

2009

# Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines

Kasthurirangan Gopalakrishnan  
*Iowa State University*, rangan@iastate.edu

Sunghwan Kim  
*Iowa State University*, sunghwan@iastate.edu

Halil Ceylan  
*Iowa State University*, hceylan@iastate.edu

Follow this and additional works at: [http://lib.dr.iastate.edu/ccee\\_conf](http://lib.dr.iastate.edu/ccee_conf)

 Part of the [Civil and Environmental Engineering Commons](#)

---

## Recommended Citation

Gopalakrishnan, Kasthurirangan; Kim, Sunghwan; and Ceylan, Halil, "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines" (2009). *Civil, Construction and Environmental Engineering Conference Presentations and Proceedings*. 13.  
[http://lib.dr.iastate.edu/ccee\\_conf/13](http://lib.dr.iastate.edu/ccee_conf/13)

This Conference Proceeding is brought to you for free and open access by the Civil, Construction and Environmental Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Civil, Construction and Environmental Engineering Conference Presentations and Proceedings by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

---

# Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines

## **Abstract**

This paper explores the feasibility of applying support vector regression (SVR) kernel-based supervised learning method to develop hot mix asphalt (HMA) dynamic modulus ( $|E^*|$ ) predictive models. SVR-based prediction models were developed using the latest comprehensive  $|E^*|$  database that is available to the researchers. The SVR model predictions were compared with the existing regression-based prediction model which is employed in the Mechanistic-Empirical Pavement Design Guide (MEPDG). The SVR based  $|E^*|$  models show better prediction accuracy compared to the existing regression models. The determination of optimal function and parameters for SVR algorithm is recommended to improve the prediction performance of SVR based  $|E^*|$  models.

## **Keywords**

Support vector regression, hot mix asphalt, mechanistic-empirical pavement design guide

## **Disciplines**

Civil and Environmental Engineering

## **Comments**

This is a manuscript of an article from *ANNIE 2009, ANNs in Engineering*, St. Louis, Missouri, November 2-4, pp. 131-138. Posted with permission.

Reference to this paper should be made as follows: Gopalakrishnan, K., Kim<sup>1</sup> S., and Ceylan, H. (2009). "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines," ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

## Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines

**KASTHURIRANGAN  
GOPALAKRISHNAN**

Department of Civil, Construction  
and Environmental Engineering  
Iowa State University  
Ames, Iowa, USA  
rangan@iastate.edu

**SUNGHWAN KIM**

Department of Civil, Construction  
and Environmental Engineering  
Iowa State University  
Ames, Iowa, USA  
sunghwan@iastate.edu

**HALIL CEYLAN**

Department of Civil, Construction  
and Environmental Engineering  
Iowa State University  
Ames, Iowa, USA  
hceylan@iastate.edu

### ABSTRACT

This paper explores the feasibility of applying support vector regression (SVR) kernel-based supervised learning method to develop hot mix asphalt (HMA) dynamic modulus ( $|E^*|$ ) predictive models. SVR-based prediction models were developed using the latest comprehensive  $|E^*|$  database that is available to the researchers. The SVR model predictions were compared with the existing regression-based prediction model which is employed in the Mechanistic-Empirical Pavement Design Guide (MEPDG). The SVR based  $|E^*|$  models show better prediction accuracy compared to the existing regression models. The determination of optimal function and parameters for SVR algorithm is recommended to improve the prediction performance of SVR based  $|E^*|$  models.

### INTRODUCTION

A conventional statistical (e.g. multivariate regression analysis) or mathematical (e.g. differential equation) approach for system modeling does not always perform well in dealing problems with complex and uncertain conditions. Rapid advances made in information processing systems in recent decades have inspired new approaches for system modeling. These new approaches attempt to imitate the human brain and perception by building intelligent systems that can learn automatically from previous experiences. A collection of these new approaches are referred to as artificial intelligence (AI) and have been applied in different areas of science, engineering, medicine, etc.

One of the broadest subfields in AI is the machine learning (ML) method which focuses on the development of data modeling techniques and algorithms that learn from data. Support Vector Machine (SVM) is one such promising ML technique derived from

**Reference to this paper should be made as follows: Gopalakrishnan, K., Kim<sup>2</sup> S., and Ceylan, H. (2009). "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines," ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.**

statistical learning theory by Vapnik and Chervonenkis (1964) in the Institute for Control Problems of the Russian Academy of Sciences. The foundations of SVM were developed by Vapnik (1995) at AT&T Bell Laboratories and SVM has been recognized as an attractive and promising tool to solve classification and regression related problems (Gunn, 1998). SVM is used as a general terminology that refers to both support vector classification (SVC) for classification problems as well as support vector regression (SVR) for regression problems.

The asphalt concrete mixture, or hot mix asphalt (HMA) is a composite material consisting of aggregate, sand, and filler, bound by asphalt binder. The HMA mechanical behavior is affected by individual component properties but shows very different response with respect to individual component responses. As a result, the prediction of HMA mechanical properties involves a high degree of complexity and uncertainty. The stiffness of HMA is an important mechanical property used in determining pavement response and performance under loading. The HMA dynamic modulus ( $|E^*|$ ), one of the stiffness measures, is the primary HMA material property input in the new Mechanistic Empirical Pavement Design Guide (MEPDG) developed under National Cooperative Highway Research Program (NCHRP) 1-37A (2004) for the American State Highway and Transportation Officials (AASHTO). The MEPDG currently employs a purely statistical regression model, namely Witczak  $|E^*|$  predictive model developed in 1999 (Andrei et al., 1999). This Witczak  $|E^*|$  model was based on conventional multivariate regression analysis of laboratory test data. The researchers at Iowa State University (ISU) (Ceylan et al., 2007; Ceylan et al., 2008; Ceylan et al., 2009) are the first to introduce AI techniques in developing  $|E^*|$  predictive models. The next-generation predictive  $|E^*|$  models developed at ISU are based on backpropagation neural networks (BPNN) approach and were found to be more accurate compared to existing multivariate regression based model (Ceylan et al., 2007; Ceylan et al., 2008; Ceylan et al., 2009).

The primary objective of this study is to explore the feasibility of employing another promising AI technique, SVR, to develop HMA stiffness predictive models. Compared to BPNN, SVR uses a set of linear functions and loss functions to carry out risk (error) minimization for nonlinear regression analysis (Tay and Cao 2001). As a consequence, SVR is able to provide efficient generalization performance, unique, optimal and nearer to global solutions. The SVR algorithm and the development and performance of SVR based  $|E^*|$  predictive models are discussed in the following sections.

### **Overview of SVr algorithm**

Comprehensive tutorials on SVR 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 SVR algorithm described here is primarily based on Gunn (1998) and Smola and Scholkopf (2004).

The SVR algorithm aims to find a function  $f(x)$  that has at most  $\epsilon$  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 the precision parameter  $\epsilon$  are accepted and errors larger than  $\epsilon$  are not accepted. The case of linear functions  $f$  for given training data can be represented by the following equation:

**Reference** to this paper should be made as follows: Gopalakrishnan, K., Kim S., and Ceylan, H. (2009). “Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines,” ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

$$f(x) = \langle w, x \rangle + b \tag{1}$$

where,  $x, w \in X$  and  $b \in \mathbb{R}^d$ ;  $\langle \cdot, \cdot \rangle$  denotes the dot product in  $X$ ; the vector  $w$  is a normal vector;  $x$  is training data; the parameter  $b/|w|$  determines the offset of the system from the origin along the normal vector  $w$ . The optimal regression function is given by minimization of function consisting of  $|w|$  and slack variables  $\xi$  by introducing loss function and using the Lagrange theory of quadratic programming. The optimal regression function for linear case is:

$$f(x) = \sum_{i=1}^n ((\alpha_i - \alpha_i^*) \langle x_i, x \rangle) + b \tag{2}$$

Based on the Karush–Kuhn–Tucker (KKT) conditions demanding that the product between the dual variables and constraints should vanish for optimality, the coefficients  $\alpha_i$  and  $\alpha_i^*$  of the training data points inside  $\epsilon$  bound of the function  $f$  are zero and the coefficient  $(\alpha_i - \alpha_i^*)$  of the training data points lying on or outside the  $\epsilon$  bound have non-zero value. The support vectors are referred to training data points with non-vanishing coefficients.

In case of non-linear regression problem, instead of fitting a nonlinear model, the training data  $x$  are mapped into a high-dimensional feature space  $\mathbb{R}^D$  by function  $\phi: \mathbb{R}^d \rightarrow \mathbb{R}^D$ . Therefore, the dot product  $\langle x_i, x_j \rangle$  in  $\mathbb{R}^d$  for the linear case is equivalent to  $\langle \phi(x_i), \phi(x_j) \rangle$  in  $\mathbb{R}^D$  for the non-linear case. The SVR training algorithm would only depend on the data through dot products in  $\mathbb{R}^D$ . If the dot product in the feature space  $\mathbb{R}^D$  is expressed by following Eq. (3), called as Kernel function, one would only need to use  $K$  in the SVR training algorithm without treating the feature space explicitly to obtain  $\phi(x_i)$ :

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \tag{3}$$

In this way, a nonlinear model in the original space can be transformed to the linear model in the new space. The optimal regression function for linear case as shown in Eq. (2) can be transformed to the one for nonlinear case with Kernel function as follows:

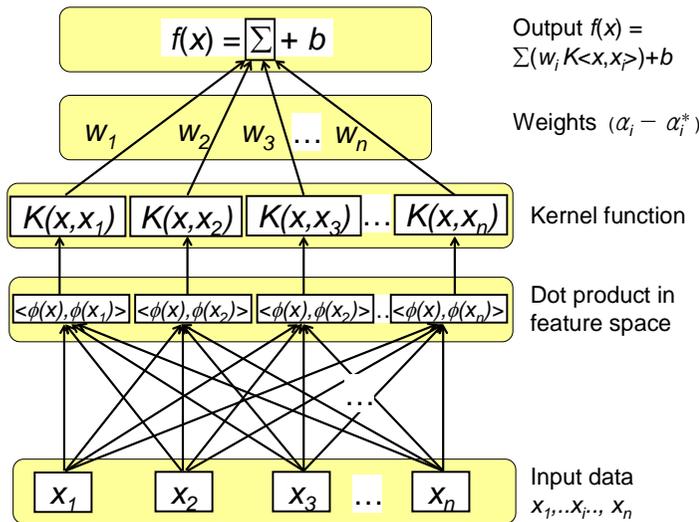
$$f(x) = \sum_{i=1}^n ((\alpha_i - \alpha_i^*) K(x_i, x)) + b \tag{4}$$

Figure 1 contains a graphical overview of the different steps in the SVR procedure. The input data for training model are introduced to SVR architecture specifying error precision (insensitive) parameter  $\epsilon$  and tolerance (capacity) parameter  $C$ . The  $\epsilon$  parameter determines a certain distance of the true value where errors can be ignored. In general, increase in  $\epsilon$  value decreases the number of support vectors and makes the representation of the solution sparser. However, a larger  $\epsilon$  can also depreciate the approximation accuracy placed on the training points. In this sense,  $\epsilon$  is a tradeoff

**Reference** to this paper should be made as follows: Gopalakrishnan, K., Kim S., and Ceylan, H. (2009). "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines," ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

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). Increase in  $C$  value makes the problem unconstrained and decrease in  $C$  value assigns more weight to regulation.

For nonlinear case, the input data are transformed into a high-dimensional feature space  $R^D$  through a function,  $\phi$ . However, instead of explicitly finding the transformation function  $\phi$  the dot product of transformed input data could be calculated through the appropriate Kernel function  $K$ , which is selected through a trial and error procedure. Solving optimal regression function with selected loss function under given parameters ( $C$ ,  $\varepsilon$  and  $K$ ) results in finding support vectors, Lagrange multipliers ( $\alpha_i$  and  $\alpha_i^*$ ) 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 SVR model can then be evaluated with new data set (testing data set). The SVR process described here is very similar to regression in conventional neural networks (NN) except that in case of SVR, the weights in the input layer are a subset of the training patterns (Smola and Scholkopf, 2004).



**Figure 1. Architecture of SVR**

## DEVELOPMENT OF SVR-BASED $|E^*|$ PREDICTION MODEL

Data used in this study were retrieved from the National Cooperative Highway Research Program (NCHRP) Report 567 DVD (CRP-CD - 46) "Simple Performance Tests: Summary of Recommended Methods and Database." (Witczak, 2005). The CRP-CD-46 included as an appendix in the NCHRP report 567 contains a total of 7,400 data records from 346 HMA mixtures (Bari and Witczak, 2006). The new pavement design guide software in US, namely the MEPDG software, employs Witczak  $|E^*|$  predictive model

**Reference** to this paper should be made as follows: Gopalakrishnan, K., Kin~~5~~ S., and Ceylan, H. (2009). “Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines,” ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

developed in 1999 as one of the user options depending on the availability of input parameter data.

The eight-input parameters for the 1999 version of Witczak  $|E^*|$  model include aggregate gradation, mixture volumetrics, viscosity of the asphalt binder ( $\eta$ ), and loading frequency ( $f$ ). The aggregate gradation variables include percent passing #200 sieves ( $\rho_{\#200}$ ), percent retained #4 sieve ( $\rho_{\#4}$ ), percent retained 9.5-mm sieve ( $\rho_{9.5\text{mm}}$ ), and percent retained 19-mm sieve ( $\rho_{19\text{mm}}$ ). The mixture volumetrics include air void ( $V_a$ ) and effective binder content ( $V_{\text{beff}}$ ). The eight input parameters of the Witczak  $|E^*|$  predictive model were used in the development of SVR model denoted as ‘SVR 1999’ with one output variable,  $|E^*|$ .

One of the important steps in SVR model development is the setting up of the appropriate Kernel function  $K$  and parameters  $C$  and  $\epsilon$  for training the SVR. This study used the Gaussian function shown in Eq. (5) as the SVR Kernel function because Gaussian kernels tend to give good performance under general smoothness assumptions (Tay and Cao, 2001).

$$K(x_i, x_j) = \exp\left(\frac{-1}{\delta^2(x_i - x_j)^2}\right) \quad (5)$$

where,  $\delta$  is the bandwidth of Gaussian Kernel. During the preliminary investigation, this study examined the performance of SVR based predictive  $|E^*|$  models with various values of  $\epsilon$  while  $C$  and  $\delta$  values were fixed at 1,000 and 0.3, respectively. The following values of  $\epsilon$  were considered: 0.01, 0.05, and 0.1. A comprehensive parametric analysis is recommended as part of the future research efforts to determine optimal selection of the Kernel function  $K$  and the parameters  $C$  and  $\epsilon$  for the SVR based predictive  $|E^*|$  models.

The data were divided randomly into two different subsets: the training data subset containing 6,900 data points and the testing data subset which consisted of 500 data points. Both datasets were normalized within the range of 0.1 to 0.9 for input and output values to prevent network saturation, which could impede the network’s performance. The training data subset was used for SVR  $|E^*|$  prediction model learning and the testing data subset was used to examine the statistical accuracy of the developed SVR model. The SVR code implemented in MATLAB was adopted from Canu et al. (2005).

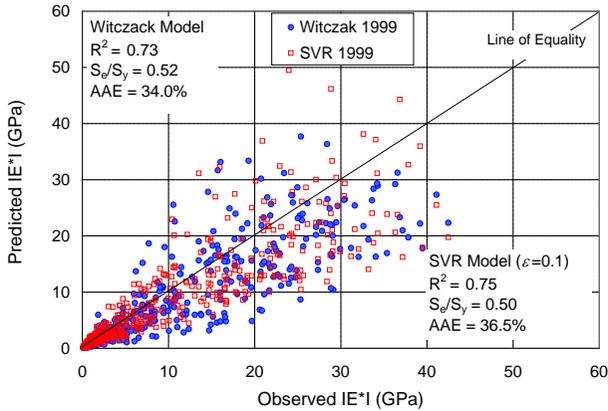
## RESULTS AND DISCUSSIONS

The *goodness-of-fit* statistics for the SVR model predictions in arithmetic scale were performed using statistical parameters such as the correlation coefficient ( $R^2$ ), the standard error of predicted values divided by the standard deviation of measured values ( $S_e/S_y$ ) and the absolute average error (AAE). The  $R^2$  is a measure of correlation between the predicted and the measured values and therefore, determines accuracy of the fitting model (higher  $R^2$  equates to higher accuracy). The  $S_e/S_y$  and the AAE indicates the relative improvement in accuracy and thus a smaller value is indicative of better accuracy.

Reference to this paper should be made as follows: Gopalakrishnan, K., King S., and Ceylan, H. (2009). "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines," ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

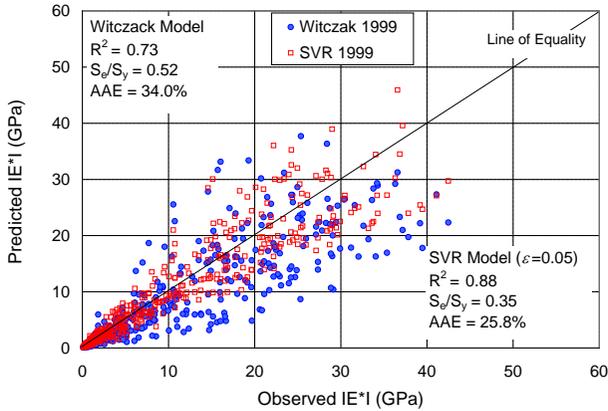
The results of statistical analysis are presented in Fig. 2 for the 500 testing data points. As mentioned previously, the 500 test vectors form an independent dataset which was not used in training the SVR and it was used to test the accuracy of the trained SVR. The predictions of SVR 1999 models with  $\epsilon$  values at 0.01 and 0.05 show better goodness-of-fit statistics compared to those of Witczak model. However, when  $\epsilon$  is increased to 0.1, the accuracy of SVR 1999 prediction decreases. In general, the accuracy of SVR model predictions seem to improve with decrease in the values of  $\epsilon$ , at least until a certain point.

Especially, the AAE obtained using SVR with  $\epsilon$  value at 0.01 is almost one-half that of Witczak's model. It is also noticed that the Witczak predictions are more scattered below the line of equality (45 degree line) with increasing  $|E^*|$  values. This indicates that the Witczak  $|E^*|$  model seems to under-predict the actual measurement. In terms of pavement performance, this prediction inaccuracy may translate into the risk of premature failure of the asphalt layer. However, SVR model predictions ( $\epsilon = 0.01$ ) are tightly scattered around the line of equality without bias and therefore there is a higher chance of preventing premature distress failure.

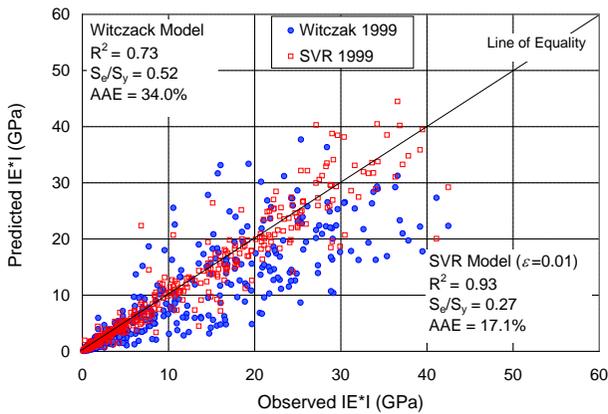


(a)

Reference to this paper should be made as follows: Gopalakrishnan, K., Kin7 S., and Ceylan, H. (2009). "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines," ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.



(b)



(c)

Figure 2. Predicted versus observed  $|E^*|$  for Witczak and SVR models: (a)  $\epsilon = 0.1$ , (b)  $\epsilon = 0.05$ , and (c)  $\epsilon = 0.01$

## CONCLUSIONS

This paper explores the feasibility of applying support vector regression (SVR) machine learning technique for developing hot mix asphalt (HMA) dynamic modulus ( $|E^*|$ ) predictive models without compromising prediction accuracy. SVR-based prediction models were developed using the latest comprehensive  $|E^*|$  database that is available to the researchers containing 7,400 data points from 346 HMA mixtures. It was found that SVR models showed significantly better performance compared to existing multivariate

**Reference** to this paper should be made as follows: Gopalakrishnan, K., Kim S., and Ceylan, H. (2009). "Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines," ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

regression-based Witczak model for  $|E^*|$  prediction. The results of this study have significant implications in the context of advancing the state of the art in mechanistic-empirical pavement analysis and design. SVR models trained over comprehensive datasets could be successfully incorporated into MEPDG as surrogates for pavement materials characterization models and pavement performance prediction models. Future research efforts will focus on determination of optimal SVR model to improve the performance of SVR based  $|E^*|$  prediction models.

## REFERENCES

- Andrei, D., Witczak, M. W., and Mirza, M. W., 1999, "Development of a Revised Predictive Model for the Dynamic (Complex) Modulus of Asphalt Mixtures," National Cooperative Highway Research Program 1-37 A Inter Team Report, University of Maryland, US.
- Bari, J. and Witczak, M. W., 2006, "Development of a New Revised Version of the Witczak  $E^*$  Predictive Model for Hot Mix Asphalt Mixtures," *Journal of the Association of Asphalt Paving Technologists*, Vol. 75, pp. 381- 423.
- Burges, C.J.C., 1998, "A tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, Vol. 2(2), pp. 121- 167.
- Canu, S., Grandvalet, Y., Guigue, V. and Rakotomamonjy, A., 2005, "SVM and Kernel Methods Matlab Toolbox", Perception Systèmes et Information, INSA de Rouen, Rouen, France.
- Ceylan, H., Kim, S and Gopalakrishnan, K., 2007, "Hot Mix Asphalt Dynamic Modulus Prediction Models Using Neural Network Approach," *Intelligent Engineering Systems through Artificial Neural Networks (the ANNIE 2007)*, American Society of Mechanical Engineers, Vol. 17, pp. 117-124.
- Ceylan, H., Gopalakrishnan, K. and Kim, S., 2008, "Advanced Approaches to Hot-Mix Asphalt Dynamic Modulus Prediction," *Canadian Journal of Civil Engineering*, Vol. 35 (7), pp. 699-707.
- Ceylan, H., Gopalakrishnan, K. and Kim, S., 2009, "Looking to the Future: the Next-Generation Hot Mix Asphalt Dynamic Modulus Prediction Models," *International Journal of Pavement Engineering*, DOI: 10.1080/10298430802342690.
- Cristianini, N. and Shawe-Taylor, J., 2000, "An Introduction to Support Vector Machines," Cambridge University Press, Cambridge, UK.
- Gill, M. K., Asefa, T., Kembrowski, M. W., and McKee, M., 2006, "Soil Moisture Prediction Using Support Vector Machines," *Journal of the American Water Resources Association*, Vol. 42(4), pp. 1033-1046.
- Gunn, S. R., 1998, "Support Vector Machines for Classification and Regression," Technical Report, Image Speech and Intelligent Systems Research Group, University of Southampton, UK. Available online at: <http://www.isis.ecs.soton.ac.uk/resource/svminfo/>.
- Herbrich, R., 2002, "Learning Kernel Classifiers: Theory and Algorithms," MIT Press, MA, US.
- National Cooperative Highway Research Program (NCHRP), 2004, "Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures," National Cooperative Highway Research Program 1-37 A Report, Transportation Research Board, Washington DC, US, Available one line at: <http://trb.org/mepdg>.
- Schölkopf, B. and Smola, A.J., 2002, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," MIT Press, MA, US.
- Smola, A. and Schölkopf, B., 2004, "A Tutorial on Support Vector Regression," *Statistics and Computing*, Vol. 14, pp. 199-222.
- Tay, F.E.H. and Cao, L., 2001, "Application of Support Vector Machines in Financial Time Series Forecasting," *Omega*, Vol. 29 (4), pp. 309-317.
- Vapnik, V., 1995, "The Nature of Statistical Learning Theory," Springer, NY, US.

**Reference** to this paper should be made as follows: Gopalakrishnan, K., King S., and Ceylan, H. (2009). “*Hot Mix Asphalt Dynamic Modulus Prediction Using Kernel Machines*,” ANNIE 2009, ANNs in Engineering, St. Louis, Missouri, November 2-4, pp. 131-138.

Vapnik, V. and Chervonenkis, A., 1964, “A Note on One Class of Perceptrons,” *Automation and Remote Control*, 25.

Witczak, M. W., 2005, “Simple performance tests: Summary of Recommended Methods and Database,” National Cooperative Highway Research Program Report 547, Transportation Research Board, Washington DC, US.