Prediction of Dynamic Modulus of Southeastern Asphalt Concrete Mixtures using Artificial Neural Networks

M. Robbins and D. Timm
Department of Civil Engineering, Auburn University, Auburn, Alabama, USA

ABSTRACT: As mechanistic-empirical pavement design comes to the forefront of design, the accurate characterization of asphalt concrete (AC) through dynamic modulus ($|E^*|$) becomes increasingly more important. $|E^*|$ captures the viscoelastic nature of AC and is essential to the accurate prediction of pavement responses under varying speed and temperature conditions. Due to the expensive and specialized equipment needed to measure $|E^*|$, predictive models have gained popularity. However, variability in the predictive capabilities of these models from study to study indicates that they may not be applicable on a global level. Thus, the use of artificial neural networks (ANN) to predict $|E^*|$ was investigated for mixtures placed at the 2006 National Center for Asphalt (NCAT) Test Track. Comparisons were drawn with the most commonly used predictive models such that an ANN was created with inputs identical to each predictive model: Witczak 1-37A, Witczak 1-40D and Hirsch. The ANNs created for comparison with the predictive models predicted measured $|E^*|$ with great success; coefficients of determination ($R^2$) ranged between 0.96 and 0.99, a notable improvement over the prediction capabilities of the predictive models. A correlation analysis was completed on a variety of input parameters to create an optimal ANN for the 2006 Test Track mixtures. An optimal ANN was created for the 2006 Test Track mixtures using only two inputs, effective binder content ($V_{be}$) and the product of dynamic shear modulus of the binder ($|G^*|$) and the sine of the associated phase angle ($\delta$). This investigation also tested the prediction capability of the newly developed ANN on an independent dataset by applying the ANN to the mixtures used in the 2009 Test Track cycle.

KEY WORDS: Asphalt concrete, dynamic modulus, artificial neural network.

1 INTRODUCTION

The advent of mechanistic-empirical (M-E) pavement design facilitates the need to accurately characterize the modulus of asphalt concrete (AC). Typically M-E design frameworks characterize AC modulus through the use of dynamic modulus ($|E^*|$). $|E^*|$ is measured in the laboratory at varying frequencies and temperatures which captures the viscoelastic nature of AC. This is essential in accurately predicting pavement responses under varying speed and temperature conditions.

$|E^*|$ can be measured in the laboratory through AASHTO TP 79-09, the update to the AASHTO TP 62-07 method. However, to complete these testing protocols requires either access to or the purchase of expensive equipment and specially-trained technicians. The alternative to measuring $|E^*|$ in the laboratory is to estimate it through the use of predictive models. The onset of M-E design frameworks which necessitate the determination of dynamic
modulus coupled with expensive laboratory testing, have helped predictive models gain popularity. The soon to be widely-adopted M-E design framework, the mechanistic-empirical design guide (MEPDG) version 1.1 (now called Pavement ME Design), requires that the user either input laboratory measured $|E^*|$ results at a broad range of frequencies and temperatures or select one of two embedded $|E^*|$ predictive models. These models are the Witczak 1-37A and Witczak 1-40D, developed in 1999 and 2006 respectively (Andrei et al. 1999, Bari and Witczak 2006). The former is a viscosity-based model, while the latter updates the Witczak 1-37A by characterizing the binder by dynamic shear modulus ($|G^*|$) rather than viscosity. Aside from the Witczak models, another commonly used model is the Hirsch model, developed by Christensen, Pellinen and Bonaquist (2003).

Numerous studies have evaluated the accuracy of these predictive models. Results have varied from study to study (Christensen et al. 2003, Dongre et al. 2005, Birgisson et al. 2005, Mohammad et al. 2007, Ping and Xiao 2007, Robbins and Timm 2011), indicating that these models are inconsistent in predicting $|E^*|$ on a global level. For instance, one study (Dongre et al. 2005) reports results for the Witczak 1-37A ranging from a coefficient of determination ($R^2$) of 0.52 to 0.98, depending on the site and aging condition of the binder. Other studies show a smaller range for the performance of the Witczak 1-37A, 0.73-0.84, also dependent on binder aging conditions (Birgisson et al. 2005). As a result, researchers have turned to artificial neural networks (ANNs), to predict $|E^*|$. ANNs do not require knowledge of the form of the relationships between input parameters and the target ($|E^*|$). This is advantageous when applied to AC, a viscoelastic material that can be difficult to model due to the numerous influencing and interacting parameters (Sakhaei Far et al., 2009). A commonly noted pitfall of an ANN, however, is that the capability of the network is limited by the values used to train the network. This is also the case with models based in statistical regression. One way to prevent this is to train the network with a large and robust dataset, although building such a database can be difficult, the training of the network is a relatively simple task.

A previous study conducted at the National Center for Asphalt Technology (NCAT) Test Track evaluated the accuracy of these three predictive models in predicting laboratory $|E^*|$ (Robbins and Timm 2011) for mixtures placed during the 2006 research cycle. Of the three models evaluated, the Hirsch model performed the best, with the model accounting for 88% of the variability in the measured data. A local calibration was performed on the Hirsch model with negligible improvement noted. Given that M-E design hinges on accurate characterization of AC, it is necessary to evaluate the use of ANNs to predict $|E^*|$ relative to these current predictive models for the same NCAT Test Track mixtures.

1.1 Objectives

The primary objective of this study was to assess the use of artificial neural networks to predict dynamic modulus for the mixtures of the NCAT Test Track 2006 research cycle. To do so, the following secondary objectives were established:

1. Compare predictions from $|E^*|$ predictive models with an ANN corresponding to the same inputs required for each predictive model.
2. Assess the possible input parameters and develop an ANN to optimize $|E^*|$ predictions for the 2006 Test Track mixtures.
3. Evaluate the developed ANN on $|E^*|$ results for mixtures in the 2009 Test Track.

2 SCOPE

Results from a previous study (Robbins and Timm 2011) on the adequacy of these three predictive models for the 2006 Test Track mixtures were utilized for comparison with ANNs.
A correlation analysis was completed in Microsoft Excel to assist in the selection of input parameters to create an ANN for optimal prediction of $|E^*|$ for Test Track mixtures. An ANN was created for comparison with each predictive model using Matlab, version 7.10.0.499.

3 METHODOLOGY

3.1 Laboratory Testing

Dynamic modulus testing was completed for the 2006 Test Track mixtures following the AASHTO TP 62-07 using an Asphalt Mixture Performance Tester (AMPT). Testing was conducted at only three temperatures, 4, 21, 37.8°C rather than the recommended 5 outlined in AASHTO TP 62-07 due to difficulties in obtaining reliable and reasonable results at the extreme temperatures. For each temperature, seven frequencies, 0.5, 1, 2, 5, 10, 20, and 25 Hz were applied. Specimens consisted of plant-produced mixtures compacted in the laboratory using a Superpave gyratory compactor. Dynamic modulus was determined for the mixtures placed during the 2009 Test Track cycle in accordance with the AASHTO TP 79-09 procedure on plant produced, lab compacted specimens.

Viscosity testing was conducted on rolling thin film oven test (RTFOT)-aged binders using the Brookfield rotational viscometer following AASHTO D4402-06 at 135 and 165°C. Dynamic shear modulus, $|G^*|$, testing, in accordance with AASHTO T 315-06, was also conducted on RTFOT-aged binders using a Dynamic Shear Rheometer (DSR). For the 2006 Test Track mixtures a frequency sweep ranging from 0.1 Hz to 25 Hz was applied at each of the four test temperatures: 4, 21, 37.8, and 54.4°C. Following the same procedure, $|G^*|$ was measured for the binders of the 2009 Test Track at a range of temperatures, 21-54.4°C, and a range of frequencies, 0.01-25 Hz.

3.2 Predictive Models

As mentioned earlier, a previous study evaluated the use of three predictive models for estimating the $|E^*|$ for the 2006 Test Track mixtures. A more complete discussion on the use of these models to predict $|E^*|$ for the 2006 mixture is documented elsewhere (Robbins and Timm 2011). The Witczak 1-37A model is a sigmoidal fit function that combines rudimentary volumetric and gradation properties with binder viscosity and frequency to estimate $|E^*|$ for a given frequency and temperature. The model, as presented in the MEPDG Part 2, Chapter 2 Appendix (ARA 2004), uses four rudimentary gradation parameters, two volumetric parameters and viscosity and frequency to predict $|E^*|$.

The Witczak 1-40D model is a modified version of its predecessor, the 1-37A model. Additional mixtures expanded the dataset and viscosity and frequency were replaced with dynamic shear modulus and its associated phase angle as displayed in (Bari and Witczak 2006). It should be noted that the frequency at which dynamic shear modulus is tested, is not equivalent to frequency in the compression mode as used for $|E^*|$ testing, rather, the frequency in shear mode is equivalent to the frequency in compression mode divided by the product of two and pi.

The Hirsch model applies the law of mixtures for composite material to AC by considering it as a three-phase system of aggregate, asphalt binder and air voids as shown in (Christensen et al. 2003). Only three inputs are required: $|G^*|$ of the binder, voids in mineral aggregate (VMA) and voids filled with asphalt (VFA). It should be noted that equivalent definitions of frequency between compression (dynamic modulus of the mix) and shear mode (dynamic shear modulus of the binder) are used.
3.3 2006 Test Track Mixtures

Mixtures placed at the NCAT Test Track during the 2006 test cycle were utilized for this investigation. These mixtures are representative of general use mixtures typically placed in the southeastern U.S. Mixtures chosen for this study had $|E^*|$ testing completed at 1, 10, 25 Hz and 4, 21, 37.8°C. To meet the requirements of the Witczak 1-40D model, these mixtures also had $|G^*|$ testing completed at 0.159, 1.59 and 3.979 Hz. Likewise, for the Hirsch model, the mixtures selected also had to have $|G^*|$ completed at 1, 10, and 25 Hz. The resulting dataset included 18 different mixtures representing 6 mix types. A total of 12 binders were included, representing five different types: PG 64-22, PG 67-22, PG 70-22 modified with Styrene Butadiene Styrene (SBS), PG 76-22 with SBS, and PG 76-28 with SBS. The resulting dataset remained robust, including unmodified and SBS modified mixtures, as well as two different Superpave mixtures with reclaimed asphalt pavement (RAP) and one stone matrix asphalt mixture (SMA). Two of the mixtures, included a 75 blow Marshall Mix designed with RAP and a dense graded granite mix designed at 50 gyrations. The ranges for nominal maximum aggregate size (NMAS), air voids, VFA and test temperatures fell within the ranges used for the Hirsch model development although the minimum VMA was slightly lower than the minimum reported in the Hirsch model development (Christensen et al. 2003).

3.4 Artificial Neural Network

An ANN was created for comparison with each predictive model, such that “ANN_37A” utilized the same inputs as the Witczak 1-37A predictive model, and likewise the complement to the Witczak 1-40D model was “ANN_1-40D” and the Hirsch was complemented by “ANN_Hirsch.” Matlab was employed to create feed-forward back propagation networks based on supervised learning. For each ANN, two hidden layers were used. Although the number of neurons was consistent between hidden layers, the number of neurons used for each ANN varied, depending on the number of inputs and the performance of each network. A common transfer function, tan-sigmoidal, was used as well as a linear output transfer function. Given the size of the dataset (537 data points), 90% of the data was used for training and 5% each was used for validation and testing, selected at random.

Due to the wide range of target values (measured $|E^*|$), and the variable ranges of the input values, it was necessary to normalize both the input and output data following Equation 1. By doing so, all input parameters and output values ranged between -1 and +1. It should be noted that the target value was normalized measured $|E^*|$, rather than log $|E^*|$ as predicted by the Witczak 1-37A and 1-40D models.

$$N_i = \frac{2(x_i - \text{min})}{\text{max} - \text{min}} - 1$$

where:

$N_i$ = $i^{th}$ Normalized value

$x_i$ = $i^{th}$ original value

The number of neurons in the hidden layers was varied until the mean squared error of the network was minimized and a high coefficient of determination, $R^2$, was reached for all levels: training, validation, and testing. Additionally, the ratio of standard error of the predictions, $S_e$, to standard deviation of the measured values, $S_y$, was used to assess the quality of the ANN. Typically, low values, less than 0.35, are considered excellent. Each network was trained 3-5 times before altering the number of neurons. In optimizing an ANN for the 2006 Test Track mixes, a correlation analysis was completed using Pearson’s product
moment correlation coefficient, r, on the available parameters to help select a minimum number of critical inputs to predict $|E^*|$.  

3.5 2009 Test Track Mixtures

The $|E^*|$ measurements determined in the laboratory for mixtures placed at the Test Track during the 2009 test cycle were used to evaluate the ANN created from $|E^*|$ measurements for the 2006 mixtures. A total of 24 different mixtures were included in the evaluation of the developed ANN. These mixtures utilized 23 different binders, including unmodified Performance Grade (PG) 67-22, PG 76-22 modified with SBS, PG 76-22 modified with ground tire rubber (GTR), PG 67-28 modified with pellets from Trinidad Lake Asphalt (TLA) and binders modified with 7.5% SBS. The mixtures included 50% RAP, 50% RAP produced with foaming technology, warm mix asphalt (WMA) produced with foaming technology, WMA produced with additives, 45% RAP, a variety of Superpave mixes and fine and coarse graded mixtures. Due to the advancement in the AASHTO testing protocol for $|E^*|$, these mixtures were tested at three temperatures with the high temperature determined by the PG of the binder: 4, 20 and 40 or 45°C. At each temperature, testing was conducted at 0.1, 1, and 10 Hz with an additional frequency of 0.01 Hz at the high temperature. The CAM model (Marasteanu and Anderson 1999) using the WLF shift factor (after Ferry 1980) were applied to the $|G^*|$ data to create a master curve for each binder. From that master curve, the $|G^*|$ at the same frequency and temperature as $|E^*|$ were selected for the evaluation.

4 RESULTS AND DISCUSSION

The results of a previous study applying these models to the same 2006 Test Track mixtures found large scatter in the Witczak 1-37A model, over-prediction by the Witczak 1-40D model and reasonable fit with the Hirsch model (Robbins and Timm 2011). Given these results and the success found by other researchers with artificial neural networks (Lacroix et al. 2008, Sakhaei Far et al. 2009, Ceylan et al. 2010), artificial neural networks were created for the 2006 Test Track mixtures. An ANN was created using the same inputs as required by each of the three predictive models. Each ANN was then compared directly with results from each predictive model, at three temperatures, 4, 21, 37.8°C and three frequencies (1, 10 and 25 Hz).

In comparing the Witczak 1-37A model, an ANN was created with 30 neurons and two hidden layers. It is speculated that a large number of neurons was required due to the large number of inputs (8) relative to the number of datapoints (537). Figure 1 illustrates that the 30-neuron ANN significantly reduced the scatter previously reported for the Witczak 1-37A model (Robbins and Timm 2011) and generally follows the line of equality. Although there are some apparent discrepancies in the ANN, it generally performed much better than the Witczak 1-37A model, with an $R^2$ of 0.96, a marked improvement over 0.60, as reported in Table 1. Likewise, improvements were seen in the slope of the overall trend line, bringing it closer to one and the $S_0/S_y$ ratio went from fair to excellent with the application of the ANN.
Figure 1: Comparison of Witczak 1-37A model and corresponding ANN.

Shown in Figure 2, the Witczak 1-40D did not perform as well as the Witczak 1-37A model, confirmed by the negative $R^2$ presented in Table 1. The negative $R^2$ value indicates that the errors in the prediction are significantly higher than the errors about the average measured $|E^*|$. This is evident by a sum of squared error (SSE) that is greater than the total sum of squares (SST), shown in Table 1. The 10 neuron-ANN created for comparison has much less scatter than the Witczak 1-40D model and generally follows the line of equality, illustrated by a slope of nearly one when a linear trend line is applied and forced through the origin. Shown in Table 1, the ANN developed for the 1-40D inputs performed the best of all three created, resulting in the highest $R^2$, lowest $S_e/S_y$ and slope closest to one; it also showed the largest improvement over its corresponding predictive model.

Figure 2: Comparison of Witczak 1-40D model and corresponding ANN.

Figure 3 shows the results of the Hirsch model and associated ANN. It is interesting to note from Table 1, that among the three predictive models the Hirsch model performed the best, with an $R^2$ of 0.88 and $S_e/S_y$ of 0.35. In Figure 3 there are obvious horizontal trends in the Hirsch model that are distinguishable by temperature. When an 8-neuron ANN was created using the same three inputs (VMA, VFA and $|G^*|$), the horizontal trends were resolved. Similar to the Witczak models, the scatter associated with this ANN is also greatest at low moduli values. Testing variability may account for some of the scatter.
The ANN that performed the best, “ANN_1-40D”, utilized eight different parameters, half of which were gradation parameters. Although this ANN returned a very high $R^2$ and low $S_y/S_y$, it is warranted to evaluate if all of these eight parameters are necessary to achieve a high performing ANN using the same data. From the perspective of an agency, a means of accurately predicting $|E^*|$ that uses the fewest, relatively easily obtained, input parameters could be of great benefit. Therefore a correlation analysis was completed on possible input parameters. These input parameters included binder properties: $|G^*|$, binder phase angle ($\delta$), $|G^*|\sin(\delta)$, $|G^*|\cos(\delta)$, $|G^*/\sin(\delta)|$, $G^*/\cos(\delta)$, viscosity, and the high and low temperatures of the performance grade (PG). Volumetric properties that were included were: VMA, VFA, air voids (Va), and effective binder content by volume ($V_{be}$). Gradation properties required for the Witczak models (percent passing the #200 sieve ($\rho_{200}$), cumulative percent retained on the ¾” sieve ($\rho_{34}$), the 3/8” sieve ($\rho_{38}$), and on the #4 sieve ($\rho_{4}$)) were also considered. Lastly, bulk specific gravity of the aggregate ($G_{sb}$), maximum specific gravity of the mix ($G_{mm}$), bulk specific gravity of the mix ($G_{mb}$), percent of binder by weight of mix ($P_b$), and $|E^*|$ testing temperature and frequency were also considered.

The parameters with the strongest correlations to measured $|E^*|$ were testing temperature and binder properties, specifically $|G^*|$, $\delta$, $|G^*|\sin(\delta)$, $|G^*|\cos(\delta)$, $|G^*/\sin(\delta)|$ and $|G^*/\cos(\delta)|$. All had correlation coefficients with an absolute value greater than 0.8. Viscosity was also found to have a moderate relationship with $|E^*|$, resulting in a correlation coefficient of 0.48. From this it appears that $|E^*|$ predictions are mostly driven by the binder properties, as these are also a function of temperature. This is to be expected as it is often postulated that stiffer binders would contribute to stiffer mixes. This notion has been supported by findings from...
(Huang et al. 2008) that showed an increase in $|E^*|$ with higher performance graded binder for certain aggregate types.

To optimize an ANN for these mixtures, parameters with the strongest correlations to measured $|E^*|$ should be selected. However, the parameters selected should be independent of each other. $|G^*|$ and $\delta$ had the strongest correlation to $|E^*|$, but because $|G^*|$ and $\delta$ were strongly correlated to one another, $|G^*|\sin(\delta)$ was selected, given that it had the next highest correlation coefficient behind testing temperature. Testing temperature and $|G^*|\sin(\delta)$ were very strongly correlated, to no surprise, since $|G^*|\sin(\delta)$ is a function of temperature, therefore, temperature was left out of the neural network. Although viscosity had a moderate correlation with $|E^*|$ and $|G^*|\sin(\delta)$, it was not included because it is also a binder property, and increasing the representation of binder by two-fold would likely introduce un-due bias to the network. Of the remaining parameters, VMA, VFA, and $V_{be}$ had the next highest correlations with $|E^*|$. However, there is a relatively strong relationship between VMA and VFA ($r = 0.69$). For this reason and that $V_{be}$ was relatively independent ($r = 0.09$) of $|G^*|\sin(\delta)$, $V_{be}$ was selected. It was found that $\rho_{3/4}$ and $\rho_{3/8}$ were strongly correlated with $V_{be}$, indicating that $V_{be}$ is driven by gradation parameters. This follows logic in that $V_{be}$ is essentially an optimized binder content derived from gradation in Superpave mix design. As a result an ANN with two hidden layers using 8 neurons was created using only two parameters, $|G^*|\sin(\delta)$ and $V_{be}$, to predict $|E^*|$ of the 2006 NCAT Test Track mixtures. The results of this ANN (ANN_2006) are shown in Figure 4. The resulting statistics shown in Figure 4 indicate that this network does not predict $|E^*|$ quite as well as any of the three previous networks, however, it still performs very well, accounting for over 96% of the variability in the dataset.

![Figure 4: Evaluation of ANN created for 2006 Test Track mixtures.](image)

The $V_{be}$ and $|G^*|\sin(\delta)$ for each mixture at the same frequency and temperature that $|E^*|$ testing was conducted (in accordance with AASHTO TP 79-09) were input into the ANN and the outputs were compared with the laboratory-determined $|E^*|$, as shown in Figure 5. Given that the 2009 dataset was much larger than the 2006 dataset and included a variety of unique mixtures and modified binders, the ANN performed quite well, accounting for 78% of the variability in the 2009 dataset. Although the ANN-predicted negative results for three data points and significantly overpredicted six others, it should be noted that these 9 data points represent a very small percentage of the 710 data points input into the ANN. Beyond the obvious difference in mix and binder technologies, the disparities can also be contributed to the difference in frequency and temperatures at which the ANN was developed and those at which it was applied. The lowest frequency used in training the ANN was only 1 Hz, whereas
the 2009 dataset included frequency as low as 0.01 Hz. Likewise, the highest temperature used in the development was 37.8°C and in application was 45°C.

Figure 5: Evaluation of ANN applied to 2009 Test Track mixes.

5 SUMMARY AND KEY FINDINGS

From this evaluation it appears that artificial neural networks are appropriate for predicting the dynamic modulus of the 2006 NCAT Test Track mixtures at intermediate temperatures (4-37.8°C). All three ANNs closely predicted measured |E*| and showed improvement over their associated predictive models, consistent with other studies (Lacroix et al. 2008, Sakhaei Far et al. 2009).

It was found that in the case of the NCAT 2006 Test Track mixtures, |E*| was primarily driven by a function of binder properties (|G*| and δ) as well as temperature. The specimens used for |E*| testing were plant-produced mixtures compacted in the laboratory. As a result, some parameters such as air voids may have a smaller range of values than if cored from existing pavement. Therefore the correlation analysis conducted is specific to the laboratory-compacted specimens. Additionally, the correlations are driven by mix type. Furthermore, it was found that only one volumetric property, Vbe, was necessary to characterize the mix properties due to its strong correlation with gradation parameters. The resulting ANN consisted of only two parameters: |G*|sin(δ) and Vbe. From this it can be concluded that an ANN can be created with minimal input parameters and still predict |E*| very well.

In applying the ANN to the 2009 Test Track mixtures, it was found that it performed moderately well with an R² of 78%. The ANN was applied to a dataset that extended beyond the range of test temperatures and frequencies used in the development of the ANN and if both datasets were combined to re-train the ANN, it is likely that its prediction capabilities would be improved due to robustness of not only the testing conditions but also the variety of mixture and binder types. This evaluation showed the advantage of using an ANN to predict |E*| in that despite the differences in mix and binder types and testing conditions, the ANN was still able to predict |E*| with relative accuracy, showing an improvement over the capability of both Witczak models in predicting |E*| for an independent dataset.

REFERENCES


