Improved predictions of asphalt concretes' dynamic modulus and phase angle using decision-tree based categorical boosting model

Fabio Rondinella [a], Fabiola Daneluz [a], Bernhard Hofko [b], Nicola Baldo [a,*]

[a] Polytechnic Department of Engineering and Architecture (DPIA), University of Udine, Via del Cotonificio 114, 33100 Udine, Italy
[b] TU Wien, Institute of Transportation, Karlsplatz 13/E230-3, 1040 Vienna, Austria

A B S T R A C T

The most suitable parameter to summarize the viscoelastic response of asphalt concrete (AC) mixtures is the complex modulus, defined by means of its two main components: the dynamic modulus $|E^*|$ and the phase angle φ. They are frequently determined by means of expensive and time-consuming laboratory procedures that require suitable equipment and high-skilled technicians. As an alternative, machine learning models can be trained to make very accurate predictions and thus, substitute at least some of these lab tests. This study proposes an innovative Categorical Boosting (CatBoost) approach for the simultaneous prediction of both $|E^*|$ and φ. Nine different AC mixtures were prepared, and an extensive 4-point bending test (4PBT) experimental campaign was carried out under ten loading frequencies and six testing temperatures. In order to thoroughly compare the developed model with two well-established empirical equations (Witczak-Fonseca and Witczak 1–37A), the same input features were selected. Pre-processing and resampling techniques were implemented to both reduce computational effort and improve model efficiency, whereas an in-depth sensitivity analysis was also performed. The entire methodology was implemented in Python 3.8.5. Six different goodness-of-fit metrics were used to robustly evaluate the performance of the developed CatBoost model and to compare it with the results of two regression-based models and a reference state-of-the-art artificial neural network (SoA-ANN). Findings showed that both machine learning (ML) models outperformed the regression-based ones, displaying significantly better performance for all metrics used. CatBoost and SoA-ANN showed roughly comparable results, characterized by a mean coefficient of determination (R^2) slightly higher than 0.98. Since goodness-of-fit metrics resulted in no marked differences between machine learning models, CatBoost approach might be preferred because of its easy implementation in Python and its high interpretability. Within the context of pavement engineering, such an advanced machine learning model could provide a useful and powerful tool for asphalt mixtures' design applications.

1. Introduction

One of the most important engineering properties to analyze the mechanical behaviour of asphalt concretes (ACs) and evaluate their performance is represented by the complex modulus E^*. This fundamental parameter allows the variation in a mixture's stiffness to be described in detail as the test conditions, namely loading frequency and testing temperature, change [1].

Complex modulus is defined by means of two main components: the dynamic modulus $|E^*|$ and the phase angle φ. The former describes the material stiffness at given temperature and frequency conditions, whereas the latter determines the corresponding elastic and viscous amounts under those test settings. By means of time–temperature superposition principle (TTSP) [2], these data are usually processed to obtain master curves. These continuous curves allow the complex modulus to be estimated for a wide range of temperature and frequency combinations. Complex modulus is a crucial parameter for both the design of asphalt pavement structures and the evaluation of pavement performance [3] and is required as input by major pavement design procedures such as the Mechanistic-Empirical Pavement Design guide (M − EPDG) [4].

Basically, two main methods are used to determine complex modulus and phase angle: experimental and computational. The former consists of performing laboratory tests on different mixture specimens. The latter

* Corresponding author.
E-mail address: nicola.baldo@uniud.it (N. Baldo).

involves using predictive empirical equations or, alternatively, mathematical models to estimate the desired parameters. However, an expensive and time-consuming procedure is necessary for the experimental determination of $|E^*|$ and φ: it is required to produce and prepare compacted specimens, to test them for wide ranges of temperature and frequency by means of suitable equipment used by high-skilled technicians [5]. The efforts of many researchers have consequently shifted to predictive methods, thus limiting the use of onerous laboratory tests and predicting $|E^*|$ and φ by means of *ad-hoc* semi-empirical equations. Early successful attempts resulted in the predictive equation introduced by Fonseca and Witczak [6], followed by the well-established viscosity based Witczak's model [7]. Features such as aggregate gradation percentages, air voids (AV) content, effective binder content ($V_{b_{eff}}$), as well as binder viscosity (η_b) and loading frequency (f_l) are required as input, with the dynamic modulus value returned as output. The resulting predictions proved to be accurate and reliable: however, since both methodologies were developed from data of standard mixtures, they showed some applicability limitations [8].

Over the years, many other models have been developed and used more or less frequently to produce AC mixture dynamic modulus predictions [9–12]. Although fewer in number, regression-based models to make phase angle predictions have also been introduced [13,14]. Nevertheless, in order to better understand and capture the relationships between the complex modulus and its several influencing variables, regression-based models have been progressively replaced by machine learning (ML) techniques [15–19]. Unlike regression-based models, these innovative techniques have the remarkable capability to model and predict even very different phenomena, whose fundamental relationships do not necessarily need to be thoroughly known.

One of the most widely practiced ML techniques is the artificial neural network (ANN), which has proven to be capable of understanding even highly nonlinear phenomena and producing remarkably accurate predictions [20–23]. However, it has some drawbacks such as complexity in identifying the best topology and hyperparameters coupled with poor interpretability of the overall model.

Other supervised ML algorithms, such as decision-tree (DT) based ones, address these problems by returning comparably accurate predictions in multiple domains [24–26] although with more easily interpretable models. In fact, such DT based approaches consist in generating a tree-like model to solve classification or regression problems by means of simple decision rules [27]. Some of their applications to solve dynamic modulus-related regression problems are fully described in the relevant literature [28,29].

The objective of this study is to develop an advanced ML model, capable of implementing an innovative DT based algorithm, namely the Categorical Boosting or CatBoost, in order to accurately predict both the dynamic modulus and phase angle of multiple AC mixtures. The CatBoost technique results from the combination of innovative feature processing algorithms focused on computational efficiency and high predictive accuracy that have made it outperform other leading gradient-boosting competitors on a wide variety of datasets [30].

An extensive four-point bending test (4PBT) experimental campaign was carried out: 1680 observations of AC mixture samples were recorded, and the resulting information served to train the model to simultaneously predict both dynamic modulus and phase angle.

This paper represents one of the first studies implementing the CatBoost approach to model the complex modulus of several AC mixtures. In order to fairly compare the developed model with state-of-the-art (SoA) ANN and two well-established empirical equations, the assigned input features remained unchanged in all four different approaches. Furthermore, an in-depth sensitivity analysis was performed to fully understand model functioning and the impact the different features respectively have on the outputs.

Considering the remarkable performance achieved by the model, it could provide a fast and reliable tool for simultaneously predicting $|E^*|$ and φ values to be implemented in the most common design procedures. Furthermore, this research confirms what is known about the correlations between compositional variables and performance parameters and could lead to an even better understanding of the existing relationships between these variables and the mechanical behaviour of the corresponding asphalt concretes. In this sense, the present study represents a contribution to the existing literature and an interesting inspiration for future applications.

2. Material and methods

2.1. Material selection and dynamic modulus testing

To create a large data base for the modelling efforts in this study, nine different asphalt mixtures were designed, produced, and tested to receive data on the viscoelastic response of the respective mixtures. A schematic representation of the entire methodology followed is provided in Fig. 1. For surface layers (AC 11, SMA 11), diabase was used as fine and coarse aggregates, while for binder and base mixtures (AC 22, AC 32), limestone was the choice. In all cases, the filler is powdered limestone. Unmodified, as well as polymer-modified binders were implemented in the study. Mineral aggregate gradation follows the respective European product standards for asphalt mixtures (EN 13108-x [31]). The optimal binder content was determined by the procedure according to Marshall. Details on mix parameters and the grading curves are shown in Table 1 and Fig. 2.

For all mixtures, asphalt mix slabs were compacted in a steel segment compactor according to EN 12697-33 [32]. Of each slab, three specimens with nominal dimensions of 60x60x50 mm were cut. After determining volumetric properties of each specimen, they were subjected to stiffness tests by 4PBT with a temperature and frequency sweep. The 4PBT was carried out according to EN 12697-26 [33] in displacement-controlled mode at temperatures ranging from –15 °C to + 45 °C and a frequency sweep from 0.1 Hz to 40 Hz at each temperature. The horizontal strain amplitude at the bottom of the beam was set to 35 μm/m to prevent any damage during testing. For each mix, at least 3 single specimens were tested. Stiffness and phase angle results used in this study are mean values from these single tests derived from the recorded test data according to the equations given in EN 12697-26 [33].

2.2. Empirical equations for $|E^*|$

Several models have been investigated and developed over the years to establish a relationship between the dynamic modulus $|E^*|$ and the main physical parameters of asphalt mixtures. One of the most scientifically sound was proposed by Witczak in 1996 and is based on the TTSP. E^* values at different test temperatures are shifted to the same reference temperature in order to develop the so-called master curve [6]. For asphalt mixtures, this curve is assumed as nonlinear sigmoidal (Equation 1) since it is able to represent the mixture's dependence on temperature:

$$\log|E^*| = \rho + \frac{\alpha}{1 + e^{\beta+\gamma\log(f_{red})}} \quad (1)$$

$|E^*|$ represents the dynamic modulus, ρ is the logarithm of $|E^*|$ minimum value, $\rho+\alpha$ is the logarithm of $|E^*|$ maximum value, β and γ are shape parameters, and f_{red} represents the reduced frequency. Master curves for AC11, AC22, AC32, and SMA11 mixtures have been reported in Fig. 3. For each mixture and for each reference temperature, the shift factor value ($\log a(T)$) required to determine the corresponding reduced frequency was given in Table 2.

Still in 1996, based on a large experimental data set (1430 observations of 149 conventional asphalt mixes), Witczak and Fonseca investigated the relationship between the sigmoidal function parameters

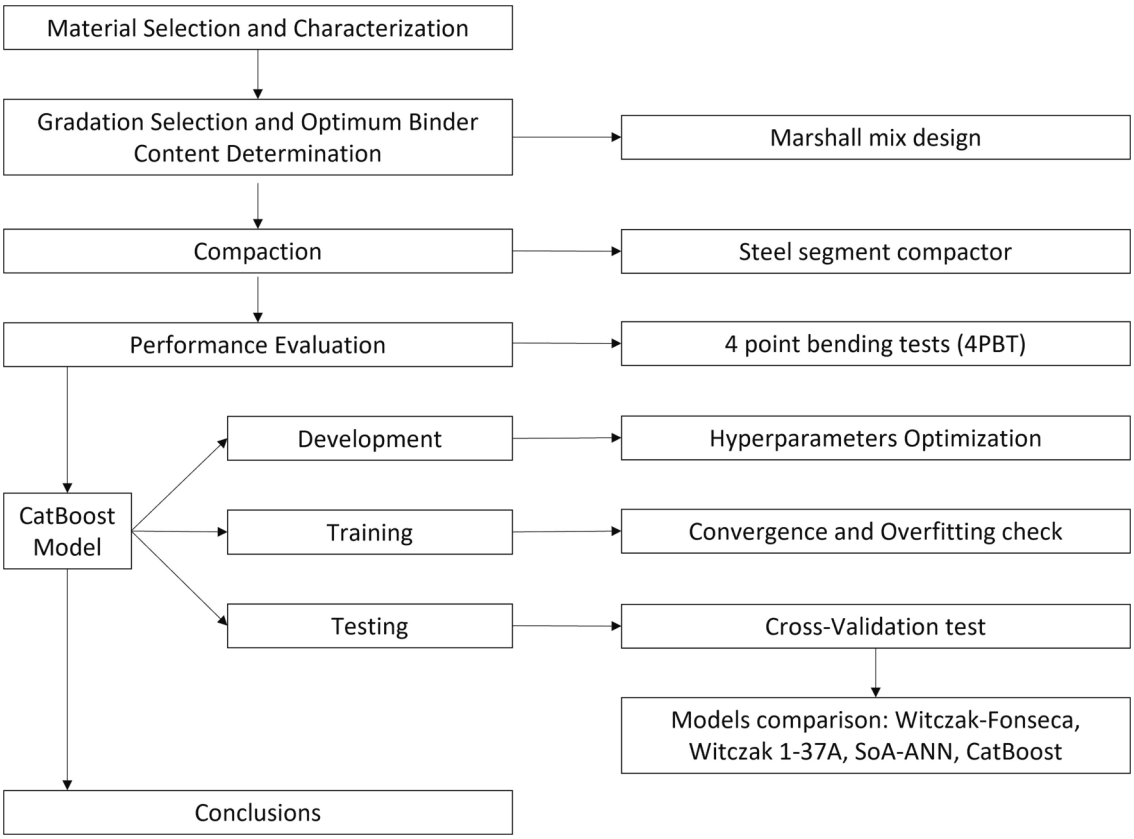

Fig. 1. Schematic representation of the methodology followed.

and the physical-volumetric properties of asphalt mixtures. The following empirical equation was then derived:

$$\log_{10}\log_{10}\eta_b = A + VTS \bullet \log_{10}(T_R) \tag{3}$$

$$\log|E^*| = -1.249937 + 0.029232 \bullet \rho_{No.200} - 0.001767 \bullet (\rho_{No.200})^2 + 0.002841 \bullet \rho_{No.4} - 0.058097 \bullet AV - 0.802208 \bullet \frac{V_{b_{eff}}}{(V_{b_{eff}} + AV)} + \frac{3.871977 + 0.0021 \bullet \rho_{No.4} + 0.003958 \bullet \rho_{3/8} - 0.000017 \bullet (\rho_{3/8})^2 + 0.005470 \bullet \rho_{3/4}}{1 + e^{(-0.603313 - 0.313551 \bullet \log(f_l) - 0.393532 \bullet \log(\eta_b))}} \tag{4}$$

$|E^*|$ is the asphalt mixture dynamic modulus (10^5 psi), $\rho_{No.200}$ is the aggregate percentage passing the No. 200 sieve, $\rho_{No.4}$ is the aggregate percentage retained on the No. 4 sieve, $\rho_{3/8}$ is the aggregate percentage retained on the $^3/_8$ inch sieve, $\rho_{3/4}$ is the aggregate percentage retained on the ¾ inch sieve; AV is the percentage by volume of air voids in the mixture, $V_{b_{eff}}$ is the percentage by volume of effective binder content, f_l is the loading frequency (Hz), and η_b is the asphalt viscosity (10^6 poises) determined as follows:

T_R is the temperature in Rankine, whereas A and VTS are regression parameters whose values have been estimated as a function of performance grade (PG) [4].

In 2006, Witczak et al. [7] combined the observations used in Fonseca and Witczak's original model with 1320 additional observations related to 56 mixes (34 of them used modified binders) developing a new model (Equation 4). This was named Witczak 1-37A and is currently used in the United States to predict the dynamic modulus of asphalt mixes in Level 2 and 3 projects, according to the M-EPDG [4].

$$\log|E^*| = -0.261 + 0.008225 \bullet \rho_{No.200} - 0.00000101 \bullet (\rho_{No.200})^2 + 0.00196 \bullet \rho_{No.4} - 0.03157 \bullet AV - 0.415 \bullet \frac{V_{b_{eff}}}{(V_{b_{eff}} + AV)} + \frac{1.87 + 0.002808 \bullet \rho_{No.4} + 0.0000404 \bullet \rho_{3/8} - 0.0001786 \bullet (\rho_{3/8})^2 + 0.0164 \bullet \rho_{3/4}}{1 + e^{(-0.716 \bullet \log(f_l) - 0.7425 \bullet \log(\eta_b))}} \tag{2}$$

Table 1
AC mixtures summary description.

Mixture	Aggregate	Mix Design	AV (%)	Stability (kN)	Bulk specific gravity
AC11_70_100	Diabase	5.6% PG 63–24	3.4%	12.6	2.56
AC11_PmB45_80_50	Diabase	5.6% PG 69–27	3.7%	13.9	2.56
AC11_PmB45_80_65	Diabase	5.6% PG 76–25	3.4%	13.3	2.56
AC22_mB160_220FT	Limestone	4.5% PG 73–22	3.8%	11.1	2.62
AC22_PmB45_80_65	Limestone	4.5% PG 76–25	4.1%	13.3	2.61
AC22_50_70	Limestone	4.5% PG 67–24	4.4%	11.2	2.62
AC32_50_70	Limestone	4.3% PG 67–24	3.2%	–	2.64
SMA11_70_100	Diabase	6.5% PG 63–24	3.6%	7.2	2.54
SMA11_PmB45_80_65	Diabase	6.5% PG 76–25	3.4%	9.9	2.52

where the parameter meanings are the same as those described above.

2.3. Model development and implementation

2.3.1. An introduction to CatBoost approach

A Machine Learning technique that in recent years has shown to be successful in addressing both regression and classification problems is Gradient Boosting. It consists in producing an ensemble predictive model by performing gradient descent in a functional space and demonstrates how a strong predictor can be obtained starting from the greedy iterative combination of several weak models or base predictors (typically decision trees [34]). This methodology has proven to be competitive in several fields, being able to handle heterogeneous features, noisy data, and complex dependencies [35–38].

The experimental dataset can be represented as $D = \{(\boldsymbol{x}_k, y_k)\}_{k=1\cdots n}$, having assumed $\boldsymbol{x}_k = (x_k^1, \cdots, x_k^m)$ as a random vector of the m features and $y_k \in \mathbb{R}$ as a target. The supervised training phase of the model is aimed at training a function $F : \mathbb{R}^m \to \mathbb{R}$ that minimizes the expected loss $\mathscr{L}$ defined as $\mathscr{L}(F) := L(y, F(\boldsymbol{x}))$, with L a smooth loss function.

The gradient boosting procedure improves the prediction of y by means of a function F^{t+1}:

$$F^{t+1} = F^t + \alpha \bullet h^{t+1} t = 0, 1, \cdots \tag{5}$$

α is defined as step size and h^{t+1} as base predictor, selected from a set of functions H in order to minimize the loss function.

$$h^{t+1} = \underset{h \in H}{argmin} \mathscr{L}(F^t + h) = \underset{h \in H}{argmin} \mathbb{E} L(y, F^t(\boldsymbol{x}) + h(\boldsymbol{x})) \tag{6}$$

Either by means of the Taylor approximation or by using the negative gradient step approach:

$$h^{t+1} = \underset{h \in H}{argmin} \mathbb{E}(-g^t(\boldsymbol{x}, y) - h(\boldsymbol{x}))^2 \tag{7}$$

having defined: $-g^t(\boldsymbol{x}, y) := \frac{\partial L(y,s)}{\partial s}|_{s=F^t(x)}$.

However, Gradient Boosting has two major critical issues called prediction shift and target leakage [39]. For this reason, within the present work it was decided to use the innovative CatBoost algorithm instead of the standard gradient boosting. In CatBoost, the basic predictors are oblivious decision trees [40,41], also referred to as decision tables [42]. The term oblivious means that the same split criterion is used for the whole tree level, resulting in a more balanced architecture, less prone to overfitting, and significantly faster during the test phase. Furthermore, by means of a modification of the standard Gradient Boosting algorithm called Ordered Boosting [39], CatBoost avoids prediction shift and target leakage. It allows model's generalization capabilities to be improved, outperforming the current best implementations of gradient boosted decision trees [30], namely XGBoost [43] and LightGBM [44]. Finally, CatBoost also allows categorical features to be automatically handled and processed [45].

2.3.2. Model implementation

A careful identification of CatBoost model relevant hyperparameters, i.e., those parameters whose values are set before the learning process begins, is necessary in order to optimize its performance. Once the

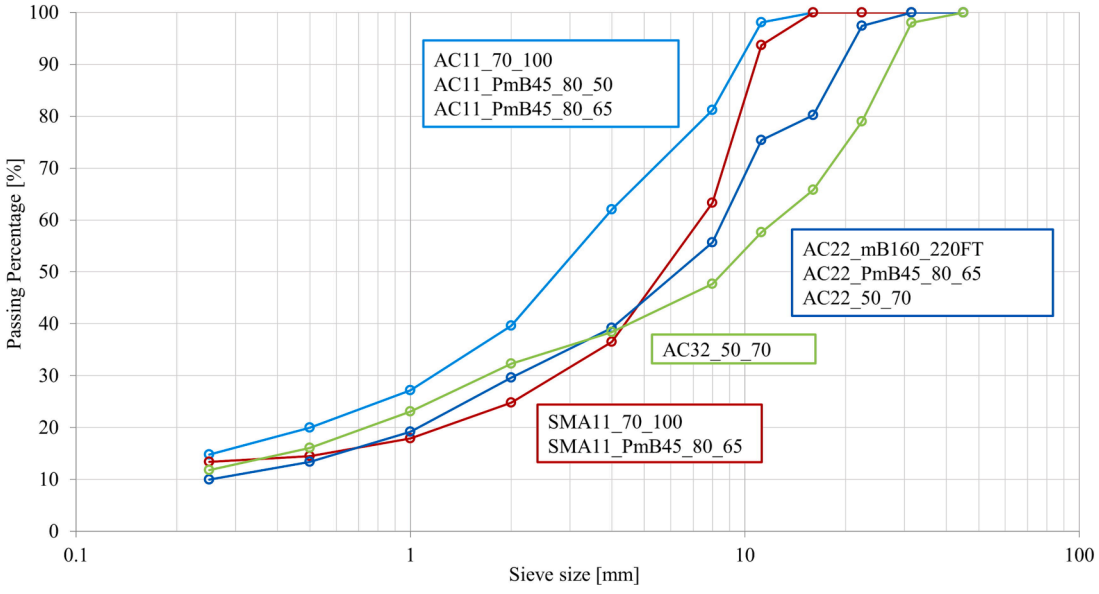

Fig. 2. Grading curves of the mixtures.

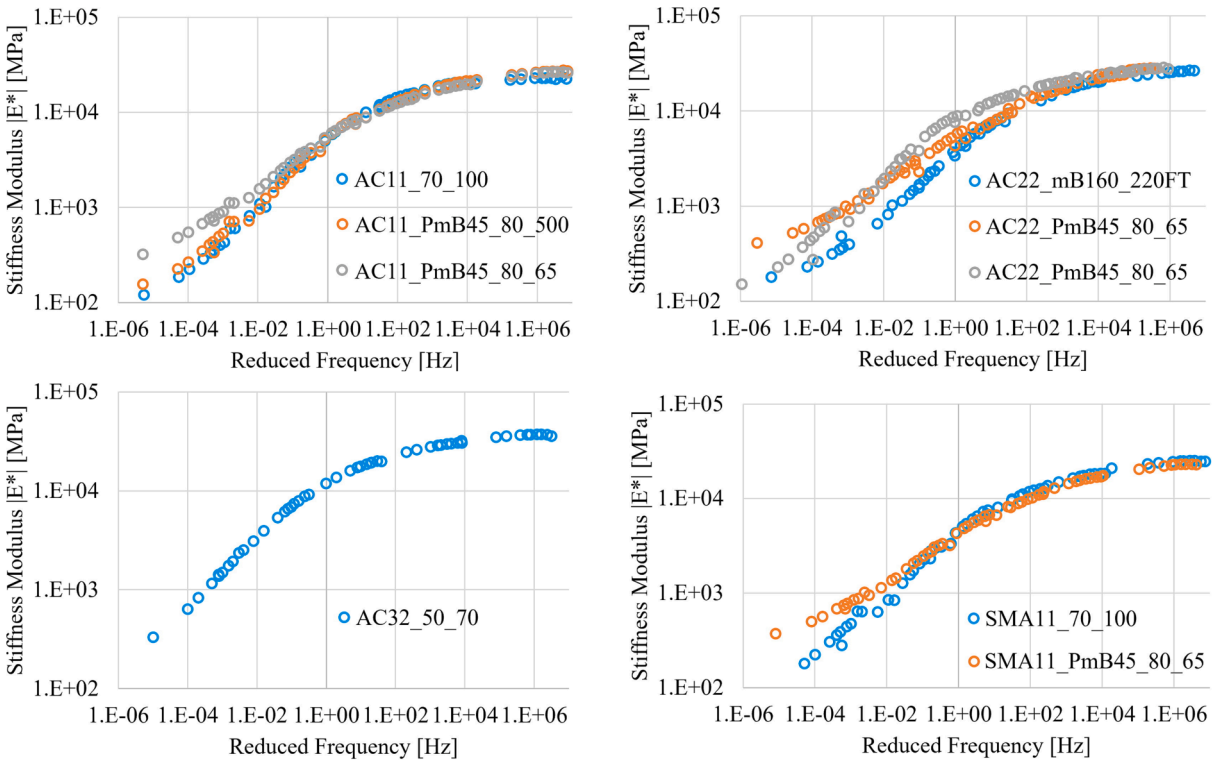

Fig. 3. Master curves for AC11 (up-left), AC22 (up-right), AC32 (down-left), and SMA11 (down-right) mixtures referred to the temperature of 15 °C.

training algorithm has been assigned, the fine-tuning process aims to define the optimal training parameters along with the best model topology. The first major parameter to be defined is the number of iterations, which represents the maximum number of trees that can be built trying to solve the considered machine learning problem. Another significant parameter is the maximum depth, defined as the maximum number of splits. This value must be chosen carefully because too low maximum depth allows for fast but inaccurate modelling, whereas too high one may result in an accurate model probably prone to overfitting [46]. Such phenomenon was addressed by randomly selecting training, validation and test subsamples and by implementing a 5-fold cross-validation technique. In addition, an overfitting detection method was also implemented in order to stop the training if overfitting occurs. Even before each new tree is built, CatBoost algorithm determines how many iterations have occurred since the one with the optimal loss function value. If this number exceeds what has been set as the upper limit for the overfitting detector, training is stopped. Such threshold has been set at its default value, i.e., 20. Lastly, an additional parameter that must be fine-tuned is the learning rate, used for reducing the gradient step during the training. An exhaustive grid search was carried out in Python 3.8.5 whose search domains are described in detail in Table 3. The optimal hyperparameters set was determined on the basis of the loss function value. In the present study, since two variables are simultaneously predicted, MultiRMSE was implemented as loss function (Equation 8). The best performing model, namely the one that achieved the lowest value of MultiRMSE, has a number of iterations equal to 5000, max depth equal to 6 and learning rate equal to 0.05.

$$MultiRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\sum_{d=1}^{D}\left(y_{T_{i,d}} - y_{P_{i,d}}\right)^2} \quad (8)$$

Table 2
Master Curves shift factors ($\log a(T)$).

Mixture	Temperature [°C]						
	45	30	20	15	10	0	–15
AC11_70_100	– 9.8	– 5.1	– 1.8	–	1.8	5.7	12.1
AC11_PmB45_80_50	– 9.9	– 5.2	– 1.8	–	1.8	5.7	12.2
AC11_PmB45_80_65	– 9.9	– 5.2	– 1.8	–	1.8	5.7	12.2
AC22_mB160_220FT	– 9.5	– 5.0	– 1.7	0.0	–	5.5	11.7
AC22_PmB45_80_65	– 10.4	– 6.3	– 3.2	–	0.0	3.5	9.2
AC22_50_70	– 11.4	– 6.8	– 3.5	–	0.0	3.8	10.0
AC32_50_70	– 9.2	– 4.8	– 1.7	0.0	–	5.3	11.3
SMA11_70_100	– 9.9	– 5.2	– 1.8	–	1.8	5.7	12.2
SMA11_PmB45_80_65	– 9.4	– 4.9	– 1.7	–	1.8	5.5	11.6

2.3.3. *Dataset and input/output variables*

CatBoost approach was used to analyze the outcomes of the experimental campaign and create a model that could simultaneously predict both the dynamic modulus and the phase angle of 9 different AC mixtures. For each mixture, three replicates of each specimen were tested at 6 testing temperatures (for two mixtures the testing temperatures were actually 7) and 10 loading frequencies thus generating 1680 observations. However, data averaged over the 3 replicates were assigned as model inputs, resulting in a total of 560 observations for each variable. The input variables implemented within the model refer to the particle size characteristics, physical properties and test conditions, in line with Witczak-Fonseca and Witczak 1-37A empirical models (Equations 2 and 4). Specifically, they have been described within the first 12 rows of Table 4: $\rho_{3/4}$, $\rho_{3/8}$, $\rho_{No.4}$, $\rho_{No.200}$, AV, $V_{b_{eff}}$, f_l, T, η_b, A, VTS, and a categorical variable distinguishing the aggregate nature (limestone or diabase). The output variables refer to the mixtures' mechanical behaviour and are represented by the last 2 rows of Table 3, namely E^* and φ.

Table 3
Summary of Grid Search.

Feature	Grid	Selected Value
Number of iterations	1000, 5000, 10,000	5000
Max depth	4, 5, 6	6
Learning Rate	0.1, 0.05, 0.01	0.05
k-fold Cross-validation	–	5
Overfitting Detector	–	20
Loss Function	–	MultiRMSE

Table 4
Statistical description of the variables considered in the CatBoost model.

Variable	Description	U.M.	Count	Mean	Std Dev
Categorical	Aggregate nature	–	560	–	–
$\rho_{3/4}$	Retained on the ¾ inch sieve	%	560	7.2	9.1
$\rho_{3/8}$	Retained on the $^3/_8$ inch sieve	%	560	25.9	12.5
$\rho_{No.4}$	Retained on the No. 4 sieve	%	560	50.5	11.2
$\rho_{No.200}$	Passing the No. 200 sieve	%	560	7.9	1.4
AV	Air voids in the mix, by volume	%	560	3.7	0.4
$V_{b_{eff}}$	Effective bitumen content, by volume	%	560	13.3	1.9
f_l	Load frequency	Hz	560	13.0	12.6
T	Temperature	°C	560	14.4	19.9
η_b	Bitumen viscosity	10^6 Poise	560	213913.3	520927.3
A	Regression constant	–	560	10	0.5
VTS	Regression constant	–	560	−3.3	0.2
$\lvert E^* \rvert$	Asphalt mixture dynamic modulus	MPa	560	10859.5	9607.3
φ	Phase angle	°	560	21.8	13.8

2.3.4. Model evaluation

Six different goodness-of-fit metrics were selected and implemented in order to evaluate the performance of the developed CatBoost model as accurately as possible. Specifically, they included: mean absolute error (MAE), mean absolute percentage error (MAPE), mean squared error (MSE), root mean squared error (RMSE), Pearson correlation coefficient (R), and the coefficient of determination (R^2). Their description along with their mathematical formulation have been reported in Table 5 with the following meaning of the terms: y_{T_i} and y_{P_i} represent the i-th target and prediction, respectively, n is the total number of observations, whereas μ and σ represent the mean and standard deviation values, respectively.

3. Results

3.1. Relationships between variables

To preliminarily explore the strength of the relationships between variables investigated, Pearson correlation [47] has been used. Pearson correlation coefficient (R) between two variables always ranges between $+1$ and -1: the $+$ sign stands for a positive correlation (if one variable increases, the other one increases as well) whereas the $-$ sign stands for a negative correlation (if one variable increases, the other one decreases). The absolute value size is a measure of the relationship's strength. The extremes of the range represent a perfect correlation, which means that it is possible to determine the exact value of one variable starting from the value of the other one. On the other hand, R equal to 0 means that there is no relationship between the two variables: knowing the value of one variable provides no help in predicting the value of the other one. The R values calculated for each variable pair is reported in a correlation matrix (Fig. 4). For modelling purposes, it is useful to notice that the dynamic modulus has a strong negative correlation with temperature [$r = -.90, n = 560, p < .0005$] and a strong positive correlation with bitumen viscosity [$r = +.63, n = 560, p < .0005$]. In contrast, phase angle has a strong positive correlation with temperature [$r = +.88, n = 560, p < .0005$] and a strong negative correlation with bitumen viscosity [$r = -.53, n = 560, p < .0005$].

3.2. CatBoost modelling results

3.2.1. Model training

Before being used as input to the model, each feature has been normalized. Each value of a specific variable was mapped to the range [0, + 1] whose lower and upper extremes correspond to the minimum and maximum values assumed by the variable, respectively (Equation 9). This data pre-processing practice is well established in ML as it helps to increase both computational speed and models efficiency.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{9}$$

Combined with normalization, a 5-fold cross-validation was implemented in order to reduce the bias in model's predictive performance. The overall dataset was randomly partitioned into two distinct subsets: the training one composed by the 80% of the total data (448 observations) and the testing one composed by the 20% of the total data (112 observations). Such a partitioning choice was in line with what reported in the relevant literature [48]. Testing set was left unchanged allowing the model to be tested on an unused portion of data, whereas training set was rearranged as follows: it was split into five different folds so that four of them could be used to train the model and the remaining one to validate it. This training process was repeated 5 times iteratively so that each time the validation fold could be varied and a performance score (i.e., a validation score) could be tracked. The average of the five validation scores returned the actual performance of the model for each

Table 5
Goodness-of-fit metrics selected.

Metric	Description	Formulation
MAE	Measures the difference between the observed and the predicted values by averaging the absolute difference over the data set.	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_{T_i} - y_{P_i} \rvert$
MAPE	Measures the error as a percentage and can be calculated by applying a slight modification to the MAE and multiplying by 100 to obtain a percentage score.	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert \frac{y_{T_i} - y_{P_i}}{y_{T_i}} \right\rvert \times 100$
MSE	Measures the difference between the observed and the predicted values, by squaring the mean difference over the data set.	$\frac{1}{n}\sum_{i=1}^{n}\left(y_{T_i} - y_{P_i}\right)^2$
RMSE	Measures the error rate as the square root of the MSE.	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{T_i} - y_{P_i}\right)^2}$
R	Measures any linearity relationship between the observed and the predicted values.	$\frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{y_{T_i} - \mu_{y_{T_i}}}{\sigma_{y_{T_i}}}\right)\left(\frac{y_{P_i} - \mu_{y_{P_i}}}{\sigma_{y_{P_i}}}\right)$
R^2	Measures the coefficient of how well predicted values fit the observed ones. Values range from 0 to 1 and are interpreted as percentages: the higher the value, the better the model.	$1 - \frac{\sum_{i=1}^{n}\left(y_{T_i} - y_{P_i}\right)^2}{\sum_{i=1}^{n}\left(y_{T_i} - \mu_{y_{T_i}}\right)^2}$

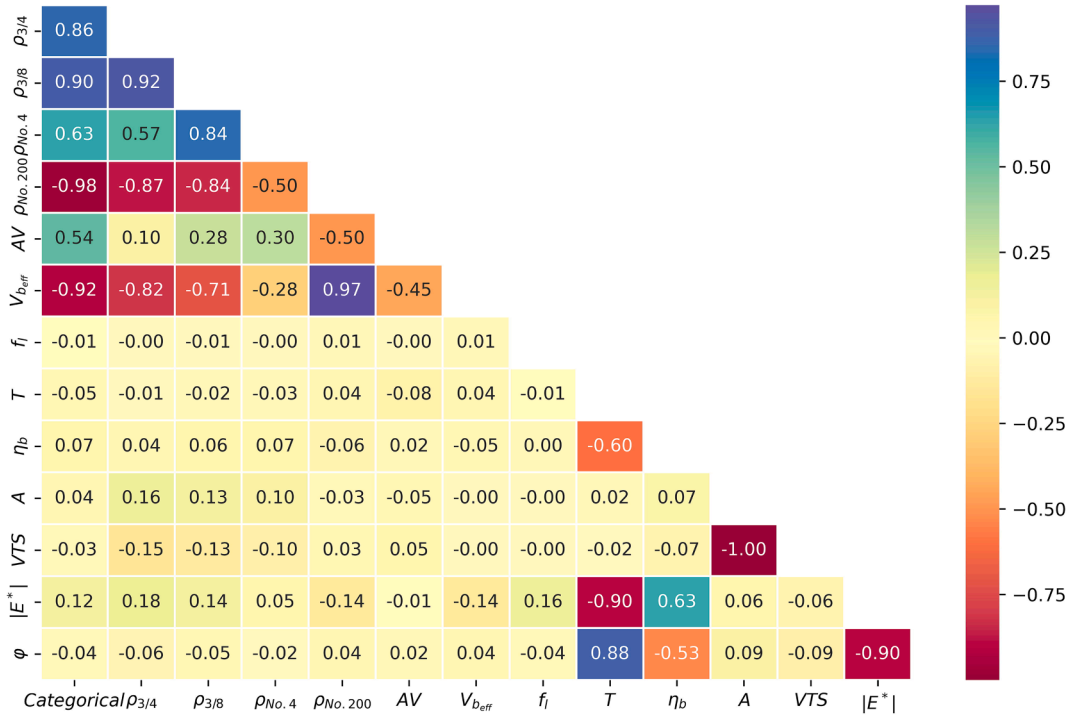

Fig. 4. Pearson Correlation Matrix.

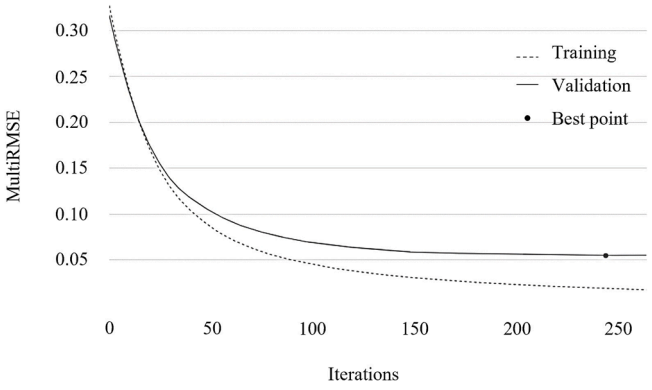

Fig. 5. Model training process with best iteration (black marker).

Table 6
Predictive performance of different models.

$\|E^*\|$ test vector	Witczak-Fonseca	Witczak 1-37A	SoA-ANN	CatBoost
MAE	3616.93	2146.30	640.05	552.21
MAPE	67.58	41.25	21.60	12.04
MSE	2.03×10^7	8.43×10^6	6.88×10^5	9.85×10^5
RMSE	4506.01	2904.13	829.21	992.36
R	0.9370	0.9694	0.9965	0.9952
R^2	0.7904	0.9129	0.9929	0.9898
φ test vector	Witczak-Fonseca	Witczak 1-37A	SoA-ANN	CatBoost
MAE	–	–	1.05	1.19
MAPE	–	–	10.90	7.84
MSE	–	–	2.17	3.68
RMSE	–	–	1.47	1.92
R	–	–	0.9942	0.9908
R^2	–	–	0.9884	0.9804

training iteration. The graphical trend of the entire training process has been shown in Fig. 5. At the beginning of the process, it can be appreciated that the error is quite high, and it gradually decreases as iterations proceed. Training stops when convergence is reached, i.e., no further reduction in the loss function is observed. However, as the training loss decreases, there should not be a concurrent increase in validation loss. Such a pattern would suggest the occurring of the overfitting phenomenon. By looking at Fig. 5, it can be appreciated how both training and validation curves decrease as iterations proceed. Furthermore, the overfitting detector set at 20 iterations allows the best validation point to be identified: beyond this point no significant decrease in validation loss ($\left|F^{t+1} - F^t\right| \leq 1 \times 10^{-4}$) can be observed. Training results can be summarized as follows: with respect to the dynamic modulus, *MAE*, *RMSE*, and R^2 were equal to 279.68 MPa, 415.12 MPa, and 0.998, respectively; with respect to the phase angle, *MAE*, *RMSE*, and R^2 were equal to 0.69°, 0.93°, and 0.995, respectively. The parameters related to the best model configuration have been fixed and kept constant before proceeding to the next testing phase.

3.2.2. *Model testing*

The same aforementioned performance metrics, along with *MAPE*, *MSE*, and *R*, have been used to describe the CatBoost model testing performance. Results have been summarized in Table 6.

To obtain further insight into the predictive capabilities of the developed methodology, Fig. 6 shows the comparisons between the experimental targets and the outputs predicted by the CatBoost model for both variables, namely dynamic modulus and phase angle. In these histogram graphs, test set values are shown in black whereas the respective CatBoost predictions are shown in gray. The horizontal axis is referred to as Test ID and serves to identify each of the 112 $|E^*| - \varphi$ pairs that make up the test set. It can be appreciated that the predictions are very close to the observed experimental data, for both parameters considered. Such findings are significant from an engineering perspective because they highlight the reliability of the developed CatBoost model along with the goodness of its predictions. Furthermore, they demonstrate how accurately and simultaneously this model can predict two fundamental parameters that characterize the mechanical behaviour of several different mixtures. To get an even deeper understanding of the model's predictive performance, two regression plots are shown, one related to the dynamic modulus and the other related to the phase

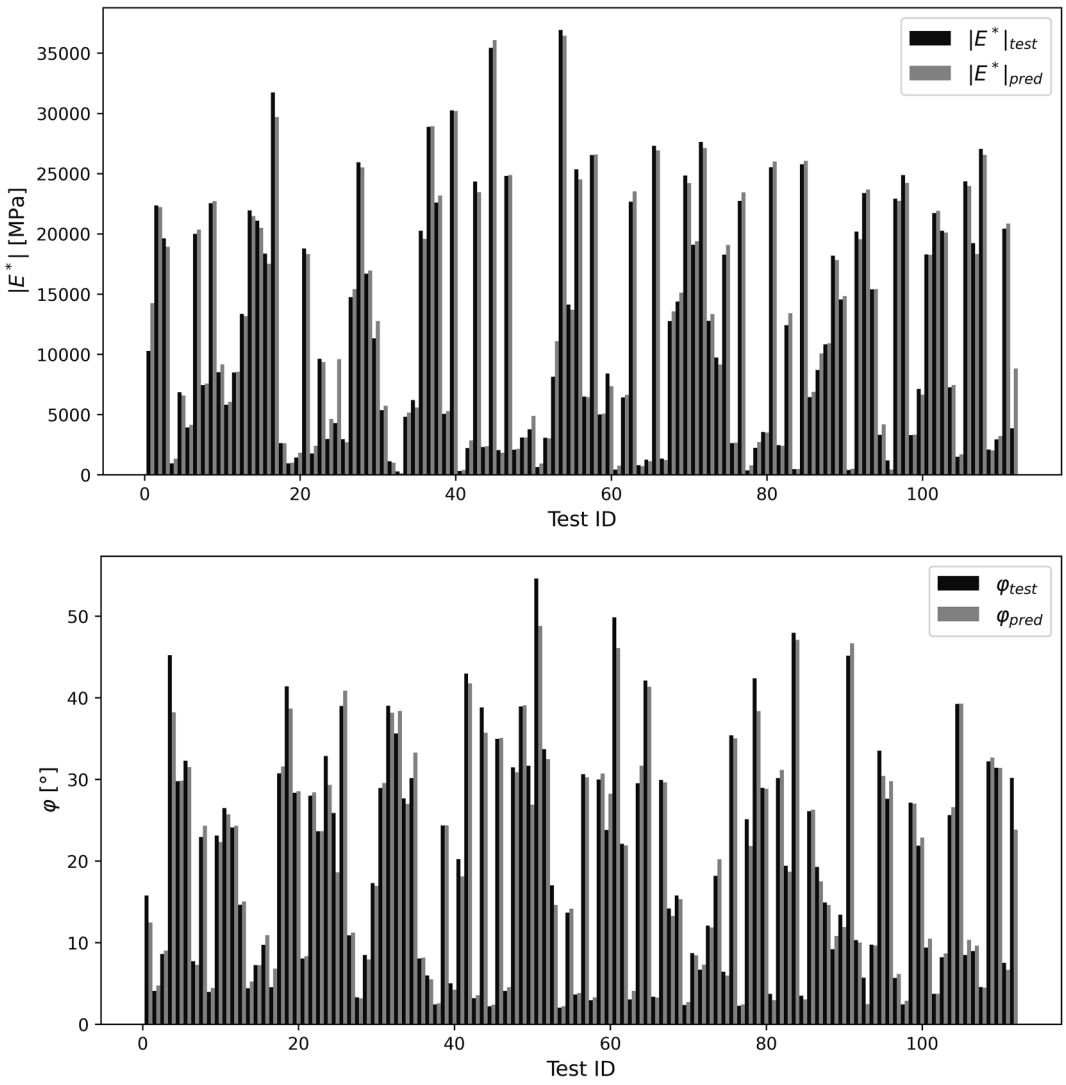

Fig. 6. Observed and CatBoost-predicted dynamic modulus (up) and phase angle (down) data.

angle (Fig. 7). The 45-degree blue solid line represents the line-of-equality, namely the line that reflects 100% accurate predictions whereas light blue circles represent CatBoost predictions. It can be observed that the predictions do not strongly differ from the line-of-equality, confirming the remarkable results described by Pearson correlation coefficients equal to 0.9952 and 0.9909 for $|E^*|$ and φ, respectively. In addition, all predictions of both variables are within the 99% confidence bounds, represented by the green solid lines. A frequency analysis was performed to determine whether the CatBoost model systematically overpredicted or underpredicted the respective values of dynamic modulus and phase angle.

3.2.3. *Features importance*

The estimation of a given feature's influence on the predictions of a ML model is often challenging. For this reason, feature importance calculation algorithm has been introduced in Python 3.8.5 to determine how much CatBoost predictions change on average if the feature values change. The higher the importance value, the greater the change in prediction value if this feature is modified. Furthermore, feature importance values were normalized so that the sum of them could be equal to 100%. This is feasible because importance values are always non-negative. Fig. 8 shows the results of the analysis performed and returns feature importance as a list of features (ID − importance) pairs ordered by the importance magnitude. It can be noticed that temperature (41.86%), binder viscosity (34.74%), and loading frequency (8.22%) are more critical than empirical regression constants, volumetric properties, and aggregates nature (15.18% all together) in predicting mechanical behaviour parameters such as dynamic modulus and phase angle.

To obtain further useful information about CatBoost model outputs, also a SHAP (SHapley Additive exPlanations) analysis has been performed. Such technique was developed to understand why a model makes certain predictions, thus focusing on the concept of prediction interpretability along with the more commonly described concept of prediction accuracy [49]. Let f be the CatBoost model to be explained and g the explanation model, assume that $f(x)$ is the prediction based on the input x. M represents the set of all features, $z' \in \{0,1\}^M$ is equal to 1 when a feature is observed (otherwise it is 0), $|z'|$ is the number of non-zero entries in z', $z' \subseteq x$ represents all z' vectors where the non-zero entries are a subset of the non-zero entries in x, and $z' \backslash i$ denotes setting $z' = 0$ [49]. The contribution provided by the i-th feature (ϕ_i) to predict the model output is determined as follows:

$$\phi_i = \sum_{z' \subseteq x} \frac{|z'|!(M - |z'| - 1)!}{M!} [f_x(z') - f_x(z' \backslash i)] \tag{9}$$

g is defined as a linear function of binary features [50] according to the following additive feature attribution method:

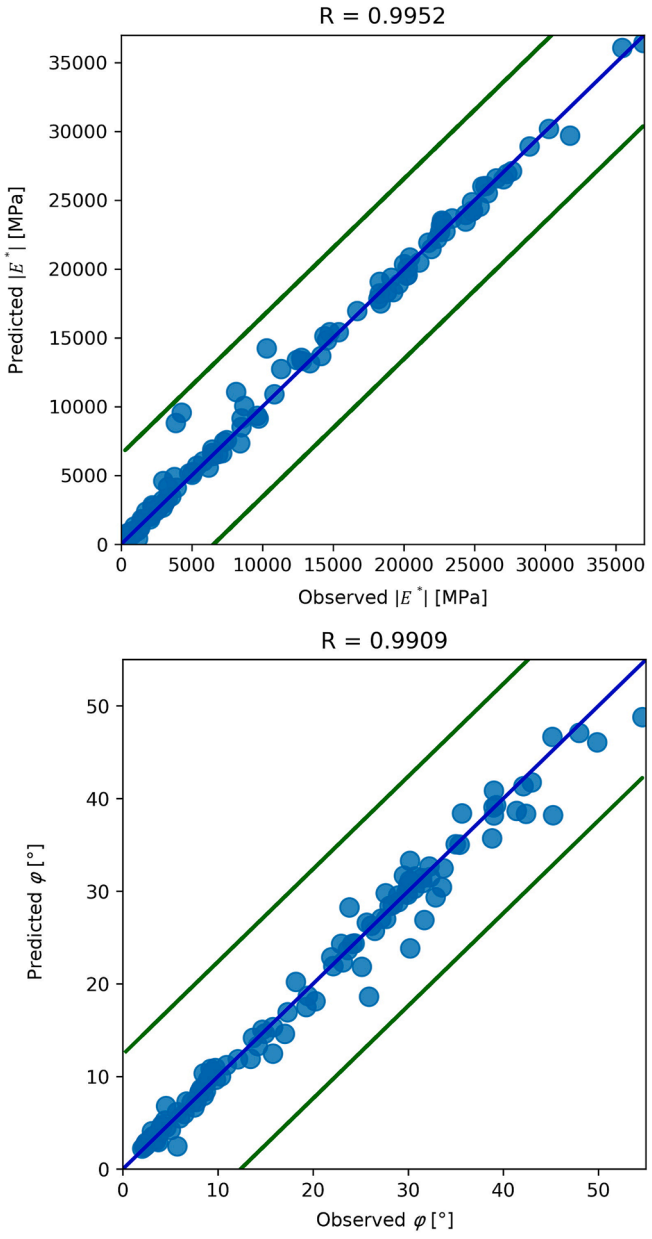

Fig. 7. CatBoost model regression plots for both $|E^*|$(up) and φ (down); blue solid line represents line-of-equality, light blue circles represent predictions, whereas green solid lines indicate 99% confidence level. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z_i' \tag{10}$$

ϕ_i values are known as Shapley values [51] and assign to each feature an importance value for a particular prediction. Two separate SHAP analyses have been performed: one for each of the two targets, namely dynamic modulus and phase angle. The results are shown in the SHAP summary plots (Fig. 9). These beeswarm diagrams rank the features from top to bottom in a descending order of importance when predicting a given output.

The temperature has the highest importance in predicting both dynamic modulus and phase angle, followed by bitumen viscosity and loading frequency, confirming what has already been observed in Fig. 8. These features are then followed by the two empirical regression constants. Volumetric properties and aggregate nature are found to have less importance in the prediction of both the outputs. It can be observed that the importance order of these last features is slightly different from that shown in Fig. 8 as their importance is now specified for the prediction of the two respective outputs. Nevertheless, in addition to the feature importance, SHAP summary plots add further information as suggested by the use of colours: the feature effect. Considering temperature variable and the summary plot referring to dynamic modulus, it can be observed that high T values (displayed in red) result in negative SHAP values and consequently in lower $|E^*|$ predictions. This trend is gradually reversed moving toward the horizontal axis where low T values (displayed in blue) result in positive SHAP values and consequently in higher $|E^*|$ predictions. This tendency helps to understand the relationship between the different features and the specific target. In general, moving from left to right, a smooth transition from blue to red stands for a direct proportionality relationship between the target and the specific feature, whereas a smooth transition from red to blue stands for an inverse proportionality relationship between them. The former can be observed for pairs ($|E^*| - \eta_b$), ($|E^*| - f_l$), and ($\varphi - T$), whereas the latter can be observed for pairs ($|E^*| - T$), ($\varphi - \eta_b$), and ($\varphi - f_l$).

3.2.4. Comparison of the models

A comprehensive comparative analysis was performed in the present study in order to compare the two previously described regression-based models (Witczak-Fonseca and Witczak 1-37A) with state-of-the-art artificial neural network (SoA-ANN) and the decision-tree based CatBoost model. A detailed discussion about SoA-ANN implemented in MATLAB® is not presented herein due to brevity and can be found in Baldo et al. [52]. Both ML models were calibrated based on the same dataset, using the same data pre-processing and resampling techniques in order to make a fair comparison. All the four mentioned models were then compared based on six different goodness-of-fit metrics, as shown in Table 6.

Based on the $|E^*|$ test vector, the following key observations can be deduced. ML models outperform regression-based ones, with CatBoost achieving both the minimum *MAE* (552.21 MPa) and *MAPE* (12.04%). In terms of the other four goodness-of-fit metrics, the comparison between SoA-ANN and CatBoost models is very competitive and returns comparable results, with ANN performing slightly better (*RMSE* equal to 829.21 MPa). Comparing only the two regression-based models, Witczak 1-37A performs significantly better than Witczak-Fonseca with respect of all six-performance metrics. A summary representation of the above-mentioned results can be observed in Fig. 10. However, one point should be emphasized: starting from the same input features, ML models have been trained to simultaneously predict not only the dynamic modulus but also the phase angle. Thereby, they are able to provide a more detailed insight about the fundamental viscoelastic properties of the asphalt mixtures analyzed. Also with respect to phase angle, both ML models' performance are remarkable and similarly comparable (Table 6) with *MAE* of approximately 1°, *MAPE* less than 11%, R and R^2 of around 0.99 and 0.98, respectively.

4. Conclusions

The present study outlines a detailed methodology in order to implement an innovative DT based ML algorithm, namely CatBoost, for the simultaneous prediction of both dynamic modulus and phase angle of nine different AC mixtures. An extensive 4PBT experimental campaign was carried out and the results obtained were used to train and test the developed model. 12 features were used as input and refer to aggregate nature and gradation percentages, air voids content, effective binder content, binder viscosity, loading frequency, and testing temperature. To evaluate model performance, six different goodness-of-fit metrics were used: these included *MAE*, *MAPE*, *MSE*, *RMSE*, R, and R^2. An in-depth sensitivity analysis was also performed to understand model

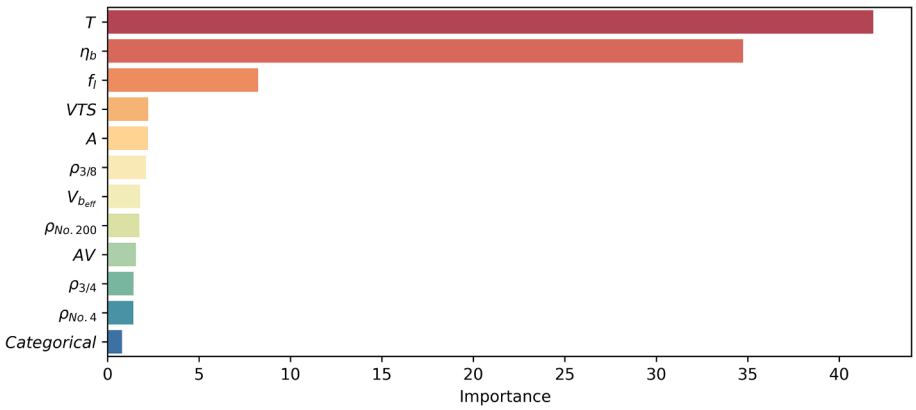

Fig. 8. Feature Importance.

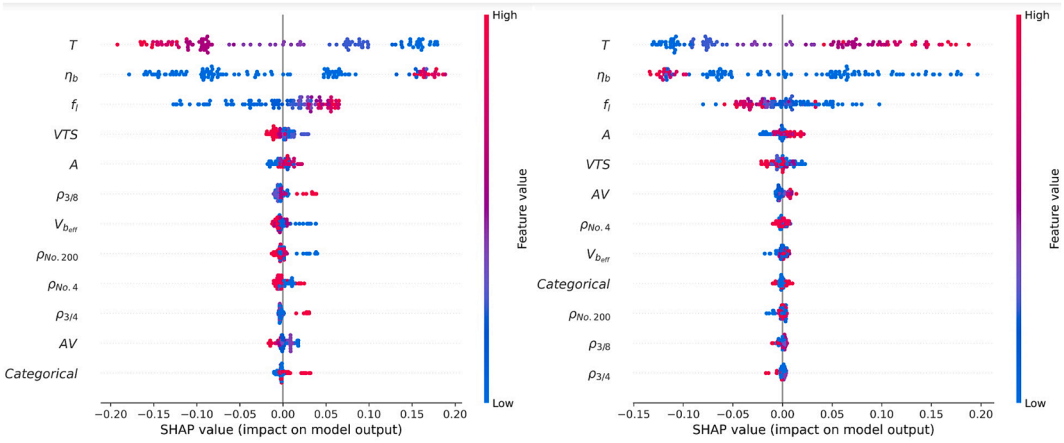

Fig. 9. SHAP summary plot for $|E^*|$ (left) and φ (right).

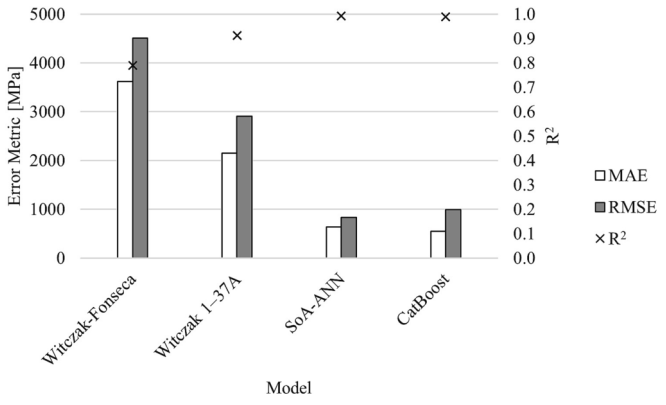

Fig. 10. Graphical representation of $|E^*|$ predictive performance.

functioning and the impact that specific features have on the two different outputs. The performance achieved by the model is very good and accounts for the remarkable accuracy of the predictions made. With regard to $|E^*|$, *MAE*, *MAPE*, *MSE*, *RMSE*, *R*, and R^2 resulted equal to 552.21 MPa, 12.04 %, 9.85×10^5 MPa^2, 992.36 MPa, 0.9952, and 0.9898, respectively. Similarly, with regard to φ, the same performance metrics resulted equal to 1.19°, 7.84 %, 3.68 $^{\circ 2}$, 1.92°, 0.9908, and 0.9804, respectively. Based on these findings, the following conclusions can be drawn:

The variables required as input by the predictive model, can be easily determined during a preliminary mix design procedure or easily derived from them, minimizing laboratory technicians' effort required.

The developed model is capable of simultaneously providing highly accurate predictions of both dynamic modulus and phase angle, avoiding the need for additional expensive and time-consuming laboratory tests to determine them experimentally.

In this way, performance parameters such as $|E^*|$ and φ can be easily estimated and then implemented within the most common asphalt pavement design procedures.

Both ML-based methods outperformed the implemented empirical equations, showing much better performance for all the goodness-of-fit metrics used.

In particular, the CatBoost and SoA-ANN models were competitive and showed roughly comparable results, with ANN performing slightly better according to the *MSE*, *RMSE*, *R*, and R^2 measures. However, the differences in goodness-of-fit metrics are not as marked, but CatBoost model can be easily implemented in Python and this, combined with its high interpretability, make it most preferable.

The outlined methodology was calibrated on data from the above-mentioned experimental campaign. For future developments, it could easily be adopted to analyze larger and more varied datasets in order to obtain an increasingly powerful and high-performance tool capable of making accurate predictions covering an increasingly conditions variety. Finally, CatBoost model has been developed and used to predict AC mixtures' dynamic modulus and phase angle, obtaining remarkable results. However, its flexibility could also inspire many further applications in pavement engineering by allowing even very different phenomena such as moisture sensitivity or permanent deformation and fatigue resistance to be investigated and modelled. Clearly any further application of the developed model different from the one proposed here should be deeply explored by implementing new calibrations and

researching for the best hyperparameters.

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.

Data availability

Data will be made available on request.

References

[1] Y. Ali, M. Irfan, S. Ahmed, S. Khanzada, T. Mahmood, Sensitivity analysis of dynamic response and fatigue behaviour of various asphalt concrete mixtures, Fatigue Fract. Eng. Mater. Struct. 38 (10) (2015) 1181–1193, https://doi.org/10.1111/ffe.12297.

[2] G.R. Chehab, Y.R. Kim, R.A. Schapery, M.W. Witczak, R. Bonaquist, Time-temperature superposition principle for asphalt concrete with growing damage in tension state, J. Assoc. Asph. Paving Technol. 71 (2002) 559–593.

[3] H. Ding, M. Yuan, H. Hu, Establishing prediction master curve of dynamic modulus of asphalt mixture considering randomness of aggregate morphology, Constr. Build. Mater. 294 (2021), 123575, https://doi.org/10.1016/j.conbuildmat.2021.123575.

[4] AASHTO, American Association of State Highway Transportation Officials (AASHTO) Guide for Design of Pavement Structures, 1993, AASHTO1993.

[5] Y. Ali, M. Irfan, S. Ahmed, S. Ahmed, Empirical correlation of permanent deformation tests for evaluating the rutting response of conventional asphaltic concrete mixtures, J. Mater. Civ. Eng. 29 (8) (2017) 04017059, https://doi.org/10.1061/(ASCE)MT.1943-5533.0001888.

[6] O.A. Fonseca, M.W. Witczak, A prediction methodology for the dynamic modulus of in placed aged asphalt mixtures, J. Assoc. Asph. Paving Technol. 65 (1996) 532–572.

[7] M.J. Bari, M.W. Witczak, Development of a new revised version of the Witczak |E*|: Predictive model for hot mix asphalt mixtures, J. Assoc. Asph. Paving Technol. 75 (2006) 381–423.

[8] R. Zhang, J.E. Sias, E.V. Dave, Using mix design information for modelling of fundamental viscoelasticity of asphalt mixtures, Constr. Build. Mater. 329 (2022), 127029, https://doi.org/10.1016/j.conbuildmat.2022.127029.

[9] O. François, H. Di Benedetto, General "2S2P1D" model and relation between the linear viscoelastic behaviours of bituminous binders and mixes, Road Mater. Pavement Des. 4 (2) (2003) 185–224, https://doi.org/10.1080/14680629.2003.9689946.

[10] H. Di Benedetto, F. Olard, C. Sauzéat, B. Delaporte, Linear viscoelastic behaviour of bituminous materials: from binders to mixes, Road Mater. Pavement Des. 5 (2004) 163–202, https://doi.org/10.1080/14680629.2004.9689992.

[11] M. W. Witczak, M. El-Basyouny, S. El-Badawy, Incorporation of the new (2005) E* predictive model in the MEPDG. NCHRP 1-40D Final Report (2007).

[12] J. Li, A. Zofka, I. Yut, Evaluation of dynamic modulus of typical asphalt mixtures in Northeast US region, Road Mater. Pavement Des. 13 (2) (2012) 249–265, https://doi.org/10.1080/14680629.2012.666641.

[13] K.P. Biligiri, K. Kaloush, J. Uzan, Evaluation of asphalt mixtures' viscoelastic properties using phase angle relationships, Int. J. Pavement Eng. 11 (2) (2010) 143–152, https://doi.org/10.1080/10298430903033354.

[14] R. Nemati, E.V. Dave, Nominal property based predictive models for asphalt mixture complex modulus (dynamic modulus and phase angle), Constr. Build. Mater. 158 (2018) 308–319, https://doi.org/10.1016/j.conbuildmat.2017.09.144.

[15] H. Ceylan, K. Gopalakrishnan, S. Kim, Advanced approaches to hot-mix asphalt dynamic modulus prediction, Can. J. Civ. Eng. 35 (7) (2008) 699–707, https://doi.org/10.1139/L08-016.

[16] J. Komba, J. W. Maina, J. K. Anochie-Boating, J. O'Connel, Analytical modeling of visco-elastic behaviour of hot-mix asphalt, in: 31st Southern African Transport Conference, July, Pretoria, South Africa, 2012.

[17] 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, https://doi.org/10.1016/j.conbuildmat.2014.11.011.

[18] P. Cao, F. Jin, D. Feng, C. Zhou, W. Hu, Prediction on dynamic modulus of asphalt concrete with random aggregate modeling methods and virtual physics engine, Constr. Build. Mater. 125 (2016) 987–997.

[19] J. Liu, K. Yan, L. You, P. Liu, K. Yan, Prediction models of mixtures' dynamic modulus using gene expression programming, Int. J. Pavement Eng. 18 (11) (2017) 971–980, https://doi.org/10.1080/10298436.2016.1138113.

[20] M. Svilar, I. Peško, M. Šešlija, Model for estimating the modulus of elasticity of asphalt layers using machine learning, Appl. Sci. 12 (20) (2022) 10536, https://doi.org/10.3390/app122010536.

[21] F. Althoey, M.N. Akhter, Z.S. Nagra, H.H. Awan, F. Alanazi, M.A. Khan, M.F. Javed, S.M. Eldin, Y.O. Özkılıç, Prediction models for marshall mix parameters using bio-inspired genetic programming and deep machine learning approaches: a comparative study, Case Stud. Constr. Mater. 18 (2023), e01774, https://doi.org/10.1016/j.cscm.2022.e01774.

[22] N. Baldo, M. Miani, F. Rondinella, C. Celauro, A machine learning approach to determine airport asphalt concrete layer moduli using heavy weight deflectometer data, Sustainability 13 (2021) 8831, https://doi.org/10.3390/su1316883.

[23] N. Baldo, M. Miani, F. Rondinella, J. Valentin, P. Vackcová, E. Manthos, Stiffness data of high-modulus asphalt concretes for road pavements: predictive modeling by machine-learning, Coatings 12 (2022) 54, https://doi.org/10.3390/coatings12010054.

[24] A. Karbassi, B. Mohebi, S. Rezaee, P. Lestuzzi, Damage prediction for regular reinforced concrete buildings using the decision tree algorithm, Comput. Struct. 130 (2014) 46–56, https://doi.org/10.1016/j.compstruc.2013.10.006.

[25] C.Q.X. Poh, C.U. Ubeynarayana, Y.M. Goh, Safety leading indicators for construction sites: a machine learning approach, Autom. Constr. 93 (2018) 375–386, https://doi.org/10.1016/j.autcon.2018.03.022.

[26] A. Upadhya, M.S. Thakur, P. Sihag, R. Kumar, S. Kumar, A. Afeeza, A. Afzal, C. A. Saleel, Modelling and prediction of binder content using latest intelligent machine learning algorithms in carbon fiber reinforced asphalt concrete, Alexandria Eng. J. 65 (2023) 131–149, https://doi.org/10.1016/j.aej.2022.09.055.

[27] S.K. Baduge, S. Thilakarathna, J.S. Perera, M. Arashpour, P. Sharafi, B. Teodosio, A. Shringi, P. Mendis, Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications, Autom. Constr. 141 (2022) 104440.

[28] D. Daneshvar, A. Behnood, Estimation of the dynamic modulus of asphalt concretes using random forests algorithm, Int. J. Pavement Eng. 23 (2) (2022) 250–260, https://doi.org/10.1080/10298436.2020.1741587.

[29] Y. Ali, F. Hussain, M. Irfan, A.S. Buller, An eXtreme Gradient Boosting model for predicting dynamic modulus of asphalt concrete mixtures, Constr. Build. Mater. 295 (2021), 123642, https://doi.org/10.1016/j.conbuildmat.2021.123642.

[30] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artif. Intell. Rev. 54 (2021) 1937–1967, https://doi.org/10.1007/s10462-020-09896-5.

[31] SIST EN 13108: Bituminous Mixtures – Material Specifications, European Committee for Standardization (2016).

[32] SIST EN 12697-33: Bituminous Mixtures – Test Methods for hot mix asphalt – Part 33: Specimen prepared by roller compactor, European Committee for Standardization (2019).

[33] SIST EN 12697-26: Bituminous Mixtures – Test Methods for hot mix asphalt – Part 26: Stiffness, European Committee for Standardization (2019).

[34] M. Kearns, L. Valiant, Cryptographic limitations on learning boolean formulae and finite automata, J. ACM 41 (1) (1994) 67–95, https://doi.org/10.1145/174644.174647.

[35] R. Caruana, A. Niculescu-Mizil. An empirical comparison of supervised learning algorithms. In Proceedings of the 23rd international conference on Machine learning. ACM (2006) 161–168. https://doi.org/10.1145/1143844.1143865.

[36] B.P. Roe, H.J. Yang, J. Zhu, Y. Liu, I. Stancu, G. McGregor, Boosted decision trees as an alternative to artificial neural networks for particle identification, Nucl. Instrum. 543 (2) (2005) 577–584, https://doi.org/10.1016/j.nima.2004.12.018.

[37] Q. Wu, C.J. Burges, K.M. Svore, J. Gao, Adapting boosting for information retrieval measures, Inf. Retr. J. 13 (3) (2010) 254–270, https://doi.org/10.1007/s10791-009-9112-1.

[38] Y. Zhang, A. Haghani, A gradient boosting method to improve travel time prediction, Transp. Res. Part C Emerg. Technol. 58 (2015) 308–324, https://doi.org/10.1016/j.trc.2015.02.019.

[39] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, A. Gulin, Catboost: unbiased boosting with categorical features. In: Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R (eds) Adv. Neural Inf. Process. Syst. 31 (2018) 6638–6648. https://doi.org/10.48550/arXiv.1706.09516.

[40] M. Ferov, M. Modrý, Enhancing lambdamart using oblivious trees. arXiv 2016, arXiv:1609.05610. https://doi.org/10.48550/arXiv.1609.05610.

[41] A. Gulin, I. Kuralenok, D. Pavlov, Winning the transfer learning track of yahoo!'s learning to rank challenge with yetirank. In Proceedings of the Learning to Rank Challenge 14(2011) 63–76.

[42] Y. Lou, M. Obukhov, Bdt: Gradient boosted decision tables for high accuracy and scoring efficiency. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining (2017) 1893–1901. https://doi.org/10.1145/3097983.3098175.

[43] T. Chen, C. Guestrin. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2016) 785–794.

[44] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T. Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Curran Associates Inc., Red Hook, NY, USA, (2017) 3149–3157.

[45] A. V. Dorogush, V. Ershov, A. Gulin. "CatBoost: gradient boosting with categorical features support." arXiv preprint arXiv:1810.11363. (2018) https://doi.org/10.48550/arXiv.1810.11363.

[46] G. James, D. Witten, T. Hastie, R. Tibshirani, An introduction to statistical learning with applications in R, Springer, New York, USA, 2013.

[47] J. Pallant, SPSS survival manual: a step by step guide to data analysis using IBM SPSS, Routledge. (2002), https://doi.org/10.4324/9781003117452.

[48] F. Hussain, Y. Ali, M. Irfan, M. Ashraf, S. Ahmed, A data-driven model for phase angle behaviour of asphalt concrete mixtures based on convolutional neural network, Constr. Build. Mater. 269 (2021), 121235, https://doi.org/10.1016/j.conbuildmat.2020.121235.

[49] S.M. Lundberg, S. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017), https://doi.org/10.48550/arXiv.1705.07874.

[50] A.B. Parsa, A. Movahedi, H. Taghipour, S. Derrible, A.K. Mohammadian, Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis, Accid. Anal. Prev. 136 (2020), 105405, https://doi.org/10.1016/j.aap.2019.105405.
[51] L. S. Shapley. A value for n-person games. In: Contributions to the Theory of Games 2(28) (1953) 307–317.
[52] N. Baldo, M. Miani, F. Rondinella, E. Manthos, J. Valentin, Road pavement asphalt concretes for thin wearing layers: a machine learning approach towards stiffness modulus and volumetric properties prediction, Period. Polytech.: Civ. Eng. 66 (4) (2022) 1087–1097, https://doi.org/10.3311/PPci.19996.