

2009

Accuracy of predictive models for dynamic modulus of hot-mix asphalt

Halil Ceylan

Iowa State University, hceylan@iastate.edu

Charles W. Schwartz

University of Maryland - College Park

Sunghwan Kim

Iowa State University, sunghwan@iastate.edu

See next page for additional authors

Follow this and additional works at: http://lib.dr.iastate.edu/ccee_pubs



Part of the [Construction Engineering and Management Commons](#)

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/ccee_pubs/39. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

Accuracy of predictive models for dynamic modulus of hot-mix asphalt

Abstract

Various models have been developed over the past several decades to predict the dynamic modulus E^* of hot-mix asphalt (HMA) based on regression analysis of laboratory measurements. The models most widely used in the asphalt community today are the Witczak 1999 and 2006 predictive models. Although the overall predictive accuracies for these existing models as reported by their developers are quite high, the models generally tend to overemphasize the influence of temperature and understate the influence of other mixture characteristics. Model accuracy also tends to fall off at the low and high temperature extremes. Recently, researchers at Iowa State Univ. have developed a novel approach for predicting HMA E^* using an artificial neural network (ANN) methodology. This paper discusses the accuracy and robustness of the various predictive models (Witczak 1999 and 2006 and ANN-based models) for estimating the HMA E^* inputs needed for the new mechanistic-empirical pavement design guide. The ANN-based E^* models using the same input variables exhibit significantly better overall prediction accuracy, better local accuracy at high and low temperature extremes, less prediction bias, and better balance between temperature and mixture influences than do their regression-based counterparts. As a consequence, the ANN models as a group are better able to rank mixtures in the same order as measured E^* for fixed (e.g., project-specific) environmental and design traffic conditions. The ANN models as a group also produced the best agreement between predicted rutting and alligator cracking computed using predicted versus measured E^* values for a typical pavement scenario.

Keywords

accuracy, asphalt pavements, mixtures, neural networks, predictions

Disciplines

Civil and Environmental Engineering | Construction Engineering and Management

Comments

This is a manuscript of an article from *Journal of Materials in Civil Engineering* 21 (2009): 286, doi: [10.1061/\(ASCE\)0899-1561\(2009\)21:6\(286\)](https://doi.org/10.1061/(ASCE)0899-1561(2009)21:6(286)). Posted with permission.

Authors

Halil Ceylan, Charles W. Schwartz, Sunghwan Kim, and Kasthurirangan Gopalakrishnan

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt

Halil Ceylan¹, Charles W. Schwartz², Sunghwan Kim³, and Kasthurirangan Gopalakrishnan⁴

Abstract: Various models have been developed over the past several decades to predict the dynamic modulus $|E^*|$ of Hot-Mix Asphalt (HMA) based on regression analysis of laboratory measurements. The models most widely used in the asphalt community today are the Witczak (1999 and 2006) predictive models. Although the overall predictive accuracies for these existing models as reported by their developers are quite high, the models generally tend to overemphasize the influence of temperature and understate the influence of other mixture characteristics. Model accuracy also tends to fall off at the low and high temperature extremes. Recently, researchers at Iowa State University (ISU) have developed a novel approach for predicting HMA $|E^*|$ using an Artificial Neural Network (ANN) methodology. This paper discusses the accuracy and robustness of the various predictive models (Witczak 1999 and 2006, and ANN-based models) for estimating the HMA $|E^*|$ inputs needed for the new Mechanistic-Empirical Pavement Design Guide (MEPDG). The ANN-based $|E^*|$ models using the same input variables exhibit significantly better overall prediction accuracy, better local accuracy at high and low temperature extremes, less prediction bias, and better balance between temperature and mixture influences than do their regression-based counterparts. As a consequence, the ANN models as a group are better able to rank mixtures in the same order as measured $|E^*|$ for fixed (e.g., project-specific) environmental and design traffic conditions. The ANN models as a group also produced the best agreement between predicted rutting and alligator cracking computed using predicted vs. measured $|E^*|$ values for a typical pavement scenario.

CE Database subject headings: Dynamic ($|E^*|$) Modulus; HMA; Artificial Neural Network; Prediction Model; MEPDG; Pavement analysis and design.

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

Introduction

Dynamic modulus ($|E^*|$) is one of the fundamental properties defining the response of Hot Mix Asphalt (HMA) mixtures in flexible pavement systems. It is also the primary HMA material property input at all three hierarchical levels 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). Moreover, it is a leading candidate for the Simple Performance Test (SPT) recommended by the NCHRP 9-19 (Witczak et al. 2002) and 9-29 (Bonaquist et al. 2003) and has been recommended as a potential quality control-quality assurance parameter (Bonaquist et al. 2003).

Various $|E^*|$ predictive models have been developed over the last several decades to estimate $|E^*|$ as an alternative to laboratory testing, which can require days of specimen preparation, temperature equilibration, and loading. The most widely used models are the Witczak predictive models (Andrei et al. 1999; Bari and Witczak 2006) based on conventional multivariate regression analysis of laboratory test data.

Level 1 of the MEDPG requires direct measurement of $|E^*|$ in the laboratory. Level 2 and 3 inputs are estimated from regression based $|E^*|$ predictive models. The early versions (v0.7 to 0.9) of the MEDPG (NCHRP. 2006a) incorporated only the $|E^*|$ predictive model developed by Witczak and colleagues in 1999 (Andrei et al. 1999). Inputs to the Witczak 1999 model include the binder viscosity at the design temperature, loading frequency (a function of design traffic speed), aggregate gradation characteristics, and mixture volumetric properties. One disadvantage of the Witczak 1999 model is its characterization of the binder in terms of conventional viscosity (η) rather than the dynamic shear modulus ($|G_b^*|$) now in common use as part of the Superpave Performance Graded (PG) binder specification. A new revised version of the Witczak $|E^*|$ predictive model (Witczak 2006 model) overcomes this disadvantage by characterizing the binder directly in terms of $|G_b^*|$ (Bari and Witczak 2006). The Witczak 2006 model, which was calibrated using a much more extensive laboratory testing database, is incorporated in the latest version of MEDPG software (version 1.000) along with the original Witczak 1999 model (NCHRP. 2006b).

Several studies have indicated that the Witczak $|E^*|$ models show significant scatter especially at the low and/or high $|E^*|$ modulus extremes (Pellinen, 2001; Schwartz 2005; Dongre et al. 2005; Bari and Witczak 2006; Al-Khateeb et al. 2006; Azari et al. 2007). There are also suggestions that the Witczak $|E^*|$ predictive models are dominated by the influence of temperature and understate the influence of other mixture parameters (Schwartz 2005). This indicates that these $|E^*|$ predictive models may not be able to predict successfully the performance differences among different HMA mixtures under a given set of project-specific environmental conditions and design traffic.

Recently, researchers at Iowa State University (ISU) developed a novel approach for predicting HMA $|E^*|$ using an Artificial Neural Network (ANN) methodology (Ceylan et al. 2007; Ceylan et al. 2008). ANN models have been developed using the same input parameters as in the Witczak 1999 and 2006 (Ceylan et al. 2007). Bari's (2005) comprehensive laboratory $|E^*|$ database containing 7,400 data records was used in the

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

development of ANN-based models and for comparing the prediction accuracies of these models against those for the corresponding Witczak models.

The primary objective of this study is to answer two fundamental questions:

- "How accurate and robust are the $|E^*|$ predictions?" This question will be addressed by: (a) reviewing the model formulations and the goodness-of-fit statistics for predictive models; and (b) evaluating overall and/or local biases in the predictions using a full data set and subsets stratified by temperature.
- "How accurate do the $|E^*|$ predictions need to be?" This question will be addressed by: (a) the ability of models to rank mixtures in the same order as for measured dynamic modulus; and (b) the variations in predicted pavement performance attributable to different prediction accuracy levels of the various models. Pavement performance will be predicted using version 1.000 of the MEPDG.

Model Formulations

Existing Regression-Based Models

The Witczak 1999 $|E^*|$ predictive model included in the earlier versions of MEPDG software is presented in Fig. 1(a) (Andrei et al. 1999). This model was developed from a large database containing 2,750 test data points from 205 un-aged laboratory blended HMA mixtures including 34 modified binders. The input variables for the 1999 version $|E^*|$ model include aggregate gradation, mixture volumetric properties, viscosity of the asphalt binder (η), and loading frequency (f). The aggregate gradation variables are the percent passing the #200 sieve ($\rho_{\#200}$), percent retained on the #4 sieve ($\rho_{\#4}$), percent retained on the 9.5 mm sieve ($\rho_{9.5\text{mm}}$), and percent retained on the 19 mm sieve ($\rho_{19\text{mm}}$). The mixture volumetric properties are the air void percentage (V_a) and effective binder percentage by volume (V_{beff}).

The revised Witczak 2006 $|E^*|$ model is shown in Fig. 1(b) (Bari and Witczak 2006). This model was developed using Bari's (2005) database of 7,400 measured $|E^*|$ values obtained from 346 different HMA mixes. The data used to develop the earlier 1999 version of the model are included in this expanded database. In addition to the expanded database, the Witczak 2006 model replaces binder viscosity (η) and loading frequency (f) with the binder dynamic shear modulus ($|G_b^*|$) and phase angle (δ_b).

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

$$\log E^* = -1.25 + 0.029\rho_{200} - 0.0018(\rho_{200})^2 - 0.0028\rho_4 - 0.058V_a - 0.822 \frac{V_{beff}}{V_a + V_{beff}}$$

$$+ \frac{3.872 - 0.0021\rho_4 + 0.004\rho_{38} - 0.000017(\rho_{38})^2 + 0.0055\rho_{34}}{1 + e^{(-0.603313 - 0.313351 \log(f) - 0.393552 \log(\eta))}}$$

where,

- E^* = dynamic modulus of mix, 10^5 psi
- η = viscosity of binder, 10^6 Poise
- f = loading frequency, Hz
- ρ_{200} = % passing #200 (0.075 mm) sieve
- ρ_4 = cumulative % retained on #4 (4.76 mm) sieve
- ρ_{38} = cumulative % retained on 3/8 in (9.5 mm) sieve
- ρ_{34} = cumulative % retained on 3/4 in (19 mm) sieve
- V_a = air void, % by volume
- V_{beff} = effective binder content, % by volume

(a)

$$\log_{10} E^* = -0.349 + 0.754 \left(|G_b^*|^{-0.0052} \right)$$

$$\times \left(6.65 - 0.032\rho_{200} + 0.0027\rho_{200}^2 + 0.011\rho_4 - 0.0001\rho_4^2 \right)$$

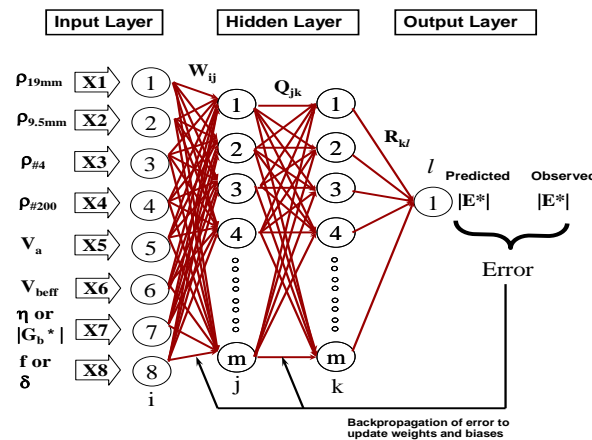
$$\times \left(+ 0.006\rho_{38} - 0.00014\rho_{38}^2 - 0.08V_a - 1.06 \left(\frac{V_{beff}}{V_a + V_{beff}} \right) \right)$$

$$+ \frac{2.56 + 0.03V_a + 0.71 \left(\frac{V_{beff}}{V_a + V_{beff}} \right) + 0.012\rho_{38} - 0.0001\rho_{38}^2 - 0.01\rho_{34}}{1 + e^{(-0.7814 - 0.5785 \log(G_b^*) + 0.8834 \log(\delta_b))}}$$

where,

- E^* = dynamic modulus, psi
- ρ_{200} = percentage of aggregates (by weight of the total aggregates) passing through no. 200 sieve, %
- ρ_4 = aggregates (by weight) retained on no. 4 sieve, %
- ρ_{38} = aggregates (by weight) retained on the 3/8 inch sieve, %
- ρ_{34} = aggregates (by weight) retained on the 3/4 inch sieve, %
- V_a = air voids (by volume of the mix), %
- V_{beff} = effective binder content (by volume of the mix), %
- $|G_b^*|$ = dynamic shear modulus of binder, psi
- δ_b = phase angle of binder associated with $|G_b^*|$, degree

(b)



(c)

Fig. 1. The $|E^*|$ predictive models evaluated in this study: (a) Witczak 1999 (Andrei et al. 1999); (b) Witczak 2006 (Bari and Witczak 2006); (c) ANN 1999 and ANN 2006 (Ceylan et al. 2007)

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

New empirical models to convert conventional viscosity temperature susceptibility parameters A and VTS to $|G_b^*|$ and δ_b are also provided for the Witczak 2006 model as fall-back in the event that Superpave binder characterization data are unavailable (Bari and Witczak, 2007). Some issues have been raised regarding the inconsistent treatment of loading frequency for mixtures and binder in the Witczak 2006 model (e.g., Christiansen, 2006). However, for the present purpose of comparing ANN vs. regression $|E^*|$ prediction models, the Witczak 2006 model is used as defined by the developers.

The ANN $|E^*|$ Models

Recently, researchers at ISU developed an approach for predicting HMA $|E^*|$ using an ANN methodology using the same input parameters as the Witczak 1999 and 2006 $|E^*|$ models (Ceylan et al. 2007). Bari's comprehensive $|E^*|$ database containing 7,400 data records used in the development of the Witczak 2006 model (Bari 2005; Bari and Witczak 2006) was also used in developing the ANN-based models.

The ANN methodology (TRB Circular 1999) encompasses a wide array of computational tools loosely patterned after biological processes. All ANNs are interconnected assemblages of mathematically simple computational elements. These computational elements contain a very limited amount of local memory and perform rudimentary mathematical operations on data passing through them. The computational power of ANN comes from parallelism – input data are concurrently operated upon (processed) by multiple computational elements. The recent adoption and use of ANN modeling techniques in the MEPDG (2004) is just one example of the successful use of neural nets in geomechanical and pavement systems. Details regarding the ANN methodology are available in Tsoukalas and Uhrig (1997).

A typical four-layered (i.e., one input- two hidden–one output layer) feed forward error-back propagation ANN architecture, as shown in Fig. 1(c), was used in development of the ANN $|E^*|$ predictive models. The eight input variables of the Witczak 1999 and Witczak 2006 equations were used in the ANN 1999 and ANN 2006 models, respectively. The predicted dynamic modulus $|E^*|$ was the sole output variable in all of the ANN models. For the ANN models, the 7400 data were divided randomly into two different subsets: a training (calibration) subset containing 6,900 data points and a testing (validation) subset containing the remaining 500 data points. The training data subset was used to train the backpropagation ANN $|E^*|$ prediction model and the testing data subset was used to examine the statistical accuracy of the developed ANN model. Note that this is in contrast to standard regression techniques, where all of the data are used to calibrate the model and no data are held back for subsequent validation. The trained ANN models were also finally evaluated using all the 7,400 data points to obtain the overall predictive accuracy and compare it with the existing $|E^*|$ predictive models. Several network architectures with two hidden layers were examined via parametric study to determine the optimum number of hidden layer nodes. Overall, the training and testing mean squared errors (MSEs) decreased with increasing number of neurons in the hidden layers. The 8-30-30-1 architecture (8 inputs, 30 and 30 hidden neurons, and 1 output neurons, respectively) was chosen as the best architecture for both the ANN

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

1999 and ANN 2006 models based on its lowest training and testing MSEs. Details of the development of ANN-based $|E^*|$ models outlined above are described in Ceylan et al. (2007).

Overall Accuracy of Models

Fig. 2(a) summarizes the predicted vs. measured $|E^*|$ values as predicted by the Witczak 1999 and ANN 1999 models. Fig. 2(b) provides similar comparisons for the Witczak 2006/ANN 2006 pairs. All of the comparisons in these figures are presented in arithmetic $|E^*|$ space. These comparisons have sometimes been presented in the literature in log $|E^*|$ space, in part because of the large range of values for $|E^*|$ and in part because the Witczak models are formulated in terms of log $|E^*|$. However, since $|E^*|$ rather than log $|E^*|$ is the direct input to mechanistic-empirical pavement design and since plotting results in log $|E^*|$ space tends to camouflage the magnitudes of the model errors, arithmetic $|E^*|$ space is preferred here.

The goodness-of-fit statistics in Fig. 2 are calculated with reference to the line of equality. Statistics about the line of equality are defined as follows:

$$S_y = \sqrt{\frac{\sum_{i=1}^n (E_{mi}^* - \bar{E}_m^*)^2}{n}} \quad (1)$$

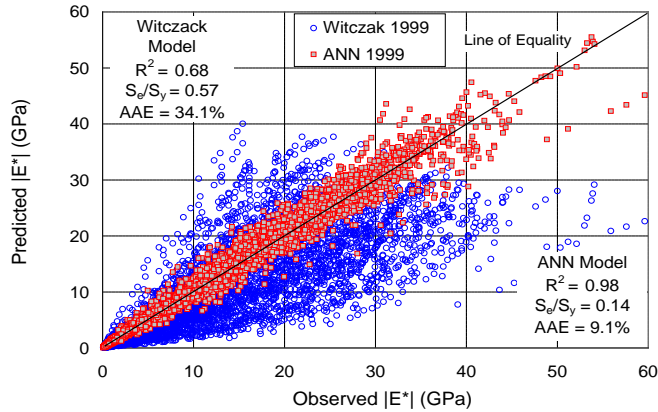
$$e_i = (E_{pi}^* - E_{mi}^*) \quad (2)$$

$$S_e = \sqrt{\frac{\sum_{i=1}^n e_i^2}{(n-p)}} \quad (3)$$

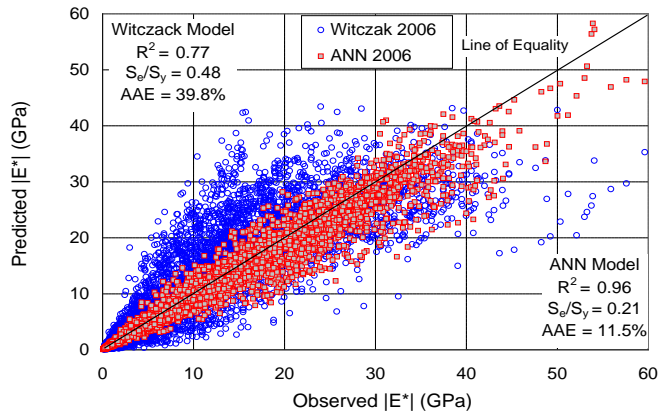
$$R^2 = 1 - \frac{(n-p)}{(n-1)} \left(\frac{S_e}{S_y} \right)^2 \quad (4)$$

where S_y = standard deviation of the measured E^* values about the mean measured E^* ; e = error between the predicted and measured E^* values; S_e = standard error (i.e., standard deviation of errors); R^2 = correlation coefficient; E_m^* = measured dynamic modulus; \bar{E}_m^* = mean value of measured dynamic modulus; E_p^* = predicted dynamic modulus; n = sample size; p = number of model parameters.

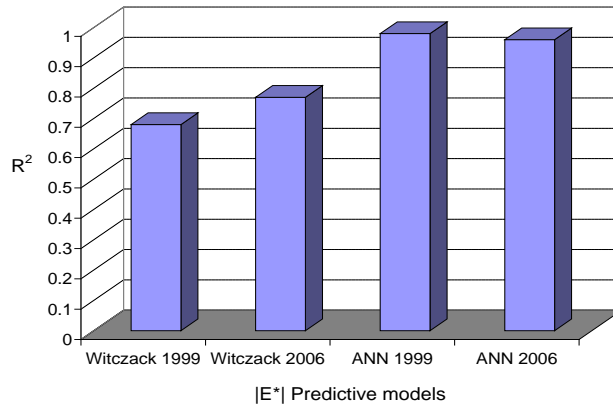
Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.



(a)



(b)



(c)

Fig. 2. Overall prediction accuracy of $|E^*|$ Models for full 7400 record data set: (a) Witczak 1999 and ANN 1999; (b) Witczak 2006 and ANN 2006; (c) R^2 values for all models

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

The goodness-of-fit statistics in Fig. 2 are based on the full 7,400 record data set from Bari (2005). As a consequence, some of the statistics in Fig. 2 are different from those reported by the model developers for calibrations to different subsets of the Bari database. The overall goodness-of-fit statistics in arithmetic space for the full 7,400 record data set for all models are summarized in Fig. 2(d). It is clear that the ANN versions of the Witczak models (ANN 1999 and ANN 2006) have the highest overall accuracy and followed somewhat more distantly by the Witczak 2006 and Witczak 1999 models. However, there is an important caveat that must be kept in mind with regard to these results. The different models have been calibrated to different subsets of the data records in the Bari (2005) data set. The Witczak 2006 model was calibrated using all 7,400 data records (346 mixtures), the ANN models (all versions) were calibrated using a training subset of 6,900 records, and the Witczak 1999 model was calibrated using a subset representing about 30% of the records. The Witczak 1999 model would likely give R^2 values closer to those of the Witczak 2006 model if it were recalibrated using the full 7400 record dataset.

Bias in Model Predictions

The results in Fig. 2 summarize the overall prediction accuracy of the models. However, overall goodness-of-fit statistics like R^2 and S_e/S_y do not necessarily tell the entire story regarding model accuracy. There may be overall and/or local biases in the predictions that can cause significant reductions in accuracy under certain conditions.

Recall that the overall goodness-of-fit statistics in Equations (1) through (4) are defined about the line of equality—i.e., a linear trend line for which the intercept is constrained to pass through the origin and the slope is constrained to unity. One measure of overall bias in the model predictions is how closely the unconstrained linear trend line matches the line of equality—i.e., how close the unconstrained intercept and slope are to 0 and 1, respectively. Another measure of overall model bias is the average error. A nonzero average error indicates a consistent over- or under-prediction by the model.

Fig. 3 summarizes the overall bias statistics for all of the models. The unconstrained trend lines all have a positive intercept ranging between 0.2 (ANN 2006) to 1.5 GPa (Witczak 2006). As partial compensation, the deviations of the slopes of the unconstrained trend lines from unity ($\Delta Slope$) are all negative (i.e., the trend line slopes are less than 1) and range between approximately 0 (ANN 1999) and -0.3 (Witczak 1999). The average error ranges between -1.9 (Witczak 1999) and 0.7 GPa (Witczak 2006). Overall, the ANN 1999 and ANN 2006 models exhibit the smallest prediction bias. This is in part expected since these models were calibrated using a substantial subset (93% or 6,900 records) of the overall data. However, the Witczak 2006 model was calibrated using 100% of the database, yet it exhibits moderate amounts of bias in arithmetic space. This may be a consequence of the calibration of the Witczak models in logarithmic space and the subsequent transformation from logarithmic to arithmetic space.

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

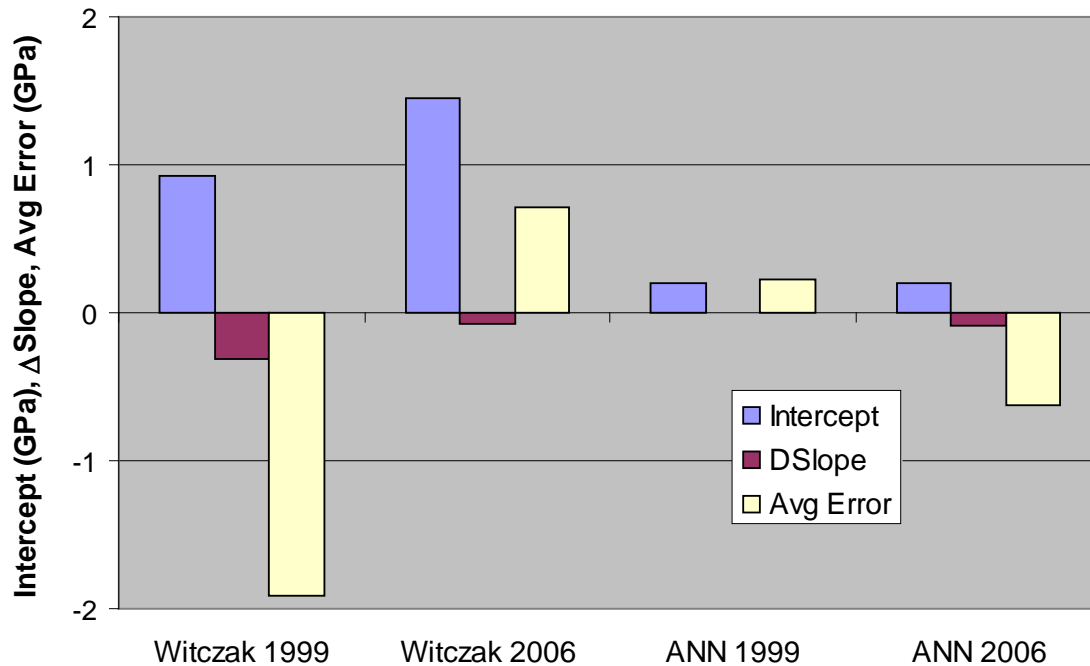


Fig. 3. Overall bias in model predictions

In addition to overall bias, there is also the potential for local bias of the models under subsets of conditions. There are several observations in the literature that the various models lose accuracy at the low and/or high temperature extremes (Pellinen, 2001; Schwartz 2005; Dongre et al. 2005; Bari and Witczak 2006; Al-Khateeb et al. 2006; Azari et al. 2007). Table 1 summarizes the average prediction errors for all of the models stratified by temperature. The Witczak 1999, Witczak 2006, and (to a lesser extent) the ANN 2006 models all exhibit greater error magnitudes at the coldest temperature conditions. The Witczak 2006 also exhibit increased error magnitudes at the highest temperature condition. This latter trend has troubling implications for the prediction of rutting in the MEPDG, where the high temperature dynamic modulus is the controlling HMA material property.

Table 1. Average Errors for Different Temperatures (Note: Highlighted cells have average error magnitude greater than 10%)

Temperature (°C)	Average Errors of Predictive Models (%)			
	Witczak 1999	Witczak 2006	ANN 1999	ANN 2006
-17.8	40.9	26.1	6.6	-13.8
4.4	-15.0	38.7	3.7	-0.8
21.1	-22.2	3.0	4.1	-6.8
37.8	-9.2	-3.1	4.7	-6.0
54.4	-5.4	15.4	5.0	-2.3

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

A more subtle local bias is the deviation of local trend lines from the line of equality. Schwartz (2005) argued that the Witczak 1999 model is dominated by temperature and understates the influence of the other mixture parameters. It is therefore worthwhile to examine the robustness of all $|E^*|$ predictive models at constant temperature in order to remove the dominating influence of temperature and highlight the influence of the non-temperature input parameters such as aggregate gradation and mixture volumetric properties.

The comparisons between predicted and measured $|E^*|$ for data segregated by temperature are illustrated in Fig. 4. Because of the interchangeability between temperature and loading rate, only data at a 1 Hz loading rate are included in the figure in order to highlight further the influence of the mixture variables. Note that data are plotted in terms of $\log |E^*|$ in this figure only for clarity of presentation. The underlying statistical analyses of the local trend lines were performed in arithmetic space, which is the reason that the local linear trend lines appear nonlinear when plotted in $\log |E^*|$ space in the figure.

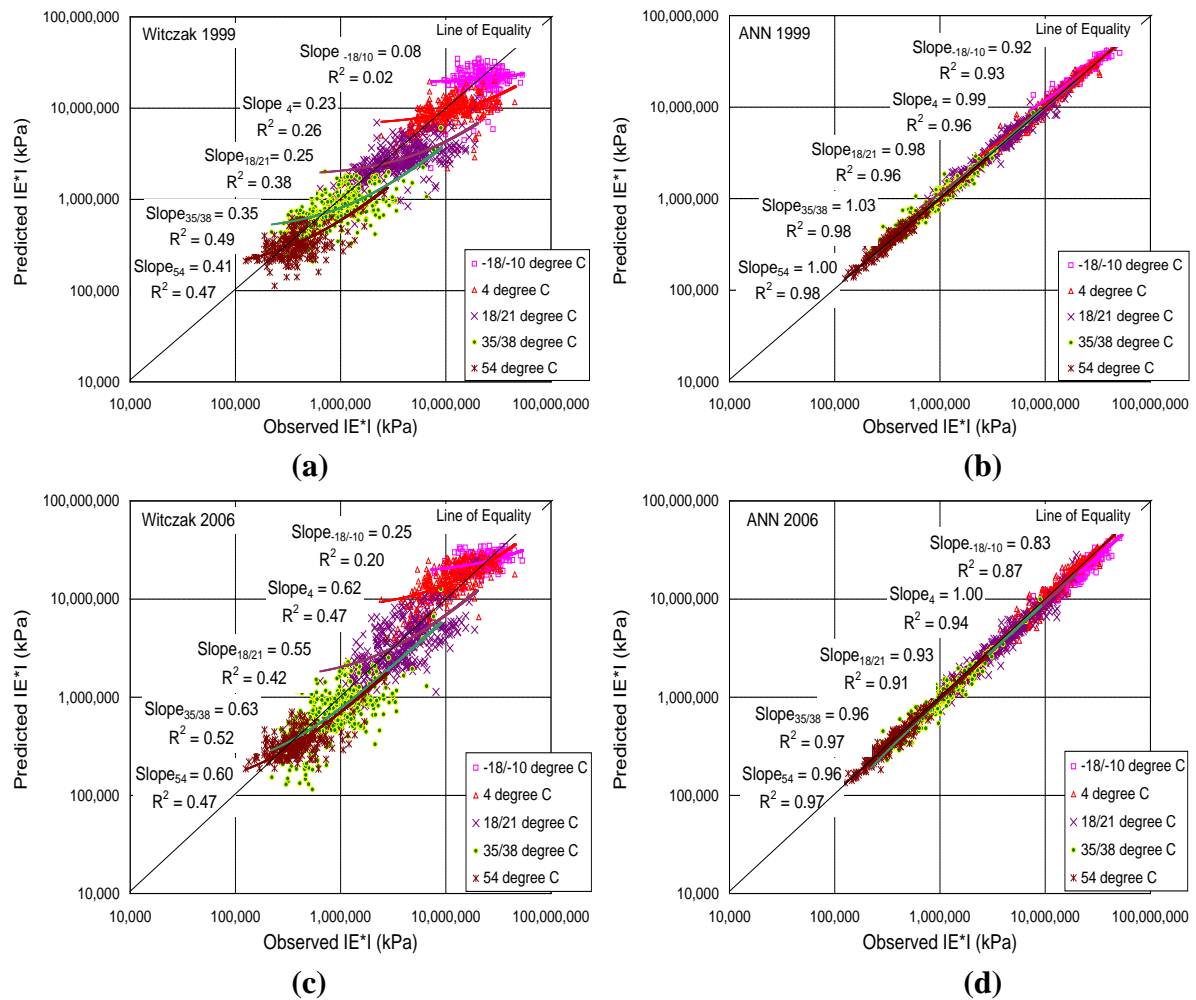


Fig. 4. Predicted vs. observed $\log |E^*|$ at 1 Hz, segregated by temperature: (a) Witczak 1999; (b) ANN 1999; (c) Witczak 2006; (d) ANN 2006

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

As suggested by the best-fit slopes through each constant-temperature subset in Fig. 4(a), only about 8 to 41% of the influences of the non-temperature inputs are captured in the Witczak 1999. In addition, the statistical strength of the local trend line at each temperature level is quite low, with R^2 values ranging only between 0.02 and 0.49. The Witczak 2006 model in Fig. 4(c) shows improvement, capturing between 25 to 63% of the influences of the non-temperature inputs. The strength of the local trend at each temperature level is still quite low, though, with R^2 values ranging only between 0.2 and 0.52.

Similar temperature-stratified analyses for the ANN 1999 and ANN 2006 models in Fig. 4(b) and 4(d) show much better model performance. The local trend lines now align much more closely with the line of equality, implying that 92 to 100% of the influences of the non-temperature inputs are captured by the ANN 1999 model and 83 to 100% by the ANN 2006 model. The statistical strength of the local trend at each temperature level is also much greater, with R^2 values ranging between 0.93 and 0.98 for the ANN 1999 model and between 0.87 and 0.97 for the ANN 2006.

The clear conclusion from these results is that the ANN predictive models do an excellent job of capturing both temperature and mixture variable influences $|E^*|$ while the Witczak models are dominated by temperature effects and only capture a portion of the influences of the non-temperature inputs. The practical implication is that the regression-based models may therefore be unable to make fine distinctions between different mixtures—e.g., an over-asphalted 12.5 mm fine mixture vs. a well-compacted 19 mm coarse mixture—under a given set of environmental and other conditions. The excellent ability of the ANN models to capture the influences of both temperature and mixture variables should provide good estimates of varying performance associated with different mixtures under a given set of environmental and other conditions.

Practical Implications

Ranking of HMA Mixtures

Pavement and bituminous materials engineers often evaluate alternative HMA mix designs in an effort to maximize pavement performance. In the MEPDG, this evaluation is done largely in terms of $|E^*|$. Predictive models for $|E^*|$ should therefore rank mixes in the same order as they would be ranked if $|E^*|$ were actually measured in the laboratory. In other words, if the measured stiffness of mix A is higher than the measured stiffness of mix B at some given temperature and loading rate, then ideally the predicted stiffness of mix A would also be greater than the predicted stiffness of mix B at the same given temperature and loading rate. If this is not the case, the predictive models will give incorrect indications of the relative performance of the alternative mixtures.

Kendall's τ rank correlation coefficient is a standard statistic for quantifying the degree of correspondence between two rankings. Consider paired lists of measured and predicted $|E^*|$ values for n HMA mixtures, with the paired items ranked in order of the first list (i.e., measured $|E^*|$) in order of decreasing magnitude. The Kendall τ coefficient for the second list (predicted $|E^*|$) is defined as (Kendall 1948):

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

$$\tau = \frac{4P}{n(n-1)} - 1 \quad (5)$$

in which P is the number of items in the second list that are also ranked correctly. A value for τ equal to 1 means that the ranking of the two lists is in perfect agreement, a value of -1 means that the rankings are in perfect disagreement (i.e., the second list is ranked in the reverse order of the first), and a value of 0 means that the rankings are completely independent. In other words, increasing positive values of τ correspond to increasing agreement between the two rankings.

The Kendall τ coefficient can be computed for the measured vs. predicted $|E^*|$ mixtures for the asphalt mixtures included in Bari's (2005) database. The data are sorted in terms of decreasing measured $|E^*|$ and the agreement of the corresponding ranking in terms of predicted $|E^*|$ is determined. Fig. 5(a) summarizes the Kendall τ values for the predicted $|E^*|$ values as determined using the various models for all of the data records in the Bari (2005) data set. As would be hoped, the τ values are all quite high, ranging between about 0.78 and 0.9 for all of the models. This indicates that the rankings by predicted $|E^*|$ values are approximately the same as the rankings by measured $|E^*|$ values for all of the models. The ANN models collectively perform only slightly better than the regression-based models (Witczak 1999 and 2006).

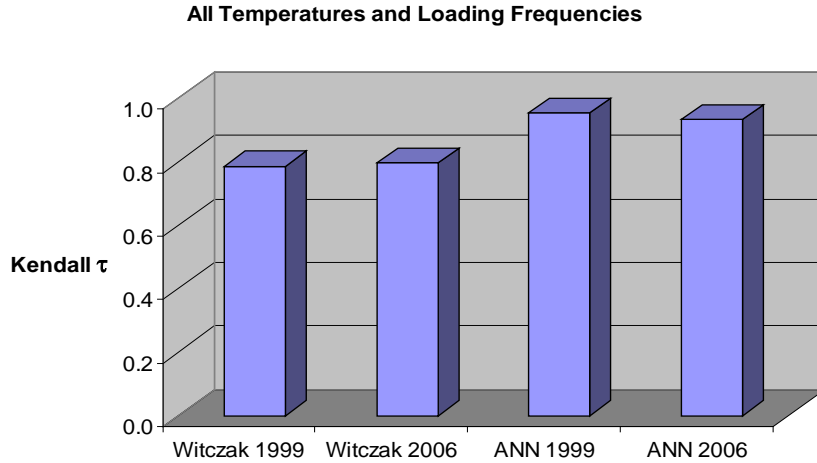
Note that the statistics in Fig. 5(a) are for the rankings of all of the test records in the database and not for rankings of individual mixtures. Each mixture has multiple test records, one for each temperature and loading frequency in the testing protocol. Rankings of individual mixtures is better examined by considering $|E^*|$ values at a single temperature and loading frequency. This corresponds more closely to real-world project level mix design and evaluation where the site environment (i.e., effective temperature) and design traffic speed (i.e., loading frequency) will be the same for all mixtures being ranked.

Fig. 5(b) summarizes the Kendall τ ranking statistics for the mixtures at constant temperatures and a fixed 1 Hz loading frequency for all of the models. Now each data set contains only one record per mixture, and the comparisons more closely resemble the types of evaluations made in project level mixture design and selection. Several observations regarding the agreement between measured and predicted $|E^*|$ rankings can be drawn from Fig. 5(b):

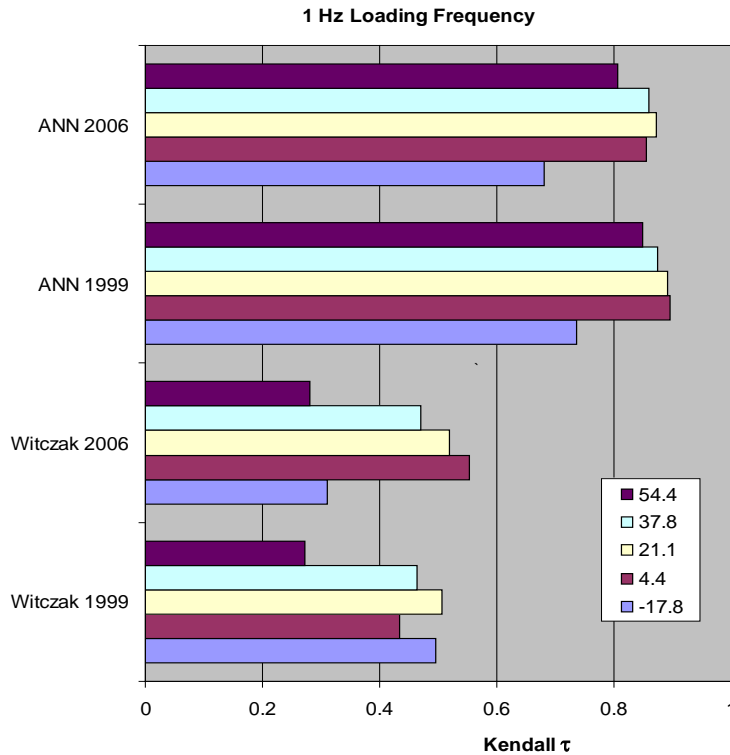
- The mixture rankings based on $|E^*|$ predictions from the ANN 1999 model are in best agreement with the rankings based on measured $|E^*|$ values (highest Kendall τ values). The ANN 2006 mixture rankings are a close second.
- The ANN models as a group display better agreement in mixture rankings (i.e., higher Kendall τ values) than do their corresponding regression-based equivalents.
- The agreement between predicted vs. measured rankings tends to be better at moderate temperatures (4.4, 21.1, and 37.8°C) than at the low (-17.8°C) and high (54.4°C) temperature extremes for all models. The decrease in ranking ability at the low and high temperature extremes is more pronounced for the regression models (with the exception of the Witczak 1999 model at low temperatures) than for the ANN models.

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

- The Kendall τ values for the ANN 1999 model exceeded 0.75 for all temperatures (0.85 for all but -17.8°C). This indicates that the $|E^*|$ predictions from the ANN 1999 model would rank all mixtures in nearly the same order as the measured $|E^*|$ values.



(a)



(b)

Fig. 5. Kendall τ estimate of correct ranking for predicted $|E^*|$ values: (a) Mixes at all loading frequency and temperatures; (b) Mixes at 1 Hz loading frequency and fixed temperature (Note: Legend indicates temperature in °C)

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

Pavement Performance

Within the context of mechanistic-empirical design, the significance of any errors in predicted $|E^*|$ are best evaluated in terms of their impact on predicted pavement performance. Version 1.000 of the MEPDG is employed here to predict pavement performance. Dynamic modulus is the principal material input for hot mix asphalt in the MEDPG. In order to minimize confounding influences, only a single pavement structure, traffic loading condition, and project location are considered. Each predictive model was used to generate a set of $|E^*|$ values over a range of temperatures and loading frequencies, which were then entered in the MEPDG software as "pseudo" Level 1 inputs. The reference mixture and binder properties and measured $|E^*|$ values are those from WesTrack section R24 in the Bari (2005) database.

The reference pavement design for this study is a three layer flexible pavement (HMA/base/subgrade) section for a hypothetical roadway in the Mid-Atlantic coastal plain region. The roadway is assumed to have two lanes in each direction and significant heavy truck traffic equivalent to 10 million ESALs. A silty sand subgrade (AASHTO A-2-5/USCS SM) with deep groundwater table and no shallow bedrock is assumed. Using typical design input parameters, 8 inches of asphalt concrete over 24 inches of crushed stone base provide a satisfactory pavement section according to the 1993 AASHTO procedure.

Fig. 6 summarizes normalized predicted AC rutting and alligator cracking performance based on predicted $|E^*|$ values from each of the models. The horizontal indicator represents the predicted performance based on the mean predicted $|E^*|$ computed as the measured $|E^*|$ adjusted by the overall prediction error from each of the models (Fig. 3). The extent of the vertical lines represents the predicted performance using the mean predicted $|E^*|$ values increased or decreased by one standard error for each model. The overall prediction error and standard error values were assumed to be constant (in percentage terms) across all temperatures and loading rates but different from model to model—i.e., no decrease in accuracy and increase in bias in predicted $|E^*|$ at the temperature extremes is considered. The trends for normalized AC rutting and alligator cracking performance are quite similar. As before, the ANN models as a group outperform their regression-based counterparts with respect to both mean prediction accuracy and lower variability of prediction. The ANN 1999 and ANN 2006 provided the best performance prediction, followed closely by the Witczak 2006 and more distantly by the Witczak 1999 models.

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

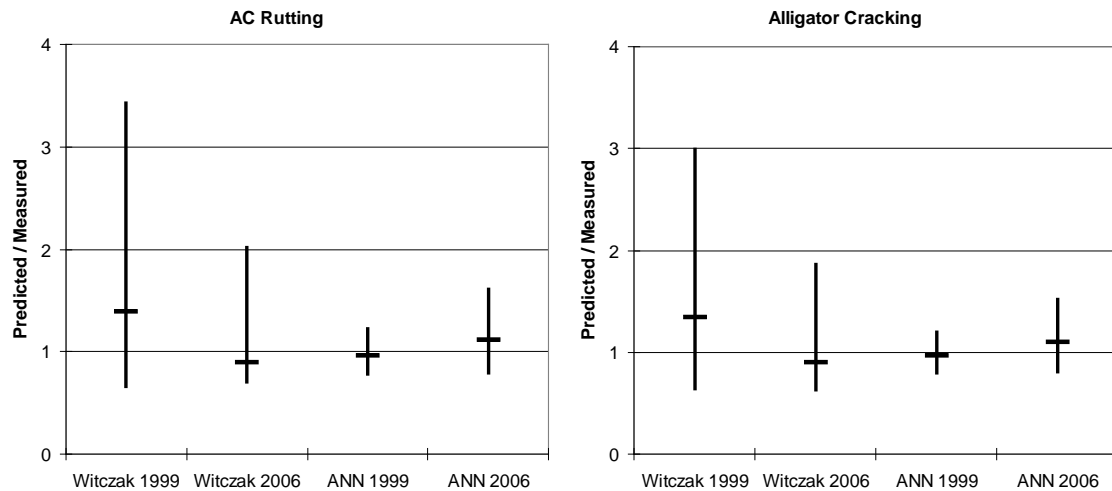


Fig. 6. Effect of predicted dynamic modulus errors on predicted pavement performance

Conclusions

The accuracy and robustness of the predictive models for estimating the HMA dynamic modulus ($|E^*|$) inputs in the new Mechanistic-Empirical Pavement Design Guide (MEPDG) have been evaluated through statistical analyses and a set of MEPDG (version 1.000) runs. The principal conclusions can be succinctly summarized in the context of the two fundamental questions raised in the introduction:

“How accurate and robust are the $|E^|$ predictions?”*

- The ANN-based $|E^*|$ models use the same input variables as the Witczak $|E^*|$ predictive models but produce $|E^*|$ predictions with significantly higher accuracy.
- Most of the regression-based $|E^*|$ predictive models exhibit bias at the lower and/or higher $|E^*|$ spectrum. The ANN-based $|E^*|$ predictive models have a much lower tendency toward this bias. This corresponds to more accurate characterization of HMA dynamic modulus at the temperature extremes and to better predictions of distresses that occur at these temperature extremes, e.g., rutting.
- The regression-based $|E^*|$ predictive models overemphasize the influence of temperature and understate the influence of mixture variables like volumetric and aggregate gradation properties. The ANN-based $|E^*|$ models do a more balanced job of capturing both temperature and other mixture influences and thus can provide better evaluation of the relative performance of different mixtures under a given set of project-specific environmental and design traffic conditions.

“How accurate do the $|E^|$ predictions need to be?”*

- The ANN models as a group are better able to rank mixtures in the same order as measured dynamic modulus than are their corresponding regression-based equivalents.

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

- Based on pavement performance predictions (AC rutting and alligator cracking) using the MEPDG (version 1.000) for a single hypothetical pavement design and an idealized HMA mixture, the ANN 1999 and ANN 2006 models provided the best performance predictions in terms of accuracy and variability, followed closely by the Witczak 2006 and more distantly by the Witczak 1999 models.

The results of this study have significant implications for advancing the state of the art in mechanistic-empirical pavement analysis and design. ANN-based models developed using comprehensive datasets could be easily and successfully incorporated into the MEPDG as alternatives to the current pavement materials characterization models. Because ANNs excel at mapping in higher-order spaces, such models could also be applied to extend the empirical distress models beyond the current univariate relationships between pavement structural response and pavement performance.

Acknowledgments

The HMA $|E^*|$ data referenced in this paper was obtained from the NCHRP 9-19 project final report DVD. The contents of this paper reflect the views of the authors who are responsible for the facts and accuracy of the data presented within. This paper does not constitute a standard, specification, or regulation.

References

- Al-Khateeb, G., Shenoy, A., Gibson, N., and Harman, T. (2006). "A new simplistic model for dynamic modulus predictions of asphalt paving mixtures." *Journal of the Association of Asphalt Paving Technologists*, 75 (CD-ROM).
- Andrei, D., Witczak, M.W., and Mirza, M.W. (1999). *Development of a revised predictive model for the dynamic (complex) modulus of asphalt mixtures*, NCHRP 1-37 A Interim Report, University of Maryland, College Park, MD.
- Azari, H., Al-Khateeb, G., Shenoy, A., and Gibson, N. (2007). "Comparison of simple performance test $|E^*|$ of accelerated loading facility mixtures and prediction $|E^*|$: use of NCHRP 1-37A and Witczak's new equations." *Transportation Research Record*, No.1998, pp. 1-9.
- Bari, J. (2005). "Development of a new revised version of the Witczak E^* predictive models for hot mix asphalt mixtures." Ph.D.Dissertation, Arizona State University, Tempe, AZ.
- 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*, 75, 381- 423.
- Bari, J. and Witczak, M.W. (2007). "New predictive models for viscosity and complex shear modulus of asphalt binders for use with mechanist-empirical pavement design guide." *Transportation Research Record*, No. 2001, pp. 9-19.

Reference to this paper should be made as follows: Ceylan, H., Schwartz, C.W., Kim, S., and Gopalakrishnan, K. (2009). "Accuracy of Predictive Models for Dynamic Modulus of Hot Mix Asphalt," ASCE Journal of Materials in Civil Engineering, Vol. 21, No. 6, pp. 286-293.

- Bonaquist, R. F., Christensen, D. W., and Stump, W. (2003). *Simple performance tester for superpave mix design development and evaluations*, NCHRP Report 513, Transportation Research Board, Washington, D.C.
- Ceylan, H., Kim, S and Gopalakrishnan, K. (2007). "Hot mix asphalt dynamic modulus prediction models using neural network approach." *Proc., the ANNIE 2007*, ASME, St. Louis, MO.
- Ceylan, H., Gopalakrishnan, K. and Kim, S. (2008) "Advanced approaches to hot-mix asphalt dynamic modulus prediction." *Canadian Journal of Civil Engineering*, 35 (7), 699-707, NRC Canada.
- Christiansen, D. (2006). Published discussion to Bari and Witczak (2006), *Journal of the Association of Asphalt Paving Technologists*, 75, 422-423.
- Dongre, R., Myers, L., D'Angelo, J., Paugh, C., and Gudimettla, J. (2005). "Field evaluation of Witczak and Hirsch models for predicting dynamic modulus of hot-mix asphalt." *Journal of the Association of Asphalt Paving Technologists*, 74 (CD-ROM).
- Kendall, M. (1948). *Rank correlation methods*. Charles Griffin and Co, London.
- NCHRP. (2004). *Guide for mechanistic-empirical design of new and rehabilitated pavement structures*, <<http://trb.org/mepdg>> (May, 2006), NCHRP 1-37A Final Report, Transportation Research Board, Washington, D.C.
- NCHRP. (2006a). *Changes to the mechanistic-empirical pavement version 0.900 - July 2006*, NCHRP 1-40D Research Results Digest 308, Transportation Research Board, Washington, D.C.
- NCHRP. (2006b). *Independent review of the mechanistic-empirical pavement design guide and software*, NCHRP 1-40 A Research Results Digest 307, Transportation Research Board, Washington, D.C.
- Pellinen, T.K. (2001). *Investigation of the Use of Dynamic Modulus as an Indicator of Hot-Mix Asphalt Performance*, Ph.D. dissertation, Arizona State University, Tempe, AZ.
- Schwartz, C. W. (2005). "Evaluation of the Witczak dynamic modulus prediction model." *Proc., the 84th Annual Meeting of the Transportation Research* (CD-ROM), Transportation Research Board, Washington, D.C
- TRB Circular. (1999). *Use of artificial neural networks in geomechanical and pavement systems*, Number E-C012, Transportation Research Board, National Research Council, Washington D.C.
- Tsoukalas, L. H. and Uhrig, R. E. (1997). *Fuzzy and neural approaches in engineering*, Wiley, NY.
- Witczak, M. W., Kaloush, K., Pellinen, T., El-Basyouny, M., and Quintus, H. V. (2002). *Simple performance test for superpave mix design*, NCHRP Report 465, Transportation Research Board, Washington, D.C.