

# Pre-trained deep learning for hot-mix asphalt dynamic modulus prediction with laboratory effort reduction

Ghada S. Moussa<sup>a,\*</sup>, Mahmoud Owais<sup>b</sup>

<sup>a</sup> Civil Engineering Department, Assiut University, Assiut, Egypt

<sup>b</sup> Civil Engineering Department, Assiut University, Assiut, Egypt, on leave to Majmaah University, KSA

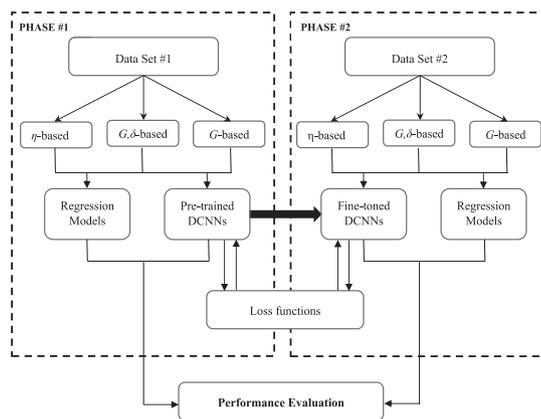


## HIGHLIGHTS

- A deep convolution neural networks technique is newly adapted for  $E^*$  prediction.
- The first pre-trained DL model for pavement materials is constructed.
- Laboratory effort reduction attained via transfer learning of pre-trained DL model.
- The robustness of the methodology is statistically verified.

## GRAPHICAL ABSTRACT

The DCNNs framework for the HMA dynamic modulus prediction



The DCNNs framework for the HMA dynamic modulus prediction

## ARTICLE INFO

### Article history:

Received 27 April 2020

Received in revised form 19 June 2020

Accepted 8 July 2020

### Keywords:

Pavements  
Dynamic modulus  
Hot asphalt mixture  
Deep learning  
Predictive models

## ABSTRACT

Evaluating the hot mix asphalt (HMA) expected performance is one of the significant aspects of highways research. Dynamic modulus ( $E^*$ ) presents itself as a fundamental mechanistic property that is one of the primary inputs for mechanistic-empirical models for pavements design. Unfortunately,  $E^*$  testing is an expensive and complicated task that requires advanced testing equipment. Moreover, a significant source of difficulty in  $E^*$  modeling is that many of the factors of variation in the HMA mixture components and testing conditions significantly influence the predicted values. For each laboratory practice, a vast number of mixes are required to estimate the  $E^*$  accurately. This study aims to extend the knowledge/practice of other laboratories to a target one in order to reduce the laboratory effort required for  $E^*$  determination while attaining accurate  $E^*$  prediction. Therefore, the transfer learning solution using deep learning (DL) technology is adopted for the problem. By transfer learning, instead of starting the learning process from scratch, previous learnings that have been gained when solving a similar problem is used. A deep convolution neural networks (DCNNs) technique, which incorporates a stack of six convolution blocks, is newly adapted for that purpose. Pre-trained DCNNs are constructed using a large data set that comes from different sources to constitute cumulative learning. The constructed pre-trained DCNNs aim to dramatically reduce the effort elsewhere (target lab) when it comes to the  $E^*$  prediction problem. Then, a laboratory effort reduction justification is investigated through fine tuning the constructed pre-trained DCNNs using

\* Corresponding author.

E-mail addresses: [ghada.moussa@aun.edu.eg](mailto:ghada.moussa@aun.edu.eg) (G.S. Moussa), [maowais@aun.edu.eg](mailto:maowais@aun.edu.eg) (M. Owais).

a limited amount of the target lab data. The performance of the proposed DCNNs is evaluated using representative statistical performance indicators and compared with well-known predictive models (e.g.,  $\eta$ -based Witczak 1-37A,  $G, \delta$ -based Witczak 1-40D and  $G$ -based Hirsch models). The proposed methodology proves itself as an excellent tool for the  $E^*$  prediction compared with the other models. Moreover, it could preserve its accurate performance with less data input using the transferred learning from the previous phase of the solution.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

One of the fundamental parameters of the hot mix asphalt (HMA), which is widely recognized nowadays, is the dynamic modulus ( $E^*$ ). It is defined as the strain response of asphalt mixtures as a function of the temperature and the loading rate. It represents the HMA's stiffness under a continuous loading [1]. It is measured by the resistance to compressive deformation when subjected to cyclic compressive loading [2]. In a frequency domain, the stress-strain relationship of continuous sinusoidal loading results in a complex number which its absolute value is  $E^*$ .

The dynamic modulus is a crucial material input, not only for Mechanistic-Empirical (ME) pavement design method [3] but also for pavement performance evaluation models associated with fatigue cracking and rutting [4]. Moreover, the dynamic modulus is considered a prospective quality assurance/quality control (QA/QC) parameter for flexible pavements in the field [5].

The Asphalt mixture is not a homogenous material, where its characteristics (e.g., stiffness, elasticity, etc.) are governed by the internal components properties such as aggregate gradation and strength, binder viscosity, and mix proportions besides some external factors such as climate [6]. For instance, the asphalt mixture non-linear behavior is prevalent at higher temperatures and low frequencies of loading, on the contrary, the mixture's binder itself behaves non-linearly at lower temperatures and higher frequencies [7]. Also, the asphalt mixture response at temperatures above 40 °C depends highly on the structure of the aggregates while this dependency disappears at low temperatures [8]. Fortunately,  $E^*$  could be considered a representative of these factors besides a predictor for other mixture potential characteristics [9].

$E^*$  is sensitive to the mixture component since the change in one of the mixtures' proportions might have a significant effect on the HMA dynamic modulus [10]. For instance, under normal conditions, the modulus for the asphalt specimen is predominantly controlled by the aggregate proportions and their internal friction [11,12]. Furthermore, it is confirmed that increases air voids result in lower  $E^*$  besides shorter fatigue life [13]. Moreover, the binder grade (PG) affects the values of  $E^*$  at temperatures higher than 30 °C. Modifiers, like lime, increase the HMA dynamic modulus values [14]. Even in reclaimed asphalt pavements, the  $E^*$  is sensitive to the percentage of fines [15].

In addition to the high cost associated with  $E^*$  laboratory determination, its procedures are complex and time-consuming. According to AASHTO TP62-07, the standard process of tests for measuring the dynamic modulus (i.e., determining  $E^*$  for one specimen mixture) is a burdensome task. It stipulates a testing machine that is able to apply a controlled haversine (sinusoidal) compressive load. It also should have load frequencies that range from 0.1 to 25 Hz and cumulative stress up to 400 psi (2800 kpa). The laboratory environment is also required to adapt the specimen to different temperatures that range from -10 to 60 °C. All this must be accompanied by the accuracy of applying and measuring [16,17].

Equipment, which is called asphalt mixture performance tester (AMPT), is developed as a part of [18] to evaluate the performance

of HMA mixtures such as; rutting resistance and fatigue [19,20]. It helps in measuring  $E^*$  over a range of loading frequencies and temperatures. As it is the assumed behavior of the HMA to follow the superposition of the time-temperature, a master curve is constructed to describe the visco-elastic action of the HMA regarding both loading frequency and temperature [21,22]. However, most of the existing studies to date have not focused on the testing method when different laboratories use mixture proportions or investigated how the variability in the various laboratories, would affect the measured  $E^*$  of the HMA for the same material [23].

To this end, the  $E^*$  modeling process is a challenging task associated with a high degree of uncertainty, complexity, and cost. Thus, developing a global predictive model that requires the minimum amount of data becomes a necessity. This study attempts to extend the prior information and models obtained previously in the dynamic modulus prediction using the deep learning architecture. The aim behind this is to reduce the required number of testing samples for  $E^*$  prediction while maintaining a suitable accuracy.

The structure of the article is as follows. Section 2 presents the literature review. Section 3 provides regression-based  $E^*$  predictive models. Section 4 defines input data sets. Section 5 presents the proposed methodology. In Section 6, results and discussion are presented. Conclusions are presented in section 7.

## 2. State of the art

The dynamic modulus of the HMA has received considerable attention in the literature due to its importance in addition to the difficulty of deriving its predictive models' accuracy. Numerous studies have been made, over the years, to develop dynamic modulus predictive models based on HMA's material proportions and properties [24,25]. Some of these studies resulted in widely used and well-known  $E^*$  predictive regression models, besides recently, more advanced techniques other than regression, which are based on complex machine learning techniques [12].

Regarding  $E^*$  predictive regression models, the most widely used models are Witczak 1-37A [26], Witczak 1-40D [27], and Hirsch [19] models, which are regression-based models built using several  $E^*$  laboratory databases. Witczak's  $E^*$  predictive models are adapted in the ME pavement design software for pavement design and analysis when using HMA level 2 or 3 inputs [3,27]. The ME pavement design software performs at three distinct levels (levels 1, 2, and 3). For level 1, laboratory measurements of dynamic modulus are required, while for levels 2 and 3, dynamic modulus values are estimated using Witczak predictive models. However, level 2 uses measured values of binder stiffness or viscosity, but level 3 uses typical values from similar mixtures based on the designer prior experience.

Numerous attempts have evaluated the performance of these regression-based  $E^*$  predictive models in terms of their variance and bias using independent data sets. Based on their evaluation, many researchers reported non-consistent performances of these predictive models with a tendency to vary according to the type of the HMA mixture and its volumetric characteristics. In [28,29],

Witczak NCHRP 1-40D and Witczak NCHRP 1-37A models are evaluated. It is concluded that the  $E^*$  for the NCHRP 1-37A in the MEPDG level-3 inputs produced the least biased and the most accurate estimates. In contrast, NCHRP 1-40D resulted in the highest biased estimates with the lowest prediction accuracy for the same phase of design.

Reviewing different lab practices, in Australian [30], Hirsch and Witczak NCHRP 1-37A under-predicted the dynamic modulus, whereas Witczak 1-40D over-predicted it. Regarding the Kingdom of Saudi Arabia (KSA) practice, in [31], the Witczak models, along with the Hirsch model, are investigated based on 25 different mixtures. All models in level 3 show the highest estimation accuracy with slight superiority to the Witczak1-37A model. On the contrary, the Hirsch model shows some bias higher than Witczak models. In Pakistan [32], the NCHRP 1-37A model for different samples is proved to underestimate the  $E^*$  values of considered mixtures continuously. The same conclusion is obtained in [33,34], where the three models are evaluated against the prediction accuracy for 15 asphalt mixtures, which are widely accepted in Greece. It could be generally concluded that these regression-based  $E^*$  predictive models exhibit noticeable bias at low and/or high-temperature spectrum [35,36].

To accommodate this inconsistency, other studies have attempted to calibrate these well-known models for their own or develop new ones. In [37], in the state of Michigan, a newly developed analytical model is calibrated from NCHRP 1-40D model that shows improved prediction accuracy. Whereas in [38], the reliability of  $E^*$  prediction is improved at high temperatures introducing parameters that represent the gradation distribution of aggregates. A linear model is produced and calibrated with 24 different mixtures and validated with 11 mixtures.

Recently, coupled with the rapid development in computer programming, machine learning (ML) approaches have attained great attention in asphalt mixtures performance prediction [39–42]. ML approaches are systems built to imitate the human brain and perception by learning automatically from previous experiences and have been applied to many civil engineering problems [39,43].

Artificial Neural Networks (ANNs) technique is one of the oldest and the most applied ML approaches developed by McCulloch and Pitts [44]. ANNs have the ability to learn and recognize the trends of data patterns without knowing the form of predictive relationships [45,46]. In recent years, ANNs have been used to predict the  $E^*$  for the HMA mixtures with high accuracy and less bias compared to the well-known regression-based predictive models [36,47–49]. In these consecutive studies [35,36,47,50], the Witczak database of extensive data records from various HMA mixtures are used within some proposed ANNs structures to suggest the best structure for these data finally.

In [51], the three well-known regression models are compared with the ANNs models for laboratory data from KSA with the same set of parameters used in the regression models where the ANNs overcome these popular models' accuracy. Similarly, in [52], three different ANNs structures are developed for predicting the  $E^*$  for the HMA using data from an online data repository found in [53]. The three predictive structures are trained with the same group of parameters used in Witczak and Hirsch models. The model with a three-layer back-propagation feedforward and a transfer function of sigmoidal type is found to achieve the highest accuracy. However, there are some shortcomings associated with using the ANNs, such as; poor generalization performance, arriving at local minimum, over-fitting problems, and trouble in the prediction in case inputs are outside the training database range.

Support Vector Machine (SVM) is another ML technique that is derived from the statistical learning theory, developed by Vapnik and Chervonenkis in 1964 [54]. It has been reported that the SVM has a better generalization performance than conventional

ANN [55]. However, the ANN-based  $E^*$  predictive model attained higher prediction accuracy than the SVM-based one, and both models achieved better accuracy than the Witczak 1-37A model [56].

Genetic Programming (GP) is a part of a broader family of ML techniques that uses a population of individuals, individuals are selected according to fitness, and then via specific operators, the genetic variation is achieved. The GP is first introduced by Koza in 1992 [57]. The GEP is a linear variant of GP, where individuals are represented as linear strings [58]. In [59], two based ANN and GEP models are developed for the  $E^*$  prediction of asphalt mixtures containing recycled shingles, and then their performance is compared with the Witczak 1-37A model. Compared with other models, the GEP is proved to avoid overtraining, achieve better generalization performance, and create a transparent-structured representation [60].

Naturally, examining new developments in the ML techniques for the  $E^*$  prediction becomes an open contribution to the problem literature. As a final example, the random forests algorithm, which is first introduced by Ho in 1995 [61] and then extended by Breiman in 2001 [62], is another ML technique used for classification and regression problems [61]. It creates a forest of an ensemble of decision trees by utilizing a different bootstrap sample of data and merges them to get an accurate prediction. It is considered more user-friendly compared to ANNs and SVMs with less sensitivity to parameters change due to the fact of owning only two parameters (i.e., number of trees in the forest and number of variables in each random node subset). In 2020, Daneshvar and Behnood [63] developed an  $E^*$  predictive model based on the random forests algorithm using a comprehensive database. Based on their results, their model improved prediction accuracy compared with the Witczak NCHRP 1-40D model.

In short, the covered literature shows confusing results regarding the prediction accuracy and bias of the well-known  $E^*$  predictive models. On the other hand, laboratory determination of the  $E^*$  is not only tedious and time-consuming but also demanding of costly, advanced equipment and skills, which are not easily accessible. Moreover, the different machine learning predictive models requires a lot of data points to achieve the aimed prediction accuracy. Furthermore, a major source of difficulty in  $E^*$  prediction is that many of the factors of variation in mixture components and testing conditions significantly influence the  $E^*$  values. In other words, most  $E^*$  predictive models are built using specific region databases, which may not be appropriate for other regions' databases.

This study contributes to the literature in two points; 1)- Adopting a new ML method to the problem. 2)- Reducing the laboratory effort needed for the  $E^*$  determination (training data) while attaining a plausible accuracy of prediction. This is done through transferring the accumulated experience in  $E^*$  determination (i.e., historical data) with collaborating deep convolution learning technology. Where, pioneer pre-trained deep convolution neural networks (DCNNs) are constructed using a comprehensive  $E^*$  database from different regions all around the world. Subsequently, laboratory effort reduction justification is investigated through fine tuning the constructed pre-trained DCNNs using a limited amount of new laboratory data. Moreover, validation of developed DCNNs models accuracy and bias is considered with a comparison with well-known  $E^*$  predictive regression models; Witczak 1-37A [26], Witczak 1-40D [27] and modified Hirsch [19], using the same input parameters.

### 3. Regression-based $E^*$ predictive models

In this section, the widely used regression-based  $E^*$  predictive models are enumerated to highlight the most commonly accepted

controlling parameters for the dynamic modulus prediction. As stated in the literature, these models are Witczak 1-37A [26], Witczak 1-40D [27] and modified Hirsch [19] and they are illustrated in the following subsections

### 3.1. Witczak 1-37A model

The Witczak 1-37A model is a revision of the original Witczak  $E^*$  predictive regression equation based on 2,750 data points from 205 mixtures, known as the 1999  $\eta$ -based National Cooperative Highway Research Program (NCHRP) 1-37A Witczak model. The NCHRP 1-37A Witczak model is a non-linear regression equation that incorporates mixture volumetric properties, aggregate gradation, and binder viscosity as follows:

$$\log_{10}|E^*| = -1.249937 + 0.02923\rho_{200} - 0.001767(\rho_{200})^2 - 0.002841\rho_4 - 0.058097V_a - 0.802208\left(\frac{V_{beff}}{V_{beff} + V_a}\right) + \frac{3.871977 - 0.0021\rho_4 + 0.003958\rho_{38} - 0.000017(\rho_{38})^2}{1 + e^{(-0.603313 - 0.313351\log f - 0.393532\log \eta)}} + \frac{0.00547\rho_{34}}{1 + e^{(-0.603313 - 0.313351\log f - 0.393532\log \eta)}} \quad (1)$$

where  $E^*$  = HMA dynamic modulus [ $6.89 \times 10^3$  MPa ( $10^5$  psi)];  $\rho_{200}$  = Percentage of aggregate passing #200 sieve;  $\rho_4$  = Percentage of aggregate retained in #4 sieve;  $\rho_{38}$  = Percentage of aggregate retained in 3/8-inch (9.56-mm) sieve;  $\rho_{34}$  = Percentage of aggregate retained in 3/4-inch (19.01-mm) sieve;  $V_a$  = Percentage of air voids (by volume of mix);  $V_{beff}$  = Percentage of effective asphalt content (by volume of mix);  $f$  = Loading frequency (Hz);  $\eta$  = Binder viscosity at temperature of interest ( $10^6$  Poise ( $10^5$  Pa.s)).

### 3.2. Witczak 1-40 model

In 2005, Bari and Witczak [64] reformulated the  $E^*$  predictive model to incorporate the Superpave binder shear modulus ( $G^*$ ) instead of the binder viscosity parameter ( $\eta$ ). The model is a non-linear regression equation based on 7,400 data points of 346 mixtures, known as 2007  $G^*$ ,  $\delta$ -based NCHRP 1-40D Witczak model. The NCHRP 1-40D Witczak model incorporates mixture volumetric properties, aggregate gradation, binder shear modulus, and binder phase angle as follows:

$$\log_{10}|E^*| = 0.02 + 0.758\left(|G_b^*|^{-0.0009}\right) \times \left( \begin{array}{l} 6.8232 - 0.03274\rho_{200} + 0.00431(\rho_{200})^2 + 0.0104\rho_4 \\ -0.00012(\rho_4)^2 + 0.00678\rho_{38} - 0.00016(\rho_{38})^2 - \\ 0.0796V_a - 1.1689\left(\frac{V_{beff}}{V_{beff} + V_a}\right) \end{array} \right) + \frac{1.437 + 0.03313V_a + 0.6926\left(\frac{V_{beff}}{V_{beff} + V_a}\right) + 0.00891\rho_{38}}{1 + e^{(-4.5868 - 0.8176\log |G_b^*| + 3.2738\log \delta)}} - \frac{0.00007(\rho_{38})^2 + 0.0081\rho_{34}}{1 + e^{(-4.5868 - 0.8176\log |G_b^*| + 3.2738\log \delta)}} \quad (2)$$

where  $E^*$  = HMA dynamic modulus [ $6.89 \times 10^{-2}$  MPa (psi)];  $|G_b^*|$  = dynamic shear modulus of asphalt binder (psi);  $\delta_b$  = binder phase angle associated with  $|G_b^*|$  (degrees); and the remaining parameters and their units are as in Eq. (1).

Since some of the mixtures in the database may not contain binders'  $|G_b^*|$  and  $\delta$  values, these values are estimated based on A-VTS values [28], using the following regression equation:

$$\eta = \left(\frac{G_b^*}{10}\right) \left(\frac{1}{\sin \delta}\right)^{4.8628} \quad (3)$$

$$\log \log \eta = A + VTS \log T_R \quad (4)$$

$$\log \log \eta_{f_s, T} = A' + VTS' \log T_R \quad (5)$$

$$A' = 0.9699 \times f_s^{-0.0527} \times A \quad (6)$$

$$VTS' = 0.9668 \times f_s^{-0.0575} \times VTS \quad (7)$$

$$\delta_b = 90 - 0.1785 \times \log(\eta_{f_s, T})^{2.3814} \times (f_s)^{(0.3507 + 0.0782VTS')} \quad (8)$$

$$|G_b^*| = 1.469 \times 10^{-9} \times \log(\eta_{f_s, T})^{12.0056} \times (f_s)^{(0.7418)} (\sin \delta_b)^{0.6806} \quad (9)$$

$$f_s = \frac{f_c}{2\pi} \quad (10)$$

where  $\eta$  = binder viscosity [cP (centipoise)];  $T_R$  = testing temperature (Rankine);  $f_s$  = loading frequency (as used in the complex shear modulus test of asphalt binders) (Hz); and  $f_c$  = loading frequency (as used in the complex dynamic modulus test of asphalt concrete mixtures) (Hz); A, VTS = regression intercept and slope of the viscosity-temperature relationship respectively;  $A'$ ,  $VTS'$  = adjusted A and VTS for loading frequency respectively;  $\eta_{f_s, T}$  = binder viscosity as a function of both loading frequency ( $f_s$ ) and temperature ( $T_R$ ) (cP);  $\delta_b$  = binder phase angle (degrees);  $|G_b^*|$  = binder shear modulus (psi).

### 3.3. Modified Hirsch model

The Hirsch model is based on the parallel model of mixtures law, which is built using 206 data points from 18 different HMA mixtures containing eight different binders. This model incorporates only the binder shear modulus and volumetric properties of the mix, as presented in the following equations:

$$E^* = P_c \left[ 4,200,000 \left( 1 - \frac{VMA}{100} \right) + 3|G^*| \left( \frac{VFA \times VMA}{10,000} \right) \right] + (1 - P_c) \times \left[ \frac{(1 - \frac{VMA}{100})}{4,200,000} + \frac{VMA}{3 \times VFA \times |G^*|} \right]^{-1} \quad (11)$$

$$P_c = \frac{\left( 20 + \frac{VFA \times 3|G^*|}{VMA} \right)^{0.58}}{650 + \left( \frac{VFA \times 3|G^*|}{VMA} \right)^{0.58}} \quad (12)$$

where  $E^*$  = dynamic modulus of the mixture [ $6.89 \times 10^{-2}$  MPa (psi)];  $P_c$  = contact factor;  $|G^*|$  =  $|G_b^*|$  = shear modulus of the binder [ $6.89 \times 10^{-2}$  MPa (psi)]; VMA = voids in the mineral aggregates (%); and VFA = voids in mineral aggregates filled with the binder (%).

## 4. Data sets

When dealing with ML techniques, the data set is a crucial element in the technical implementation. Variability and consistency in the offered data manage the technique achieving better performances in the training and prediction stages. The laboratory test data implemented in this work comprises two sets; data set #1, and data set #2. While the data set #1 is used as the prior information found before attempting to predict  $E^*$  in the target lab, data set #2 is used to justify the potential effort reduction in the needed data for  $E^*$  prediction. We should note the main objective of this work is to reduce the amount of data needed from any data sets

(e.g., data set #2) to predict  $E^*$  resorting to previous/cumulative experience (e.g., data set #1).

Table 1 presents descriptive statistics of the implemented data sets. The data set #1 is a combination of three valuable databases from different regions all around the world; Arizona State University (ASU) in the USA, Australia, and the Kingdom of Saudi Arabia (KSA) databases.

Arizona State University (ASU) database contains a comprehensive set of test data from various studies. The majority of this database comprises laboratory tests that were conducted as part of TRB's National Cooperative Highway Research Program (NCHRP) studies at the University of Maryland and then continued at Arizona State University (ASU) [27,65]. Also, the ASU database contains tests results from several material-characterization studies conducted for the Arizona Department of Transportation (ADOT). The ASU database that is used in this study contains a total of 2490  $E^*$  test results from 83 mixtures with different gradations, volumetric characteristics, and binder types. The range of temperatures used for  $E^*$  testing varies from  $-10\text{ }^\circ\text{C}$  to  $54.4\text{ }^\circ\text{C}$ , with stress ranged from 0.1 Hz to 25 Hz. A more detailed description of the ASU database can be found in [20].

The Australia database contains 1002  $E^*$  measurements of 22 different Australian asphalt mixes produced by Australia's leading asphalt producers [66,67]. Investigated mixtures cover most commonly used asphalt materials in Australia's major projects. Out of the 22 combinations, 12 blends have a Nominal Maximum Aggregate Size (NMAS) of 14 mm and 10 mixes with NMAS of 20 mm. Mixtures have a Rap content ranges from 0 to 30%.  $E^*$  tests were carried out on the asphalt mixtures at four temperatures (5, 20, and 35, 50  $^\circ\text{C}$ ) and five different frequencies (0.5, 1, 5, 10, and 25 Hz).

KSA database contains comprehensive  $E^*$  data (2568 measurements) for 25 Superpave mixtures that are widely used in the KSA [31,68]. The KSA mixtures have different aggregate gradations and binder performance grades that cover the different KSA's climatic regions. Binders used in these mixtures are modified either with crumb rubber or different polymers.  $E^*$  tests were carried out on the asphalt mixes at four temperatures ( $-10, 4.4, 21.1, \text{ and } 54.4\text{ }^\circ\text{C}$ ) and six loading frequencies (0.1, 0.5, 1, 5, 10, and 25 Hz).

On the other hand, the data set #2 is a database from a study conducted in [69]. In which, 13 mixtures, from four different demonstration projects in Indiana, Iowa, Minnesota, and Missouri, tested at Iowa State University at three different temperatures (4, 21, and 37  $^\circ\text{C}$ ) and nine different frequencies (0.1, 0.2, 0.5, 1, 2, 5, 10, 20, and 25 Hz). The database contains 1,701 experimental data points of  $E^*$ .

### 5. Methodology

One of the fastest-growing methods in the field of ML, which has gained popularity in the last years due to its outstanding results in numerous engineering application domains, is deep learning architecture (DL) [70,71]. DL architecture is an artificial neural network that contains multiple layers (deep networks) between input and output layers [72]. Multiple layers allow the architecture to progressively extract high-level features from the raw input data [73,74]. Therefore, DL architecture has the potential to learn complicated features and functions of input data than an artificial neural network does [75]. One of the most successful DL architectures is the deep convolutional neural networks (DCNNs) [76]. The name "convolutional" implies that the network uses mathematical operation called convolution, in which sliding convolutional filters (kernels) are applied to the input [77]. DCNNs have gained considerable attention due to its capability for automatic feature extraction, hierarchical learning, multi-tasking, and weight sharing [78,79]. However, very limited research has adapted the DCNNs architecture in the field of pavement engineering and infrastructural materials, and no study, to the best of our knowledge, has used it in the  $E^*$  prediction of the HMA. In the proposed methodology, two consequent phases are constructed; first phase: pre-trained DCNNs, and second phase: fine-tuned DCNNs.

In the pre-trained DCNNs phase, three DCNNs are built, trained, and tested using a global  $E^*$  database from different regions all around the world (data set #1). In the fine-toned DCNNs phase, the three pre-trained DCNNs are fine-toned with a limited portion of a new laboratory database (data set #2), then tested with the rests of it. The main reason for the fine-toned DCNNs stage is the

**Table 1**  
Summary of the implemented databases.

Data Set	Variable	Range	Minimum	Maximum	Mean	Standard Deviation	
Data Set #1	Aggregate gradation	$\rho_{3/4}$ (%)	0	26.10	4.10	6.65	
		$\rho_{3/8}$ (%)	5.0	58.2	26.6	11.7	
		$\rho_4$ (%)	34.0	74.0	49.9	8.0	
		$\rho_{200}$ (%)	1.8	8.5	4.9	1.3	
	Mixture Volumetric	$V_a$ (%)	0.7	18.1	6.3	2.1	
		$V_{beff}$ (%)	5.5	25.1	9.8	2.8	
	Binder	$\eta^I (x10^6 \text{ poise})$	1.70E-05	2.70E + 04	5.64E + 03	1.03E + 04	
		$ G_b^* (psi)^II$	1.45E-02	4.08E + 04	2.21E + 03	4.58E + 03	
		$\delta^{III}$ (degree)	0.14	87.7	50.6	23.3	
		$f_c$ (Hz)	0.1	25	7.1	8.8	
	Data Set #2	Loading Frequency	$ E^* $ (psi)	9.43E + 03	8.64E + 06	1.42E + 06	1.41E + 06
			Dynamic modulus				
Aggregate gradation		$\rho_{3/4}$ (%)	0.0	0.8	0.1	0.2	
		$\rho_{3/8}$ (%)	3.4	27.0	13.4	5.8	
		$\rho_4$ (%)	24.3	61.0	38.3	10.6	
		$\rho_{200}$ (%)	0.0	1.5	0.9	0.4	
Mixture Volumetric		$V_a$ (%)	6.5	7.3	6.9	0.2	
		$V_{beff}$ (%)	9.6	11.5	10.5	0.5	
Binder		$\eta^I (x10^6 \text{ poise})$	4.46E-01	2.70E + 04	6.61E + 03	1.07E + 04	
		$ G_b^* (psi)^II$	1.37	1.41E + 04	2.22E + 03	3.49E + 03	
		$\delta^{III}$ (degree)	11.4	69.7	43.9	17.2	
		$f_c$ (Hz)	0.1	25.0	7.1	8.8	
Loading Frequency	Dynamic modulus	$ E^* $ (psi)	1.75E + 04	2.80E + 06	9.30E + 05	7.37E + 05	

I: Variable only used in  $\eta$ -based Witczak 1-37A and  $\eta$ -based DCNNs

II: Variables only used in G &  $\delta$ -based Witczak 1-40D, G, $\delta$ -based DCNNs, G-based Hirsch, and G-based DCNNs

III: Variables only used in G &  $\delta$ -based Witczak 1-40D, G, $\delta$ -based DCNNs

laboratory effort reduction justification, i.e., only a small amount of laboratory data is needed for valuable  $E^*$  prediction.

The input parameters for the constructed DCNNs are the same as those used for the widely used  $E^*$  predictive regression models; section 3, at binder input Level 2 according to the ME pavement design guide. Table 2 illustrates the data variables used for regression and DCNNs  $E^*$  predictive models.

For the two phases (pre-trained DCNNs, and fine-toned DCNNs), inputs and outputs of the two data sets (data set #1, and data set #2) are normalized to values between 0 and 1, and then each data set is randomly divided into two groups: training and testing. In the pre-trained DCNNs phase, out of the 6,060 data points in the data set #1, 80% are used for training, and the remaining 20% is used for testing. In the fine-toned CNNs phase, out of the 1,701

data points in the data set #2, only 20% is used for training, and the remaining 80% is used for testing. The research methodology flowchart is summarized in Fig. 1.

5.1. Deep convolution neural networks architecture

The architecture of the proposed DCNNs (pre-trained DCNNs and fine-toned DCNNs) consists of an input layer followed by a stack of six convolution blocks, then a fully connected layer and ending with an output layer, as illustrated in Fig. 2. Each convolution block comprises of three consequent layers: a convolution layer, a batch normalization layer, and a ReLU activation layer, as shown in Fig. 3. The description of the DCNNs components is as follows:

**Table 2**  
Variables used in  $E^*$  predictive regression and DCNNs models.

Variables			$E^*$ predictive Models					
			$\eta$ -based		$G, \delta$ -based		$G$ -based	
			Witczak 1-37A	DCNNs	Witczak 1-40D	DCNNs	Hirsch	DCNNs
Inputs	Aggregate	$\rho_{3/4}$ (%)	X	X	X	X	—	—
		$\rho_{3/8}$ (%)	X	X	X	X	—	—
		$\rho_4$ (%)	X	X	X	X	—	—
		$\rho_{200}$ (%)	X	X	X	X	—	—
	Binder	$\eta(cP)$	X	X	—	—	—	—
		$G_b^*$	—	—	X	X	X	X
		$\delta$ (degree)	—	—	X	X	—	—
	Mix	$V_a$ (%)	X	X	X	X	—	—
		$V_{beff}$	X	X	X	X	—	—
		VMA (%)	—	—	—	—	X	X
VFA (%)		—	—	—	—	X	X	
Temperature	$T$ ( $^{\circ}C$ )	—	—	—	—	—	—	
	Stress	$f_c$ (Hz)	X	X	—	—	—	
Output		$E^*$ (psi)	X	X	X	X	X	

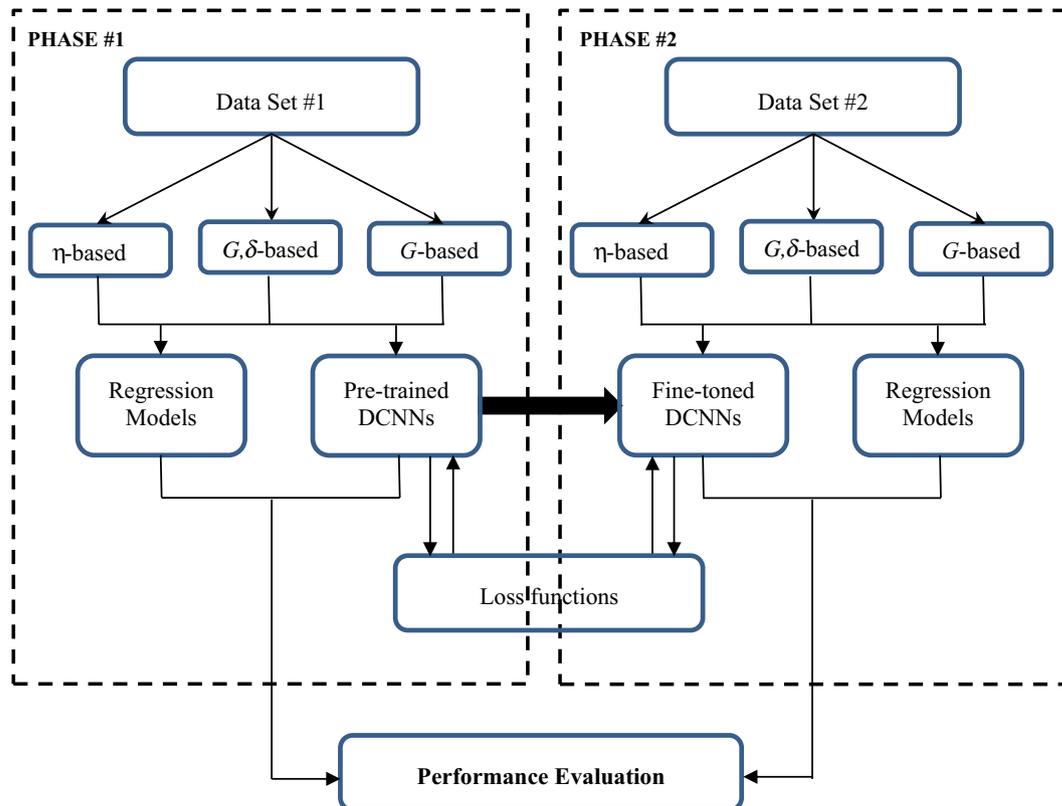


Fig. 1. The research methodology flowchart.

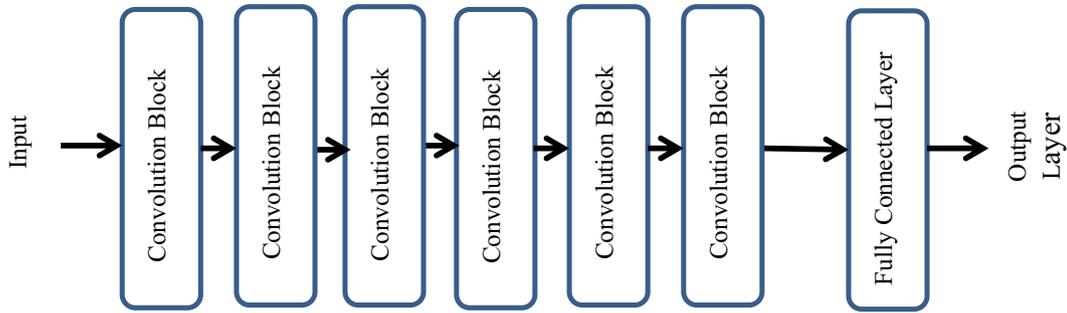


Fig. 2. Proposed Deep Convolutional Neural Networks (DCNNs) architecture.

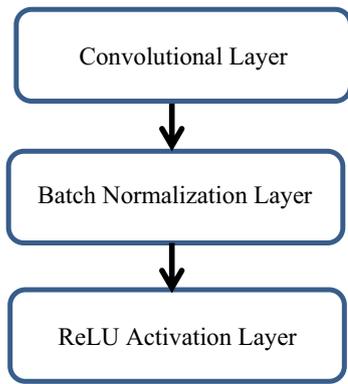


Fig. 3. Main Convolutional Block of the proposed DCNNs architecture.

- *First, the input layer:* it feeds the array of the input data to the network. It normalizes the input data by subtracting the data set's mean from each data input.
- *Second, the convolution blocks:* each convolution block includes a convolution layer, a batch normalization layer, and the ReLU activation layer.
- *Third, the fully connected layer:* it is mostly used at the end of any DCNNs just before the output layer for prediction purposes.
- *Finally, the output layer:* is the last layer in the DCNNs that follows the fully connected layer.

The convolutional layer is the core building layer of the DCNNs. Its primary purpose is to extract distinct features from the input data. The convolutional layer consists of the sequence of convolutional filters (learnable kernels) that extract local features from the input data. In which, each filter (kernel) calculates a feature map by computing the dot product of the weights and the input and then adding a bias term [80].

One of the main challenges for all deep learning networks is the overfitting, which occurs when there is a large gap between the training error and the testing error. One of the methods to avoid overfitting is applying regularization during training, which can be achieved via normalization [81]. A batch normalization layer is implemented in our DCNNs, between the convolutional layer and the ReLU activation layer. The batch normalization layer is used not only for regularization purpose but also it allows each segment of the network to learn independently from other layers. Moreover, the batch normalization layer speeds up the training process and reduces the sensitivity to the DCNNs initialization [82]. The batch normalization layer normalizes each input channel (from the input layer) across a mini-batch in two steps.

In the first step, the batch normalization inputs ( $x_i$ ) are normalized by subtracting the mini-batch mean ( $\mu_B$ ) and dividing by the

mini-batch variance ( $\sigma_B^2$ ) over the mini-batch and over each input channel. In the second step, the batch normalization calculates the normalized activations as;

$$\hat{x}_i = \left( \frac{(x_i - \mu_B)}{\sqrt{(\sigma_B^2 + \epsilon)}} \right) \quad (13)$$

where;  $\epsilon$  is the property Epsilon that is used to improve numerical stability when the mini-batch variance is minimal. Then, the batch normalization layer shifts the input by a learnable offset ( $\beta$ ) and scales it by a learnable scale factor ( $\gamma$ ) as;

$$y_i = \gamma \hat{x}_i + \beta \quad (14)$$

where;  $\beta$  &  $\gamma$  are offset and scale properties, respectively, that are learnable parameters and updated during network training. After finishing the training process, the batch normalization layer calculates the mean and variance over the fully trained data and stores them as trained-mean and trained-variance properties. Afterward, when the DCNNs are used for predictions (using new data), the batch normalization layer uses the trained-mean and variance instead of the mini-batch mean and variance for normalizing the activations.

The Rectified Linear Unit (ReLU) is a non-linear activation layer, which works as a decision function ( $f(x) = \max(0, x)$ ), where any input value less than zero is set to zero. ReLU introduces the non-linearity into the DCNNs that helps in learning more complex patterns [83]. There are other non-linear activation functions such as; the saturating hyperbolic tangent ( $f(x) = \tanh(x)$ ,  $f(x) = |\tanh(x)|$ ) and the sigmoid function ( $\sigma(x) = (1 + e^{-x})^{-1}$ ). However, the ReLU is often preferred to other non-linear activation functions because it overcomes the vanishing gradient problem and trains the DCNNs much faster than other functions without significantly affecting the generalization accuracy [71].

For the fully connected layer, as the name implies, all neurons in the fully connected layer have connections to all the neurons in the previous layer. It takes the input from feature extraction stages and multiplies it by a weight matrix then adds a bias vector. The fully connected layer can combine all of the features (local information) learned by the previous layers to identify the more significant patterns for  $E^*$  prediction.

The output layer is a regression layer that computes the mean squared error and the loss function for the  $E^*$  prediction problem. The loss function is the half-mean-squared-error of the predicted  $E^*$  for each time step, not normalized by  $n$  as presented in the following equations:

$$MSE = \frac{\sum_{i=1}^n (E_{pi}^* - E_{mi}^*)^2}{n} \quad (15)$$

$$Loss = \frac{1}{S} \sum_{i=1}^S \sum_{j=1}^n (E_{pi}^* - E_{mi}^*)^2 \tag{16}$$

where;  $n$  is the number of values of  $E^*$ , the subscripts  $p$  &  $m$  indicate predicted, and measured values of  $E^*$ , respectively, and  $S$  is the sequence length.

5.2. Model performance evaluation

The evaluation process is considered as a foundation stone in any prediction in order to answer the following questions: i) Does the predictive model fit the data? ii) Which predictive model has the best performance? And iii) How similar are the predictive models?

In this study, the performance of the proposed and the commonly used  $E^*$  predictive models are evaluated using the following aspects: 1) Goodness-of-fit statistics, 2) Overall bias indicators, and 3) Regression error characterization curve.

A. The goodness-of-fit statistics

They are quantitative assessments of the predictive model's accuracy, i.e., how close the model's predicted values are to their corresponding measured values. They are performed using some statistical parameters; the coefficient of determination ( $R^2$ ), and the standard error of predicted values divided by the standard deviation of measured values ( $Se/Sy$ ). These goodness-of-fit statistics parameters are calculated for each predictive model, using the following equations [84]:

$$S_e = \sqrt{\left(\frac{\sum_{i=1}^n (E_{pi}^* - E_{mi}^*)^2}{n - p}\right)} \tag{17}$$

$$S_y = \sqrt{\left(\frac{\sum_{i=1}^n (E_{mi}^* - \bar{E}_m^*)^2}{n - 1}\right)} \tag{18}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (E_{pi}^* - E_{mi}^*)^2}{\sum_{i=1}^n (E_{mi}^* - \bar{E}_m^*)^2} \tag{19}$$

**Table 3**  
Statistical parameters evaluation Criteria, Pellinen 2002 [84].

Model Accuracy	$R^2$	$Se/Sy$
Excellent	$\geq 0.9$	$\leq 0.35$
Good	0.7–0.89	0.36–0.55
Fair	0.4–0.69	0.56–0.75
Poor	0.2–0.39	0.76–0.9
Very Poor	$\leq 0.19$	$\geq 0.9$

**Table 4**  
Goodness-of-fit statistics of the DCNNs models and regression models for phase #1& phase #2.

Predictive model		Phase #1		Model Accuracy*	Phase #2		Model Accuracy
		Goodness-of-fit statistics			Goodness-of-fit statistics		
		$R^2$	$Se/Sy$		$R^2$	$Se/Sy$	
Regression models	$\eta$ -based Witczak 1-37A	0.54	0.68	Fair	0.85	0.38	Good
	$G, \delta$ -based Witczak 1-40D	0.73	0.52	Good	0.53	0.69	Fair
	$G$ -based Hirsch	0.52	0.70	Fair	0.89	0.33	Good
DCNNs models	$\eta$ -based DCNN	0.95	0.21	Excellent	0.98	0.13	Excellent
	$G, \delta$ -based DCNN	0.96	0.19	Excellent	0.99	0.10	Excellent
	$G$ -based DCNN	0.80	0.42	Good	0.92	0.26	Excellent

\* Model accuracy according to Pellinen 2002 standard evaluation criteria

where;  $n$  is the number of values of  $E^*$ , the subscripts  $p$  &  $m$  indicate predicted and measured values of  $E^*$ , respectively, and  $\bar{E}_m^*$  is the mean of the measured values of  $E^*$ .

While  $R^2$  is a measure of the correlation between predicted and measured values, the  $Se/Sy$  is a direct indicator of uncertainty in the predictive model. Therefore,  $R^2$  and  $Se/Sy$  are considered as indicators of the predictive model's accuracy. Moreover, a higher  $R^2$  and lower both  $Se/Sy$  values are desired for a better predictive model's accuracy. In [84], Pellinen developed standard evaluation criteria for predictive models' efficiency, which are presented in Table 3. These criteria have been adapted in many research considering  $E^*$  prediction [48,59,65,69,85]. Therefore, they are implemented in this study.

B. Overall bias indicators

The goodness-of-fit statistics ( $R^2$  &  $Se/Sy$ ) alone do not conclusively define the model's prediction accuracy. Where the existence of overall model bias may significantly affect its prediction accuracy. Therefore, global bias indicators (slope, intercept, and average error) for each predictive model have to be considered. By plotting measured versus predicted  $E^*$  values along with the line of equality (LOE), the model's slope, intercept, and average error can be estimated by fitting an unconstrained linear trend line to the measured-predicted plot [86].

The reliable model would have matching points clustered along the LOE line, i.e., has a slope, intercept, and average error values close to 1, 0, and 0, respectively. The deviation of the slope from the unity indicates the dependence of the prediction errors to the measured values [6]. On the other hand, the non-zero intercept and average error suggest the model's over- or under-prediction [18,86].

C. Regression Error Characterization (REC) Curve.

REC curve, proposed in [87], is a graphical tool that can be utilized for visualizing predictive models' performance. Its drawn curve is a generalization of the Receiver Operating Characteristic (ROC) curve that is widely incorporated into classification problems. The REC curve is a two-dimensional graph, in which the x-axis represents the absolute deviation (error tolerance), and the y-axis represents the accuracy of the predictive model.

The REC curves are very informative and can provide a much more convincing presentation of the predictive models' performance than other statistics since they consider the whole error distribution of the models rather than just a single indicator of error. Also, they offer a specialized analysis and valuable information about the performances of comparative  $E^*$  predictive models. The predictive model is said to perform well if its REC curve climbs rapidly towards the upper left corner [88].

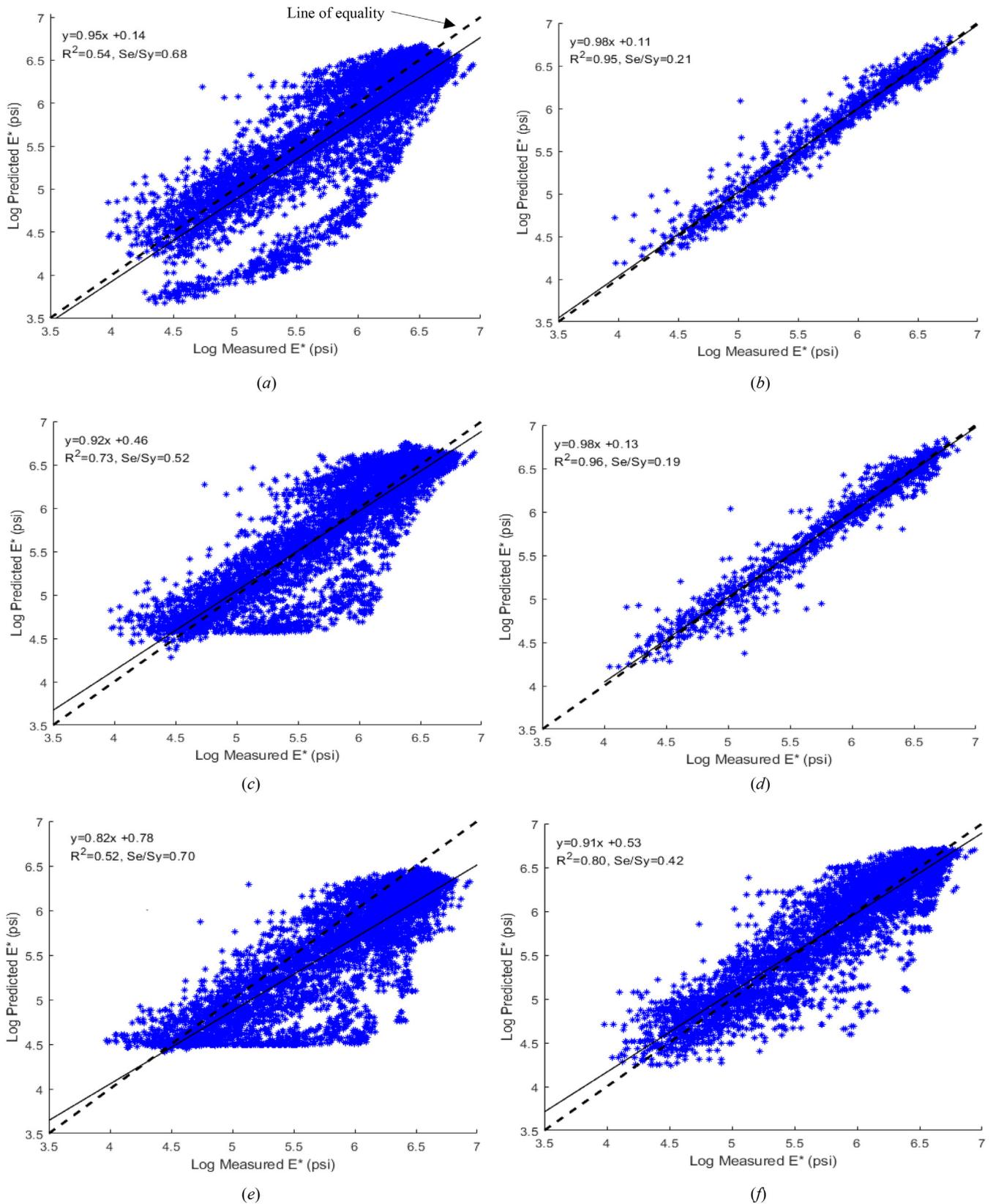
6. Results and discussion

The whole DCNNs architecture is implemented in MAT-LAB R2018b software [89] that is run on a PC with Intel Core I7

2.8-GHz processor, and 12 gigabytes of RAM. A multi-criteria evaluation process of the proposed DCNNs and well-known regression-based  $E^*$  predictive models has conducted for the two phases, including the aspects described in section 5.2, i.e.,

i) goodness-of-fit statistics ii) overall bias indicators, and iii) REC curves.

While, in the pre-trained DCNNs phase a total of 1212 data points (20% of data set#1) are used to assess the performance of



**Fig. 4.** Measured-predicted  $E^*$  values plot using (a)  $\eta$ -based Witczak 1-37A, (b)  $\eta$ -based DCNNs, (c)  $G,\delta$ -based Witczak 1-40D, (d)  $G,\delta$ -based DCNNs, (e)  $G$ -based Hirsch, and (f)  $G$ -based DCNNs (phase #1, using 1212 data points).

the predictive models, in the fine-tuned DCNNs phase a total of 1361 data points (80% of data set#2) are used. It is worth noting that, in the two phases, testing data are not used in the training process for the independent validation process.

6.1. The pre-trained DCNNs

In this stage, the DCNNs are validated against the data set #1. The importance of this step is that the DCNNs try to learn the relationships among the different controlling variables of the  $E^*$ . It is believed that deep learning architecture is able to learn the latent features of the dynamic modulus and the asphalt mixture components, where we could transfer this learning to the next phase of the solution. Also, this step is essential in judging the performance of the DCNNs for the  $E^*$  prediction as a new ML tool for the problem.

Details of the goodness-of-fit statistics for all proposed DCNNs and well-known regression-based  $E^*$  predictive models, in terms of the correlation coefficient,  $R^2$  and  $Se/Sy$ , are presented in Table 4. It is noticeable that DCNNs models exhibit higher prediction accuracy compared with regression-based models in terms of  $R^2$  and  $Se/Sy$

and  $Se/Sy$ . Both  $G, \delta$ -based DCNNs, and  $\eta$ -based DCNNs models achieve the most superior performance. They attain the highest  $R^2$  of 0.96 & 0.95 respectively and the lowest  $Se/Sy$  of 0.19 & 0.21 respectively. According to Pellinen [84], the evaluation criteria can be rated as an excellent fit to data. In the 3rd place comes the  $G$ -based DCNNs model with  $R^2$  and  $Se/Sy$  of 0.8 & 0.42 respectively, and a good accuracy according to the Pellinen criteria. Then comes the  $G, \delta$ -based Witczak 1-40D regression model of  $R^2$  and  $Se/Sy$  values of 0.73 & 0.52, respectively, with good accuracy. On the other hand, low  $R^2$  and large  $Se/Sy$  values of both  $\eta$ -based Witczak 1-37A and  $G$ -based Hirsch regression models indicate that they do not fit the data well. Consequently, their accuracy can be rated as fair according to the same criteria.

For an overview of how close predictions match measured data, the measured-predicted plot for each  $E^*$  predictive model, with its best-fitting line compared with the line of equality (LOE), is revealed, see Fig. 4. As can be seen in the figure,  $G, \delta$ -based DCNNs, and  $\eta$ -based DCNNs models' fitting lines are the closest to their corresponding LOE, which indicates that they match well with the measured data. On the other hand, both  $\eta$ -based Witczak 1-

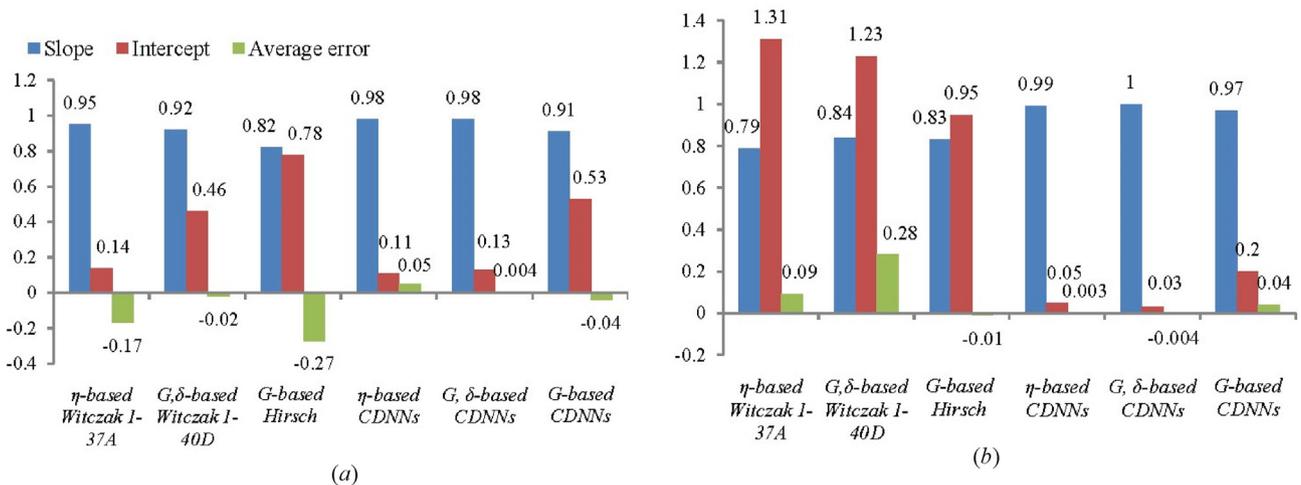


Fig. 5. Overall bias indicators for all  $E^*$  predictive models (a-phase #1 & b-phase #2).

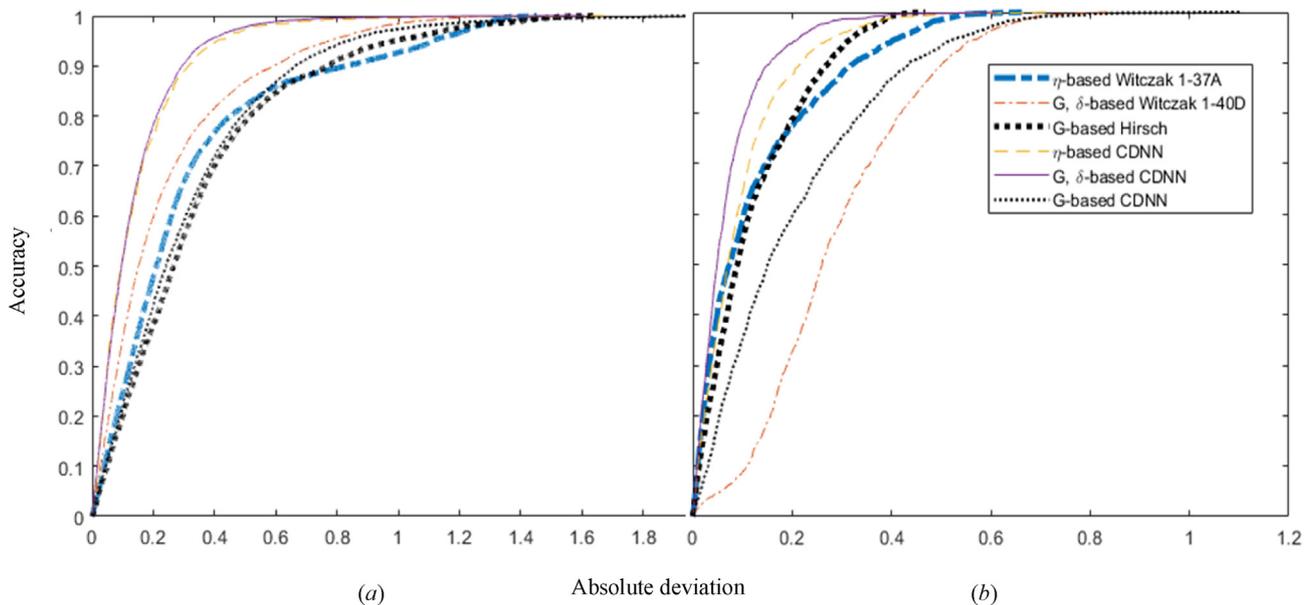
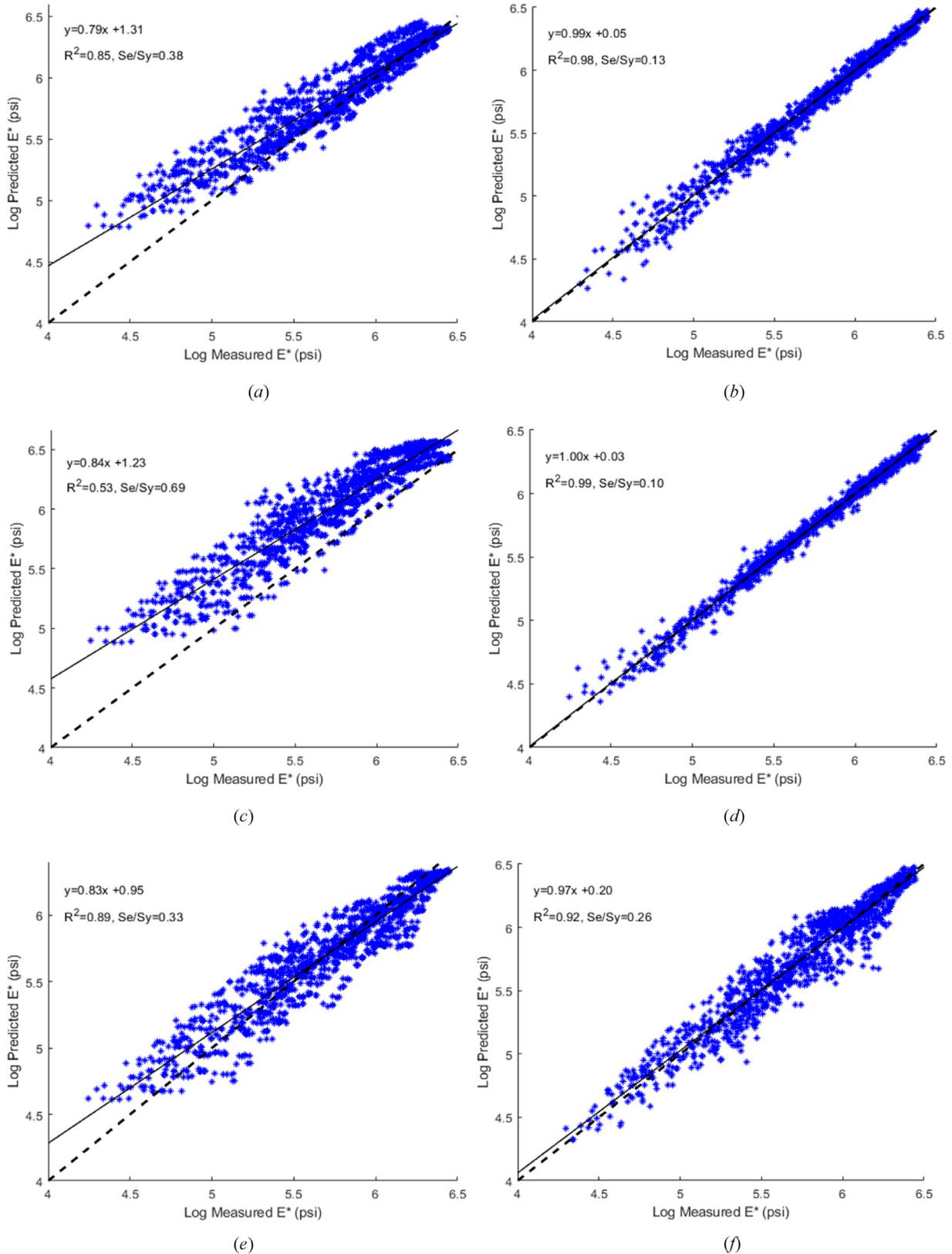


Fig. 6. REC curves for all  $E^*$  predictive models (a-phase #1 & b-phase #2).

37A and G-based Hirsch regression models fitting lines are the farthest to their corresponding LOE and also bellow them, which implies their bad matches and underprediction of the measured data. However, both G-based DCNNs and G,δ-based Witczak

1-40D regression models fitting lines are close to their corresponding LOE.

To get more insight into the models' levels of over- or under-prediction, the aforementioned overall bias indicators (slope,



**Fig. 7.** Measured-predicted  $E^*$  values plot using (a)  $\eta$ -based Witczak 1-37A, (b)  $\eta$ -based DCNNs, (c)  $G,\delta$ -based Witczak 1-40D, (d)  $G,\delta$ -based DCNNs, (e) G-based Hirsch, and (f) G-based DCNNs (phase #2, using 1361 data points).

intercept, and average error) are depicted in Fig. 5.a. Slopes, intercepts, and average errors of  $G,\delta$ -based DCNNs and  $\eta$ -based DCNNs models are substantially the best since they are closer to 1, 0, and 0, respectively, in comparison with the other predictive models. The results confirm their exposure to the smallest prediction bias. In contrast, the  $G$ -based Hirsch regression model's overall bias indicators are the worst relative to other predictive models, with a slope, intercept, and an average error of 0.82, 0.78, and  $-0.27$ , respectively. Moreover, both  $\eta$ -based Witczak 1-37A and  $G$ -based Hirsch regression models under-predicted  $E^*$ , with an average error of  $-0.17$ , and  $-0.27$ , respectively.  $G$ -based DCNNs and  $G,\delta$ -based Witczak 1-40D regression models exhibit overall bias indicators relatively close to each other.

The REC curves for all proposed DCNNs and well-known regression-based  $E^*$  predictive models are depicted in Fig. 6. a, which reveals that  $G,\delta$ -based DCNNs, and  $\eta$ -based DCNNs models outperform other predictive models, as their REC curves climbed rapidly towards the upper left corner. For the remaining four predictive models,  $G,\delta$ -based Witczak 1-40D regression model performs better than the other ones as its REC curve is all the time above their corresponding REC curves. Moreover, the  $\eta$ -based Witczak 1-37A regression model seems to perform better than  $G$ -based DCNNs and  $G$ -based Hirsch regression models for small values of absolute deviations (errors). Still, the opposite is the case as the absolute deviation becomes larger (larger than 0.58). Furthermore, the  $G$ -based DCNNs model performed better than the  $G$ -based Hirsch regression model.

## 6.2. The Fine-toned DCNNs

In this next stage, the second contribution of the study is examined. For the laboratory effort reduction justification, it is assumed that the available conducted experimental points are 20% of the data set #2. The remaining 80% of the data set is used for testing.

In this step (phase #2), we use the same performance indicators in phase #1. Table 4, Fig. 5.b & 6.b, and Fig. 7 draw a similar analysis (as in sub-section 6.1.) for the models under study; however, more conclusions could be derived in the comparison form. The structure of data set #2 is more homogenous (i.e., comes from one source) than data set #1 (i.e., multiple sources). Therefore, regression-based models achieve more stability in prediction than phase #1. On the contrary, DCNNs based models, in general, absorb the variability in data and thus no noticeable change in the prediction performance between the two phases. Interestingly, the DCNNs models maintain the high accuracy in phase #2, despite reducing the input data. Since, as known in ML learning practice, increasing the training data leads to better results. The fine-toned DCNNs managed to transfer the learning gained by the pre-trained DCNNs using the data set #1 to the next phase.

## 7. Conclusions

In this study, the laboratory effort reduction is undertaken as the central theme for the HMA dynamic modulus prediction problem. A powerful ML technique is introduced for that purpose as a new solution methodology. The main goal is to reduce the laboratory effort needed for the  $E^*$  determination using a pre-trained Deep Learning (DL) model via a transfer learning. For the  $E^*$  prediction problem, such a pre-trained DL model is not available. Therefore, the problem solution is justified by adapting deep convolution learning technology (DCNNs) through two consequent phases; pre-trained DCNNs and fine-toned DCNNs. Two different data sets are used for that purpose, namely; data set #1 and data set #2. While the data set #1 is used in the pre-trained DCNNs phase, data set #2 is used in the fine-toned DCNNs phase. Utilizing the same input

variables, the proposed DCNNs show a substantial superior predictive performance, particularly  $G,\delta$ -based DCNNs, and  $\eta$ -based DCNNs models, compared with the well-known regression-based models. Also, the results confirm with the literature regarding the inconsistency and bias associated with the well-known regression-based predictive models ( $\eta$ -based Witczak 1-37A,  $G,\delta$ -based Witczak 1-40D, and  $G$ -based Hirsch). Where both  $\eta$ -based Witczak 1-37A and  $G$ -based Hirsch regression models under-predict  $E^*$  values when using the data set#1, while both  $\eta$ -based Witczak 1-37A and  $G,\delta$ -based Witczak 1-40D regression models over-predict  $E^*$  values using data set#2. On the contrary,  $G,\delta$ -based DCNNs, and  $\eta$ -based DCNNs models achieve high stable performance indicators regarding the two data sets. The DCNNs implementation in this paper opens the gate to investigate further their applications in other complicated material properties research in the context of pre-trained techniques. We believe that the proposed pre-trained DCNNs can be adapted in future research not only for  $E^*$  prediction but also for any similar problem via transfer learning.

## CRediT authorship contribution statement

**Ghada S. Moussa:** Conceptualization, Methodology, Software, Validation, Writing - original draft. **Mahmoud Owais:** Formal analysis, Writing - review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## references

- [1] L. Yongliang, K. Xiangming, Z. Yanrong, Y. Peiyu, Static and dynamic mechanical properties of cement-asphalt composites, *J. Mater. Civ. Eng.* 25 (10) (2013) 1489–1497.
- [2] K.H. McGhee, NCHRP synthesis of highway practice 334: automated pavement distress collection techniques, Transportation Research Board of the National Academies, Washington, DC, 2004.
- [3] R.L. Carvalho, C.W. Schwartz, Comparisons of Flexible Pavement Designs: AASHTO Empirical Versus NCHRP Project 1-37A Mechanistic-Empirical, *Transp. Res. Rec.* 1947 (1) (2006) 167–174.
- [4] X. Shu, B. Huang, Micromechanics-based dynamic modulus prediction of polymeric asphalt concrete mixtures, *Compos. B Eng.* 39 (4) (2008) 704–713.
- [5] X. Shu, B. Huang, Predicting dynamic modulus of asphalt mixtures with differential method, *Road materials and pavement design* 10 (2) (2009) 337–359.
- [6] J. Li, A. Zofka, I. Yut, Evaluation of dynamic modulus of typical asphalt mixtures in Northeast US region, *Road materials and pavement design* 13 (2) (2012) 249–265.
- [7] E. Masad, H. Bahia, EFFECTS OF LOADING CONFIGURATION AND MATERIAL PROPERTIES ON NON-LINEAR RESPONSE OF ASPHALT MIXTURES (WITH DISCUSSION), *J. Assoc. Asphalt Paving Technol.* 71 (2002).
- [8] B. Birgisson, G. Sholar, R. Roque, Evaluation of a predicted dynamic modulus for Florida mixtures, *Transp. Res. Rec.* 1929 (1) (2005) 200–207.
- [9] T. Aashto, Standard method of test for determining dynamic modulus of hot-mix asphalt concrete mixtures, ed.: American Association of State Highway and Transportation Officials ..., 2011.
- [10] G.M. Rowe, S.H. Khoei, P. Blankenship, K.C. Mahboub, Evaluation of aspects of  $E^*$  test by using hot-mix asphalt specimens with varying void contents, *Transp. Res. Rec.* 2127 (1) (2009) 164–172.
- [11] B. Dołżycki, J. Judycki, Behaviour of asphalt concrete in cyclic and static compression creep test with and without lateral confinement, *Road materials and pavement design* 9 (2) (2008) 207–225.
- [12] E. Rahmani, M.K. Darabi, R.K.A. Al-Rub, E. Kassem, E.A. Masad, D.N. Little, Effect of confinement pressure on the nonlinear-viscoelastic response of asphalt concrete at high temperatures, *Constr. Build. Mater.* 47 (2013) 779–788.
- [13] Y. Seo, O. El-Haggan, M. King, S. Joon Lee, Y. Richard Kim, Air void models for the dynamic modulus, fatigue cracking, and rutting of asphalt concrete, *J. Mater. Civ. Eng.* 19 (10) (2007) 874–883.
- [14] J. Bari, M.W. Witczak, Evaluation of the Effect of Lime Modification on the Dynamic Modulus Stiffness of Hot-Mix Asphalt: Use with the New Mechanistic-Empirical Pavement Design Guide, *Transp. Res. Rec.* 1929 (1) (2005) 10–19.

- [15] Y. Kim, H.D. Lee, M. Heitzman, Validation of new mix design procedure for cold in-place recycling with foamed asphalt, *J. Mater. Civ. Eng.* 19 (11) (2007) 1000–1010.
- [16] M. H. King, "Determination of dynamic moduli in uniaxial compression for North Carolina hot mix asphalt concrete," 2004.
- [17] T. A. Bennert, "Dynamic modulus of hot mix asphalt," 2009.
- [18] H. Azari, G. Al-Khateeb, A. Shenoy, N. Gibson, Comparison of Simple Performance Test| E\*| of Accelerated Loading Facility Mixtures and Prediction| E\*| Use of NCHRP 1–37A and Witzczak's New Equations, *Transp. Res. Rec.* 1998 (1) (2007) 1–9.
- [19] D. Christensen Jr, T. Pellinen, R. Bonaquist, Hirsch model for estimating the modulus of asphalt concrete, *J. Assoc. Asphalt Paving Technol.* 72 (2003).
- [20] M.W. Witzczak, *Simple performance tests: Summary of recommended methods and database*. Washington, D.C.: National Cooperative Highway Research Program Report 547, Transportation Research Board (2005).
- [21] N.H. Gibson, C.W. Schwartz, R.A. Schapery, M.W. Witzczak, Viscoelastic, viscoplastic, and damage modeling of asphalt concrete in unconfined compression, *Transp. Res. Rec.* 1860 (1) (2003) 3–15.
- [22] R. Bonaquist, D.W. Christensen, Practical procedure for developing dynamic modulus master curves for pavement structural design, *Transp. Res. Rec.* 1929 (1) (2005) 208–217.
- [23] A. Designation, R30, Mixture Conditioning of Hot Mix Asphalt (HMA), ed., American Association of State Highway and Transportation Officials . . . , 2005.
- [24] R. Dongre, L. Myers, J. D'Angelo, C. Paugh, J. Gudimetla, Field evaluation of Witzczak and Hirsch models for predicting dynamic modulus of hot-mix asphalt (with discussion), *J. Assoc. Asphalt Paving Technol.* 74 (2005).
- [25] J. Zhang, Z. Fan, H. Wang, W. Sun, J. Pei, D. Wang, Prediction of dynamic modulus of asphalt mixture using micromechanical method with radial distribution functions, *Mater. Struct.* 52 (2) (2019) 49.
- [26] M. Witzczak, D. Andrei, and W. Mirza, "Development of revised predictive model for the dynamic (complex) modulus of asphalt mixtures," Inter-team Technical Report prepared for the NCHRP 1-37A Project, 1999.
- [27] M. Witzczak, M. El-Basyouny, and S. El-Badawy, "Incorporation of the New (2005) E\* Predictive Model in the MEPDG," NCHRP 1-40D Final Report, 2007.
- [28] M.W. Witzczak, Specification criteria for simple performance tests for rutting, Transportation Research Board (2007).
- [29] S. El-Badawy, F. Bayomy, A. Awed, Performance of MEPDG dynamic modulus predictive models for asphalt concrete mixtures: local calibration for Idaho, *J. Mater. Civ. Eng.* 24 (11) (2012) 1412–1421.
- [30] S. Yousefdoost, B. Vuong, I. Rickards, P. Armstrong, B. Sullivan, "Evaluation of dynamic modulus predictive models for typical Australian asphalt mixes," in *Delivering New Age Solutions*, 15th AAPA International Flexible Pavements Conference (2013) 22–25.
- [31] A.M. Khattab, S.M. El-Badawy, M. Elmwafi, Evaluation of Witzczak E\* predictive models for the implementation of AASHTOWare-Pavement ME Design in the Kingdom of Saudi Arabia, *Constr. Build. Mater.* 64 (2014) 360–369.
- [32] Y. Ali, M. Irfan, S. Ahmed, S. Khanzada, T. Mahmood, Investigation of factors affecting dynamic modulus and phase angle of various asphalt concrete mixtures, *Mater. Struct.* 49 (3) (2016) 857–868.
- [33] K. Georgouli, C. Plati, A. Loizos, Assessment of dynamic modulus prediction models in fatigue cracking estimation, *Mater. Struct.* 49 (12) (2016) 5007–5019.
- [34] K. Georgouli, A. Loizos, C. Plati, Calibration of dynamic modulus predictive model, *Constr. Build. Mater.* 102 (2016) 65–75.
- [35] H. Ceylan, K. Gopalakrishnan, S. Kim, Advanced approaches to hot-mix asphalt dynamic modulus prediction, *Can. J. Civ. Eng.* 35 (7) (2008) 699–707.
- [36] H. Ceylan, S. Kim, and K. Gopalakrishnan, "Hot mix asphalt dynamic modulus prediction models using neural networks approach," 2007.
- [37] A. Jamrah, M. E. Kutay, and H. I. Ozturk, "Characterization of asphalt materials common to Michigan in support of the implementation of the mechanistic-empirical pavement design guide," 2014.
- [38] G. Harran, A. Shalaby, Improving the prediction of the dynamic modulus of fine-graded asphalt concrete mixtures at high temperatures, *Can. J. Civ. Eng.* 36 (2) (2009) 180–190.
- [39] A. Behnood, E.M. Golafshani, Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves, *J. Cleaner Prod.* 202 (2018) 54–64.
- [40] T.S. Ozsahin, S. Oruc, Neural network model for resilient modulus of emulsified asphalt mixtures, *Constr. Build. Mater.* 22 (7) (2008) 1436–1445.
- [41] A. Fathi, M. Mazari, M. Saghafi, A. Hosseini, S. Kumar, Parametric Study of Pavement Deterioration Using Machine Learning Algorithms, *Airfield and Highway Pavements* (2019) 31–41.
- [42] H. Majidifard, B. Jahangiri, W.G. Buttlar, A.H. Alavi, New machine learning-based prediction models for fracture energy of asphalt mixtures, *Measurement* 135 (2019) 438–451.
- [43] Y.B. Dibike, S. Velickov, D. Solomatine, M.B. Abbott, Model induction with support vector machines: introduction and applications, *J. Comput. Civil Eng.* 15 (3) (2001) 208–216.
- [44] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics* 5 (4) (1943) 115–133.
- [45] P. Singh, B. Borah, Indian summer monsoon rainfall prediction using artificial neural network, *Stoch. Env. Res. Risk Assess.* 27 (7) (2013) 1585–1599.
- [46] I.-C. Yeh, Exploring concrete slump model using artificial neural networks, *J. Comput. Civil Eng.* 20 (3) (2006) 217–221.
- [47] H. Ceylan, K. Gopalakrishnan, S. Kim, Looking to the future: the next-generation hot mix asphalt dynamic modulus prediction models, *Int. J. Pavement Eng.* 10 (5) (2009) 341–352.
- [48] D. Singh, M. Zaman, S. Commuri, Artificial neural network modeling for dynamic modulus of hot mix asphalt using aggregate shape properties, *J. Mater. Civ. Eng.* 25 (1) (2013) 54–62.
- [49] M.S. Sakhaeifar, Y.R. Kim, P. Kabir, New predictive models for the dynamic modulus of hot mix asphalt, *Constr. Build. Mater.* 76 (2015) 221–231.
- [50] H. Ceylan, C.W. Schwartz, S. Kim, K. Gopalakrishnan, Accuracy of predictive models for dynamic modulus of hot-mix asphalt, *J. Mater. Civ. Eng.* 21 (6) (2009) 286–293.
- [51] S.M. El-Badawy, A.M. Khattab, A.A. Al Hazmi, Using artificial neural networks (ANNs) for hot mix asphalt E\* predictions, *Geo-China* (2016, 2016,) 83–91.
- [52] M.S.S. Far, B.S. Underwood, S.R. Ranjithan, Y.R. Kim, N. Jackson, Application of artificial neural network for estimating dynamic modulus of asphalt concrete, *Transp. Res. Rec.* 2127 (1) (2009) 173–186.
- [53] Long Term Pavement Performance [Online]. Available: <https://infopave.fhwa.dot.gov/>
- [54] V. Vapnik, A.Y. Chervonenkis, A class of algorithms for pattern recognition learning, *Avtomat. i Telemekh* 25 (6) (1964) 937–945.
- [55] F.E. Tay, L. Cao, Application of support vector machines in financial time series forecasting, *omega* 29 (4) (2001) 309–317.
- [56] K. Gopalakrishnan, S. Kim, Support vector machines approach to HMA stiffness prediction, *J. Eng. Mech.* 137 (2) (2011) 138–146.
- [57] J.R. Koza, J.R. Koza, Genetic programming: on the programming of computers by means of natural selection, MIT press (1992).
- [58] C. Ferreira, Gene expression programming in problem solving, in: *Soft computing and industry*, Springer, 2002, pp. 635–653.
- [59] L. Jun, K. Yan, X. Zhao, Predicting the Dynamic Modulus of Asphalt Mixtures Containing Recycled Asphalt Shingles using Artificial Neural Networks, *DEStech Transactions on Engineering and Technology Research*, no. ictim (2016).
- [60] J. Yan, Z. Zhang, H. Zhu, F. Li, Q. Liu, Experimental study of hot recycled asphalt mixtures with high percentages of reclaimed asphalt pavement and different recycling agents, *J. Test. Eval.* 42 (5) (2014) 1183–1190.
- [61] T.K. Ho, Random decision forests vol. 1 (1995) 278–282.
- [62] L. Breiman, Random forests, *Machine learning* 45 (1) (2001) 5–32.
- [63] D. Daneshvar, A. Behnood, Estimation of the dynamic modulus of asphalt concretes using random forests algorithm, *Int. J. Pavement Eng.* (2020) 1–11.
- [64] J. Bari, Development of a new revised version of the Witzczak E\* predictive models for hot mix asphalt mixtures, Arizona State University Tempe, 2005
- [65] M.W. Witzczak, Simple performance test for superpave mix design, Transportation Research Board (2002).
- [66] B. W. Sullivan, "Development of a long life design procedure for Australian asphalt pavements," Curtin University, 2015.
- [67] S. Yousefdoost, E. Denneman, A. Beecroft, B. Vuong, Development of a national database of asphalt material performance properties in support of perpetual pavement design implementation in Australia, *Constr. Build. Mater.* 188 (2018) 68–87.
- [68] A. Khattab, "Dynamic modulus predictive models for superpave asphalt concrete mixtures," M. Sc, Thesis, Public Works Engineering, Mansoura University, Egypt, 2015.
- [69] J. Yu, "Modification of dynamic modulus predictive models for asphalt mixtures containing recycled asphalt shingles," 2012.
- [70] D. Ciregan, U. Meier, J. Schmidhuber, in: Multi-column deep neural networks for image classification, *IEEE*, 2012, pp. 3642–3649.
- [71] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [72] M. Owais, G.S. Moussa, K.F. Hussain, Robust Deep Learning Architecture for Traffic Flow Estimation from a Subset of Link Sensors, *Journal of Transportation Engineering, Part A: Systems* 146 (1) (2020) 04019055.
- [73] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning," *nature* 521 (7553) (2015) 436–444.
- [74] Y. Bengio, "Learning deep architectures for AI," *Foundations and trends*<sup>®</sup>, *Machine Learning* 2 (1) (2009) 1–127.
- [75] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.
- [76] M.A. Wani, F.A. Bhat, S. Afzal, A.I. Khan, Supervised deep learning in face recognition, in: *Advances in Deep Learning*, Springer, 2020, pp. 95–110.
- [77] Y. Bengio, I. Goodfellow, A. Courville, *Deep learning*, MIT press (2017).
- [78] M. Meng, Y.J. Chua, E. Wouterson, C.P.K. Ong, Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks, *Neurocomputing* 257 (2017) 128–135.
- [79] C. Affonso, A.L.D. Rossi, F.H.A. Vieira, A.C.P. de Leon Ferreira, Deep learning for biological image classification, *Expert Syst. Appl.* 85 (2017) 114–122.
- [80] S. Krig, *Computer vision metrics: Survey, taxonomy, and analysis*, Apress (2014).
- [81] J. Zang, L. Wang, Z. Liu, Q. Zhang, G. Hua, N. Zheng, in: Attention-based temporal weighted convolutional neural network for action recognition, *Springer*, 2018, pp. 97–108.
- [82] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.
- [83] C. Szegedy et al. Going deeper with convolutions, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

- [84] T. K. Pellinen, "Investigation of the use of dynamic modulus as an indicator of hot-mix asphalt performance," 2002.
- [85] S. El-Badawy, R. Abd El-Hakim, A. Awed, Comparing artificial neural networks with regression models for Hot-Mix asphalt dynamic modulus prediction, *J. Mater. Civ. Eng.* 30 (7) (2018) 04018128.
- [86] G. Al-Khateeb, A. Shenoy, N. Gibson, T. Harman, A new simplistic model for dynamic modulus predictions of asphalt paving mixtures, *J. Assoc. Asphalt Paving Technol.* 75 (2006).
- [87] J. Bi, K.P. Bennett, Regression error characteristic curves, in: in *Proceedings of the 20th international conference on machine learning (ICML-03)*, 2003, pp. 43–50.
- [88] J. Hernández-Orallo, ROC curves for regression, *Pattern Recogn.* 46 (12) (2013) 3395–3411.
- [89] M. Matlab, "MATLAB R2018b," *The MathWorks: Natick, MA, USA*, 2018.