Journal of Civil Engineering (IEB), 38 (2) (2010) 173-190

Journal of Civil Engineering IEB

Support vector machines for nonlinear pavement backanalysis

Kasthurirangan Gopalakrishnan, Sunghwan Kim

Dept. of Civil, Construction & Environmental Engineering
Iowa State University, Ames, IA 50011-3232,USA

Received 28 June 2010

Abstract

Backanalysis or backcalculation of in-service pavement mechanical properties (such as elastic modulus) from pavement Non-Destructive Test (NDT) deflection data is a routine practice carried out by highway engineers for pavement structural condition evaluation, remaining life calculations, and mechanistic-based analysis. Owing to the complexity of this ill-conditioned inverse modeling problem, numerous backcalculation routines have been developed and implemented over the years ranging from simple deflection-basin matching programs to intelligent and soft computing based methodologies and each has its own pros and cons. This paper presents an efficient off-line pavement backcalculation system based on support vector machines (SVM) and compares its performance with another popular machine learning technique, multi-layer perceptrons (MLP). Both systems are trained and tested using synthetic deflection basins generated using a two-dimensional axisymmetric finite element software covering a wide range of in-service pavement scenarios. The results show that the effectiveness of SVM approach in pavement backanalysis is comparable to MLP approach, in general, and better in some specific cases.

Keywords: Machine learning; inverse analysis; pavement rehabilitation; falling weight deflectometer (fwd); pavement layer moduli; backcalculation

Notation

The following symbols are used in this paper.
b = bias term;
C = a pre-specified SVM tolerance parameter;
f = function of;
f_0 = empirical risk minimizer;
H = function class;
$R[f]$ = risk functional;
$R_{emp}[f]$ = empirical risk functional;
x = input vector (of independent variables);
y = output variable;
α_i , α_i^* = Lagrange multipliers;
ε = precision parameter;
ξ , ξ^* = slack variables;
w = a normal vector;

K = Kernel function;
L = loss function;
l = number of examples in training set;
P = probability of;
R^d = d-dimensional real vector space;
R^D= a high-dimensional feature vector space;
<· , · > = dot product;
δ^2 = bandwidth of Gaussian Kernel;
y_i^t = target modulus values;
y_i^p = predicted modulus values;
$\overline{y_i^t}$ = mean of the target modulus values;
$\overline{y_i^p}$ = mean of the predicted modulus values
Φ = function to be minimized;
$W(\alpha_i, \alpha_i^*)$ = dual of Lagrange functional;
$\phi(x)$ = input vector mapped in feature vector space;
y_x^p = the number of predicted modulus values (out of n total predicted) for which the absolute relative error less x%.
E_{AC} = asphalt concrete young's modulus;
υ = poisson ratio;
M_R = resilient modulus;
θ = bulk stress;
K, n = statistical parameters of K-θ model;
M_{Ri} = breakpoint resilient modulus (psi);
σ_d = applied deviator stress (psi);
K_1 and K_2 = statistical coefficients of bilinear model;

1. Introduction

A conventional asphalt concrete pavement typically consists of three layers: a surface layer paved with Asphalt Concrete (AC) mix, a base or/and subbase layer made up of crushed stone, and a subgrade layer made up of natural soil. Evaluating the structural condition of existing, in-service asphalt concrete pavement roads and their component layers is a part of the routine maintenance and rehabilitation activities for sustainable infrastructure. The use of nondestructive testing (NDT) methods has increased in evaluating the structural condition of new or in-service pavements because NDT testing is faster than destructive tests and do not entail the removal of pavement materials. Among the available NDT methods used in pavement testing is falling weight deflectometer (FWD) based deflection basin testing technique which has become a well-known and widely accepted procedure since 1980s (Alavi et al. 2008).

When a load is applied on a pavement surface, the pavement layers deflect nearly vertically to form a basin. The FWD measures a basin of pavement surface deflections in response to a stationary dynamic load, similar to a passing wheel load (Alavi et al. 2008). Since the deflection of a pavement represents an overall "system response" of the pavement layers to an applied load, the deflected shape of the basin is predominantly a function of the thickness of the pavement layers, the moduli of individual layers, and the magnitude of the load. Based on this mechanical concept, the in situ moduli of individual layers can be estimated from FWD measurements through appropriate analysis methods. This procedure called as a pavement layer modulus backcalculation or backanalysis. The backcalculation problem of layer modulus of asphalt pavement has been recognized as a complex problem (Sharma and Das 2008).

A number of models and software have been developed for the backcalculation of asphalt pavement layer moduli and are still evolving to estimate moduli quickly and accurately. These models can be classified into static, dynamic, and adaptive models (Goktepe et al. 2006). The detail of these models and studies on comparative performance are presented in literature (Ali and Khosla 1987, Ullidtz and Coetzee 1995, Shoukry and William 1999, Fwa and Thakur 2005, Goktepe et al. 2006, Sharma and Das 2008, Alavi et al. 2008). Goktepe et al. (2008) explained that the main difference among these models is related to the type of the asphalt pavement response model and the optimization procedure employed in model development. The types of asphalt pavement response model employed are either layered elastic analysis (LEA) theory or finite element method (FEM) with linear or nonlinear elastic material behaviors. An optimization procedure is used to determine appropriate layer modulus values at given deflection measurements. The optimization procedures in existing backcalaution models can be preformed using the least-squares (parameter identification), database search, and limited number of artificial intelligence (AI) based techniques (Goktepe et al. 2008) such as artificial neural networks (ANN) (Meier and Rix 1994, Gopalakrishnan 2004, Ceylan et al. 2007), neuro-fuzzy systems (Goktepe et al. 2006, Goktepe et al. 2008), and Genetic Algorithms (GAs) (Fwa et al. 1997, Terzi et al. 2003, Alkasawneh 2007). Particle Swarm Optimization (Gopalakrishnan 2010), Differential Evolution (Gopalakrishnan and Khaitan 2010), etc.

The rapid development of computer hardware and software has increased the level of AI application in data processing and mining (modeling) for knowledge discovery in database (KDD), which is a nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data (Fayyad et al, 1996). The AI is a collection of several intelligent analytical tools which attempt to simulate the working natural intelligent systems and build systems which could learn (Miradi, 2009). Among AI tools is artificial neural networks (ANN) which has been successfully used to solve complicated pattern recognition and classification problems in many fields (Panakka and Adeli, 2009). The ANN has also been applied for the problem of asphalt pavement layer moduli backcalculation. Once ANN is trained well with a large amount of data, the ANN can provide quick computation of solution and considerable accuracy and precision compared with backcaluation methods employing classic optimization process and neuro-fuzzy systems (Goktepe et al. 2008, Sharma and Das, 2008).

Support Vector Machines (SVM) is another promising AI technique derived from statistical learning theory by Vapnik and Chervonenkis (1964) in the Institute for Control Problems of the Russian Academy of Sciences. An interesting property of SVM compared to conventional ANN is that SVM employ structural risk minimization (SRM) principle minimizing a bound on a generalized risk (error) while conventional ANN employ empirical risk minimization (ERM) principle minimizing only the error the mean square error over the data set (Gunn 1998). This eventually results in better generalization performance than conventional ANN (Dibike et al 2001, Tay and Cao 2001). Another merit of SVM is that the solution of SVM is more unique, optimal and absent from local minima rather than conventional ANN (Tay and Cao 2001). This is reason that the training of SVM is equivalent to solving a linearly constrained quadratic programming as the training of conventional ANN requires non-linear optimization.

Some recent SVM applications in civil engineering domain include remote sensing images analysis and rainfall-runoff model (Dibike et al 2001), document classification for construction information systems (Caldas and Soibelman 2003), soil moisture prediction (Gill et al 2006), conceptual cost estimates in construction projects (An et al 2007), seismic liquefaction assessment (Goh and Goh 2007), slope reliability analysis (Zhao 2008), studying settlement of shallow foundations (Samui 2008), and model for contractor prequalification (Lam et al 2009). This paper explores the feasibility of SVM application in backcalculation of asphalt pavement layer moduli. The SVM theory and procedure are briefly reviewed. The development of SVM based asphalt pavement layer moduli backcalculation models is presented and the performance of SVM models are compared with ANN models.

2. Support vector machine (SVM)

One of the broadest subfield in artificial intelligence is the machine learning (ML) focusing on the development of data modeling techniques and algorithms that learn from data. SVM is one of the ML techniques derived from statistical learning theory by Vapnik and Chervonenkis in the sixties (1964). The foundations of SVM have been developed by Vapnik (1995) at AT&T Bell Laboratories and are recognized as attractive and promising tool to solve classification and regression related problems (Gunn 1998). Initial work of SVM as a classifier focused on optical character recognition and object recognition tasks (Smola and Scholkopf 2004). SVM have also provided excellent performances in regression and time series prediction applications (Smola and Scholkopf 2004). Compared to regression methods by conventional ANN, SVM in regression approximation has three distinct characteristics as follows (Tay and Cao 2001):

- SVM use a set of linear functions defined in a high dimensional space,
- SVM carry out risk minimization using loss functions,
- SVM use a risk function consisting of the empirical error and a regularization term which is derived from the SRM.

Comprehensive tutorials on the use of SVM in regression are available in many sources (Burges 1998, Gunn 1998, Cristianini and Shawe-Taylor 2000, Herbrich 2002, Scholkopf and Smola 2002, Smola and Scholkopf 2004). A brief summary of the SVM in regression is given here primary based on Gunn (1998) and Smola and Scholkopf (2004).

2.1 SVM algorithm in regression approximation

Errors in choosing a model from the hypothesis space arise from two cases. A poor choice of the model space results in a large approximation error which is a consequence of the hypothesis space being smaller than the target space. Estimation error is related to the learning procedure which results in a technique selecting the non-optimal model from the hypothesis space. Combination of these errors forms the generalization error.

Suppose that training data are given as $\{(x_1, y_1), \ldots, (x_l, y_l)\}$ □ $X \times R^d$, where X denotes the space of the input patterns (e.g. $X = R^d$). The goal of SVM is to find a function f minimizing the expected risk (Vapnik 1982).

$$R[f] = \int L(x, y, f(x))\, dP(x, y) \tag{1}$$

Where, $R[f]$ is risk functional; $L(x, y, f(x))$ denotes a loss function determining how we will penalize estimation errors based on the empirical data **X**; $P(x, y)$ denotes distribution. However, $P(x, y)$ is unknown. A possible approximation consists in replacing the integration by the empirical estimate, to get the so called empirical risk functional as follows:

$$R_{emp}[f] \cong \frac{1}{n}\sum_{i=1}^{n} L(x_i, y_i, f(x_i)) \tag{2}$$

Where, $R_{emp}[f]$ is empirical risk functional. A first attempt would be to find the empirical risk minimizer as follows:

$$f_0 \cong \arg\min_{f \in H} R_{emp}[f] \tag{3}$$

Where, f_0 denotes empirical risk minimizer; H is some function class. Empirical risk minimization makes sense only if,

$$\lim_{l\to\infty} R_{emp}[f] = R[f] \tag{4}$$

Which is true from the law of large numbers. However, it must also satisfy the following equation:

$$\lim_{l\to\infty} \min_{f\in H} R_{emp}[f] = \min_{f\in H} R[f] \tag{5}$$

Which is only valid when H is ‘small’ enough. This condition is less intuitive and requires that the minima also converge. Based on risk functional conditions, the SVM algorithm aims to find a function $f\,(x)$ that has at most ε deviation from the actually obtained targets y_i for all the training data, and at the same time is as flat as possible (Smola and Scholkopf 2004). In other words, errors less than precision parameter ε are accepted and errors larger than ε are not accepted. The case of linear functions f for given training data can be represented by the following equation:

$$f(x) = \langle w, x \rangle + b \tag{6}$$

Where, $w\in X$ and b∈ R; $<\cdot\,,\,\cdot>$ denotes the dot product in X; the vector w is a normal vector; The parameter $\frac{b}{|w|}$ determines the offset of the system from the origin along the normal vector w. The optimal regression function is given by the minimum of the following functional Φ:

$$\Phi(w,\xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{7}$$

Where, C is a pre-specified SVM tolerance parameter; ξ and ξ^* are slack variables determining the degree to which data points will be penalized if the error is larger than precision parameter ε. A loss function can be introduced to functional Φ to penalize estimation errors. This step is called as soft margin loss setting (Scholkopf and Smola 2002). Some loss functions commonly used include ε-insensitive, Laplacian, Gaussian, Huber’s robust loss, Polynomial and Piecewise polynomial. A detailed description of these loss functions and corresponding density models are provided in Smola and Scholkopf (2004). The loss function most often referred is the ε-insensitive because it can produce sparseness in the support vectors. Fig. 1 illustrates ε-insensitive loss function setting to given data in SVM and the ε-insensitive loss function can be represented by the following equation:

$$L_\varepsilon = \begin{cases} 0 & if\ |f(x)-y| < \varepsilon \\ |f(x)-y| - \varepsilon & otherwise \end{cases} \tag{8}$$

The minimum of the functional Φ problem can be considered as quadratic programming (QP) problem. Using the Lagrange theory of quadratic programming, the solution is given by following equation:

$$\max_{\alpha,\alpha^*} W(\alpha,\alpha^*) = \max_{\alpha,\alpha^*}\left\{-\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*) < x_i, x_j > -\varepsilon\sum_{i=1}^{n}(\alpha_i+\alpha_j^*) + \sum_{i=1}^{n} y_i(\alpha_i-\alpha_i^*)\right\} \tag{9}$$

Where, α_i and α_i^* are Lagrange multipliers. Equation (9) is subjected to following condition:

$$\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \quad and \quad \alpha_i, \alpha_i^* \in [0, C] \tag{10}$$

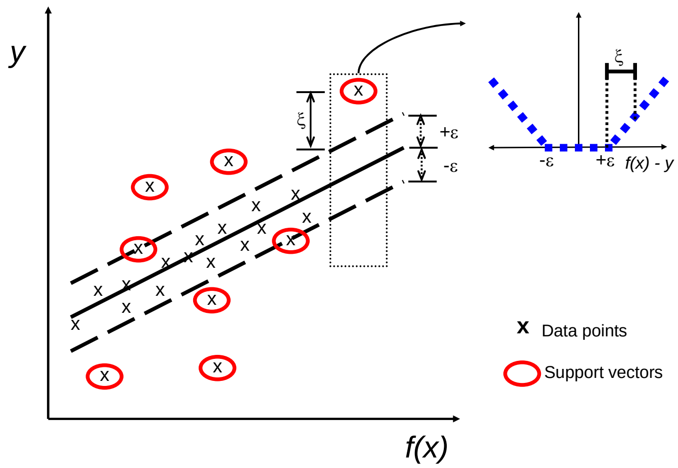


Figure 1. ε-Insensitive loss function setting for SVM in regression.

Solving equation (9) with constraints equation (10) determines the Lagrange multipliers (α_i and α_i^*), and the w and b of regression function given by equation (6) are finally obtained as follows:

$$w = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)x_i, \quad thus\ f(x) = \sum_{i=1}^{n}((\alpha_i - \alpha_i^*) < x_i, x >) + b \tag{11}$$

$$b = \frac{1}{n}\sum_{i=1}^{n}(y_i - \sum_{j=1}^{n}(\alpha_j - \alpha_j^*) < x_i, x_j >) \tag{12}$$

Based on the Karush–Kuhn–Tucker (KKT) conditions demanding that the product between the dual variables and constraints should vanish for optimality, the coefficients α_i and α_i^* of the data points inside ε bound of the function f are zero and the coefficient (α_i - α_i^*) of the data points lying on or outside the ε bound have no-zero value. The support vectors are referred to data points with non-vanishing coefficients (see Fig. 1). According to equation (11), it is evident that support vectors are only used in determining the decision function since the coeffiecient (α_i - α_i^*) of other data points are all equal to zero. In case of non-linear regression, instead of trying to fit a nonlinear model, the training patterns x_i are preprocessed into a high-dimensional feature space R^D by a mapping $\phi : R^d \rightarrow R^D$. Therefore, the dot product $<x_i, x_j>$ in R^d for the linear case is equivalent to $<\phi(x_i), \phi(x_j)>$ in R^D for the non-linear case. The SVM training algorithm would only depend on the data through dot products in R^D. If the dot product in the feature space R^D is expressed by following equation called as Kernel function, one would only need to use K in the SVM training algorithm without treating the feature space explicitly to obtain $\phi(x_i)$;

$$K(x_i, x_j) = < \phi(x_i), \phi(x_j) > \quad (13)$$

Typical Kernel functions described in Gunn (1998) include polynomial, gaussian radial basis function, exponential radial basis function, multi-layer perceptron, and so on. In this way, a nonlinear model in the original space can be transformed to the linear model in the new space. The regression function for linear case as shown in Equations (11) and (12) can be transformed to the one for nonlinear case with Kernel function as follows:

$$f(x) = \sum_{i=1}^{n}((\alpha_i - \alpha_i^*)K(x_i, x)) + b \quad (14)$$

$$b = \frac{1}{n}\sum_{i=1}^{n}(y_i - \sum_{j=1}^{n}(\alpha_j - \alpha_j^*)K(x_i, x_j)) \quad (15)$$

2.2 SVM procedure in regression approximation

Fig. 2 contains a graphical overview of the different steps used in the SVM procedure in regression approximation. The input data for training model are introduced to SVM architecture by specifying error precision (insensitive) parameter ε and SVM tolerance (capacity) parameter C. The ε parameter determines a certain distance of the true value where errors can be ignored. In general, the increase in ε value decreases the number of support vectors and then make the representation of the solution sparser. However, a larger ε can also depreciate the approximation accuracy placed on the training points. In this sense, ε is a tradeoff between the sparseness of the representation and closeness of data (Tay and Cao 2001). The parameter C controls the tradeoff between levels of constraints and complexity of system regulation (Gill et al 2006). The increase in C value makes the problem unconstrained and the decrease in C value assigns more weight to regulation.

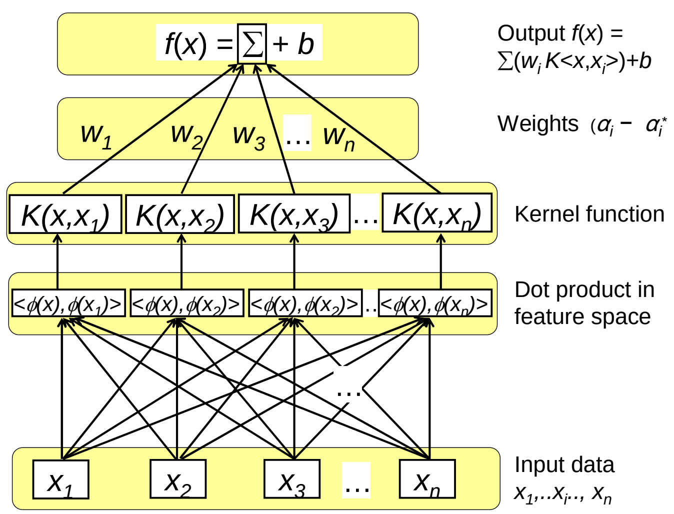


Figure 2. Architecture of SVM in regression.

For the nonlinear case, the input data are transformed into a high-dimensional feature space R^D through a function, ϕ. However, instead of explicitly finding the transformation function ϕ, the dot product of transformed input data could be calculated through the

Kernel function *K* selected through a trial and error procedure. Solving optimal regression function with selected loss functional under given parameters (*C*, ε and *K*) results in finding support vectors, Lagrange multipliers (α_i and α_i^*) and weights $w_i = \alpha_i - \alpha_i^*$. The calculated dot products are added up using the weights w_i. The primal bias term variable *b* is also recovered and formulated to the final predictive model as output. The constructed SVM regression model can be evaluated with new data set (testing data set). The SVM process described is very similar to regression in a NN except that in the SVM case the weights in the input layer are a subset of the training patterns (Smola and Scholkopf 2004).

3. Development of SVM based nonlinear backcalculation models

3.1 Data generation using finite-element program

Data used in this study were generated from a two-dimensional axi-symmetric pavement finite-element (FE) software (ILLI-PAVE) developed at the University of Illinois at Urbana-Champaign (Raad and Figueroa 1980). It incorporates stress-sensitive material models and it provides a more realistic representation of the AC pavement structure and its response to loading. Numerous research studies have validated that this pavement FE model provides a realistic AC pavement structural response prediction for highway and airfield pavements (Thompson and Elliot 1985, Thompson 1992, Gomez-Ramirez et al. 2002, Garg et al. 1998). Therefore, it was used in this study as the main validated nonlinear structural model for analyzing conventional flexible pavements.

A typical FWD test is performed by dropping a 40-kN (9,000-lb) load on the top of circular plate with a radius of 152 mm (6 inches) resting on the surface of the pavement. Typically deflections are measured at offsets of 0 (D_0), 300 (D_{300}), 600 (D_{600}), 900 (D_{900}), 1200 (D_{1200}), and 1500 (D_{1500}) mm from the center of the loading plate. The effect of FWD loading was simulated in the pavement FE program.

The AC surface layer was treated as a linear elastic material with Young's Modulus, E_{AC}, and Poisson ratio, υ. Stress-dependent elastic models along with Mohr-Coulomb failure criteria were applied for the unbound aggregate base and fine-grained soil subgrade layers. The 'stress-hardening' K-θ model (Hicks and Monismith 1971) was used for the base layer:

$$M_R = \frac{\sigma_D}{\varepsilon_R} = K\theta^n \tag{16}$$

Where M_R is resilient modulus (psi), θ is bulk stress (psi) and K and n are statistical parameters. Based on extensive testing of unbound aggregate materials, Rada and Witczak (1981) proposed the following relationship between K and n (R^2 = 0.68, Standard Error of Estimate [SEE] = 0.22):

$$\text{Log10}(K) = 4.657 - 1.807n \tag{17}$$

The 'stress-softening' bilinear model (Thompson and Robnett 1979) was used for the subgrade layer:

$$\begin{aligned} M_R &= M_{Ri} + K_1.(\sigma_d - \sigma_{di}) \quad for\ \sigma_d < \sigma_{di} \\ M_R &= M_{Ri} + K_2.(\sigma_d - \sigma_{di}) \quad for\ \sigma_d > \sigma_{di} \end{aligned} \tag{18}$$

Where M_R is resilient modulus (psi), M_{Ri} is breakpoint resilient modulus (psi), σ_d is applied deviator stress (psi), K_1 and K_2 are statistically determined coefficients from laboratory tests

Asphalt concrete modulus E_{AC}, granular base K-θ model parameter K, and the subgrade soil break point deviator stress M_{Ri} in the bilinear model were used as the layer stiffness inputs for all the different conventional flexible pavement FE runs. The 40-kN (9-kip) wheel load was applied at a uniform pressure of 552 kPa (80 psi) over a circular area of radius 152 mm (6 in.). The thickness and moduli ranges used in developing the synthetic solutions are summarized in Table 1.

Table 1
Pavement geometry and material property/model inputs for FE solutions

Material type	Layer thickness	Material model	Layer modulus inputs	Poisson's ratio
Asphalt concrete	h_{AC} = 76 to 381 mm (3 to 15 in.)	Linear elastic	E_{AC} = 690 to 13,800 MPa (100 to 2,000 ksi)	$\nu = 0.35$
Unbound aggregate base	h_{GB} = 102 to 559 mm (4 to 22 in.)	Nonlinear K-θ model	$M_R = K\theta^n$ "K" = 20.7 to 82.7 MPa (3 to 12 ksi) "n" from Equation 17	$\nu = 0.35$ for $K \geq 34.5$ MPa (5 ksi) $\nu = 0.40$ for $K < 34.5$ MPa (5 ksi)
Fine-grained soil subgrade	7,620 mm (300 in.) minus total pavement thickness	Nonlinear bilinear model	$M_R = f(M_{Ri})$; M_{Ri} = 6.9 to 96.5 MPa (1 to 14 ksi)	$\nu = 0.45$

A total of 8,500 FE runs were conducted by randomly choosing the pavement layer thicknesses and input variables within the given ranges in Table 1 to generate a knowledge database for both SVM and ANN trainings. The total analysis depth of the pavement system was taken as 7,620 mm (300 in.). The subgrade thicknesses were calculated by subtracting the thicknesses of the AC and the base from the total analysis depth. The outputs recorded were the pavement surface deflection basin and the critical pavement responses, radial strain at the bottom of the AC layer (ε_{AC}), vertical strain on top of the subgrade (ε_{SG}), and the deviator stress on top of the subgrade layer (σ_D).

3.2 SVM modeling

The six deflections computed at radial offset values from the load center which define the deflection basin, the thickness of AC surface layer (T_{AC}), and the thickness of the base layer (T_b) together formed the eight input features for the pavement layer moduli backcalculation models. The natural subgrade is assumed to be of infinite thickness and was not considered. The modulus of the AC surface layer, E_{AC}, the nonlinear modulus parameter of the unbound aggregate base layer K_b, and the nonlinear elastic modulus of the fine-grained soil subgrade layer, M_{Ri}, represent the three output vectors. The eight input parameters were used in the development of SVR models denoted as 'SVM E_{AC}' with one output variable of E_{AC}, and 'SVM M_{Ri}' with one output variable of M_{Ri}. The prediction of base layer modulus parameter, K_b is not straightforward as confirmed by previous studies (Ceylan et al. 2007). Therefore, both the SVM predicted E_{AC} and M_{Ri} were used as inputs in addition to AC surface layer thickness, base layer thickness, and the first four surface deflections in developing 'SVM K_b' model with one output variable of K_b. The input and output parameters for the developed SVM models are summarized in Table 2.

One of the important steps in SVM model development is the setting up of the appropriate Kernel function *K* and parameters *C* and *ε* for training the SVM. This study used the Gaussian function shown in Equation (19) as the SVM Kernel function because Gaussian kernels tend to give good performance under general smoothness assumptions (Tay and Cao, 2001).

$$K(x_i, x_j) = \exp\left(\frac{-1}{\delta^2 (x_i - x_j)^2}\right) \tag{19}$$

Table 2
Summary of input and output parameters used in developing SVM backcalculation models

SVM Model	Output	Input
SVM E_{AC}	E_{AC}	T_{AC}, T_b, D_0, D_{300}, D_{600}, D_{900}, D_{1200}, and D_{1500}
SVM M_{Ri}	M_{Ri}	T_{AC}, T_b, D_0, D_{300}, D_{600}, D_{900}, D_{1200}, and D_{1500}
SVM K_b	K_b	E_{AC}, M_{Ri}, T_{AC}, T_b, D_0, D_{300}, D_{600}, and D_{900}

Note: E_{AC} – elastic modulus of AC surface layer (MPa); K_b – modulus parameter of unbound base layer (MPa); M_{Ri} – elastic modulus of unbound subgrade layer (MPa); T_{AC} – thickness of AC surface layer (mm); T_b – thickness of unbound base layer (mm); D_0, D_{300}, D_{600}, D_{900}, D_{1200}, and D_{1500} – surface deflections measured at radial offsets of 0, 300, 600, 900, 1200, and 1500 mm, respectively, from the center of the FWD loading plate.

Where δ^2 is the bandwidth of Gaussian Kernel. The values for *C*, *ε*, and δ^2 were selected based on preliminary parametric sensitivity analysis as well as based on previous experience (Gopalakrishnan et al. 2009). The SVM model parameters used in this study are summarized in Table 3.

Table 3
Summary of SVM model parameters

SVM Model Parameter	Value
C, SVM tolerance parameter	50
ε, precision parameter	0.001
K, kernel function	Gaussian
δ^2, kernel bandwidth	0.3

The data were divided randomly into two different subsets as the training data subset and the testing data subset in such a way that they are representative of same statistical population. Ninety percent (i.e., 7,500 data points) of the data were used for training and 10% (1,000 data points) were used for testing. Both datasets were normalized within the range of 0.1 to 0.9. A grid of scatterplot matrices between pairs of variables used in SVM backcalculation modeling is displayed in Fig. 3. These scatterplot matrices present comparisons among many variables visually by presenting orderly collections of bivariate graphs. A 95% bivariate normal density ellipse is imposed on each scatterplot and the variables are considered to be uncorrelated if the ellipse is fairly round and is not diagonally oriented. It can be easily observed that AC surface layer modulus, E_{AC}, is better correlated to surface deflections nearer to the loading plate starting with D_0 whereas subgrade layer modulus, M_{Ri}, is strongly correlated to surface deflections farther from the loading plate. Base layer modulus parameter, K_b, is poorly correlated to all the surface deflections. This is further confirmed by the sample contour plots depicted in Fig. 4 between input and output variables in the synthetic database.

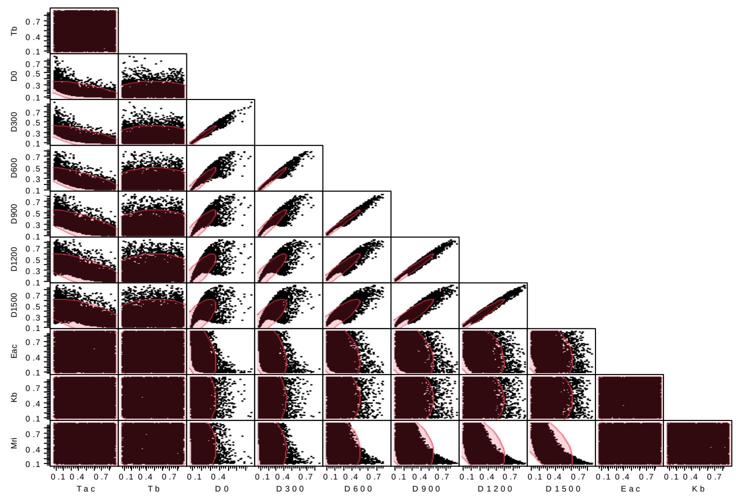


Figure 3. Grid of scatterplot matrices between variables used in SVM modeling.

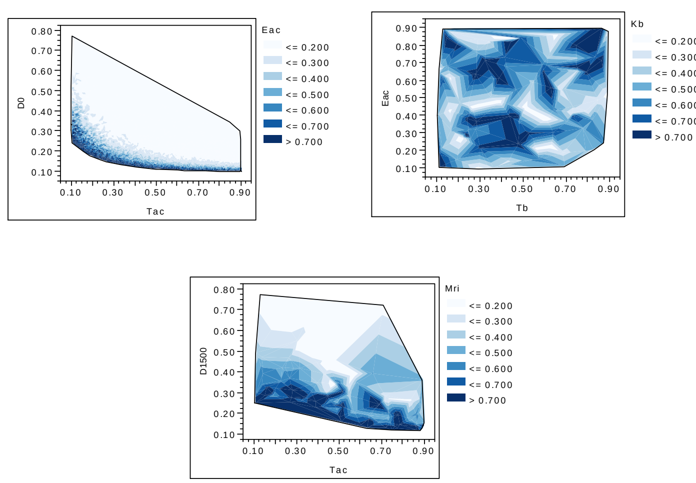


Figure 4. Contour plots showing relationship between normalized input and output variables in the synthetic database: (a) elastic modulus of AC surface layer (E_{AC}), thickness of AC surface layer (T_{AC}), and surface deflection measured at 0-mm radial offset from the center of the loading (D_0); (b) elastic modulus parameter of base layer (K_b), thickness of base layer (T_b), and elastic modulus of AC surface layer (E_{AC}); and (c) elastic modulus of subgrade layer (M_{Ri}), thickness of AC surface layer (T_{AC}), and surface deflection measured at 1500-mm radial offset from the center of the loading (D_{1500})

The training data subset was used for SVR model learning and the testing data subset was used to examine the statistical accuracy of the developed SVR models. As stated earlier, three kinds of models were developed in this study, namely, SVM E_{AC}, SVM K_b and SVM M_{Ri}. The MATLAB toolbox for SVM (Rakotomamonjy and Canu 2008) was adapted for developing the SVM backcalculation models.

4. Evaluation of SVM models

4.1 Evaluation criteria

Quantitative assessments of the degree to how close the models could predict the actual outputs are used to provide an evaluation of the models' predictive performances. A multi-criteria assessment with various goodness-of-fit statistics was performed using the 1,000 test vectors as an independent dataset which was not used in training the models. The criteria that were employed for evaluation of models' predictive performances were the coefficient of correlation (*R*), the coefficient of determination (R^2) with reference to the line of equality, root-mean-square error (*RMSE*) between the actual and predicted values, the absolute average error (*AAE*), and threshold statistic (*TS*). The definitions of these evaluation criteria are provided in Table 4.

The *R* and R^2 are a measure of correlation between the predicted and the measured values and therefore, determines accuracy of the prediction model (higher *R* and R^2 equates to higher accuracy). The *RMSE* and the *AAE* indicate the relative improvement in accuracy and thus a smaller value is indicative of better accuracy. The threshold statistic (*TS*) or cumulative frequency employed as a local statistics criterion provides the distribution of prediction errors (Jain et al. 2001, Padmini et al. 2008). For a level of error *x* %, the threshold statistic, represented as *TSx*, is a measure of the consistency in forecasting errors from a particular model. Thus, a higher number of *TSx* at a given level error *x* % indicate higher accuracy.

4.2 Accuracy of SVM models

A summary of the model performance statistics for the developed models are summarized in Table 5 for both SVM and ANN. The ANN MLP models used in this study are discussed in detail by Gopalakrishnan (2004). The linear correlation between the actual and predicted pavement layer moduli values is shown in Fig. 5 with the corresponding values of R^2. The correlation statistic (*R*) between the actual and predicted modulus is as high as 0.9 for all the models. The value of R^2 that evaluates the performance of the model in predicting the modulus values scattered around the line of equality without bias is found to be more than 0.9 for the all the models except ANN K_b model with about 0.6 of R^2.

A *RMSE* value of 245.5 MPa and *AAE* value of 1.24 % for ANN E_{AC} model are lower than the 1,434.8 MPa of *RMSE* and 9.68 % of *AAE* values for SVM E_{AC}. However, 10.2 MPa of *RMSE* and 14.21 % of *AAE* values for ANN K_b model are higher than 5.1 MPa of *RMSE* and 8.08 % of *AAE* values for SVM $E_{AC.}$ This indicates that the ANN MLP model provides better performance in predicting the modulus of the AC surface layer (E_{AC}) while the SVM provides better prediction performance in the case of unbound aggregate base layer nonlinear modulus parameter (K_b). Little difference in performance statistics between SVM M_{Ri} and ANN M_{Ri} gives an indication that both methods provide similar prediction performance for the nonlinear elastic modulus of the fine-grained soil subgrade layer (M_{Ri}).

SVM predicted nonlinear pavement layer moduli response surfaces are illustrated in Figs. 6, 7, and 8, for E_{AC}, K_b, and M_{Ri}, respectively as functions of input variables to which they are closely correlated.

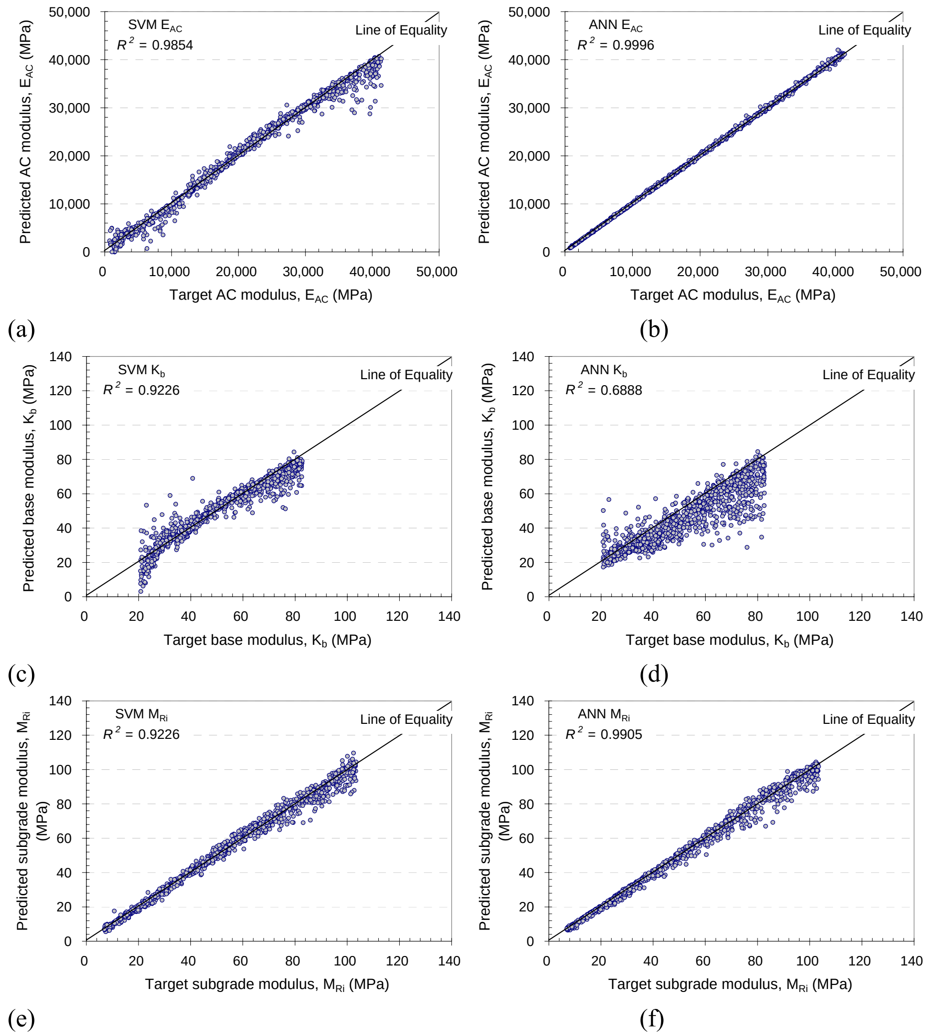


Figure 5. Targeted vs. predicted pavement layer modulus values.

The previously discussed statistical performance evaluation criteria *(R*, R^2, *RMSE*, and *AAE)* provide relevant information on the overall or global performance of the models. The threshold statistic (*TS*) or cumulative frequency was employed to evaluate local performance of the models. The *TS* not only provides the performance index in terms of predicting modulus but also the distribution of prediction errors (Jain et al. 2001, Padmini et al. 2008). This criterion can be expressed for different levels of absolute relative error from the model. For E_{AC} predictions as shown in Fig. 9 (a), the ANN model predicted 98 % of the total number of testing data with less than 5 % relative error while the corresponding value of SVM model is 69 %. However, the SVM model in Fig. 9 (b) shows much better model performance for K_b predictions. Only 24 % of the total number of testing data with less than 5 % relative error was predicted by ANN K_b model but the corresponding value of SVM K_b model is 69 %. Little difference in *TS* values of SVM M_{Ri} and ANN M_{Ri} models were observed in Fig. 9 (c).

Table 4
SVM model performance evaluation criteria

Evaluation criteria	Definition
Coefficient of correlation (*R*)	$R=\frac{\sum_{i=1}^{n}(y_i^t-\overline{y^t})(y_i^p-\overline{y^p})}{\sqrt{\sum_{i=1}^{n}(y_i^t-\overline{y^t})^2}\sqrt{\sum_{i=1}^{n}(y_i^p-\overline{y^p})^2}}$
Coefficient of determination (R^2)	$R^2=1-\frac{\sum_{i=1}^{n}(y_i^t-y_i^p)^2}{\sqrt{\sum_{i=1}^{n}(y_i^t-\overline{y^t})^2}}$
Root-mean-square error (*RMSE*)	$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(y_i^t-y_i^p)^2}{n}}$
Absolute average error (*AAE*)	$AAE=\frac{\sum_{i=1}^{n}\left\lvert\frac{y_i^t-y_i^p}{y_i^t}\times 100\right\rvert}{n}$
Threshold statistic (TS)	$TS_x=\frac{y_x^p}{n}\times 100$

Note: y_i^t and y_i^p are the target and predicted modulus values, respectively, y_i^t and y_i^p are the mean of the target and predicted modulus values corresponding to *n* patterns. y_x^p is the number of predicted modulus values (out of n total predicted) for which the absolute relative error less *x*% from the model.

Table 5
Comparison of model performance

Performance index	Predictive parameters					
	E_{AC}		K_b		M_{Ri}	
	SVM E_{AC}	ANN E_{AC}	SVM K_b	ANN K_b	SVM M_{Ri}	ANN M_{Ri}
R	0.9930	0.9998	0.9616	0.8933	0.9947	0.9955
R^2	0.9854	0.9996	0.9226	0.6888	0.9891	0.9905
RMSE, MPa	1,434.8	245.5	5.1	10.2	3.0	2.8
AAE, %	9.68	1.24	8.08	14.21	3.96	3.50

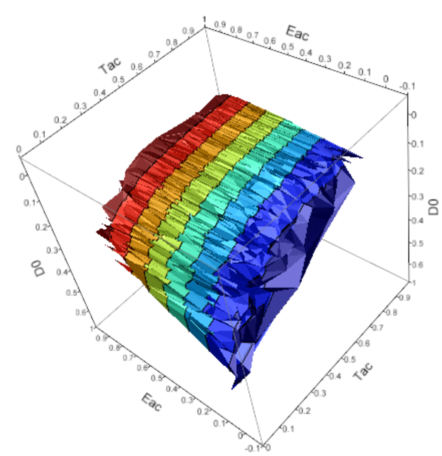


Figure 6. SVM predicted AC surface layer modulus (E_{AC}) response surface as a function of AC surface layer thickness (T_{AC}) and surface deflection at 0-mm radial offset from the center of the loading plate (D_0).

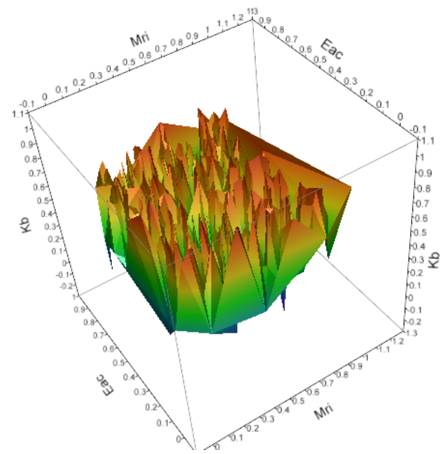


Figure 7. SVM predicted base layer modulus parameter (K_b) response surface as a function of AC surface layer elastic modulus (E_{AC}) and subgrade layer elastic modulus (M_{Ri}).

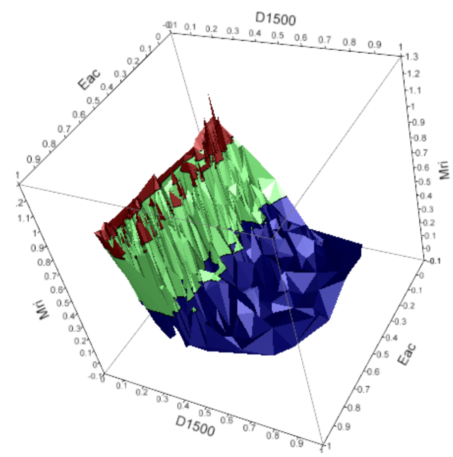
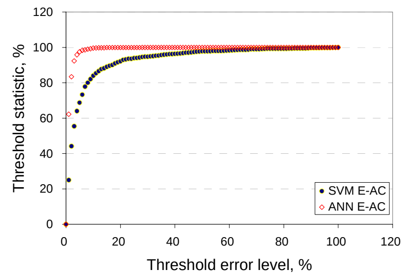


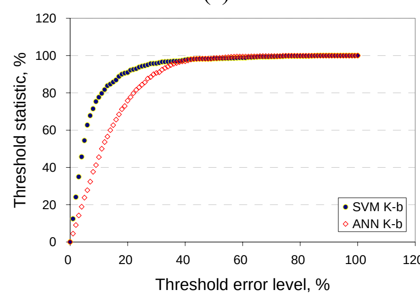
Figure 8. Subgrade layer elastic modulus (M_{Ri}) as a function of AC surface layer elastic modulus (E_{AC}) and surface deflection measured at 1500-mm radial offset from the center of the loading (D_{1500}).

5. Concluding remarks

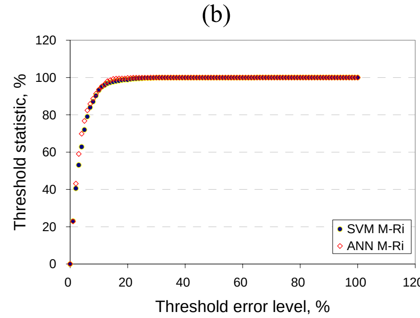
Evaluating the structural condition of existing, in-service asphalt concrete pavement roads using Non-Destructive Test (NDT) is a part of the routine maintenance and rehabilitation activities for sustainable infrastructure. Backcanalysis or backcalculation of pavement mechanical properties (such as elastic modulus) from pavement NDT deflection data is carried out by highway engineers for pavement structural condition evaluation, remaining life calculations, and mechanistic-based analysis. This paper presented an efficient off-line nonlinear pavement backcalculation system based on support vector machines (SVM) and compares its performance with another popular machine learning technique, multi-layer perceptrons (MLP). The results show that the effectiveness of SVM approach in pavement backanalysis is comparable to MLP approach, in predicting Asphalt Concrete (AC) surface layer and nonlinear subgrade layer moduli, and better in predicting unbound nonlinear base layer modulus parameter.



(a)



(b)



(c)

Figure 9. Threshold statistic (TS): SVM Vs ANN models.

However, ANN MLP model development require a large number of controlling parameters to optimize and relatively larger training database while SVM require only three controlling parameters (C, ε, and δ^2) with little dependency on the magnitude of training datasets required. These advantages of SVM can make it a promising alternative to MLP considering the availability of limited and non-representative data frequently encountered in civil infrastructure real-life decision-making situations.

References

Alavi, S., Lecates, J. F., and Tavares, M. P. (2008). Falling weight deflectometer usage, NCHRP Synthesis 381, Transportation Research Board, Washington, DC.

Ali, N. A., and Khosla, N. P. (1987). “Determination of layer moduli using a falling weight deflectometer.” Transportation Research Record: Journal of the Transportation Research Board, 1117, National Research Council, Washington, DC., 1-10.

Alkasawneh, W. (2007). “Backcalculation of pavement moduli using genetic algorithms.” Ph.D. Dissertation, Dept. Civil Eng., Univ. of Akron, Akron, Ohio.

An, S. H., Park, U. Y., Kang, K. I., Cho, M. Y., and Cho, H. H. (2007). “Application of support vector machines in assessing conceptual cost estimates.” ASCE Journal of Computing in Civil Engineering, 21 (4), 259–264.

Burges, C. J. C. (1998). “A tutorial on support vector machines for pattern recognition.” Data Mining and Knowledge Discovery, 2(2), 121– 167.

Caldas, C. H., and Soibelman, L. (2003). "Automating hierarchical document classification for construction management information systems." Automation in Construction, 12, 395–406.

Ceylan, H., Guclu, A., Bayrak, M. B., and Gopalakrishnan, K. (2007). Nondestructive evaluation of Iowa pavements-phase I, CTRE Project 04-177, CTRE, Iowa State University, Ames, Iowa.

Cristianini, N., and Shawe-Taylor, J. (2000). An introduction to support vector machines, Cambridge University Press, Cambridge, UK.

Dibike, Y. B., Velickov, S., Solomatine, D., and Abbot, M. B. (2001). "Model induction with support vector machines: introduction and applications." ASCE Journal of Computing in Civil Engineering, 15 (3), 208–216.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). "From data mining to knowledge discovery in databases." AI Magazine, 17, 37-54.

Fwa, T. F., and Thakur, R. S. (2005). "Seed modulus generation algorithm for backcalculation of flexible pavement moduli." Transportation Research Record: Journal of the Transportation Research Board, 1905, National Research Council, Washington, DC., 117-127.

Fwa, T. F., Tan, C. Y., and Chan, W. T. (1997). "Backcalculation analysis of pavement-layer moduli using genetic algorithms." Transportation Research Record: Journal of the Transportation Research Board, 1570, National Research Council, Washington, DC., 134-142.

Garg, N., Tutumluer, E., and Thompson, M. R. (1998). "Structural modeling concepts for the design of airport pavements for heavy aircraft." Proceedings of the 5th International Conference on the Bearing Capacity of Roads and Airfields, Trondheim, Norway.

Gill, M. K., Asefa, T., Kemblowski, M. W., and McKee, M. (2006). "Soil moisture prediction using support vector machines." Journal of the American Water Resources Association, 42(4), 1033-1046.

Goh, A. T. C., and Goh, S. H. (2007). "Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data." Computers and Geotechnics, 34(5), 410-421.

Goktepe, A. B., Agar, E., and Lav, A. H. (2006). "Advances in backcalculating the mechanical properties of flexible pavements." Advances in Engineering Software, 37, 421–431.

Goktepe, A. B., Agar, E., and Lav, A. H. (2006). "Comparison of multilayer perceptron and adaptive neuro-fuzzy system on backcalculating the mechanical properties of flexible pavements." ARI: The Bulletin of the Istanbul Technical University, 54 (3), 1-13.

Gomez-Ramirez, F., Thompson, M. R., and Bejarano, M. (2002). "ILLI-PAVE based flexible pavement design concepts for multiple wheel heavy gear load aircraft." Proceedings of the 9th International Conference on Asphalt Pavements, Copenhagen, Denmark.

Gopalakrishnan, K. (2004). "Performance analysis of airport flexible pavement subjected to new generation aircraft." Ph.D. Dissertation, University of Illinois at Urbana-Champaign, Illinois.

Gopalakrishnan, K., Kim, S., and Ceylan, H. (2009). "Hot mix asphalt dynamic modulus prediction using kernel machines." Proceedings of ANNIE 2009 Conference, ASME, St. Louis., Missouri.

Gopalakrishnan, K. (2010). Neural Network–Swarm Intelligence Hybrid Nonlinear Optimization Algorithm for Pavement Moduli Back-Calculation. Journal of Transportation Engineering, 136(6), p.528–536.

Gopalakrishnan, K. & Khaitan, S.K., (2010) Development of an intelligent pavement analysis toolbox. Proceedings of the ICE-Transport, 163(4), p.211–221.

Gunn, S. R. (1998). Support vector machines for classification and regression, Technical Report, Image Speech and Intelligent Systems Research Group, University of Southampton, UK.

Herbrich, R. (2002). Learning Kernel classifiers: Theory and Algorithms, MIT Press, MA, US.

Hicks, R. G., and Monismith, C. L. (1971). "Factors influencing the resilient properties of granular materials." Transportation Research Record: Journal of the Transportation Research Board, 345, National Research Council, Washington, DC., 15-31.

Jain, A., Varshney, A. K., and Joshi, U. C. (2001). "Short term water demand forecast modeling at IIT Kanpur using artificial neural networks." Water Resour Manage, 15(5), 299–321.

Lam, K. C., Palaneeswaran, E., and Yu, C. (2009). "A support vector machine model for contractor prequalification." Automation in Construction, 18(3), 321-329.

Meier, R. W., and Rix, G. J. (1995). "Backcalculation of flexible pavement moduli from dynamic deflection basins using artificial neural networks" Transportation Research Record: Journal of the Transportation Research Board, 1473, National Research Council, Washington, DC., 72-81.

Miradi, M. (2009). "Knowledge discovery and pavement performance - intelligent data mining." Ph.D. Dissertation, Delft University of Technology, Nederlands.

Padmini, D., Ilamparuthi, K., and Sudheer, K. P. (2008). "Ultimate bearing capacity prediction of shallow foundations on cohesionless soils using neurofuzzy models." Computers and Geotechnics, 35, 33–46.

Panakkat, A., and Adeli, H. (2009). "Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators." Computer-Aided Civil and Infrastructure Engineering, 24, 280–292.

Raad, L., and Figueroa, J. L. (1980). "Load response of transportation support systems." Transportation Engineering Journal, ASCE, 106(1), 111-128.

Rada, G., and Witczak, M. W. (1981). "Comprehensive evaluation of laboratory resilient moduli results for granular material." Transportation Research Record: Journal of the Transportation Research Board, 810, National Research Council, Washington, DC., 23-33.

Rakotomamonjy, A., and Canu, S. (2008). "SVM and Kernel methods Matlab toolbox", UFR Sciences, Dept Physics, University of Rouen, French, <http://asi.insa-rouen.fr/enseignants/~arakotom/toolbox/index.html> (accessed on Aug. 1, 2009).

Samui, P. (2008). "Support vector machine applied to settlement of shallow foundations on cohesionless soils." Computers and Geotechnics, 35 (3), 419–427.

Scholkopf, B., and Smola, A. J. (2002). Learning with Kernels: support vector machines, regularization, optimization, and beyond, MIT Press, MA, US.

Sharma, S., and Das, A. (2008). "Backcalculation of pavement layer moduli from falling weight deflectometer data using an artificial neural network." Canadian Journal of Civil Engineering, 35 (1), 57-66.

Shoukry, S. N., and William, G. W. (1999). "Performance evaluation of backcalculation algorithms through three-dimensional finite-element modeling of pavement structures." Transportation Research Record: Journal of the Transportation Research Board, 1655, National Research Council, Washington, DC., 150-160.

Smola, A., and Scholkopf, B. (2004). "A tutorial on support vector regression." Statistics and Computing, 14, 199–222.

Tay, F. E. H., and Cao, L. (2001). "Application of support vector machines in financial time series forecasting." Omega, 29 (4), 309-317.

Terzi, S., Saltan, M., and Yildirim, T. (2003) "Optimization of deflection basin by genetic algorithm and neural network approach." Proc. Int. Conf. on Neural Networks, (ICANN/ICONIP), Springer Berlin / Heidelberg, 662-669.

Thompson, M. R. (1992). "ILLI-PAVE based conventional flexible pavement design procedure." Proceedings of the 7th International Conference on Asphalt Pavements, Vo.1, Nottingham, U.K., 222–30.

Thompson, M. R., and Elliott, R. P. (1985). "ILLI PAVE based response algorithms for design of conventional flexible pavements." Transportation Research Record: Journal of the Transportation Research Board, 1043, National Research Council, Washington, DC., 50-57.

Thompson, M. R., and Robnett, Q. L. (1971). "Resilient properties of subgrade soils." Transportation Engineering Journal, ASCE, 105 (1), 71–89.

Ullidtz, P., and Coetzee, N. F. (1995). "Analytical procedures in nondestructive testing pavement evaluation." Transportation Research Record: Journal of the Transportation Research Board, 1482, National Research Council, Washington, DC., 61-66.

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

Vapnik, V. (1995). The Nature of statistical learning theory, Springer, NY.

Vapnik, V., and Chervonenkis, A. (1964). "A note on one class of perceptrons." Automation and Remote Control, 25.

Zhao, H. B. (2008). "Slope reliability analysis using a support vector machine." Computers and Geotechnics, 35 (3), 459–467.