On Support Vector Regression to Predict Poisson's Ratio and Young's Modulus of Reservoir Rock

A.F. Al-Anazi and I.D. Gates

Abstract Accurate prediction of rock elastic properties is essential for wellbore stability analysis, hydraulic fracturing design, sand production prediction and management, and other geomechanical applications. The two most common required material properties are Poisson's ratio and Young's modulus. These elastic properties are often reliably determined from laboratory tests by using cores extracted from wells under simulated reservoir conditions. Unfortunately, most wells have limited core data. On the other hand, wells typically have log data. By using suitable regression models, the log data can be used to extend knowledge of core-based elastic properties to the entire field. Artificial neural networks (ANNs) have proven to be successful in many reservoir characterization problems. Although nonlinear problems can be well resolved by ANN-based models, extensive numerical experiments (training) must be done to optimize the network structure. In addition, generated regression models from ANNs may not perfectly generalize to unseen input data. Recently, support vector machines (SVMs) have proven successful in several real-world applications for its potential to generalize and converge to a global optimal solution. SVM models are based on the structural risk minimization principle that minimizes the generalization error by striking a balance between empirical training error and learning machine capacity. This has proven superior in several applications to the empirical risk minimization (ERM) principle adopted by ANNs that aims to reduce the training error only. Here, support vector regression (SVR) to predict Poisson's ratio and Young's modulus is described. The method uses a fuzzy-based ranking algorithm to select the most significant input variables and filter out dependency. The learning and predictive capabilities of the SVR method is compared to that of a backpropagation neural network (BPNN). The results demonstrate that SVR has similar or superior learning and prediction capabilities to that of the BPNN. Parameter sensitivity analysis was performed to

A.F. Al-Anazi · I.D. Gates (✉)
Department of Chemical and Petroleum Engineering, Schulich School of Engineering, University of Calgary, Calgary, Canada
e-mail: ian.gates@ucalgary.ca

C. Cranganu et al. (eds.), *Artificial Intelligent Approaches in Petroleum Geosciences*,
DOI 10.1007/978-3-319-16531-8_5

investigate the effect of the SVM regularization parameter, the regression tube radius, and the type of kernel function used. The result shows that the capability of the SVM approximation depends strongly on these parameters.

Keywords Poisson's ratio · Young's modulus · Artificial neural networks · Support vector machines · Log data · Core data

Nomenclature

AAE	Absolute average error
b	Bias term
BPNN	Backpropagation neural networks
C	Regularization parameter
E	Young's modulus, psi
f	An unknown function
g	Overburden stress, psi
G	Shear modulus, psi
h	Vapnik–Chervonenkis dimension
K	Bulk modulus, psi
L	Lagrangian equation for a dual programming problem or Loss function
MAE	Maximum absolute error
r	Correlation coefficient
RBF	Radial basis function
RMSE	Root mean square error
R_{emp}	Empirical risk
R	Structural risk
P	Pressure, psi
SVMs	Support vector machines
SVR	Support vector regression
x	Input variable
y	Output variable
$\hat{y}$	Estimated output value
v	Velocity
w	Weight vector

Greek Symbols

α,α^*	Lagrangian multiplier to be determined
ε	Error accuracy/strain
φ	Porosity, fraction
φ	Mapping function from input space into a high-dimensional feature space
η,η^*	Lagrangian multipliers
κ, ϑ	Sigmoid function parameters
μ	Poisson's ratio, dimensionless
ρ	Density, g/cc

σ Stress/variance
σ^2 Standard deviation
ξ,ξ^* Slack variables

Subscripts and Superscripts
a Axial
b Bulk
h Horizontal
k, l Indices
N Number of samples
n Input space dimension
p Pore/compressional
r Radial
s Shear

1 Introduction

Elastic properties of reservoir rock are important geomechanical data required in hydraulic fracture design, wellbore stability analysis, reservoir dilation, surface heave (especially for high-pressure processes such as cyclic steam stimulation), and sand production anticipation as in cold heavy oil production. Rock properties including Poisson's ratio, μ, shear modulus, G, Young's modulus, E, and bulk modulus, K, can be determined under static test conditions by using triaxial stress cells or under dynamic conditions by measuring compressional and shear velocities and density of reservoir core samples measured by acoustic or sonic logging tools (Gatens et al. 1990; Barree et al. 2009; Khaksar et al. 2009).

Core-measured elastic properties of reservoir rock obtained from detailed laboratory analysis are often considered to be the most direct and accurate method (Ameen et al. 2009). However, due to the costs of obtaining and handling core samples, most often, a limited number of samples are analyzed and often correlations are established between core and log data to estimate elastic properties of other reservoir rocks where core data are not available. Acoustic log measurements provide compressive and shear wave velocities which can be combined with density log and elasticity theory to estimate elastic properties of reservoir rock (Gatens et al. 1990; Abdulraheem et al. 2009). In complex reservoirs, acoustic theory may be insufficient to describe its actual behavior, thus limiting the accuracy of the derived rock correlations (Barree et al. 2009; Khaksar et al. 2009). Therefore, an integrated knowledge of the logging tool responses and understanding of underlying geology in addition to the use of advanced statistical techniques are necessary to determine an interpretation model that can effectively map the dependency between well log data and elastic rock properties. Due to several different factors that control log

responses and the construction of the mapping function between log data and the elastic properties, this can be a challenging task since the mapping can be nonlinear. To deal with this nonlinearity, artificial neural networks (ANNs), fuzzy logic (FL), and functional networks (FN) approaches to predict Poisson's ratio and Young's modulus have been used in the literature with varying degrees of success (Widarsono et al. 2001; Abdulraheem et al. 2009).

A further complication is that the data available to train artificial intelligence methods is scarce. The smaller the training data set, the greater the risk of poor estimation of rock properties. This is especially the case if data are missing from specific intervals along a well. Recently, support vector machines (SVMs) have been recognized as an efficient and accurate tool with strong learning and prediction capabilities (Al-Anazi and Gates 2010a, b, c, d). The SVM learning machine approach is novel and is based on the principle of structural risk minimization (SRM), which aims to minimize an upper bound of the generalization error. On the other hand, ANNs follow the principle of empirical risk minimization (ERM) which attempts to minimize the training error. The SRM principle is based on bounding the generalization error by minimizing the sum of the training error and a confidence interval term depending on the Vapnik–Chervonenkis (VC) dimension (Vapnik and Chervonenkis 1974; Vapnik 1982, 1995). To accomplish this, in support vector regression (SVR), a regularization term is used to determine the trade-off between the training error and VC confidence term. Consequently, the SVM approach provides a promising tool to generalize to unseen data. To capture nonlinear behavior of the mapping, kernel functions are used to project the input space into a higher dimensional feature space where a linear regression hyperplane is devised (Kecman 2005).

The accuracy and robustness of the SVM approach has been robustly demonstrated in many real-world applications; for example, face recognition, object detection, hand writing recognition, text detection, speech recognition and prediction, porosity and permeability determination from log data, and lithology classification (Li et al. 2000; Lu et al. 2001; Choisy and Belaid 2001; Gao et al. 2001; Kim et al. 2001; Ma et al. 2001; Van Gestel et al. 2001; Al-Anazi and Gates 2010a, b, c, d). In this study, the potential of SVR to establish an interpretation model that relates core and log data to elastic rock properties is evaluated from data originating from a well in a hydrocarbon reservoir. Two separate interpretation models were constructed: one for Poisson's ratio and the other for Young's modulus. The first model for Poisson's ratio was developed by using the density and compressive and shear wave velocities, whereas the second one for Young's modulus was devised by using variables chosen from a fuzzy selection scheme. For the well used in this study, the available core-derived input variables are porosity, minimum horizontal stress, pore pressure, and overburden stress, whereas the available well log data include bulk density and compressional and shear velocities.

In this research, the performance of SVM was compared with that of a back-propagation neural network (BPNN) to evaluate its potential to predict elastic rock properties under scarce data conditions. For SVM nonlinear approximation, a radial

basis function (RBF) kernel function is used. During the machine learning stage of the SVM, the kernel function parameter, insensitivity tube parameter, ε, and penalty parameter were selected through grid search and pattern search schemes. To avoid overfitting the data, a ten-fold cross-validation was used to select the optimal parameter to control the trade-off between the bias and variance of the model. Error analysis was done by examining the correlation coefficient, *r*, root mean square error (RMSE), absolute average error (AAE), and maximum absolute error (MAE) between target and predicted values. The study also investigated the impact of the penalty parameter, insensitivity tube radius, and the type of kernel function on the SVM approximation capability.

2 Backpropagation Neural Network (BPNN)

Backpropagation multilayer perceptron neural networks (BPNN) have been extensively used to interpret rock physical and elastic properties in hydrocarbon reservoirs (Rogers et al. 1995; Huang et al. 1996, 2001; Fung et al. 1997; Helle and Ursin 2001; Helle and Bhatt 2002; Abdulraheem et al. 2009). BPNN can approximate any continuous nonlinear function over a compact interval to any desirable accuracy depending on the number of hidden layers. They use activation functions in the processing neurons to facilitate the modeling of nonlinear mapping functions (Suykens et al. 2002). During learning, in BPNNs, the input patterns are propagated forward through hidden layers toward the output, while error is backpropagated toward the input layer. Here, a conjugate gradient algorithm is used to train the BPNN by minimizing the square of the residuals between target and training data. One well-known issue with such training algorithms is that it may become trapped in local minima since it is sensitive to the starting weight values (Hastie et al. 2001). Here, an Nguyen–Widrow algorithm was used to select the initial range of weight values and the conjugate gradient algorithm was then used to optimize the weights. Optimization is done several times with different random weight values to ensure convergence to the global optimal solution. One shortcoming of BPNNs is its susceptibility to overfit training data which results in poor generalization capability to interpret new input data. In this work, three layers (input, hidden, and output) were used with neurons being automatically optimized. A ten-fold cross-validation technique was used to stop training (DTREG v9.1 2009).

3 Support Vector Regression

SVMs are learning algorithms originally developed to solve classification problems. By using Vapnik's ε-insensitive loss function, SVMs can be used to solve nonlinear regression problems by using kernel functions (Vapnik 1995). Given a data set

$(y_k,\ \mathbf{x}_k)$ of dimension N where y_k is the output value at input variable values expressed in the vector $\mathbf{x}_k$, the relationship between the input and output variables can be expressed by the following linear regression function:

$$f(\mathbf{x}_k) = \mathbf{w}^T\mathbf{x}_k + b \tag{1}$$

where $\mathbf{w}$ is a set of weights and b is an offset or bias. The empirical risk (to be minimized) can be expressed by

$$R_{\text{emp}} = \frac{1}{N}\sum_{k=1}^{N} v\left(y_k - \mathbf{w}^T\mathbf{x}_k - b\right) \tag{2}$$

where $v(\cdot)$ is the Vapnik's ε-insensitive loss function defined by

$$v(y - f(x)) = \begin{cases} 0 & \text{if} \quad |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon & \text{otherwise} \end{cases} \tag{3}$$

The optimization problem in the primal space that has to be solved to achieve an optimal linear function in terms of the weights, $\mathbf{w}$, and bias, b, is given by

$$\min\ J_P(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} \tag{4}$$

$$\text{subject to} \begin{cases} y_k - \mathbf{w}^T\mathbf{x}_k - b \le \varepsilon,\ k = 1, \ldots, N \\ \mathbf{w}^T\mathbf{x}_k + b - y_k \le \varepsilon,\ k = 1, \ldots, N \end{cases}$$

The value of ε in Vapnik's loss function, v, characterizes the radius of an approximation tube which in turn controls the accuracy of the model. Slack variables, ξ_k, ξ_k^* for $k = 1, \ldots, N$, are introduced to the optimization problem given by Eq. 4:

$$\min\ J_P(\mathbf{w}, \xi, \xi^*) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{k=1}^{N}\left(\xi_K + \xi_K^*\right) \tag{5}$$

$$\text{subject to} \begin{cases} y_k - \mathbf{w}^T\mathbf{x}_k - b \le \varepsilon + \xi_k, & k = 1, \ldots, N \\ \mathbf{w}^T\mathbf{x}_k + b - y_k \le \varepsilon + \xi_k^*, & k = 1, \ldots, N \\ \xi_k, \xi_k^* \ge 0, & k = 1, \ldots, N \end{cases}$$

The regularization (penalty) constant, C, is positive and determines how large the deviation from the desired accuracy is tolerated. The Lagrangian form of the problem is expressed by

$$L(\mathbf{w}, b, \xi, \xi^*; \alpha, \alpha^*, \eta, \eta^*) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{k=1}^{N}\left(\xi_k + \xi_k^*\right) - \sum_{k=1}^{N}\alpha_k\left(\varepsilon + \xi_k - y_k + \mathbf{w}^T\mathbf{x}_k + b\right) - \sum_{k=1}^{N}\alpha_k^*\left(\varepsilon + \xi_k^* + y_k - \mathbf{w}^T\mathbf{x}_k - b\right) - \sum_{k=1}^{N}\left(\eta_k\xi_k + \eta_k^*\xi_k^*\right) \tag{6}$$

where $\alpha_k, \alpha_k^*, \eta_k, \eta_k^*$ are positive Lagrange multipliers. The solution is obtained by solving a saddle point problem: The Lagrangian, L, must be minimized with respect to $\mathbf{w}, b, \xi, \xi^*$ and maximized with respect to $\alpha, \alpha^*, \eta, \eta^*$. The following conditions are satisfied at the saddle point:

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \displaystyle\sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)\mathbf{x}_k \\ \dfrac{\partial L}{\partial b} = 0 \rightarrow \displaystyle\sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)\mathbf{x}_k = 0 \\ \dfrac{\partial L}{\partial \xi_k} = 0 \rightarrow C - \alpha_\mathrm{k} - \eta_k = 0 \\ \dfrac{\partial L}{\partial \xi_k^*} = 0 \rightarrow C - \alpha_k^* - \eta_k^* = 0 \end{cases} \tag{7}$$

The primal optimization problem can be re-formulated as a dual problem as follows:

$$\max J_D(\alpha, \alpha^*) = -\frac{1}{2}\sum_{k,l=1}^{N}\left(\alpha_k - \alpha_k^*\right)\left(\alpha_l - \alpha_l^*\right)x_k^T x_l - \varepsilon\sum_{k=1}^{N}\left(\alpha_k + \alpha_k^*\right) + \sum_{k=1}^{N} y_k\left(\alpha_k - \alpha_k^*\right) \tag{8}$$

$$\text{subject to } \begin{cases} \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right) \\ \alpha_k, \alpha_k^* \in [0, c] \end{cases}$$

By solving Eq. 8, the optimal Lagrange multiplier pairs can be found and the linear regression is then given by

$$f(\mathbf{x}_k) = \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)\mathbf{x}_k^T\mathbf{x}_k + b \tag{9}$$

with

$$\mathbf{w} = \sum_{k=1}^{N} \left(\alpha_k - \alpha_k^*\right)\mathbf{x}_k \tag{10}$$

The training points with nonzero α_k values allow the calculation of the bias term, b (Kecman 2005). Generalization of the SVM approach to nonlinear regression estimation is accomplished by using kernel functions. Typical kernel functions that are used are linear ones, Gaussian RBF, and polynomial and sigmoid functions as listed in Table 1. In the primal weight space, the regression model is given by

$$f(\mathbf{x}) = \mathbf{w}^T\varphi(\mathbf{x}) + b \tag{11}$$

with given training data $\{\mathbf{x}_k, y_k\}_{k=1}^N$ and $\varphi(\cdot)$: $R^n \rightarrow R^{n_h}$ is a kernel mapping function which projects the input space to a higher dimensional feature space. The primal problem is then formulated as follows:

$$\min J_P(\mathbf{w}, \xi, \xi^*) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{k=1}^{N}\left(\xi_k + \xi_k^*\right) \tag{12}$$

$$\text{subject to} \begin{cases} y_k - \mathbf{w}^T\varphi(\mathbf{x}_k) - b \leq \varepsilon + \xi_k, & k = 1, \ldots, k \\ \mathbf{w}^T\varphi(\mathbf{x}_k) + b - y_k \leq \varepsilon + \xi_k^*, & k = 1, \ldots, k \\ \xi_k, \xi_k^* \geq 0, & k = 1, \ldots, N \end{cases}$$

The dual problem is then formulated as follows:

$$\begin{aligned} \max J_D(\alpha, \alpha^*) = &-\frac{1}{2}\sum_{k,l=1}^{N}\left(\alpha_k - \alpha_k^*\right)\left(\alpha_l - \alpha_l^*\right)K(\mathbf{x}_k, \mathbf{x}_l) \\ &- \varepsilon\sum_{k=1}^{N}\left(\alpha_k + \alpha_k^*\right) + \sum_{k=1}^{N} y_k\left(\alpha_k - \alpha_k^*\right) \end{aligned} \tag{13}$$

$$\text{subject to} \begin{cases} \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right) = 0 \\ \alpha_k, \alpha_k^* \in [0, c] \end{cases}$$

Table 1 Common kernel function and corresponding mathematical expression

Kernel function	Mathematical expression
Linear	$k(\mathbf{x}_i, \mathbf{x}) = \langle \mathbf{x}_i, \mathbf{x} \rangle$
Gaussian radial basis function	$k(\mathbf{x}_i, \mathbf{x}) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}\|^2}{2\sigma^2}}$
Sigmoid	$k(\mathbf{x}_i, \mathbf{x}) = \tanh(\kappa\langle \mathbf{x}_i, \mathbf{x} \rangle + \vartheta)$

Table 2 Error measures used for accuracy assessment

Accuracy measure	Mathematical expression
Correlation coefficient, r	$\frac{\sum_{i=1}^{l}(y_i-\bar{y}_i)\left(\hat{y}_i-\bar{\hat{y}}_i\right)}{\sqrt{\sum_{i=1}^{l}(y_i-\bar{y}_i)^2\sum_{i=1}^{N_P}\left(\hat{y}_i-\bar{\hat{y}}_i\right)^2}}$
Root mean square error, RMSE	$\sqrt{\frac{1}{l}\sum_{i=1}^{l}(y_i-\hat{y}_i)^2}$
Average absolute error, AAE	$\frac{1}{l}\sum_{i=1}^{l}\lvert y_i-\hat{y}_i\rvert$
Maximum absolute error, MAE	$\max \quad \lvert y_i-\hat{y}_i\rvert, \quad i=1,\ldots,l$

By setting $K(x_k, x_l) = \varphi(x_k)^T\varphi(x_l)$ for $k, l = 1, \ldots, N$, the explicit calculation of the kernel function is avoided. The nonlinear regression representation of the dual problem is given by

$$f(\mathbf{x}) = \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)K(\mathbf{x}, \mathbf{x}_k) + b \tag{14}$$

where α_k, α_k^* are the solution of Eq. 13 and the bias term b is calculated as an average value over the support vectors corresponding to the training data set. The solution to Eq. 13 is unique and is a global minimum so long as the kernel function is positive definite (Suykens et al. 2002).

4 Bounds on the Generalization Error

The VC theory underlying the SVM formulation characterizes the generalization error instead of training (empirical) error. Given a set of functions $f(\mathbf{x}, \boldsymbol{\theta})$ characterized by different adjustable parameter vector $\boldsymbol{\theta}$ with a training data set $\{(\mathbf{x}_k, y_k)\}_{k=1}^{N}$ where $\mathbf{x}_k \in R^{N\times n}$ and $y_k \in R^n$. The empirical error is defined as follows:

$$R_{\text{emp}}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{k=1}^{N}(y_k - f(\mathbf{x}, \boldsymbol{\theta}))^2 \tag{15}$$

whereas the generalization error is defined by

$$R(\boldsymbol{\theta}) = \int (y_k - f(\mathbf{x}, \boldsymbol{\theta}))^2 p(\mathbf{x}, y)\mathrm{d}x\mathrm{d}y \tag{16}$$

measures the error over all patterns that are extracted from an underlying probability distribution $p(\mathbf{x}, y)$ which is typically unknown in practical applications. However, from Eqs. 15 and 16, the upper bound on the generalization error is given as follows:

$$R(\boldsymbol{\theta}) \leq R_{\text{emp}}(\boldsymbol{\theta}) + \left(\frac{1}{1 - c\sqrt{\frac{h(\ln(aN/h)+1) - \ln(\eta)}{N}}} \right)_+ \tag{17}$$

where h is the VC dimension of the set of approximating functions and the notation $(x)_+$ indicates that $(x)_+ = x$ if $x > 0$ and 0 otherwise. The upper bound given by Eq. 17 holds for probability $1 - \eta$. This is the probability (or level of confidence) to approximate functions at which the generalization bound holds (Schölkopf and Smola 2002; Kecman 2005). The confidence term (second one in Eq. 17) also depends on the VC dimension which in turn characterizes the capacity of the set of approximating functions which in turns reflects model complexity (Vapnik 1998; Schölkopf and Smola 2002; Suykens et al. 2002). In this research, $a = c = 1$ (Cherkassky and Shao 2001).

5 SVR Parameter and Model Selection

To determine the optimal set of parameters (C, ε, and kernel function shape factors, e.g., the variance of the Gaussian RBF) for the SVR model, iterative grid and pattern searches are used. The grid search aims to use values extracted from a specified range controlled by geometric steps, whereas pattern search is based on the idea that a range is specified and that the search starts at the center of the range and takes trial steps in each direction for each parameter. If the parameters at the new point enhance the fit of the regression function, the search center moves to the new point and the process is repeated. Otherwise, the step size is reduced and the search is resumed. This iterative process is stopped after the step size drops below a pre-defined tolerance. Grid search is computationally demanding since the model must be evaluated at many points within the grid for each parameter. This limitation may be exaggerated if cross-validation has been adopted as a model selection technique. Pattern search requires far fewer evaluations of the model than that of grid search. However, pattern search can potentially converge to a local instead of a global optimum. Here, both search methods are used to overcome the shortcomings of each method: The optimization process starts with grid search seeking to locate a region close to the global optimal point. Next, pattern search is done over a narrow range that surrounds the best point located by the grid search.

6 Elastic Properties Prediction Methodology

6.1 Well Log and Core Data Description

Figure 1 displays a subset of the core and log data versus depth used in this study. The input variables used to model the Poisson's ratio, μ, and Young's modulus, E,

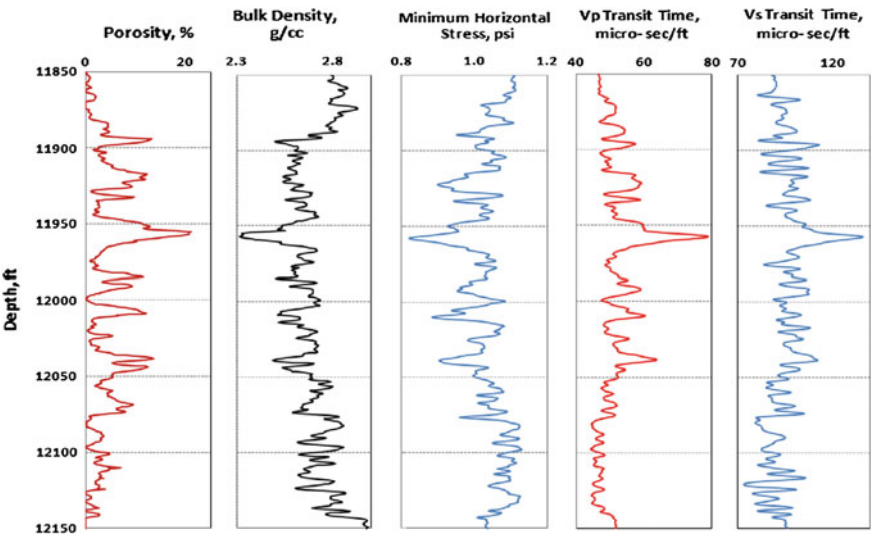

Fig. 1 Subset of raw log data

at each depth, $\mathbf{x}_k$, are core-derived porosity ϕ, minimum horizontal stress σ_h, pore pressure P_p, and overburden stress g, and log data including bulk density ρ_b, compressional wave velocity v_p, and shear wave velocity v_s. Elastic rock properties were extracted from samples tested in a laboratory-based triaxial pressure test cell. The input data set consists of 601 multidimensional data points spanning over 300 ft. of wellbore length. To evaluate the methods, the input data are randomly separated into training subsets consisting of 10, 20, 30, 40, 50, and 60 % of the total data set. The complements of the training data subsets are the testing subsets.

Acoustic logging tools measure the characteristic propagation speed of the P (compression) and S (shear) waves which are related to the elastic properties of the formation (Serra 1984). For core samples, elastic rock properties are determined from the stress–strain relationship that results from changes of stress with changes of the core strain (Montmayeur and Graves 1985, 1986):

$$E = \frac{\mathrm{d}\sigma_a}{\mathrm{d}\varepsilon_r}$$

where σ_a is the axial stress applied to the core sample and ε_r is the radial strain, and

$$\mu = \frac{\mathrm{d}\varepsilon_r}{\mathrm{d}\varepsilon_a}$$

where μ is the Poisson's ratio and ε_a is the axial strain.

7 Modeling of the Poisson's Ratio, μ

Feature Selection

Here, a model is developed that uses core-derived porosity ϕ, minimum horizontal stress σ_h, pore pressure P_p, and overburden stress g, and log data including bulk density ρ_b, compressional wave velocity v_p, and shear wave velocity v_s. Following Al-Anazi and Gates (2009), a Two-stage fuzzy ranking algorithm is used to identify information-rich core and log measurements and filter out data dependencies. The result of fuzzy ranking analysis, listed in Table 3, shows that the significant input variables, in order of importance, are σ_h, v_s, v_p, ρ_b, P_p, and ϕ. Therefore, these variables are the best correlators of Poisson's ratio and will be used in SVM model construction.

Poisson's Ratio Training and Prediction

The BPNN and SVR methods were compared by using different data training fractions (10, 20, 30, 40, 50, and 60 %) of the total available data to examine the impact of data scarcity on the predictive capabilities of the generated models. For example, a 10 % data training fraction means that 10 % of the total data available is selected randomly and used to train the model. The remaining 90 % of the data are used to test the correlation. For the SVR model, the Gaussian RBF kernel function was used.

Table 4 lists the error measures for the BPNN model results at different data training fractions when it is used to reproduce the training data set values. Learning performance is measured by correlation coefficient (r) and error statistics including RMSE, average absolute error (AAE), and MAE as defined in Table 2. The results reveal that the correlation coefficient is low and that as the data training fraction is enlarged, the capability of the BPNN to reproduce the training data set does not improve; in other words, the errors do not diminish as the training fraction grows. Table 5 lists the results of the SVR method. Even with training fraction as low as 10 %, the correlation coefficient is essentially equal to 1 and the errors are much smaller than that of the BPNN. In other words, the SVR method fully reproduces the training data set. These results demonstrate that for this data set, the SVR has superior learning capability to that of the BPNN.

Table 6 lists the error measures for the BPNN model results at different training fractions when it is used to predict the testing data set. The results show that the method does not provide a good prediction of the testing data set. The SVR results for the predicted testing data set are presented in Table 7. The results demonstrate an excellent prediction performance by the SVM revealing that the trained correlation models have captured the underlying relationships and have the potential to generalize accurately to new data. It can also be observed that consistent prediction capabilities are maintained over all data training fractions.

Table 3 Two-stage fuzzy ranking analysis for Poisson's ratio

Fuzzy curves ranking for Poisson's ratio data				
Input x^j	P_{c^i}	P_{c^i}/P_{c^R}	$P_{y_c^i}$	$P_{v_c^i}$
σ_h	0.6494	0.6512	0.6812	0.0489
v_s	0.8643	0.8667	0.8776	0.0154
v_p	0.8748	0.8773	0.8859	0.0127
ϕ	0.9194	0.9220	0.9279	0.0092
ρ_b	0.9540	0.9567	0.9723	0.0192
P_p	0.9740	0.9767	0.9794	0.0056
g	0.9740	0.9767	0.9794	0.0056

Second fuzzy surfaces ranking performance for Poisson's ratio data					
Input x^j	$P_{s^{1,j}}$	$P_{s^{1,j}}/P_{s^{1,R}}$	$P_{s^{1,j}}/P_{c^1}$	$P_{y_s^{1,j}}$	$P_{v_s^{1,j}}$
v_s	0.3530	0.5475	0.5436	0.4110	0.1642
ρ_b	0.5759	0.8931	0.8869	0.6169	0.0711
P_p	0.5997	0.9300	0.9235	0.6496	0.0831
g	0.5997	0.9300	0.9235	0.6496	0.0831
ϕ	0.6282	0.9742	0.9674	0.6685	0.0641
v_p	0.6340	0.9832	0.9763	0.6763	0.0667

Third fuzzy surfaces ranking performance for Poisson's ratio data					
Input x^j	$P_{s^{2,j}}$	$P_{s^{2,j}}/P_{s^{2,R}}$	$P_{s^{2,j}}/P_{c^2}$	$P_{y_s^{2,j}}$	$P_{v_s^{2,j}}$
v_p	0.6325	0.7365	0.7319	0.6665	0.0538
ϕ	0.7135	0.8308	0.8256	0.7468	0.0467
g	0.7878	0.9173	0.9116	0.8048	0.0215
P_p	0.7878	0.9173	0.9116	0.8048	0.0215
ρ_b	0.8085	0.9414	0.9355	0.8436	0.0434

Fourth fuzzy surfaces ranking performance for Poisson's ratio data					
Input x^j	$P_{s^{3,j}}$	$P_{s^{3,j}}/P_{s^{3,R}}$	$P_{s^{3,j}}/P_{c^3}$	$P_{y_s^{3,j}}$	$P_{v_s^{3,j}}$
ρ_b	0.8380	0.9621	0.9578	0.8580	0.0239
P_p	0.8489	0.9746	0.9703	0.8708	0.0258
g	0.8489	0.9746	0.9703	0.8708	0.0258
ϕ	0.8495	0.9753	0.9710	0.8665	0.0200

Fourth fuzzy surfaces ranking performance for Poisson's ratio data					
Input x^j	$P_{s^{5,j}}$	$P_{s^{5,j}}/P_{s^{5,R}}$	$P_{s^{5,j}}/P_{c^5}$	$P_{y_s^{5,j}}$	$P_{v_s^{5,j}}$
P_p	0.8789	0.9294	0.9212	0.9190	0.0457
g	0.8789	0.9294	0.9212	0.9190	0.0457
ϕ	0.8794	0.9299	0.9218	0.9000	0.0235

Fifth fuzzy surfaces ranking performance for Poisson's ratio data					
Input x^j	$P_{s^{6,j}}$	$P_{s^{6,j}}/P_{s^{6,R}}$	$P_{s^{6,j}}/P_{c^6}$	$P_{y_s^{6,j}}$	$P_{v_s^{6,j}}$
ϕ	0.8461	0.8785	0.8687	0.8768	0.0363

Table 4 Comparison of partition-based training error performance of BPNN using minimum horizontal stress, P and S wave velocities, RHOB, pore pressure, and porosity for prediction of Poisson's ratio

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.6375	0.5394	0.5269	0.5940	0.5202	0.5628
RMSE	0.0291	0.0278	0.0289	0.0266	0.0295	0.0261
AAE	0.0239	0.0227	0.0224	0.0212	0.0229	0.0203
MAE	0.0852	0.0847	0.0984	0.0852	0.1043	0.1055

Table 5 Comparison of partition-based training error performance of SVR using minimum horizontal stress, *P* and *S* wave velocities, RHOB, pore pressure, and porosity for prediction of Poisson's ratio

	10 %	20 %	30 %	40 %	50 %	60 %
r	1.000	1.000	1.000	1.000	1.000	1.000
RMSE	0.0003	0.0003	0.0002	0.0003	0.0002	0.0002
AAE	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002
MAE	0.0006	0.0006	0.0008	0.0012	0.0006	0.0006

Table 6 Comparison of partition-based testing error performance of BPNN using minimum horizontal stress, *P* and *S* wave velocities, RHOB, pore pressure, and porosity for prediction of Poisson's ratio

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.5247	0.5330	0.5554	0.5005	0.5519	0.4997
RMSE	0.0277	0.0216	0.0271	0.0286	0.0256	0.0294
AAE	0.0214	0.0216	0.0213	0.0220	0.0204	0.0231
MAE	0.1062	0.1004	0.1050	0.1061	0.1014	0.1014

Table 7 Comparison of partition-based testing error performance of SVR using minimum horizontal stress, *P* and *S* wave velocities, RHOB, pore pressure, and porosity for prediction of Poisson's ratio

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.9992	0.9998	1.000	0.9999	1.000	1.000
RMSE	0.0014	0.0007	0.0003	0.0004	0.0002	0.0003
AAE	0.0007	0.0004	0.0002	0.0003	0.0002	0.0002
MAE	0.0090	0.0051	0.0021	0.0030	0.0008	0.0013

8 Modeling of Young's Modulus, *E*

Feature Selection

Similar to the analysis on Poisson's ratio described above, a two-stage fuzzy ranking analysis was done to identify the most important variables to predict Young's modulus. The results, listed in Table 8, reveal that the ranking of input data, in order of importance, are v_s, ρ_b, v_p, P_p, ϕ, g, and σ_h. As expected, the ranking of the variables make it clear that the acoustic and density logs are most important with respect to the Young's modulus.

Young's Modulus Training and Prediction

As above, Gaussian RBFs was the kernel function used in the SVM nonlinear regression. Tables 9 and 10 list the results of the learning capabilities (the capability of the methods to reproduce the training data set) of the BPNN and SVR, respectively, at different training data fractions. For the BPNN, the results show that the correlation coefficient is high for all training data fractions. However, the RMSE and AAE do not exhibit a strong reducing trend as the training data fraction is enlarged and the MAE, in fact, shows a growth trend as the training data fraction increases. For the SVR, the correlation coefficients are high and are similar to that of the BPNN, but there is a decreasing trend of the RMSE, AAE, and MAE as the size of the training data fraction grows. Beyond 30 % training data set, all of the error measures of the SVR are lower than that of the BPNN. The results suggest that the SVM approach has higher learning capability than that of the BPNN for the data set used here.

Tables 11 and 12 list the correlation coefficient and error measures for the BPNN and SVR methods, respectively, at different training data fraction to predict the testing data subset. The results reveal excellent prediction performance can be observed by both techniques in terms of correlation coefficient. However, the analysis of the error performance indicates that the errors of the trained SVR models decline faster versus the size of the training data fraction than that of the BPNN and that the errors are lower than that of the BPNN when the training data fraction is 40 % or higher.

9 SVM Parameter Sensitivity Analysis

The generalization capability of SVMs depends on the regularization parameter, C, (see Eq. 5) that controls the trade-off between the training error (empirical error) and the VC dimension (complexity) of the regression model. If the value of C is very small, the training error is the primary error that is minimized, whereas if its value is very large, the estimate on the prediction error dominates and the training error plays a lesser role to construct the SVR model. Additional inputs of the SVR model include the cost function parameter (the regression tube radius ε of the

Table 8 Two-stage fuzzy ranking analysis for Young's modulus

Fuzzy curves ranking for Young's modulus data				
Input x^j	P_{c^i}	P_{c^i}/P_{c^R}	$P_{y_c^i}$	$P_{v_c^i}$
v_s	0.2484	0.2507	0.3054	0.2295
v_p	0.4010	0.4048	0.4620	0.1521
ρ_b	0.4676	0.4720	0.5206	0.1135
ϕ	0.4868	0.4913	0.5468	0.1234
σ_h	0.5026	0.5073	0.5589	0.1120
P_p	0.5848	0.5903	0.6181	0.0571
g	0.5848	0.5903	0.6181	0.0571

Second fuzzy surfaces ranking performance for Young's modulus data					
Input x^j	$P_{s^{1,j}}$	$P_{s^{1,j}}/P_{s^{1,R}}$	$P_{s^{1,j}}/P_{c^1}$	$P_{y_s^{1,j}}$	$P_{v_s^{1,j}}$
ρ_b	0.1344	0.5432	0.5412	0.1823	0.3559
v_p	0.1447	0.5847	0.5826	0.1967	0.3596
g	0.1497	0.6051	0.6029	0.1993	0.3307
P_p	0.1497	0.6051	0.6029	0.1993	0.3307
σ_h	0.1514	0.6117	0.6095	0.2065	0.3642
ϕ	0.1567	0.6332	0.6309	0.2122	0.3540

Third fuzzy surfaces ranking performance for Young's modulus data					
Input x^j	$P_{s^{3,j}}$	$P_{s^{3,j}}/P_{s^{3,R}}$	$P_{s^{3,j}}/P_{c^3}$	$P_{y_s^{3,j}}$	$P_{v_s^{3,j}}$
v_p	0.2415	0.5325	0.5166	0.3052	0.2636
P_p	0.3136	0.6912	0.6706	0.3856	0.2298
g	0.3136	0.6912	0.6706	0.3856	0.2298
σ_h	0.3141	0.6924	0.6717	0.3842	0.2231
ϕ	0.3551	0.7827	0.7594	0.4259	0.1995

Fourth fuzzy surfaces ranking performance for Young's modulus data					
Input x^j	$P_{s^{2,j}}$	$P_{s^{2,j}}/P_{s^{2,R}}$	$P_{s^{2,j}}/P_{c^2}$	$P_{y_s^{2,j}}$	$P_{v_s^{2,j}}$
P_p	0.2770	0.6983	0.6909	0.3477	0.2551
g	0.2770	0.6983	0.6909	0.3477	0.2551
ϕ	0.3054	0.7698	0.7616	0.3808	0.2469
σ_h	0.3436	0.8661	0.8569	0.4229	0.2307

Fourth fuzzy surfaces ranking performance for Young's modulus data					
Input x^j	$P_{s^{6,j}}$	$P_{s^{6,j}}/P_{s^{6,R}}$	$P_{s^{6,j}}/P_{c^6}$	$P_{y_s^{6,j}}$	$P_{v_s^{6,j}}$
ϕ	0.2909	0.5072	0.4975	0.3645	0.2531
σ_h	0.3599	0.6276	0.6155	0.4370	0.2140
g	0.5175	0.9023	0.8850	0.5717	0.1046

Fourth fuzzy surfaces ranking performance for Young's modulus data					
Input x^j	$P_{s^{4,j}}$	$\frac{P_{s^{4,j}}}{P_{s^{4,R}}}$	$\frac{P_{s^{4,j}}}{P_{c^4}}$	$P_{y_s^{4,j}}$	$P_{v_s^{4,j}}$
g	0.2909	0.6115	0.5977	0.3645	0.2531
σ_h	0.3595	0.7556	0.7385	0.4371	0.2160

Fifth fuzzy surfaces ranking performance for Young's modulus data					
Input x^j	$P_{s^{7,j}}$	$P_{s^{7,j}}/P_{s^{7,R}}$	$P_{s^{7,j}}/P_{c^7}$	$P_{y_s^{7,j}}$	$P_{v_s^{7,j}}$
σ_h	0.3599	0.6276	0.6155	0.4370	0.2140

Table 9 Comparison of partition-based training error performance of BPNN using minimum horizontal stress, *P* and *S* wave velocities, pore pressure, overburden stress, and porosity for prediction of Young's modulus

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.9995	0.9995	0.9996	0.9995	0.9995	0.9995
RMSE	0.2336	0.2415	0.2635	0.2705	0.2864	0.2594
AAE	0.1903	0.1806	0.1990	0.2032	0.2181	0.1792
MAE	0.5223	0.9198	1.2780	1.2645	1.4517	2.0796

Table 10 Comparison of partition-based training error performance of SVR using minimum horizontal stress, *P* and *S* wave velocities, pore pressure, overburden stress, and porosity for prediction of Young's modulus

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.9944	0.9993	0.9999	0.9999	0.9999	1.0000
RMSE	0.9389	0.3098	0.1619	0.1390	0.0954	0.0759
AAE	0.5270	0.1584	0.0845	0.0764	0.0460	0.0419
MAE	4.6358	1.7137	0.9692	1.0244	1.1114	0.4942

Table 11 Comparison of partition-based testing error performance of BPNN using minimum horizontal stress, *P* and *S* wave velocities, pore pressure, overburden stress, and porosity for prediction of Young's modulus

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.9981	0.9989	0.9997	0.9987	0.9994	0.9994
RMSE	0.5426	0.4209	0.249	0.4658	0.2922	0.3159
AAE	0.3421	0.2731	0.1964	0.2609	0.2337	0.2002
MAE	4.0785	3.2514	1.3903	3.4581	1.5281	2.0031

Table 12 Comparison of partition-based testing error performance of SVR using minimum horizontal stress, *P* and *S* wave velocities, pore pressure, overburden stress, and porosity for prediction of Young's modulus

	10 %	20 %	30 %	40 %	50 %	60 %
r	0.9905	0.9986	0.9996	0.9994	0.9998	0.9999
RMSE	1.5067	0.5016	0.2405	0.3153	0.1547	0.1445
AAE	0.922	0.2747	0.1337	0.1378	0.0764	0.0786
MAE	7.5939	3.596	1.9554	2.2864	1.4729	1.0244

ε-insensitivity cost function) and the type of kernel function used and its associated input parameters (for example, the variance of the Gaussian RBF kernel function). In the above analysis, these parameters were chosen by using the cross-validation method. Here, the impact of these parameters on the SVR model is investigated. Here, 10 % data training fraction subset (with 90 % remaining for testing) is used.

10 Impact of *C* on SVM Regression Performance

The effect of the value of *C* on the SVR training regression model for Poisson's ratio and Young's modulus were investigated with RBF kernel function with both approximation tube radius ε and the variance of the kernel function σ kept constant at values selected in the above analysis. Tables 13 and 14 list the training performance as examined by correlation coefficient and error statistics to predict Poisson's ratio and Young's modulus. The results indicate that the constructed regression model fits the data perfectly as the values of the regularization parameter, *C*, increases. The selection of higher values of *C*, however, does not always improve performance of the SVR especially if the data itself does not describe the underlying function that relates the input vector to the output scalar as would be the case with data polluted with severe noise. In this case, the SVR would model the noise rather than the true relationship between the input and output data.

11 Impact of ε on SVM Regression Performance

The radius of the regression tube, ε, within which the regression function must lay, is a measure of the error tolerance of the predictive capability of the regression model. If a predicted value lies within the tube radius, that is, its absolute value is less than ε and the loss (error) is set equal to zero. For a predicted value lying outside the ε-tube, the loss (error) equals the difference between the predicted value and the radius of the tube ε. Here, to investigate the effect of the regression tube radius on the capability of the SVR to generalize to unseen data, the regularization

Table 13 Comparison of partition-based training error performance of SVM over different regularization constant, *C* values for prediction of Poisson's ratio ($\sigma = 0.1877$, $\varepsilon = 9 \times 10^{-5}$)

	1×10^{-4}	1×10^{-3}	1×10^{-2}	1×10^{-1}	1	1×10^{1}	1×10^{2}
r	0.6849	0.6849	0.7297	0.9865	0.9994	0.9999	1.0000
RMSE	0.0356	0.0351	0.0312	0.0092	0.0014	0.0004	0.0003
AAE	0.0287	0.0283	0.0250	0.0067	0.0009	0.0003	0.0003
MAE	0.0964	0.0959	0.0895	0.0292	0.0067	0.0010	0.0006

Table 14 Comparison of partition-based training error performance of SVM over different regularization constant, *C* values for prediction of Young's modulus ($\sigma = 0.1208$, $\varepsilon = 4 \times 10^{-5}$)

	1×10^{3}	1×10^{4}	1×10^{5}	1×10^{6}	1×10^{7}	1×10^{8}	1×10^{9}
r	0.0000	0.7908	0.8246	0.9331	0.9944	0.9995	1.0000
RMSE	7.5444	7.4695	6.8751	3.8890	0.9389	0.2379	0.0432
AAE	5.8183	5.7562	5.2528	2.8527	0.5270	0.1476	0.0262
MAE	25.0131	24.7374	21.9113	12.1491	4.6358	0.9137	0.2444

Table 15 Comparison of partition-based training error performance of SVM over different approximation tube radius, ε values for prediction of Poisson's ratio (σ = 0.1877, C = 45.66)

	1×10^{-4}	1×10^{-3}	1×10^{-2}	1×10^{-1}
r	1.0000	0.9997	0.9907	0.000
RMSE	0.0003	0.0009	0.0060	0.0378
AAE	0.0003	0.0007	0.0051	0.0319
MAE	0.0005	0.0014	0.0104	0.0776

Table 16 Comparison of partition-based training error performance of SVM over different approximation tube radius, ε values for prediction of Young's modulus (σ = 0.1208, $C = 1 \times 10^7$)

	1×10^3	1×10^4	1×10^5	1×10^6	1×10^7
r	0.9944	0.9945	0.9934	0.9155	0.0000
RMSE	0.9389	0.9228	0.9868	3.8853	7.9617
AAE	0.5270	0.5303	0.7088	3.2992	6.0465
MAE	4.6358	4.5669	4.2223	8.4958	22.5317

parameter, C, and the variance, σ, of the RBF kernel function were fixed at the values determined above. The results listed in Tables 15 and 16 reveal that the constructed training model of Poisson's ratio and Young's modulus fits perfectly the data as the values of radius of the regression tube decreases as indicated by the correlation coefficient and error statistics. On the contrary, as the size of the insensitivity tube increases, the accuracy of the prediction model drops. The input data for the Poisson's ratio and Young's modulus data are quite clustered, and there are few outliers. As a result, when the radius of the regression tube enlarges, a fewer number of training points is used, thus the number of support vectors drop, and thus, the quality of the SVR model degrades. A further increase in the radius causes the SVR model to overshoot the test data. The results reveal that an increase in the radius of the insensitivity tube has a smoothing effect on modeling, whereas a decrease in the size may lead to overfitting the data.

12 Impact of Kernel Function on SVM Regression Performance

The kernel function is an integral part of the SVM formulation because it is necessary to solve nonlinear regression problems. The kernel function maps the nonlinear input space into a high-dimensional feature space where a linear SVM formulation can be applied. As such, the kernel function is assumed to have the capability to provide or approximately provide the nonlinear mapping. Thus, the choice of the kernel function will depend on the nature of the regression problem being solved. Here, three kernel functions are investigated including the linear,

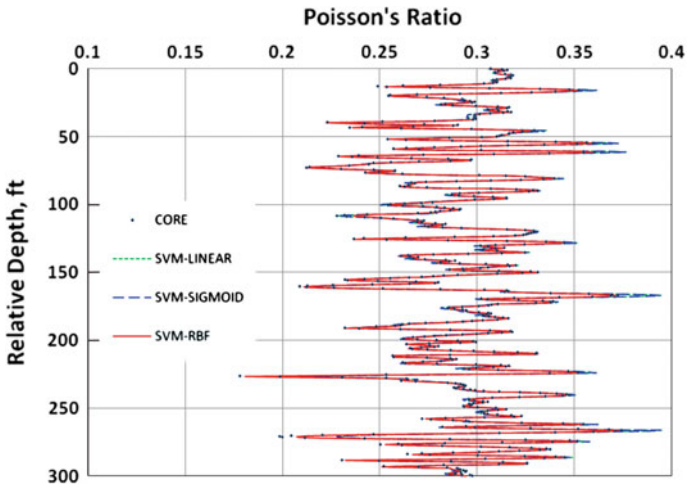

Fig. 2 Comparison of SVM-LINEAR, SVM-SIGMOID, and SVM-RBF predictions of Poisson's ratio using 10 % of the data for training and 90 % for testing. The dots are data obtained from triaxial tests of core samples

sigmoid, and Gaussian RBF functions. The prediction capabilities of SVR approximation models based on these kernels to predict Poisson's ratio and Young's modulus are compared are shown in Figs. 2 and 3, respectively. The corresponding correlation coefficients and error statistics are listed in Tables 17 and 18. The results reveal that all three kernel functions exhibit reasonably good capabilities to

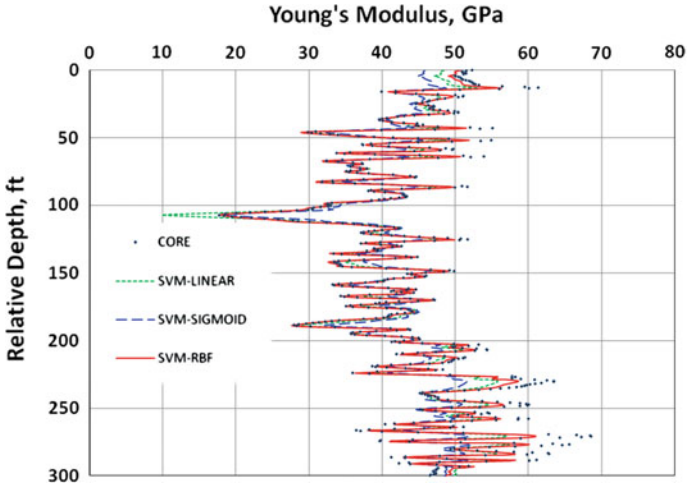

Fig. 3 Comparison of SVM-LINEAR, SVM-SIGMOID, and SVM-RBF predictions of Young's modulus using 10 % of the data for training and 90 % for testing. The dots are data obtained from triaxial tests of core samples

Table 17 Comparison of correlation coefficient, RMSE, AAE, and MAE of the errors between SVM-LINEAR, SVM-SIGMOID, and SVM-RBF predictions of Poisson's ratio using 10 % of the data for training and 90 % for testing

	Linear	Sigmoid	RBF
r	0.9935	0.9937	1.0000
RMSE	0.0040	0.0040	0.0003
AAE	0.0029	0.0029	0.0003
MAE	0.0193	0.0184	0.0006

Table 18 Comparison of correlation coefficient, RMSE, AAE, and MAE of the errors between SVM-LINEAR, SVM-SIGMOID, and SVM-RBF predictions of Young's modulus using 10 % of the data for training and 90 % for testing

	Linear	Sigmoid	RBF
r	0.9612	0.9203	0.9905
RMSE	2.8075	4.5299	1.5067
AAE	1.9659	3.3711	0.9219
MAE	11.6000	16.8300	7.6000

approximate the core-measured Poisson's ratio values. The best results are achieved with the RBF kernel function. Similarly, all three kernel functions performed well to predict the core-measured Young's modulus, but the RBF function performs slightly better than the linear and sigmoid kernel functions.

13 Conclusions

SVMs integrated with a fuzzy-based curve and surface input variable ranking analysis has demonstrated its potential applicability to develop interpretation models of the Poisson's ratio and Young's modulus under limited core data conditions. The main conclusions are as follows:

1. The SVM model is easier to construct than that of an artificial neural network model.
2. The SVMs formulation facilitates unique global solution compared to BPNN which often suffers multiple local minima.
3. The SVMs formulation offers natural means to deal with sparse data given that the number of support vectors used is equal to the number of training data points.
4. The SVMs shows high learning capability for both Poisson's ratio and Young's modulus under the presence of limited core data.
5. The results show that SVR yields a better model to predict Poisson's ratio than a backpropagation neural networks model.

6. The results demonstrate that the prediction errors for the Young's modulus obtained from the SVR decrease faster as the training data size grows than that of the backpropagation neural network.
7. Linear, sigmoid, and Gaussian RBF kernel functions show high prediction generalizability for Poisson's ratio and Young's modulus. The RBF function exhibits slightly better performance than the other two kernel functions.

References

Abdulraheem A, Ahmed M, Vantala A, Parvez T (2009) Prediction of rock mechanical parameters for hydrocarbon reservoirs using different artificial intelligence techniques. Paper SPE 126094 presented at the SPE Saudi Arabia section technical symposium, Alkhobar, Saudi Arabia, 9–11 May 2009

Al-Anazi A, Gates ID (2009) Fuzzy logic data-driven permeability prediction for heterogeneous reservoirs. Paper SPE 121159 presented at the 2009 SPE EUROPEC/EAGE annual conference and exhibition, Amsterdam, The Netherlands, 8–11 June 2009

Al-Anazi A, Gates ID (2010a) On the capability of support vector machines to classify lithology from well logs. Nat Resour Res 19(2):125–139. doi:10.1007/s11053-010-9118-9

Al-Anazi A, Gates ID (2010b) Support vector regression for permeability prediction in a heterogeneous reservoir: a comparative study. SPE Res Eval Eng 13(3):485–495. SPE-126339-PA. doi:10.2118/126339-PA

Al-Anazi A, Gates ID (2010c) A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs. Eng Geol 114:267–277. doi:10.1016/j.enggeo.2010.05.005

Al-Anazi A, Gates ID (2010d) Support vector regression for porosity prediction in a heterogeneous reservoir: a comparative study. Comput Geosci 36(12):1494–1503

Ameen MS, Smart Brian GD, Mc J, Somerville Sally Hammilton, Naji Nassir A (2009) Predicting rock mechanical properties of carbonates from wireline logs (a case study: Arab-D reservoir, Ghawar field, Saudi Arabia). Mar Petrol Geol 26(4):430–444

Barree RD, Gilbert JV, Conway MW (2009) Stress and rock property profiling for unconventional reservoir stimulation. Paper SPE 118703 presented at the SPE hydraulic fracturing technology conference, The Woodlands, Texas, 19–21 Jan 2009

Cherkassky V, Shao X (2001) Signal estimation and denoising using VC-theory. Neural Netw 14:37–52

Choisy C, Belaid A (2001) Handwriting recognition using local methods for normalization and global methods for recognition. In: Proceedings of sixth international conference on document analysis and recognition, pp 23–27

DTREG (2009) Predictive modeling software user's manual, version 9.1 (http://www.dtreg.com)

Fung C, Wong K, Eren H (1997) Modular artificial neural network for prediction of petrophysical properties from well log data. IEEE Trans Instrum Measure 46(6):1295–1299

Gao D, Zhou J, Xin L (2001) SVM-based detection of moving vehicles for automatic traffic monitoring. IEEE Intell Transp Syst 745–749

Gatens JM III, Harrison CW III, Lancaster DE, Guldry FK (1990) In-situ stress tests and acoustic logs determine mechanical properties and stress profiles in the devonian shales. SPE Formation Eval 5(3):248–254

Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning: data mining, inference, and prediction. Springer, New York

Helle H, Bhatt A (2002) Fluid saturation from well logs using committee neural networks. Petrol Geosci 8:109–118

Helle H, Ursin B (2001) Porosity and permeability prediction from wireline logs using artificial neural networks: a North Sea case study. Geophys Prospect 49:431–444

Huang Y, Gedeon TD, Wong PM (2001) An integrated neural-fuzzy-genetic-algorithm using hyper-surface membership functions to predict permeability in petroleum reservoirs. J Eng Appl Artif Intell 14:15–21

Huang Z, Shimeld J, Williamson M, Katsube J (1996) Permeability prediction with artificial neural network modelling in the venture gas field, offshore eastern Canada. Geophysics 61:422–436

Kecman V (2005) Support vector machines—an introduction. In: Wang L (ed) Support vector machines: theory and applications, Chap. 1. Springer, Berlin, pp 1–47

Khaksar A, Taylor PG, Fang Z, Kayes T, Salazar A, Rahman K (2009) Rock strength from core and logs, where we stand and ways to go. Paper SPE 121972 presented at the EUROPEC/EAGE conference and exhibition, Amsterdam, The Netherlands, 8–11 June 2009

Kim K, Jung K, Park S, Kim HJ (2001) Support vector machine-based text detection in digital video. Pattern Recognit 34:527–529

Li Z, Weida Z, Licheng J (2000) Radar target recognition based on support vector machine. In: Proceedings of 5th international conference on signal processing, vol 3, pp 1453–1456

Lu J, Plataniotis K, Ventesanopoulos A (2001) Face recognition using feature optimization and v-support vector machine. IEEE neural networks for signal processing, vol 11, pp 373–382

Ma C, Randolph M, Drish J (2001) A support vector machines-based rejection technique for speech recognition. In: Proceedings of IEEE international conference on acoustics, speech, and signal processing, vol 1, pp 381–384

Montmayeur H, Graves RM (1985) Prediction of static elastic/mechanical properties of consolidated and unconsolidated sands from acoustic measurements: basic measurements. Paper SPE 14159 presented at the 60th SPE annual technical conference and exhibition, Las Vegas, NV, 22–25 Sept 1985

Montmayeur H, Graves RM (1986) Prediction of static elastic/mechanical properties of consolidated and unconsolidated sands from acoustic measurements: correlations. Paper SPE 15644 presented at the 61st SPE annual technical conference and exhibition, Orleans, LA, 5–8 Oct 1986

Rogers SJ, Chen HC, Kopaska-Merkel DC, Fang JH (1995) Predicting permeability from porosity using artificial neural networks. AAPG Bull 79:1786–1797

Schölkopf B, Smola AJ (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge, MA

Serra O (1984) Fundamentals of well-log interpretation. Elsevier, Amsterdam

Suykens JAK, Van Gestel T, Brabanter J, De Moor B, Vandewalle J (2002) Least squares support vector machines. World Scientific, Singapore

Van Gestel T, Suykens J, Baestaens D, Lambrechts A, Lanckriet G, Vandaele B, De Moor B, Vandewalle J (2001) Financial time series prediction using least squares support vector machines within the evidence framework. IEEE Trans Neural Netw 12(4):809–821

Vapnik V, Chervonenkis A (1974) Theory of pattern recognition [in Russian]. Nauka, Moscow (German trans: Wapnik W, Tscherwonenkis A., Theorie der Zeichenerkennung, Akademie, Berlin, 1979)

Vapnik VN (1982) Estimation of dependences based on empirical data. Springer, Berlin

Vapnik V (1995) The nature of statistical learning theory. Springer, New York

Vapnik V (1998) Statistical learning theory. Wiley, New York

Widarsono B, Wong PM, Saptono F (2001) Estimation of rock dynamic elastic property profiles through a combination of soft computing, acoustic velocity modeling and laboratory dynamic test on core samples. Paper SPE 68712 presented at the SPE Asia Pacific oil and gas conference and exhibition, Jakarta, Indonesia, 17–19 Apr 2001