization (ERM); Conventional well logs; Rock Mechanics

## 1. Introduction

Sonic measurements in hydrocarbon wells provide precious information for rock mechanical and geophysical studies. Compressional wave velocity is easily recorded and is available for all wells. However, measurement of shear wave velocity is more complicated and these measurements are not available in old wells owing to lack of technologic tools in those days. Run of recent tools in old wells is not practical for most of them due to prevailing casing completion. Therefore, a quantitative formulation between conventional well logs (available in all wells) and shear wave velocity eliminates the mentioned problems and makes it possible to perform geophysical and geomechanical studies. Combination of shear and compressional wave velocities measurements adds up to invaluable source of information for lithology identification (Pickett 1963), rock mechanical properties calculation (Eaton 1972; Kumar 1976; Chang et al. 2006; Ameen et al. 2009), and pore type identification (Eberli et al. 2003). Due to significance of subject several researchers have tried to establish empirical correlations estimating shear wave velocity (Pickett, 1963; Tosaya and Nur, 1982; Castagna et al., 1985; Han, 1986; Eberhart-

Phillips, 1989; Castagna et al., 1993; Anselmetti and Eberli, 1993; Eskandari et al., 2004; Brocher, 2005).

Recent studies have proved the superiority of intelligent systems to empirical and statistical approaches in geosciences and petroleum related problems. A growing tendency is observed among researchers to utilize intelligent systems in solving their problems of various fields (Mohaghegh et al., 2000; Saggaf and Nebrija 2003; Artun et al., 2005; Kadkhodaei-Illkchi et al., 2008; Asoodeh and Bagheripour, 2012a). Several researchers suggested estimation of shear wave velocity from conventional well logs using traditional learning algorithm such as neural networks which use empirical risk minimization (ERM) principle (Rezaee et al. 2006; Rajabi et al. 2009; Asoodeh and Bagheripour, 2012b). In this study, shear wave velocity is estimated from conventional well log data using support vector regression (SVR). SVR utilizes structural risk minimization (SRM) in conjunction with ERM. Therefore, it produces more reliable results compared with neural networks that solely use ERM principle. SVR model was compared with neural network and four well-known empirical correlations. Results confirm superiority of SVR to other methods. This methodology was successfully implemented to Asmari carbonate reservoir rocks, the major reservoir of Iranian Oil Fields. Top of the reservoir formation is varying in range of 2983m to 2996m in field of our study. Therefore, there is a compaction knowing approximate 1 psi/ft overburden pressure.

## **2. Method: Support Vector Regression**

Support vector regression was introduced as a machine learning technique by Vapnik (1995). SVR has been deemed as an arresting tool featuring promising applications owing to its superior capability in successfully solving large variety of nonlinear regression problems. SVR method

utilizes structural risk minimization principle in addition to supplanted empirical risk minimization principle that traditionally has been used by neural networks with a view to developing an accurate model (Al-Anazi and Gates, 2010a; El-Sebakhy, 2009; Jiang and Zhao 2013; Liao et al. 2011; Ustun et al. 2005; Wu and Law 2010). An elaboration on SVR underlying structure is brought as follows. In SVR regression, the ultimate goal is to find linear relation between  $n$ -dimensional input vectors  $x \in R^n$  and output variables  $y \in R$  as follow:

$$f(x) = w^T x + b \quad (1)$$

Where,  $w$  and  $b$  are the slope and offset of the regression line respectively. For determining the values of  $b$  and  $w$ , it is necessary to minimize following equation:

$$R = \frac{1}{2} \|w\|^2 + \frac{C}{l} \sum_{i=1}^l |y_i - f(x_i)|_\epsilon \quad (2)$$

Loss function, used in SVR is  $\epsilon$ -insensitive which has been introduced by Vapnik (1995) as below:

$$|y_i - f(x_i)|_\epsilon = \begin{cases} 0 & \text{if } |y_i - f(x_i)| \leq \epsilon \\ |y_i - f(x_i)| - \epsilon & \text{Otherwise} \end{cases} \quad (3)$$

This problem can be reformulated in a dual space by:

$$\text{Maximize } L_p(\alpha_i, \alpha_i^*) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_i^T x_j - \epsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i \quad (4)$$

$$\text{Subject to } \begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l \\ 0 \leq \alpha_i^* \leq C, \quad i = 1, \dots, l \end{cases} \quad (5)$$

Where,  $\alpha_i, \alpha_i^* \geq 0$  are positive Lagrange multipliers.  $C$  is regulated positive parameter which determines trade-off between approximation error and the weight vector norm  $\|w\|$ . After calculation of Lagrange multipliers  $\alpha_i$  and  $\alpha_i^*$ , training data points, those meeting the conditions  $\alpha_i - \alpha_i^* \neq 0$ , will be employed to construct the decision function. Hence, the best linear hyper surface regression is given by:

$$f(x) = w_o^T x + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i^T x + b \quad (6)$$

Which desired weight vector of the regression hyper plane is given by:

$$w_o = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i \quad (7)$$

In nonlinear regression, Kernel function is employed for mapping input data onto higher dimensional feature space in order to generate a linear regression hyper plane. Polynomial, radial basis function (RBF), and sigmoid are the common kernel functions in SVR. In the case of the nonlinear regression, the learning problem is again formulated in the same way as in a linear case, i.e., the nonlinear hyperplane regression function becomes:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (8)$$

In above equation,  $K(x_i, x)$  is kernel function which is defined as follow:

$$k(x_i, x_j) = \Phi^T(x_i) \Phi(x_j) \quad i, j = 1, \dots, l \quad (9)$$

Where,  $\Phi(x_i)$  and  $\Phi(x_j)$  are projection of the  $x_i$  and  $x_j$  in feature space respectively. For simplicity, a brief description of SVR was explained. More detailed studies about SVR are

provided in several papers and reviews which readers can refer to (Al-Anazi and Gates, 2010b; Kecman 2005, 2006; Mousavi et al. 2013; Vogt and Kecman, 2005).

### 3. Input Selection by Sensitivity Analysis

Using a back-propagation neural network, Dutta and Gupta (2010) suggested a stable method based on partial derivative of output with respect to  $i^{th}$  input to find relative contribution of each input in estimating output. Partial derivative of output with respect to  $i^{th}$  input is evaluated using following equation:

$$\frac{\partial V_s}{\partial x_i} = \sum_j W_{oj} (1 - h_j^2) W_{ji} \quad (10)$$

Where,  $\frac{\partial V_s}{\partial x_i}$  is partial derivative of shear wave velocity with respect to  $i^{th}$  input,  $W_{oj}$  is weight between output neuron and  $j^{th}$  hidden neuron and  $h_j$  is the response of  $j^{th}$  neuron in the hidden layer. Relative contribution of back-propagation neural network inputs is calculated by sum of the squares of the partial derivatives ( $S$ ) as follow:

$$S_i = \sum_{j=1}^N \left[ \left( \frac{\partial V_s}{\partial x_i} \right) \right]_j \quad (11)$$

$$RCi = \frac{S_i}{\sum_i S_i} \times 100 \quad (12)$$

Where,  $RCi$  is relative contribution of  $i^{th}$  input.

To achieve influence of each input in estimation of shear wave velocity, an improved strategy was followed and subsequently optimal number of inputs was evaluated. Firstly, a feed forward

back-propagation neural network was constructed using all available well logs and a sensitivity analysis was performed to compute RC value for each input as is shown in Table 1. In spite of correlation coefficient which is a qualitative criterion for illustrating dependency between inputs and output, sensitivity results are quantitative norms and are more reliable. In next step, RC values were used for ranking inputs. In SVR model, optimal number of introduced inputs is a crucial design factor. Therefore, conventional well logs were introduced into SVR model one by one according to their RC values and performance of SVR model was evaluated for each set of inputs. Results indicated that optimal SVR model is achieved when four inputs, including compressional wave slowness, neutron porosity, bulk density, and true resistivity are used. The mentioned procedure is summarized in a flowchart, shown in Fig.1. Readers unfamiliar to back-propagation neural networks (NNs) can refer to a work by Mohaghegh (2000) for more detailed study about NNs. In this study, we were to model the simplest way of formulating conventional well log data to shear wave velocity. Knowing that conventional well logs implicitly records effects of lithology changes, we didn't include lithology as input. Several studies have been done to show conventional well log data contain invaluable lithology information in their records (e.g., Delfiner et al., 1987; Cabello et al., 2010; Sfidari et al., 2014).

#### 4. Results and Discussion

An epsilon support vector regression ( $\epsilon$ -SVR) algorithm was employed for construction a model meant to estimate shear wave velocity from conventional well log data, including compressional wave slowness, neutron porosity, bulk density, and true resistivity. The primary task which should be done before SVR model construction is data normalization in range of  $[-1 \ 1]$ . This task reduces confusion to SVR model due to better performance of kernel functions and enhances the accuracy of final prediction. A group of 2879 data points belonging to one well was used for



model construction. Previous works demonstrates that radial basis functions (RBF) are the most appropriate choice for kernel function owing to fewer parameters to be tuned and low computational cost (Keerthi and Lin, 2003). Therefore, RBF was used as kernel function for SVR model construction. Performance of SVR model is strongly governed by involved parameters in SVR model and kernel function (C, Gamma, and Epsilon). Therefore, a thorough survey for determining such parameters is desired. You et al. (2010) suggests carrying out this survey through combination of grid search and pattern search techniques such that the grid search determines the area surrounding the optimal point and pattern search finds global optimal point within the found area by grid search. The specified search range for “C”, “Gamma”, and “Epsilon” were [0.1 500000], [0.000001 20], and [0.0001 100] respectively, while the extracted optimal points from these ranges are 125814.37412, 0.198179, and 0.014732, correspondingly. For this study, 2879 data points were employed as training data for model construction. Hence 2879 Lagrange multiplier pairs  $(\alpha_i, \alpha_i^*)$  is determined during the training the model. Among all extracted multiplier pairs, 2740 Lagrange multiplier pairs which meet the condition  $\alpha_i - \alpha_i^* \neq 0$  will be employed to build the decision function. In other word, number of support vectors used by this model is 2740. Moreover, in SVR algorithm the dimension of Hessian matrix is two times of number of input data for training model, i.e., (5758, 5758). After model construction, 1176 data points from another well were employed for assessment of SVR model. After model construction two powerful concepts, including correlation coefficient and error distribution were employed for assessment of the SVR performance. Fig.2 shows crossplot of measured shear wave velocity versus predicted values. Correlation coefficient for SVR prediction is equal to 0.9716(R-square=0.944) which verifies robustness of SVR Model. Fig.3 allows more statistical analysis of SVR performance using error distribution information. Mean and standard deviation

of error distribution are in turn equal to 0.0000 and 0.0734 which are relatively small values. It means error of 68% of samples is in range of  $0.0000 \pm 0.0734$  that is an acceptable error for shear wave velocity. Fig.4 provides an opportunity to compare measured shear wave velocity versus predicted values for all samples. This figure states SVR predicted values are in good agreement with reality. Eventually, relative error (error percentage) for each sample is evaluated and demonstrated in Fig.5. Relative error for most data points is located in range of  $[-5\% \ 5\%]$ , which is an acceptable value.

## 5. Comparison among SVR Model, Neural Network, and Empirical correlations

In the latter stage of this study, a comparison among SVR model, neural network, and empirical correlations, proposed by Pickett (1963), Castagna et al. (1993), Eskandari et al. (2004), and Brocher (2005) was performed. Following equations show the used empirical correlations.

Pickett (1963):

$$v_s = v_p / 1.9 \quad (13)$$

Castagna et al. (1993):

$$v_s = -0.05509v_p^2 + 1.0168v_p - 1.0305 \quad (14)$$

Eskandari et al. (2004):

$$v_s = -0.1236v_p^2 + 1.612v_p - 2.3057 \quad (15)$$

Brocher (2005):

$$v_s = 0.7858 - 1.2344v_p + 0.7949v_p^2 - 0.1238v_p^3 + 0.0064v_p^4 \quad (16)$$

Where,  $v_p$  refers to compressional wave velocity. Different statistical concepts, including correlation coefficient (R), average relative error (ARE), average absolute relative error (AARE), and root mean square error (RMSE) were employed to carry out this comparison. Following equations show the mentioned statistical tools:

$$R = \sqrt{1 - \frac{\sum_{i=1}^N [(V_e - V_m)_i]^2}{\sum_{i=1}^N [(V_e)_i - \bar{V}_m]^2}} \quad (17)$$

$$ARE = \frac{\sum_{i=1}^N \left( \frac{V_e - V_m}{V_m} \right)}{N} \times 100 \quad (18)$$

$$AARE = \frac{\sum_{i=1}^N \left| \frac{V_e - V_m}{V_m} \right|}{N} \times 100 \quad (19)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N \left[ \left( \frac{V_e - V_m}{V_m} \right)_i \right]^2}{N}} \times 100 \quad (20)$$

Where,  $V_e$ ,  $V_m$ ,  $\bar{V}_m$  and  $N$  are estimated shear wave velocity, measured shear wave velocity, mean of measured shear wave velocity, and number of testing data points, respectively. Results of this comparison are shown in Table 2. As it is seen in Table 2, SVR outperformed other methods owing to higher R and lower ARE, AARE, and RMSE.

## 6. Conclusions

Shear wave velocity can provide valuable data for reservoir characterization, and geomechanical and geophysical studies when it is used in conjunction with compressional wave velocity. Due to significance of calling for shear wave velocity knowledge, several researchers attempted to

determine shear wave velocity through empirical correlations and/ or traditional intelligent systems. Nonetheless, the quest for precision as much as possible offers looking for high accuracy methods. In this study, support vector regression method was employed for responding to this quest. SVR was utilized to formulate conventional well log data, including compressional wave slowness, bulk density, true resistivity, and neutron porosity into shear wave velocity. Results indicated SVR model performed satisfyingly and it was capable of mining hidden knowledge about shear wave velocity from conventional well logs. Since SVR utilizes structural risk minimization (SRM) in conjunction with ERM, it was expected that SVR model performs better than traditional learning algorithm such as neural network. A comparison between SVR and previous works, including neural network and four well-known empirical correlations verified superiority of SVR model. The comparison showed SVR model has a higher correlation coefficient and lower average relative error, absolute average relative error, and root mean square error. Finally, implementation of proposed methodology can produce shear wave velocity for elderly and/ or cased holes where no shear wave measurement is done. Applying SVR model for new wells can significantly reduces costs and saves time.

## References

- Al-Anazi, A.F., Gates, I.D., 2010a. Support vector regression for porosity prediction in a heterogeneous reservoir: A comparative study. *Computer and Geosciences* 36, 1494-1503.
- Al-Anazi, A.F., Gates, I.D., 2010b. A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. *Engineering Geology* 114, 267-277.

- Ameen, M.S., Smart, B.G.D., Somerville, J.M.C., Hammilton, S., Naji, N.A., 2009. Predicting rock mechanical properties of carbonates from wireline logs (A case study: Arab-D reservoir, Ghawar field, Saudi Arabia). *Journal of Marine and Petroleum Geology* 26, 430–444.
- Anselmetti, F.S., Eberli, G.P., 1993. Controls on Sonic Velocity in Carbonates. *Pure and Applied Geophysics* 141, 287–323.
- Artun, E., Mohaghegh, S., Wilson, T., 2005. Reservoir Characterization using Intelligent Seismic Inversion. SPE Paper presented in Eastern regional meeting held in Morgantown, 14-16 September, SPE 98012.
- Asoodeh, M., Bagheripour, P., 2012a. Prediction of Compressional, Shear, and Stoneley Wave Velocities from Conventional Well Log Data Using a Committee Machine with Intelligent Systems. *Rock Mech Rock Eng* 45, 45-63.
- Asoodeh, M., Bagheripour, P., 2012b. Estimation of Bubble Point Pressure from PVT Data Using a Power-Law Committee with Intelligent Systems. *J Pet Sci Eng* 90-91: 1-11.
- Brocher, T.M., 2005. Empirical Relations between Elastic Wave Speeds and Density in the Earth's Crust. *Bulletin of the Seismological Society of America* 95 (6), 2081–2092.
- Cabello, P., Falivene, O., Lopez-Blanco, M., Arbues, P., Ramos, E., 2010. Modelling facies belt distribution in fan deltas coupling sequence stratigraphy and geostatistics: The Eocene Sant Llorenç del Munt example (Ebro foreland basin, NE Spain). *Marine and Petroleum Geology* 27(1), 254-272.
- Castagna, J.P., Batzle, M.L., Eastwood, R.L., 1985. Relationship between Compressional and Shear Wave Velocities in Silicate Rocks. *Geophysics* 50, 571–581.

Castagna, J.P., Batzle, M.L., Kan, T.K., 1993. Rockphysics -the Link between Rock Properties and AVO Response. In: Castagna JP, Backus MM (Eds.), Offset- Dependent Reflectivity—Theory and Practice of AVO Analysis. Society of Exploration Geophysicists, Tulsa, pp. 135–171.

Chang, C., Zoback, M.D., Khaksar, A., 2006. Empirical Relations between Rock Strength and Physical Properties in Sedimentary Rocks. *Journal of Petroleum Science and Engineering* 51, 223–237.

Delfiner, P., Olivier, P., Oberto, S., 1987. Automatic Determination of Lithology From Well Logs. *SPE Formation Evaluation* 2(3), 303-310.

Dutta, S., Gupta, J. P., 2010. PVT correlations for Indian crude using artificial neural networks. *Journal of Petroleum Science and Engineering* 72, 93-109.

Eaton, B.A., 1972. Graphical Method Predicting Pressure Worldwide. *World oil* 185, 51-56.

Eberhart-Phillips, D.M., 1989. Investigation of Crustal Structure and Active Tectonic Processes in the Coast Ranges, Central California. PH.D. dissertation, Stanford University.

Eberli, G.P., Baechle, G.T., Anselmetti, F.S., Incze, M.L., 2003. Factors controlling elastic properties in carbonate sediments and rocks. *The Leading Edge* 22 (7), 654–660.

El-Sebakhy, E.A., 2009. Forecasting PVT properties of crude oil systems based on support vector machines modeling scheme. *Journal of Petroleum Science and Engineering* 64, 25-34.

Eskandari, H., Rezaee, M.R., Mohammadnia, M., 2004. Application of Multiple Regression and Artificial Neural Network Techniques to Predict Shear Wave Velocity from Wireline Log Data

for a Carbonate Reservoir, South Iran. Canadian Society of Exploration Geophysicists (CSEG) Recorder 29, 42–48.

Han, D.H., 1986. Effects of porosity and clay content on acoustic properties of sandstone and unconsolidated sediments. Ph.D. dissertation, Stanford University.

Jiang, B.T., Zhao, F.Y., 2013. Combination of support vector regression and artificial neural networks for prediction of critical heat flux. International Journal of Heat and Mass Transfer. 62, 481-494.

Kadkhodaei-Illkchi, A., Rahimpour-Bonab, H., Rezaee, M.R., 2008. A committee machine with intelligent systems for estimation of total organic carbon content from petrophysical data: an example from Kangan and Dallan reservoirs in South Pars Gas Field, Iran. Computer and Geosciences 35, 459-474.

Kecman, V., 2005. Chapter ' Support vector machines-an introduction ', in a Springer-Verlag book, ' Support Vector Machines: Theory and Applications', Ed. L. Wang, Series: Studies in Fuzziness and Soft Computing, Vol. 177, pp. 1-47.

Kecman, V., 2006. High Dimensional Function Approximation (Regression, Hypersurface Fitting) by an Active Set Least Squares Learning Algorithm. Faculty of Engineering, University of Auckland, New Zealand.

Keerthi, S.S., Lin, C.J., 2003. Asymptotic behavior of support vector machines with Gaussian kernel. Neural Computing. 15(7), 1667-1689.

Kumar, J., 1976. The Effect of Poisson's Ratio on Rock Properties. In: 51st Annual Fall Technical Conference and Exhibition of the Society of Petroleum Engineers of AIME (New Orleans, Oct. 3-6), SPE 6094.

Liao, R., Zheng, H., Grzybowski, S., Yang, L., 2011. Particle swarm optimization-least squares support vector regression based forecasting model dissolved gases in oil-filled power transforms. *Electric Power Systems Research* 81.2074-2080.

Mohaghegh, S.D., 2000. Virtual-Intelligence Applications in Petroleum Engineering: Part 1—Artificial Neural Networks. *Journal of Petroleum Technology* 52 (9), 64-73.

Mohaghegh, S.D., Goddard, C., Popa, A., Ameri, S., Bhuiyan, M., 2000. Reservoir Characterization through Synthetic Logs. SPE Eastern Regional Meeting (Morgantown, West Virginia, 17–19 October), SPE 65675.

Mousavi, S.M., Tavakkoli-Moghaddam, R., Vahdani, B., Hashemi, H., Sanjari, M.J., 2013. A new support vector model-based imperialist competitive algorithm for time estimation in new product development projects 29. 157-168.

Pickett, G.R., 1963. Acoustic character logs and their applications in formation evaluation. *Journal of Petroleum Technology* 15, 650-667.

Rajabi, M., Bohloli, B., Ahangar, E., 2009. Intelligent approach for prediction of compressional, shear and stoneley wave velocities from conventional well log data: A case study from the Sarvak carbonate reservoir in the Abadan plain (Southern Iran). *Computer and Geosciences* 36, 647-664.



Rezaee, M.R., Kadkhodaie-Ilkhchi, A., Barabadi, A., 2006. Prediction of shear wave velocity from petrophysical data utilizing intelligent systems: an example from a sandstone reservoir of Canarvon Basin, Australia. *Journal of Petroleum Science and Engineering* 55, 201–212.

Saggaf, M.M., Nebrija, E.d.L., 2003. A fuzzy approach for the estimation of facies from wireline logs. *AAPG Bulletin* 87, 1223–1240.

Sfidari, E., Kadkhodaie-Ilkhchi, A., Rahimpour-Bonab, H., Soltani, B., 2014. A hybrid approach for lithofacies characterization in the framework of sequence stratigraphy: A case study from the South Pars gas field, the Persian Gulf basin. *J Petrol Sci Eng* 121, 87-102.

Tosaya, C., Nur, A., 1982. Effects of diagenesis and clays on compressional velocities in rocks. *Geophys Res Lett* 9, 5-8.

Ustun, B., Melssen, W.J., Oudenhuijzen, M., Buydens, L.M.C., 2005. Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization. *Analytica Chimica Acta* 544, 292-305.

You, Z., Yin, Z., Han, K., Huang, D., Zhou, X., 2010. A semi-supervised learning approach to predict synthetic genetic interactions by combining functional and topological properties of functional gene network. *BMC Bioinformatics* 11, 343-355.

Vapnik, V., 1995. *The Nature of Statistical Learning Theory*. Springer, New York.

Vogt, M., Kecman, V., 2005. Chapter 'Active-Set Methods for Support Vector Machines', in a Springer Verlag book, 'Support Vector Machines: Theory and Applications', Ed. L. Wang, Series: *Studies in Fuzziness and Soft Computing*, Vol. 177, pp. 133-158

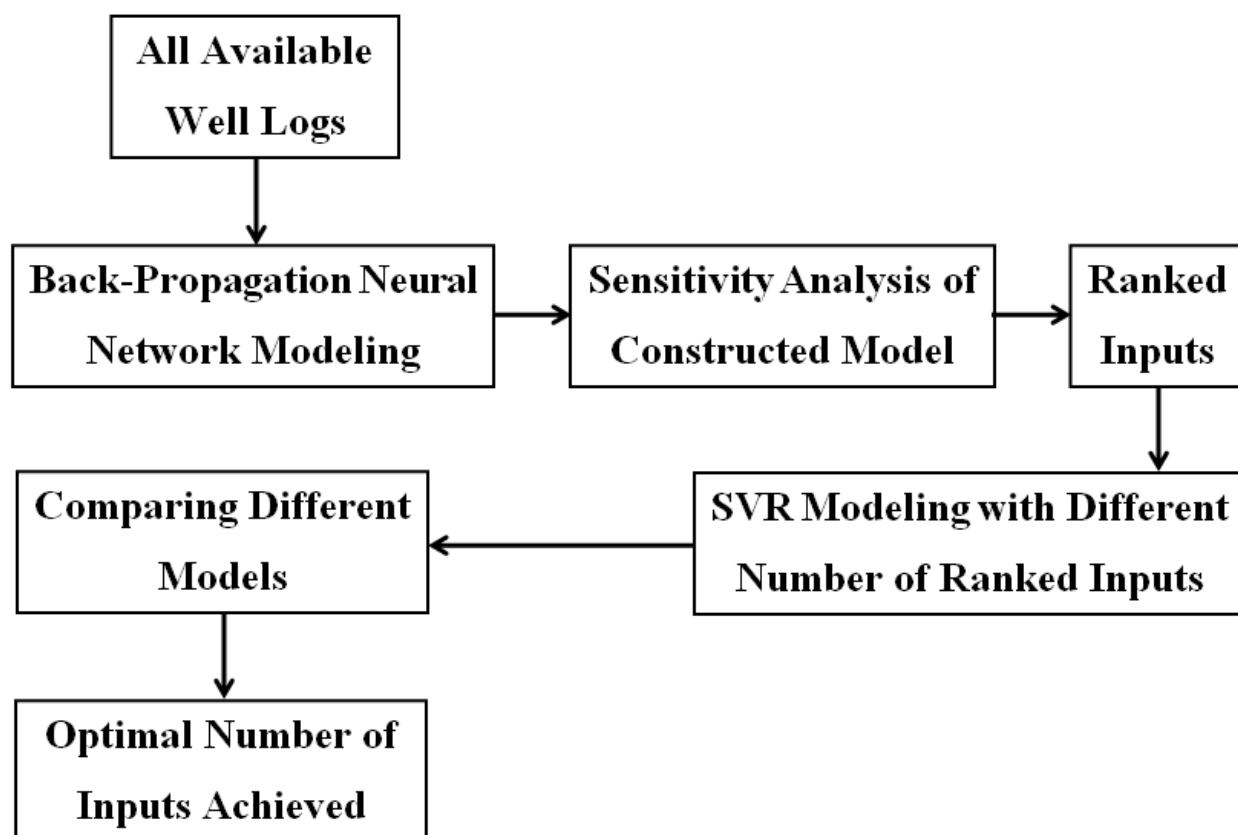
Wu, Q., Law, R., 2010. Fuzzy support vector regression machine with penalizing Gaussian noises on triangular fuzzy number space. *Expert Systems with Applications* 37, 7788-7795.

**Table 1:** Relative contribution of each input in shear wave velocity estimation, based on sensitivity analysis and correlation coefficient concept.

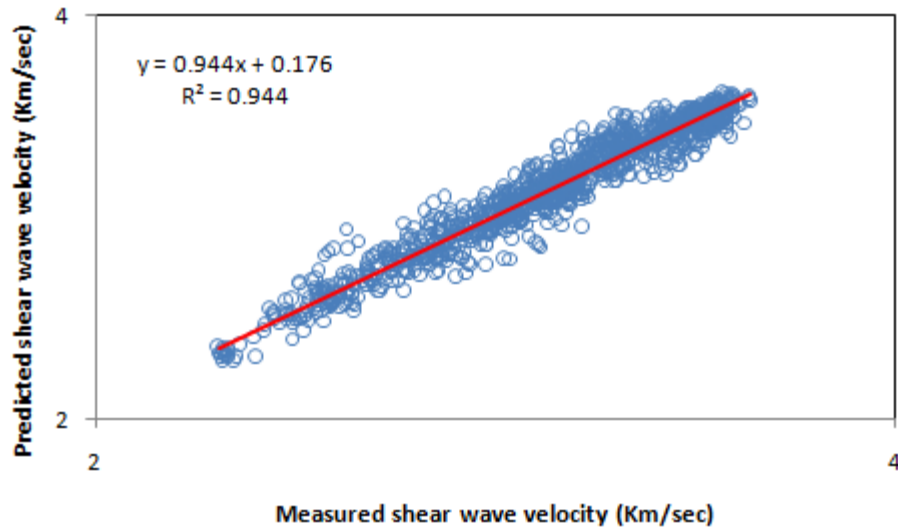
Conventional well logs	Relative Contribution	Correlation Coefficient
Compressional wave slowness (DT)	41.03%	0.76
Bulk density (RHOB)	23.73%	0.51
Neutron porosity (NPHI)	18%	0.31
True resistivity (RT)	12.41%	0.19
Photoelectric factor (PEF)	2.53%	0.23
Shallow resistivity (RS)	1.84%	0.04
Gamma ray (GR)	0.46%	0.11

**Table 2:** Comprising SVR model with neural network and four well-known empirical correlations based on correlation coefficient (R), average relative error (ARE), absolute average relative error (AARE), and root mean square error (RMSE).

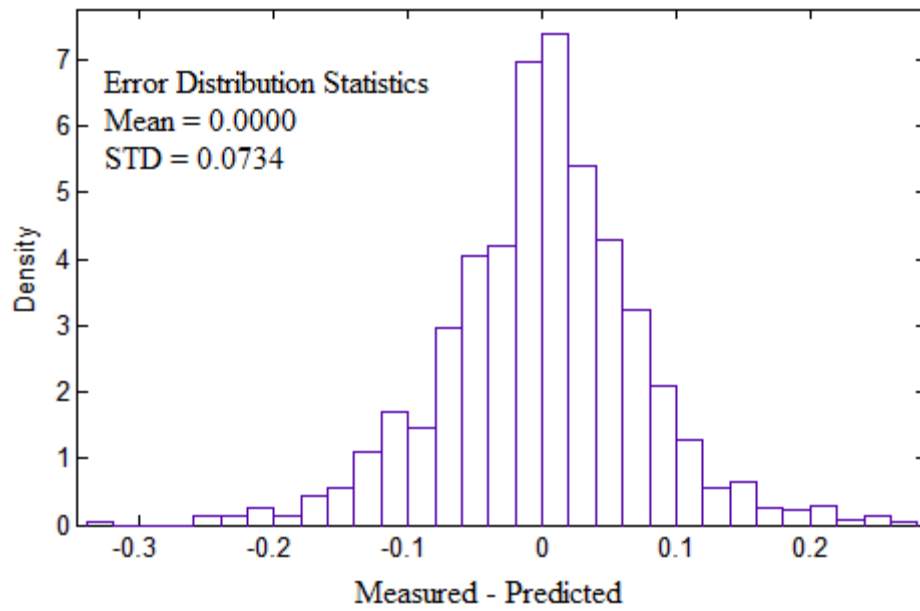
Method	R	ARE	AARE	RMSE
SVR	0.9716	-0.0567	1.7595	0.0733
NN	0.956	-0.0682	1.8691	0.0793
Pickett	0.947	-3.0211	3.3028	0.1145
Castagna et al	0.941	-4.6908	4.6955	0.1761
Eskandari et al	0.8819	-9.2848	9.2848	0.3550
Brocher	0.9359	8.6817	8.9723	0.9241



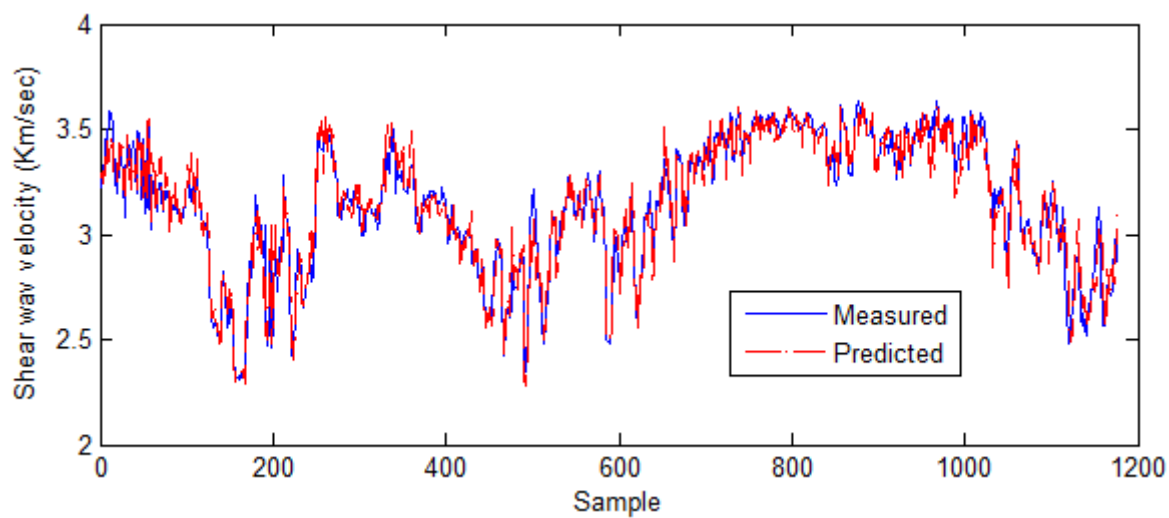
**Figure 1:** flowchart showing input selection via sensitivity analysis.



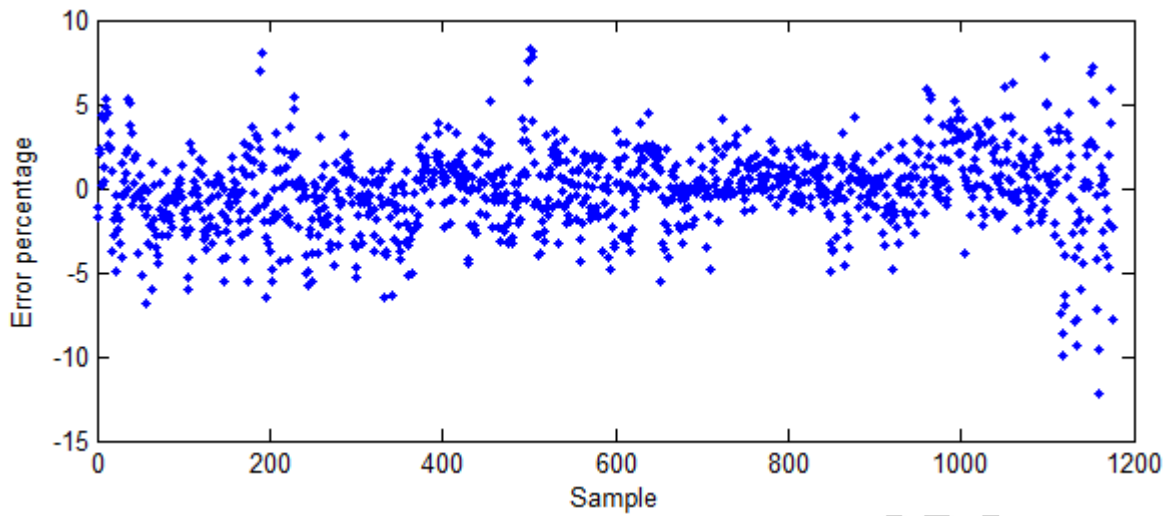
**Figure 2:** Crossplot showing correlation coefficient between measured and SVR predicted shear wave velocity. High value of correlation coefficient, i.e. 0.9716 ( $R$ -square=0.944) proves the robustness of SVR modeling.



**Figure 3:** Error distribution statistics for SVR model meant to predict shear wave velocity. Small values of mean and standard deviation (STD) reveal high performance of SVR modeling. Error distribution indicates 68% of predicted values have errors in range of  $0.0000 \pm 0.0734$ .



**Figure 4:** A comparison between measured and SVR predicted shear wave velocity versus different samples. Results indicate high match between measured and predicted values.



**Figure 5:** Relative error (error percentage) of SVR model in prediction of shear wave velocity for each sample. Results indicate relative error is located in range of  $[-5\% \ 5\%]$  for most data points.

- In this study, shear wave velocity ( $v_s$ ) was predicted from conventional well log data.
- Support vector regression (SVR) algorithm was used for model construction.
- Results of SVR was compared with those of neural network and empirical correlations
- Comparison showed superiority of SVR algorithm to other methods.
- Implementation of SVR model in wells with no  $v_s$  data reduces costs and saves time