15TH AAPA INTERNATIONAL FLEXIBLE PAVEMENTS CONFERENCE

Evaluation of Dynamic Modulus Predictive Models for Typical Australian Asphalt Mixes

Saeed Yousefdoost[1], Binh Vuong[2], Ian Rickards[3], Peter Armstrong[4], Bevan Sullivan[5]
1: PhD Candidate, Faculty of Engineering and Industrial Sciences, Swinburne University of Technology, Australia, saeed.yousefdoost@yahoo.com
2: Associate Professor, Faculty of Engineering and Industrial Sciences, Swinburne University of Technology, Australia, binhvuong.services@gmail.com
3: Technical Consultant, Australian Asphalt Pavement Association (AAPA), Australia, ian_rickards@bigpond.com
4: National Technical Manager, Fulton Hogan Pty Ltd, Australia, Peter.Armstrong@fultonhogan.com.au
5: National Technical Manager, Fulton Hogan Pty Ltd, Australia, Bevan.Sullivan@fultonhogan.com.au

Keywords: Hot Mix Asphalt**,** Dynamic Modulus, Predictive Models, Hirsch, Alkhateeb, Witczak, Goodness-of-Fit Statistics, Sensitivity Analysis, Tornado Plot, Extreme Tail

1. Abstract

Dynamic modulus is a fundamental property of asphalt mixes, which is required as material input in most mechanistic-empirical pavement design systems. In the US, various database of dynamic modulus for asphalt mixes have been established and a number of models (Witczak 1-37A, Witczak 1-40D, Hirsch and Alkhateeb) have been developed to predict dynamic modulus of asphalt materials from various intrinsic characteristics of the mix. However, there is concern about whether these models are appropriate for Australian asphalt mixes. In this study, dynamic moduli of 28 different Australian asphalt mixes were tested over a spectrum of temperatures and loading frequencies, and intrinsic characteristics of the mixes were also determined to establish an Australian database that allows comparison with the US databases and validation of the nominated US models. Results in this study showed that Witczak 1-37A, Hirsch and Alkhateeb models generally under predicted dynamic moduli of the mixes while Witczak 1-40D over predicted the values. The level of bias and error observed in Hirsch, Alkhateeb and Witczak 1-40 was found to be high. The performance of all the models was also found to be inconsistent across different temperatures. Sensitivity analyses of the correlations between dynamic modulus and intrinsic characteristics used in the studied models showed that intrinsic characteristics related to binder properties have the most significant effect on the predicted dynamic modulus. The overall performance of the studied models suggested that they are not robust enough for Australian asphalt mixes.

2. Introduction

Dynamic modulus (|E*|) is a fundamental property of asphalt mixes, which is required as a material input in most mechanistic-empirical pavement design systems, typically the Mechanistic Empirical Pavement Design Guide (MEPDG) in the US. As a result of the National Cooperative Highway Research Program (NCHRP) project 9-19 (Superpave Support and Performance Models Management) and project 9-29 (Simple Performance Tester for Superpave Mix Design), the Asphalt

Materials Performance Tester (AMPT), formerly known as the Simple Performance Tester (SPT), was developed to measure dynamic modulus and evaluate the performance of Superpave HMA mixes [1]. By implementing the time-temperature superposition principle, dynamic modulus master curves can be developed to account for the effects of temperature and loading frequency from which dynamic modulus can be determined across any desired temperature and loading rate [2]. However, dynamic modulus testing and construction of the master curves can be time consuming and costly and it requires trained staff [3-6]. Consequently, a number of databases of dynamic modulus for US asphalt mixes have been established and several models (Witczak 1-37A, Witczak 1-40D, Hirsch and Alkhateeb) have been developed to predict dynamic modulus of asphalt materials from various intrinsic characteristics of the mix such as volumetric properties, aggregate gradation and binder properties, for pavement design purposes.

Currently, the US pavement design MEPDG uses measured dynamic modulus values (determined from laboratory testing) for level one design scheme in the hierarchical design approach and allows estimated dynamic modulus values, which are predicted using one of the approved predictive models, for levels two and three of the design. There has been growing interest amongst road authorities in Australia in adopting dynamic modulus as a more rational and realistic representative of materials properties than currently used resilient modulus in the pavement design procedure [7]. However, there is concern about whether the US dynamic modulus predictive models are appropriate for Australian asphalt mixes.

As part of a comprehensive research program to improve the effective deployment of Long Life Asphalt Pavement (LLAP) structures within Australian highway construction practice [1], a database of dynamic moduli of Australian asphalt mixes suitable for LLAP has been established by the Australian Asphalt Pavement Association (AAPA). A PhD research project titled "Perpetual Pavement – Design Enhancement" at Swinburne University of Technology has also investigated the performance of four common dynamic modulus predictive models (Witczak 1-37A, Hirsch, Witczak 1-40D and Alkhateeb) using the developed Australian database.

This paper provides brief descriptions of the Australian dynamic modulus database and evaluations of the accuracy of four dynamic modulus predictive models (Witczak 1-37A, Witczak 1-40D, Hirsch and Alkhateeb) in estimating the dynamic modulus for typical asphalt mixes in Australia. The sensitivity of the nominated models to their input parameters (mix and binder properties) is also discussed.

3. Dynamic Modulus Database of Australian Asphalt Mixes

The database includes a total number of 1344 data points of dynamic modulus, which is suitable for assessing the performance of the nominated predictive models. Brief descriptions of the mix and binder properties and test methods used to determine the dynamic modulus are given below.

3.1. Description of Materials

This study focused on 28 standard dense graded asphalt materials produced by Australia's major asphalt producers. The asphalt samples consisted of 15 mixes with nominal maximum aggregate size (NMAS) of 14mm and 13 mixes with NMAS of 20mm. The asphalt samples were carefully selected and nominated to make sure that the spectrum of Australia's generally used aggregates and binders in major projects was covered. Studied mixes contained nine types of aggregates (Honfels, Granite, Greywacke, Limestone, Dolomite, Dolomatic Siltstone, Basalt, Latite and Dacite) and five types of binders (C320, C600, AR450, Multigrade and A15E PMB). The Rap content of the mixes varied from nil up to 30 percent.

3.2. Sample Preparation

Asphalt samples were taken in loose form from plant production and were then cooled and delivered to the laboratory. The mix design of the studied asphalt samples was based on two common methods currently being used in Australia: a) Marshall mix design method (Marshall compaction) and the Austroads mix design method (Gyropac compaction). After reheating, samples were compacted into 450mm x 150mm x 180mm blocks using PReSBOX compactor. PReSBOX compactor applies a constant vertical axial load and a cyclic horizontal shear force with a constant maximum shear angle to compact asphalt blocks. It is believed that the shearing action of the PReSBOX compactor closely replicates the conditions under which asphalt is placed in the field, producing blocks with uniform air void distribution, particle orientation and density [8, 9]. Compacted blocks were then cored and trimmed to obtain three cylindrical specimens 102mm in diameter and 150mm in height. Specimens were compacted at 4±0.5% target air voids. A total number of 28 blocks were compacted from which 56 cores (two replicates per mix) were prepared for dynamic modulus testing.

3.3. Test Methods

To meet the objectives of this study, dynamic modulus laboratory testing was conducted on plant produced laboratory compacted asphalt mixes. Dynamic Shear Rheometer (DSR) testing was also conducted on the corresponding binders. The Asphalt Mixture Performance Tester (AMPT) at Fulton Hogan R&D asphalt laboratory was used to perform dynamic modulus tests in accordance with AASHTO TP 79-11 (Determining the dynamic modulus and flow number for hot mix asphalt using the asphalt mixture performance tester). In the dynamic modulus test, an asphalt specimen is subjected to a controlled sinusoidal compressive stress at various temperatures, frequencies and confinement pressures. The dynamic modulus and the phase angle of the specimen are then calculated based on the applied stresses and the axial response strains with respect to time. Testing was carried out at four temperatures (5, 20, 35 and 50°C), six loading frequencies (0.1, 0.5, 1, 5, 10 and 25 Hz) and four levels of confinement pressures (Nil, 50, 100 and 200 KPa). Only results from unconfined testing condition are reported in this study. Results of the dynamic modulus tests were then compared with the values predicted by the four nominated predictive models (Witczak 1-37A, Witczak 1-40D, Hirsch and Alkhateeb) and the accuracy of the models was evaluated.

DSR test equipment at Fulton Hogan binder laboratory was used to test the rheological properties of RTFO aged binders (Complex shear modulus and Phase Angle) in accordance with Fulton Hogan's DSR test procedure 121110 (equivalent to AASHTO T 315-12). The dynamic shear rheometer applies sinusoidal waveform oscillatory shear force to a binder sample formed between two parallel plates at preselected frequencies and temperatures to calculate complex shear modulus (G^*) and the phase angle (δ) of the binder. Testing was carried out at temperatures from 20 to 60°C with increments of five degrees and 11 frequencies ranging from 0.1 to 10 Hz.

3.4. Development of Master Curves

Since the test temperature and loading frequency used for dynamic modulus and DSR tests were different, master curves were developed for the DSR test results (complex shear modulus and phase angle) for all the binder samples so that shear moduli and phase angles could be obtained from developed master curves for any desired temperature and loading frequency.

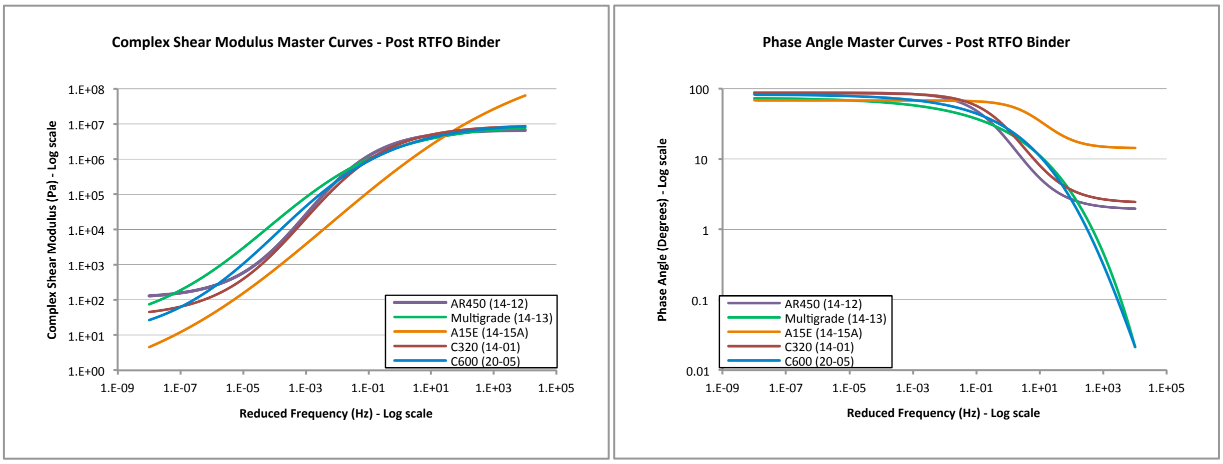
Master curves are constructed based on the time-temperature superposition principle. The binders' complex shear modulus and phase angle data at various temperatures were shifted in line with respect to the loading frequency until the curves merged into a single smooth function. Master curves describe the time dependency of materials. The temperature dependency of materials is defined by the amount of shifting required at each temperature to form the master curve [2].

Christensen-Anderson, Christensen-Anderson-Marasteanu, Alqadi-Elseifi and a sigmoidal function models were nominated to construct the G* and δ master curves [10-12]. The sigmoidal function and a second order polynomial shift factor provided the best fit with minimum errors and therefore were selected to be used for developing the master curves. The general forms of the utilized master curve and the shift factor functions are given in equations 1 and 2 respectively.

$$\log P = \delta + \frac{(\alpha)}{1+e^{\beta+\gamma \log f_r}} \quad \textbf{(1)}$$

$$\log f_r = \log f + a_1(T_R - T) + a_2(T_R - T)^2 \quad \textbf{(2)}$$

P: The calculated Parameter (complex shear modulus in Pa or phase angle in degrees)
α, β, δ, ϒ, a_1 and a_2: Fitting parameters
f_r: Reduced frequency at the reference temperature, Hz
f: Loading frequency at the test temperature, Hz
T_R: Reference temperature, °C (20°C in this study)
T: Test temperature, °C



(a) (b)

Figure1. Examples of master curves developed for C320, C600, AR450, Multigrade and A15E binders (a) complex shear modulus (b) phase angle

Microsoft Excel Solver was then used to find the fitting parameters by minimizing the sum of squared errors between the measured complex shear modulus and phase angle at each temperature/frequency combination and the calculated values by equation 1. Figure 1 shows examples of developed master curves for the complex shear modulus and the phase angle of C320, C600, AR450, Multigrade and A15E RTFO aged binders.

4. Review of the US Dynamic Modulus Database and Predictive Models

4.1. Witczak 1-37A Model

Andrei Witczak et al. developed this model based on a database of 2750 data points representing dynamic modulus test results for 205 asphalt mixes tested over 30 years in the laboratories of the Asphalt Institute, the University of Maryland, and the Federal Highway Administration [2]. The original version of the model was developed by Shook and Kalas in 1969 and was further modified and refined by Fonseca and Witczak in 1996 [13]. Witczak 1-37A was developed based on a combined database of the original Fonseca and Witczak model (1430 data points, 149 conventional asphalt mixes) and an additional 1320 data points of 56 new mixes including 34 mixes with modified binders [14]. This model (given in equation 3) is based on nonlinear regression analysis and is currently being utilized to predict dynamic modulus of the asphalt mixes in level 2 and 3 designs of the MEPDG in the US [2]. It incorporates basic volumetric properties and grading of the asphalt mix, binder viscosity and loading frequency into a sigmoidal function to predict the dynamic modulus over a spectrum of temperature and frequency. Summary details of the Witczak 1-37A database are presented in table 1[15].

$$\begin{aligned} log\,E = & -1.249937 + 0.029232.p_{200} - 0.001767.(p_{200})^2 - 0.002841.p_4 - 0.058097.V_a \\ & - 0.802208.\frac{V_{beff}}{(V_{beff}+V_a)} \\ & + \frac{3.871977 - 0.0021.p_4 + 0.003958.p_{38} - 0.000017.(p_{38})^2 + 0.005470.p_{34}}{1+e^{(-0.603313-0.313551.log(f)-0.393532.log(\eta))}} \end{aligned} \quad \textbf{(3)}$$

Where:
E: Asphalt mix dynamic modulus, in 10^5 psi
η: Bitumen viscosity, in 10^6 poise (at any temperature, degree of aging)
f: Load frequency in Hz
V_a: % Air voids in the mix, by volume
V_{beff}: % Effective bitumen content, by volume
p_{34}: % Retained on the 3/4 inch sieve, by total aggregate weight (cumulative)
p_{38}: % Retained on the 3/8 inch sieve, by total aggregate weight (cumulative)
p_4: % Retained on the No. 4 sieve, by total aggregate weight (cumulative)
p_{200}: % Passing the No.200 sieve, by total aggregate weight

Witczak 1-37A requires an established linear viscosity–temperature relationship to calculate the viscosity of the binder at a desired temperature using equation 4, in which η is the viscosity of binder (centipoise), T_R is temperature (Rankine) and A and VTS are regression constants.

$$\log\log\eta \;=\; \mathrm{A} + \mathrm{VTS}\log T_R \quad \textbf{(4)}$$

Depending on the availability of the data and the required level of accuracy, the binder viscosity can be obtained directly from laboratory measurements or can be estimated using empirical relationships with a series of conventional binder tests such as ring and ball, penetration and kinematic viscosity; typical values for A and VTS are also provided in the MEPGDG based on performance grade, viscosity grade and penetration grade of the binder [2, 16]. In this study, A and VTS were calculated by establishing a linear regression for laboratory measured viscosities of the binders at 5, 20, 35 and 50°C to comply with the planned dynamic modulus testing temperature regime.

Table1. Summary of Witczak 1-37A Dataset

Temperature Range	0 to 130°F (-17.7 to 54.4°C)
Loading Frequency	0.1 to 25 Hz
Binder	9 unmodified 14 Modified
Aggregate	39 types
Asphalt Mix	171 with unmodified binder 34 with modified binder
Compaction	Kneading and Gyratory
Specimen Size	Cylindrical 4 x 8 in (10 .2 x 20.3 cm): Kneading Compaction Cylindrical 2.75 x 5.5 in(7 x 14 cm): Gyratory Compaction
Specimen Aging	Un-aged
Total Data Points	2750

4.2. Witczak 1-40D Model

Bari and Witczak conducted further dynamic modulus testing on asphalt mixes which resulted in a revised version of Witczak's former model based on 7400 data points acquired from 346 different HMA mixes. Their expanded database consisted of HMA mixes with different aggregate gradations, binder types (conventional, polymer modified and rubber modified), mix types (conventional un-modified and lime or rubber modified) and aging conditions (no aging, short-term oven aging, plant aging and field aging) [14]. The test specimens in their new database had a cylindrical size of 2.75 x 5.5 in (7 x 14 cm) and were compacted by gyratory compaction. Dynamic modulus tests were conducted from 0 to 130°F (-17.7 to 54.4°C) temperatures and 0.1 to 25 Hz loading frequencies [17].

In their new model, the sigmoidal structural form of the original model was maintained and the same inputs of volumetric and gradation properties of the mix were used. However, DSR test results were incorporated instead of viscosity and frequency to reflect the rheology of the binder with changing temperature and load rate [16, 17]. The revised model is expressed as equation 5.

$$\begin{aligned} log\, E = {} & -0.349 + 0.754(|G_b{}^*|^{-0.0052}) \\ & \times \left(6.65 - 0.032p_{200} + 0.0027{p_{200}}^2 + 0.011p_4 - 0.0001{p_4}^2 + 0.006p_{38} \right. \\ & \left. - 0.00014{p_{38}}^2 - 0.08V_a - 1.06\left(\frac{V_{beff}}{V_a + V_{beff}}\right)\right) \\ & + \frac{2.56 + 0.03V_a + 0.71\left(\frac{V_{beff}}{V_a + V_{beff}}\right) + 0.012p_{38} - 0.0001{p_{38}}^2 - 0.01p_{34}}{1 + e^{(-0.7814 - 0.578585 log|G_b{}^*| + 0.8834 log\delta_b)}} \end{aligned} \qquad \textbf{(5)}$$

Where $|G_b{}^*|$ is the dynamic shear modulus of the binder (in psi) and δ_b is the phase angle of the binder associated with $|G_b{}^*|$ (in degrees) [17].

As part of the model development, Bari and Witczak assumed that the angular loading frequency in dynamic compression mode (f_c in Hz) as used in the dynamic modulus test and the loading frequency in dynamic shear mode (f_s in Hz) as used in the DSR test, are not equal but related as $f_c=2\pi f_s$ [17]. Therefore, in this study loading frequencies of 3.979, 1.592, 0.796, 0.159, 0.080 and 0.016 Hz (equivalent to 25, 10, 5, 1, 0.5 and 0.1 Hz in compression mode) were used in the DSR test to estimate shear modulus ($|G_b^*|$) and phase angle (δ_b) as the inputs into the dynamic modulus predictive model.

4.3. Hirsch Model

Christensen et al. applied the Hirsch model, an existing law of mixtures which combines series and parallel elements of phases, to asphalt mixes and developed this model based on a database that includes results from testing on 18 asphalt mixes using eight different binders and five different aggregate sizes and gradations (Table 2). A total of 206 observations were included in the dataset for each shear and compression test. The Hirsch model incorporates volumetric properties of the mix (VMA and VFB) and dynamic shear modulus of the binder (G_b^*). The binder shear moduli of the Hirsch database were measured using the Superpave Shear Tester (SST) frequency sweep procedure on asphalt specimens. The model is given in equations 6 and 7 [18].

$$|E^*|_{mix} = P_c \times \left[4{,}200{,}000\left(1-\frac{VMA}{100}\right) + 3|G^*|_{binder}\left(\frac{VFB \times VMA}{10{,}000}\right)\right] + \frac{(1-P_c)}{\left[\frac{1-\frac{VMA}{100}}{4{,}200{,}000} + \frac{VMA}{3VFA|G^*|_{binder}}\right]} \quad \textbf{(6)}$$

$$P_c = \frac{\left(20 + \frac{VFB \times 3|G^*|_{binder}}{VMA}\right)^{0.58}}{650 + \left(\frac{VFB \times 3|G^*|_{binder}}{VMA}\right)^{0.58}} \quad \textbf{(7)}$$

Where:
|Gb*|: Absolute Value of the binder complex shear modulus (psi)
|E*|: Absolute value of asphalt mix dynamic modulus (psi)
VMA: Voids in mineral aggregates in compacted mix (%)
VFB: Voids filled with binder in compacted mix (%)

Table2. Summary of Hirsch Dataset

Factor	ALF[1]	MN/Road	West Track	Total
Design Method	Marshall	Marshall	Superpave	2
Binders	AC-5, 10, 20 SBS Modified PE-Modified	AC-20 120/150-Pen	PG 64-22	8
Aggregate Size and Gradation	19mm Dense 37.5mm Fine	9.5mm Fine	19mm Fine 19mm Coarse	5
Asphalt Mix	7	5	6	18
Total Data Point	78	59	69	206
For Complete Database				
Air Voids (%)	5.6 to 11.2	Complex Shear Modulus (MPa)		20 to 3,880
VMA (%)	13.7 to 21.6	Phase Angle (degrees)		8 to 61
VFB (%)	38.7 to 68	Temperature (°C)		4, 21 and 38
Dynamic Modulus (MPa)	183 to 20900	Loading Frequency		0.1 and 5

1: FHWA Accelerated Loading Facility

4.4. Alkhateeb Model

Alkhateeb et al. used law of mixtures and combined behavior of a three-phase system of aggregate, binder and air voids in parallel arrangement with each other to formulate their predictive model as described in equation 8. Alkhateeb et al. tested the dynamic modulus of plant-produced lab-compacted (PPLC), lab-produced lab-compacted (LPLC) and field cored specimens for their study. The LPLC specimens were fabricated with 6 binders: unmodified PG-70-22, PG 70-28 air blown, PG 70-28 Styrene-Butadiene-Styrene Linear-Grafted (SBS LG), PG 76-28 crumb rubber blended at the terminal (CR-TB), PG 70-28 Ethylene Terpolymer, and PG 70-34 binder containing a blend of SB and SBS. The aggregates used to prepare the specimens were crushed Diabase stone with NMAS of 12.5mm. Details of the aggregate specifications are presented in table 3.

Table3. Summary of Alkhateeb Dataset

Aggregate and Mix Properties		Aggregate Grading	
Bulk Dry Specific Gravity (t/m^3)	2.947	Sieve size (mm)	%Passing
Bulk Saturated Surface Dry Gravity (t/m^3)	2.965	37.5	100
Apparent Specific Gravity (t/m^3)	3.001	12.5	93.6
Absorption (%)	0.6	9.5	84.6
Los Angeles Abrasion (%)	19	4.75	56.7
Sand Equivalent (%)	75	2.36	34.9
NMAS (mm)	12.5	1.18	24.8
LPLC Binder Content (%)	5.3	0.6	18.2
LPLC and PPLC Compaction	Gyratory	0.3	13.1
Targeted Air Voids (%)	7±0.5	0.15	9.3
Specimen Cylinder Size(mm)	100 x 150	0.075	6.7
Dynamic Modulus Test Temperature (°C)	4, 9, 31, 46, 58		
Dynamic Modulus Test Frequency (Hz)	0.1, 0.5, 1, 5, 10		

The LPLC specimens were short-term oven aged (4 hours at 135°C) with a binder content of 5.3% and targeted air voids of 7±0.5 percent. Gyratory compaction was used to compact the LPLC and PPLC specimens (100 x 150mm). The dynamic modulus test was performed on the specimens at 4, 9, 31, 46 and 58°C and 0.1, 0.5, 1, 5, 10 Hz loading frequencies. Alkhateeb's model was calibrated with four polymer modified asphalts and one plant produced neat asphalt. The calibrated model was then validated against the other mixes of the dataset [19].

$$|\mathrm{E}^*|_\mathrm{m} = 3\left(\frac{100-\mathrm{VMA}}{100}\right)\left(\frac{\left(90+1.45\frac{|\mathrm{G}^*|_\mathrm{b}}{\mathrm{VMA}}\right)^{0.66}}{1100+\left(0.13\frac{|\mathrm{G}^*|_\mathrm{b}}{\mathrm{VMA}}\right)^{0.66}}\right)|\mathrm{G}^*|_\mathrm{g} \tag{8}$$

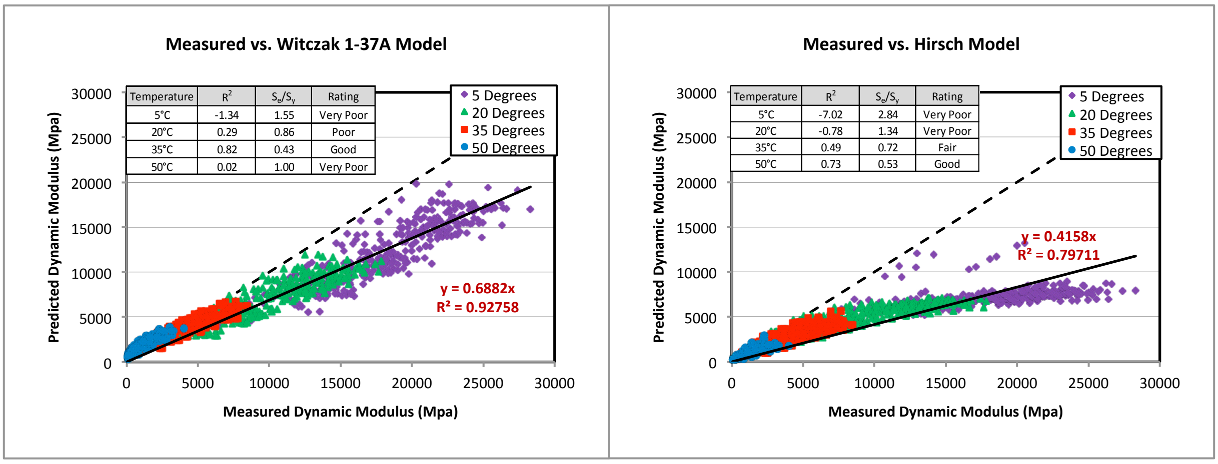
$|G_b^*|$ is the complex shear modulus of binder in Pa and $|G_g^*|$ is the complex shear modulus of binder in glassy state in Pa, which is assumed to be 10^9 Pa.

5. Evaluation of the Accuracy of the US Models for Australian Asphalt Mixes

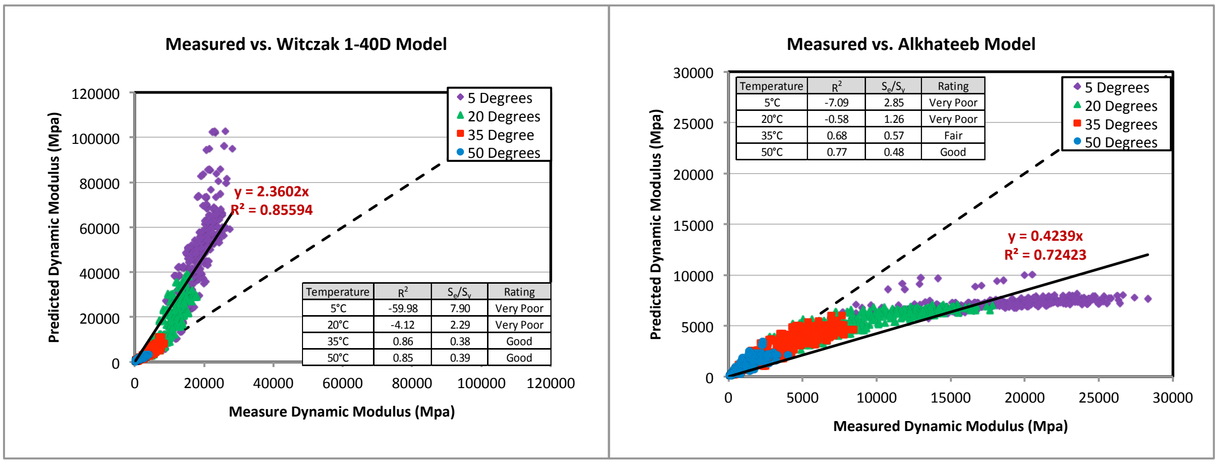
Four approaches were adopted to assess the reliability of the nominated models: (i) Plotting measured versus predicted dynamic modulus values and comparing them with the line of equality (LOE), (ii) evaluating goodness-of-fit statistics, (iii) analyzing the residuals and (iv) analyzing bias statistics. A total number of 1344 data points were used to assess the performance of the nominated predictive models. In some research, logarithmic scale is chosen to present the results, most likely because of a relatively wide range of dynamic modulus values, also due to the fact that both Witczak models (1-37A and 1-40D) estimate the log |E*| rather than the actual values [3, 5, 13, 16]. In this study the dynamic modulus values illustrated in the figures and tables are in arithmetic scale as it was believed that the logarithmic scale may not realistically reflect the level of errors developed by the models [20].

5.1. Measured Versus Predicted Dynamic Modulus

First, the predicted dynamic modulus values were plotted against the measured values along with LOE to examine the scattering of the data. A good predictive model will produce results following the LOE in an oval shape [21]. Figure 2 shows the comparison between measured and predicted dynamic modulus values by the nominated models. Plots were segregated by temperature to evaluate the robustness of the models to predict dynamic moduli at each test temperature. The goodness-of-fit statistics of the model predictions at different temperatures are also listed in figure 2.



(a) (b)



(c) (d)

Figure2. Predicted versus measured dynamic modulus for four nominated models: (a) Witczak 1-37A, (b) Hirsch, (c) Witczak 1-40D and (d) Alkhateeb

As can be seen from figure 2, Witczak 1-37 A, Hirsch and Alkhateeb generally under predicted dynamic modulus values; on the other hand Witczak 1-40D over predicted the |E*| to a great extent. The magnitude of over/under prediction of the models varies with temperature and generally broadens at lower temperatures (higher dynamic modulus values). Note that due to relatively larger values predicted by the Witczak 1-40D model, the axes scales for this model are set differently from the other three in figure 2.

To get an approximate overview of the level of over or under prediction of the models, linear constrained trendlines (lines with zero intercepts) were fitted to data in the predicted-measured space (figure 2). The conformity of trendlines with LOE (with the slope of unity) shows how the estimations match the actual data. The overall performance of Hirsch and Alkhateeb models at different temperatures were relatively similar; they both under predicted |E*| by approximately 58%. Witczak 1-37A predictions were the closest to the line of equality, yet still under predicted the values by approximately 31%. On the other hand the second Witczak model (1-40D) overestimated dynamic moduli by about 136%. Dongre et al. found that Hirsch and Witczak 1-40D models tend to under predict |E*| when the air voids or the binder content of the mix is more than the mix design [22].

5.2. Statistics of Goodness-of-Fit

To evaluate the goodness-of-fit statistics, standard error of estimate (S_e), standard deviation of the measured values (S_y), the coefficient of determination (R^2) and error between predicted and measured dynamic modulus (e) were calculated for each model using equations 9-12 [23]. S_e is an indicator of likely error in the prediction and R^2 is a measure of the model accuracy. The standard error ratio (S_e/S_y) was also determined, to facilitate the evaluation of the nonlinear models [1, 20]. Witczak et al. and Singh et al. have used S_e/S_y and R^2 criteria in their studies to subjectively classify the performance of the models for their dataset [1, 3] which is presented in table 4. In equations 9-12, E^*_m is the measured dynamic modulus, E^*_p is the estimated dynamic modulus, $\overline{E}^*$ is the average of measured dynamic moduli, n is the size of the sample and k is the number of regression coefficients of the model.

$$S_y = \sqrt{\frac{\sum_{i=1}^{n}(E^*_{mi} - \overline{E}^*_m)^2}{(n-1)}} \quad \textbf{(9)}$$

$$S_e = \sqrt{\frac{\sum_{i=1}^{n}\left(E^*_{pi} - E^*_{mi}\right)^2}{(n-k-1)}} \quad \textbf{(10)}$$

$$e = \left(E^*_{pi} - E^*_{mi}\right) \quad \textbf{(11)}$$

$$R^2 = 1 - \frac{(n-k-1)}{(n-1)}\left(\frac{S_e}{S_y}\right)^2 \quad \textbf{(12)}$$

Details of the goodness-of-fit statistics for the studied asphalt mixes are listed in table 5. As can be seen from the table, Witczak 1-37A has the highest R^2 and the least sum of squared error (SSE), which indicates that its performance is relatively superior to the other models, however, according to Witczak et al. classification criteria, it still cannot be classified as an “Excellent” fit to data. A substantial level of sum of squared error (SSE) and S_e/S_y, and low R^2 values calculated for the Witczak 1-40D, Hirsch and Alkhateeb models indicate that their overall performance can be rated as poor or very poor. Negative average error for the Witczak 1-37A, Hirsch and Alkhateeb models reflects the fact that predictions of these models were consistently lower than the measured values that could be visually observed from LOE plots (figure 2). Similarly, the positive average error of Witczak 1-40D indicates the over predictive nature of this model.

Table 4. Criteria for subjective classification of the goodness-of-fit statistics

Criteria	R^2 (%)	S_e/S_y
Excellent	>90	<0.35
Good	70-89	0.36-0.55
Fair	40-69	0.56-0.75
Poor	20-39	0.76-0.90
Very Poor	<19	>0.90

The negative R^2 of the Witczak 1-40D model implies that for this model, sum of squared errors (SSE) is more than total sum of squares (SST); in other words horizontal line of y= $\overline{E}^*_m$ (average measured dynamic modulus) fits the data better and produces less error than the investigated nonlinear model. This suggests that the utilized model is not necessarily a good fit for the particular selected dataset. Robbins and Timm also encountered similar issues with the Witczak 1-40D model in their studies [16].

Table 5. Goodness-of-fit statistics of the original predictive models

Criteria		Witczak 1-37A	Witczak 1-40D	Hirsch	Alkhateeb
S_e/S_y	Overall	0.49	2.29	0.88	0.87
	5 Degrees	1.55	7.90	2.84	2.85
	20 Degrees	0.86	2.29	1.34	1.26
	35 Degrees	0.43	0.38	0.72	0.57
	50 Degrees	1.00	0.39	0.53	0.48
R^2	Overall	0.76	-4.19	0.23	0.24
	5 Degrees	-1.34	-59.98	-7.02	-7.09
	20 Degrees	0.29	-4.12	-0.78	-0.58
	35 Degrees	0.82	0.86	0.49	0.68
	50 Degrees	0.02	0.85	0.73	0.77
Other Statistics	SSE	1.75E+10	3.79E+11	5.59E+10	5.51E+10
	Average $\|E^*_p\|$	5953	16555	3812	4000
	Average Error	-2007	8595	-4147	-3959
	Slope	0.618	2.688	0.342	0.332
	Intercept	1030.8	-4837.5	1090.0	1357.8
	Rating	Good	Very Poor	Poor	Poor

Average measured $|E^*_m|$: 7959
Total Sum of Squares (SST): 7.29E+10

Unlike the overall performance of the models, the robustness of the models was considerably different at various levels of temperature. As the goodness-of-fit statistics in figure 2 and table 5 suggest, although the overall performance of Witczak 1-37A can be classified as good, only its predictions at 35°C are consistent with the measured values (R^2 = 0.82) and the model performance at other temperatures is significantly inaccurate. Similarly, Hirsch and Alkhateeb models seem to be applicable only at 50°C, providing R^2 of 0.73 and 0.77 respectively; their under prediction becomes significant at lower temperatures. Witczak 1-40D overall goodness-of-fit statistics show a very poor performance, however its accuracy at 35 and 50°C is significantly more reliable than at lower temperatures, providing R^2 of 0.86 and 0.85 respectively, which in fact is higher than the other models. However, its substantial over predictions at 5 and 20°C influence its overall performance and make it unreliable on the whole. Singh et al. demonstrated that Hirsch and Alkhateeb models' predictions were reliable at low temperatures whereas both Witczak models (1-37A and 1-40D) performed well at high temperatures only. They noted that all models were inaccurate at low temperatures and high air voids [3].

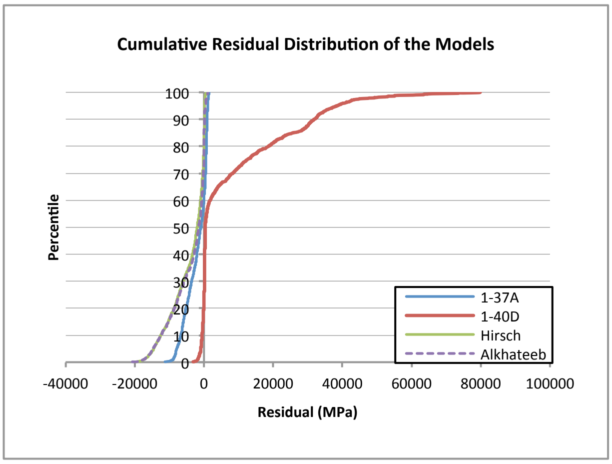
5.3. Residual Distributions

Figure 3(a) illustrates the cumulative distribution of residuals for each model. Ideally, residuals are expected to be of a small width, to be symmetrical and to center at nil [16]. Witczak 1-37A residuals range from -11.3 to 1.5 GPa giving a span of 12.8 GPa, which is quite different from the later model (1-40D) that gives approximately 6.4 and 3.9 times greater width than the Witczak 1-37 and Hirsch models respectively. Hirsch and Alkhateeb's residual distributions show a similar trend, which range from -20 to 1 GPa. More than 70% of Witczak 1-40D residuals are in the positive area, which is an indication of its considerable over estimation. The 50^{th} percentile of Witczak 1-40 was the smallest one (375 MPa), followed by Witczak 1-37A with -887 MPa and that of Hirsch was the furthest from nil (-1957 MPa). Overall, it appeared that none of the models had the residuals cumulative distribution of a robust model.

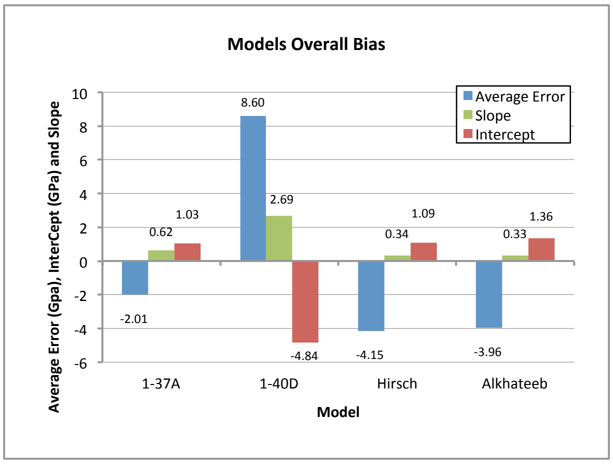
5.4. Overall Bias of the Models

Another measure to evaluate the performance of the models is to look at their overall bias by comparing the slope and intercept of non-constrained linear fits to the predicted versus measured dynamic modulus data. The slope and the intercept of a reliable model would be close to 1 and zero respectively. Figure 3(b) shows the average error, slope and intercept of non-constrained linear trendlines for each model. The divergence between slopes and the unity suggests the dependence of prediction errors to actual values [13]. The larger the intercepts, depending on their sign, the more the estimations are over or under predicted. Slopes ranges are as low as 0.33 for the under predictive Hirsch and Alkhateeb models to as high as 2.69 for the considerably over predictive Witczak 1-40D model. The intercepts of trendlines with slopes below unity are positive (Witczak 1-37A, Hirsch and Alkhateeb) but for the Witczak 1-40D model the intercept is negative to make up for the steep incline. Overall it seems that all models exhibit a significant amount of bias in their predictions. Ceylan et al. compared the overall bias of both Witczak models with two artificial neural network (ANN) models and found a relatively moderate amount of bias in Witczak models compared to studied ANN based models [20]. From the results obtained in the first stage of this study, it can be concluded that the dynamic moduli estimated by the four nominated models are not accurate or robust enough to be used to predict dynamic moduli of typical Australian asphalt mixes, which warrants some modification to the models to improve their performance for the studied material.

Part of the inaccuracy and errors developed in the models' predictions can be attributed to different databases based on which the studied models were developed and calibrated, i.e. binder type and aging condition; aggregate size, type and gradation; shear modulus and dynamic modulus test setup (equipment, temperature and loading rate) and specimen size; mix design and compaction method; and volumetric properties of the mixes (air voids, VMA and VFB) could partially be held responsible for the discrepancies observed in the $|E^*|$ predictions and laboratory measurements for the Australian asphalt materials. To find out the key input parameters into the models a sensitivity analysis was carried out which is discussed in more details in the next section.



(a)



(b)

Figure3. (a) Cumulative distribution of Residuals (b) Overall bias in Witczak 1-37A, Witczak 1-40D, Hirsch and Alkhateeb models

6. Sensitivity of Model Input Parameters

Sensitivity analysis is a technique to determine the influence of the input parameters of a model on the outputs; it also reveals the correlation between the uncertainties between the inputs and the outputs of the model. Therefore it can help determine the input parameters of a model that are required to be measured more precisely to improve the accuracy of the model. Spearman rank correlation index and the extreme tail coefficient were used to evaluate the sensitivity of the models to their input parameters.

Spearman rank correlation index (ρ) is a non-parametric technique that tests the strength of a relationship between two sets of data, which is calculated by using equation 13. The absolute value of ρ quantifies the strengths of the correlations between two variables; when ρ approaches to unity, the variable has the maximum impact on the model output and when it is closer to zero, the effect is marginal. Also the sign of ρ indicates that the variable is either directly or inversely proportional to the outcome; a positive value indicates that the correlation is directly proportional and a negative value shows an inversely proportional correlation. In equation 13, ρ is Spearman rank correlation index, d_i is the difference in the ranks between the input and the output values in the same data pair and n is the number of data [24].

$$\rho = 1 - \left(\frac{6\sum d_i^2}{n(n^2-1)}\right) \quad \textbf{(13)}$$

Tornado plots of the Spearman index of the studied models are illustrated in figure 4. As can be seen from the figure, the binder properties inputs into the model (i.e. viscosity, complex shear modulus and phase angle) had the most significant effect on the predicted dynamic moduli for all the predictive models. The high positive ρ values of viscosity and complex shear modulus indicate that as the viscosity and G* of the binder increases, the predicted dynamic modulus value will also intensify. Phase angle had a high yet negative ρ value for Witczak 1-40D which indicates that as the viscous behavior of the binder increases the predicted dynamic modulus decreases; for a particular type of binder, this can happen due to an increase in temperature or a drop in loading rate. Figure 4 also shows that the aggregate and volumetric properties of the mixes had relatively marginal effects on the models predictions. V_a (%Air Voids) and p_{200} (Passing the No.200 sieve) had the lowest ρ value in both Witczak models and thus had a minimal influence on the predicted E*. Also the sensitivity of Hirsch and Alkhateeb's models' predictions to VMA and VFB are insignificant.

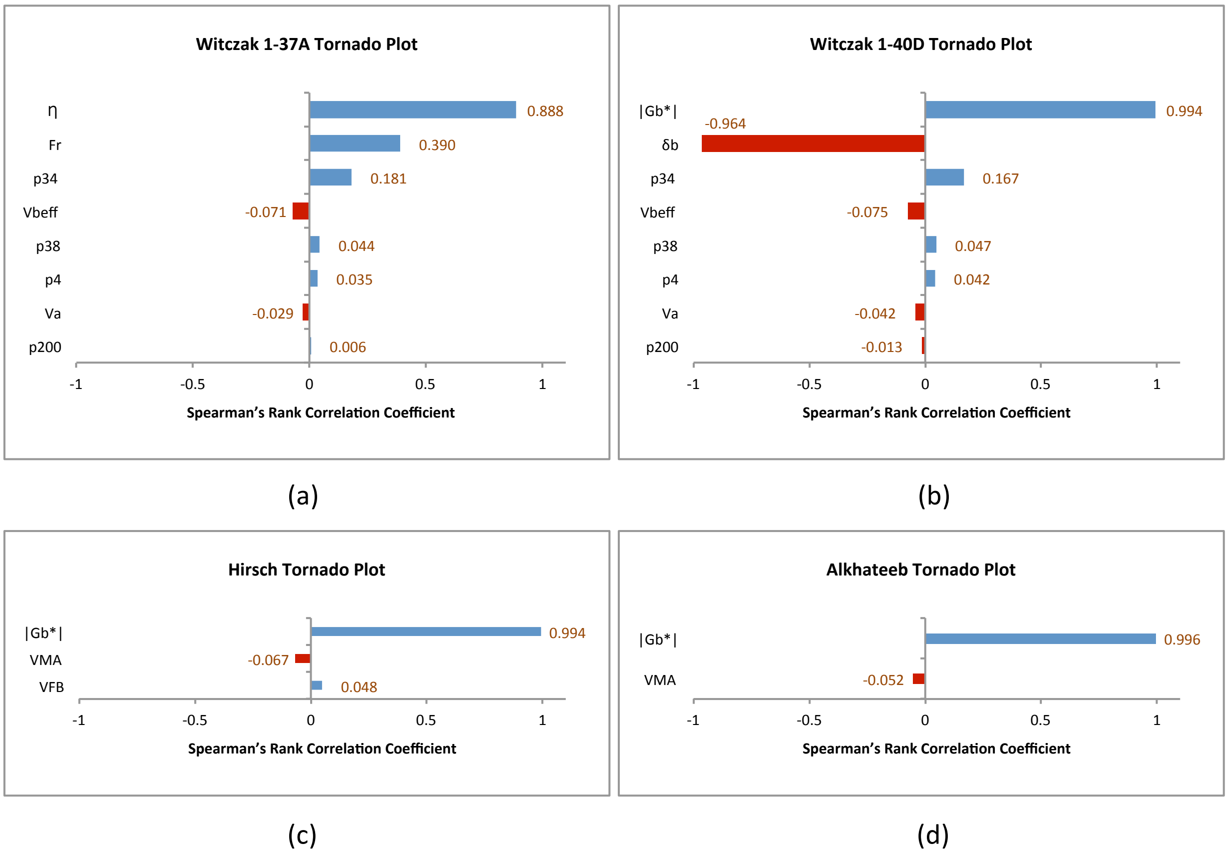


Figure4. Tornado plots of Spearman rank correlation index (a) Witczak 1-37A model, (b) Witczak 1-40D model, (c) Hirsch model and (d) Alkhateeb model

Extreme tail analysis is a statistical tool to evaluate the extent to which the input parameters contribute to the extreme values of the predictions; therefore it can assist with improving the accuracy of the model by controlling the uncertainties of the parameters that cause the tail of the models output distribution. To identify the main contributors to the extreme values of the predictions, the normalized α coefficient was calculated for each input parameter of each model using equation 14. The top high 5% and bottom low 5% of the predicted dynamic modulus values were considered as the right and left tails respectively. The effect of parameters with |α| values of greater than 0.5 are considered as significant. In equation 14, $Median_{Group}$ is the median of the input in the group, $Median_{Total}$ is the median of the input in the total dataset and σ_{Total} is the standard deviation of the input in the total dataset [24].

$$\alpha = \left(\frac{Median_{Group} - Median_{Total}}{\sigma_{Total}}\right) \quad \textbf{(14)}$$

The results of the extreme tail analysis (table 6) confirm the results from tornado plots. For Witczak 1-37A, viscosity and frequency are the primary contributors to the high values of dynamic modulus; also V_a seems to have an insignificant effect on extreme values in both Witczak models. Furthermore, complex shear modulus is the main contributor to extreme values of Witczak 1-40D, Hirsch and Alkhateeb predictions. Extreme tail analysis results also suggest that generally, input parameters do not have a significant effect on forming the low extreme values of dynamic modulus; the only exception is δ_b in Witczak 1-40D which significantly contribute to extremely low $|E^*|$ predictions.

Table6. Extreme tail analysis details of models' input parameters

Input	Witczak 1-37A		Input	Witczak 1-40D		Input	Hirsch		Alkhateeb	
	Left Tail	Right Tail		Left Tail	Right Tail		Left Tail	Right Tail	Left	Right
η	-0.0109	2.4012	$\|G_b^*\|$	-0.2238	2.5831	$\|G_b^*\|$	-0.4654	1.9326	-0.4654	1.9726
f	-0.3301	2.5039	δ_b	0.8564	-1.8269	VMA	0.2177	-0.7360	0.0933	-0.5079
V_a	0.0921	-0.2762	V_a	0.1204	-0.4746	VFB	-0.0561	0.8941		
V_{beff}	0.4093	-0.1488	V_{beff}	0.4093	-0.1488					
p_{34}	0.0000	0.3685	p_{34}	0.0000	0.3685					
p_{38}	-0.2625	0.2625	p_{38}	0.1312	0.7874					
p_4	-0.1109	0.3326	p_4	0.1109	0.7760					
p_{200}	0.0000	0.0409	p_{200}	0.0000	-0.6817					

Results from sensitivity analysis suggest that binder properties such as viscosity, complex shear modulus, and binder phase angle play an important role in the predicted $|E^*|$ values. This can partially explain the discrepancies between the measured and the predicted dynamic moduli by the nominated models as the binder databases based on which the models were developed have been entirely different.

7. Summary and Conclusion

In this study, 1344 data points of dynamic modulus were determined for 28 Australian typical dense graded asphalt mixes and used to evaluate the accuracy and feasibility of four commonly used dynamic modulus predictive models (Witczak 1-37A, Witczak 1-40D, Hirsch and Alkhateeb). It was found that the overall performance of Witczak 1-37A was more accurate than the other studied models yet it under predicted the dynamic modulus of the mixes by approximately 31%. The overall predictions of Hirsch and Alkhateeb at different temperatures were similar; they both under predicted $|E^*|$ by approximately 58%. On the other hand the second Witczak model (1-40D) substantially overestimated dynamic moduli by 136%.

Unlike the overall performance of the models, robustness of the models at different temperatures varied considerably. Generally, it was found that the models perform fairly poorly at low temperatures (5 and 20°C). Witczak 1-37A can only reliably predict dynamic modulus at 35°C. Hirsch and Alkhateeb predictions could only be classified as "good" at 50°C. Although the overall goodness-of-fit statistics of Witczak 1-40D was very poor, its predictions at 35 and 50°C were the most accurate of all.

The level of bias and error developed in the models suggested that the application of the nominated models to predict $|E^*|$ for Australian mixes is impractical. In general, this study has found that the four nominated models performed fairly poorly in predicting dynamic modulus for the studied asphalt materials. It is believed that part of the inaccuracy and errors observed in the models prediction can be attributed to the different material databases based on which the studied models were developed and calibrated.

Sensitivity analyses on the nominated models showed that the models' predictions are highly dependent on the inputs related to binder characteristics (complex shear modulus, phase angle and viscosity). P_{200} and V_a were found to have minimal influence on both Witczak models' predictions. Also the effect of volumetric properties of the mix (VMA and VFB) in Hirsch and Alkhateeb models were marginal.

The overall performance of the studied models suggested that they are not robust enough for Australian asphalt mixes. Current research at SUT has focused on the development of improved dynamic modulus predictive models to enable more accurate predictions of dynamic modulus for Australian asphalt mixes.

8. Acknowledgement

This project was sponsored by Australian Asphalt Pavement Association (AAPA) and Fulton Hogan Pty Ltd was commissioned to conduct the laboratory testing. The Authors would like to thank Mr. Tom Gabrawy from Fulton Hogan R&D Asphalt laboratory, Mr. Alireza Mohammadinia from Swinburne University of Technology and Mr. Mohammadjavad Abdi from University of Birmingham for providing technical support and consultation. The presented results and discussions in this paper are the authors' opinion and do not necessarily disclose the sponsoring organizations' point of view.

References

1. Witczak, M.W., et al., *Simple performance test for superpave mix design - NCHRP report 465*. 2002, Transportation Research Board- National Research Council.
2. ARA Inc. ERES Consultant Devision, *Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures - National Cooperative Highway Research Program (NCHRP1-37A Final Report)*. 2004, Transportation Research Board: Washington, DC.
3. Singh, D., M. Zaman, and S. Commuri, *Evaluation of Predictive Models for Estimating Dynamic Modulus of Hot-Mix Asphalt in Oklahoma.* Transportation Research Record: Journal of the Transportation Research Board, 2011. **2210**(-1): p. 57-72.
4. Obulareddy, S., *Fundamental characterization of Louisiana HMA mixtures for the 2002 mechanistic-empirical design guide*. 2006, Louisiana State University.
5. Birgisson, B., G. Sholar, and R. Roque, *Evaluation of a predicted dynamic modulus for Florida mixtures.* Transportation Research Record, 2005. **1929**(1929): p. 200-207.
6. Azari, H., et al., *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: Journal of the Transportation Research Board, 2007. **1998**(-1): p. 1-9.
7. Rickards, I. and P. Armstrong. *Long term full depth asphalt pavement performance in Australia*. in *24th ARRB Conference*. 2010. Melbourne, Australia.
8. Qiu, J., et al., *Evaluating laboratory compaction of asphalt mixtures using the shear box compactor.* Journal of Testing and Evaluation, 2012. **40**(5).
9. Molenaar, A., M. van de Ven, and I. Rickards, *Laboratory compaction of asphalt samples using the shear box compactor*, in *13th AAPA International Flexible Pavements Conference*. 2009: Surfers Paradise, Queensland, Australia. p. 9p (day 3, session 6).
10. Christensen, D.W. and D.A. Anderson, *Interpretation of Dynamic Mechanical Test Data for Paving Grade Asphalt Cements (With Discussion).* Journal of the Association of Asphalt Paving Technologists, 1992. **61**.
11. Marasteanu, M. and D. Anderson. *Improved model for bitumen rheological characterization*. in *Eurobitume Workshop on Performance Related Properties for Bituminous Binders*. 1999. European Bitumen Association Brussels, Belgium.
12. Elseifi, M.A., et al., *Viscoelastic Modeling of Straight Run and Modified Binders Using the Matching Function Approach.* International Journal of Pavement Engineering, 2002. **3**(1): p. 53-61.
13. Li, J., A. Zofka, and I. Yut, *Evaluation of dynamic modulus of typical asphalt mixtures in Northeast US Region.* Road Materials and Pavement Design, 2012. **13**(2): p. 249-265.
14. Andrei, D., M. Witczak, and M. Mirza, *Development of a Revised Predictive Model for the Dynamic (Complex) Modulus of Asphalt Mixtures*. 1999, NCHRP.
15. Garcia, G. and M. Thompson, *HMA Dynamic Modulus Predictive Models–a Review*. 2007, Illinois Center for Transportation. p. 61801.
16. Robbins, M. and D. Timm, *Evaluation of Dynamic Modulus Predictive Equations for Southeastern United States Asphalt Mixtures.* Transportation Research Record: Journal of the Transportation Research Board, 2011. **2210**(-1): p. 122-129.
17. Bari, J. and M.W. Witczak, *Development of a New Revised Version of the Witczak E* Predictive Model for Hot Mix Asphalt Mixtures (With Discussion).* Journal of the Association of Asphalt Paving Technologists, 2006. **75**.
18. Christensen Jr, D.W., T. Pellinen, and R.F. Bonaquist, *Hirsch model for estimating the modulus of asphalt concrete.* Journal of the Association of Asphalt Paving Technologists, 2003. **72**: p. 97-121.
19. Al-Khateeb, G., et al., *A new simplistic model for dynamic modulus predictions of asphalt paving mixtures.* Journal of the Association of Asphalt Paving Technologists, 2006. **75**.

20. Ceylan, H., et al., *Accuracy of Predictive Models for Dynamic Modulus of Hot-Mix Asphalt.* Journal of Materials in Civil Engineering, 2009. **21**(6): p. 286-293.
21. Kim, Y.R., et al., *LTPP computed parameter: dynamic modulus*. 2011, Federal Highway Administration (FHWA).
22. Dongre, R., et al., *Field Evaluation of Witczak and Hirsch Models for Predicting Dynamic Modulus of Hot-Mix Asphalt (With Discussion).* Journal of the Association of Asphalt Paving Technologists, 2005. **74**.
23. Francis, G. and F. Glenda, *Multiple regression*. HMA279 Design and Measurement 3, ed. C. SUT. Lilydale and S. SUT. School of Mathematical. 2003, Hawthorn, Vic.]: Hawthorn, Vic. : Swinburne University of Technology, Lilydale.
24. Abu Abdo, A.M., *Sensitivity analysis of a new dynamic modulus (| E*|) model for asphalt mixtures.* Road Materials and Pavement Design, 2012. **13**(3): p. 549-555.