

## Article

# Stiffness Data of High-Modulus Asphalt Concretes for Road Pavements: Predictive Modeling by Machine-Learning

Nicola Baldo <sup>1,\*</sup>, Matteo Miani <sup>1</sup>, Fabio Rondinella <sup>1</sup>, Jan Valentin <sup>2</sup>, Pavla Vackcová <sup>2</sup>  
and Evangelos Manthos <sup>3</sup>

<sup>1</sup> Polytechnic Department of Engineering and Architecture (DPIA), University of Udine, Via del Cottonificio 114, 33100 Udine, Italy; matteo.miani@phd.units.it (M.M.); fabio.rondinella@phd.units.it (F.R.)

<sup>2</sup> Faculty of Civil Engineering, Czech Technical University, Thákurova 7, 166 29 Prague, Czech Republic; jan.valentin@fsv.cvut.cz (J.V.); pavla.vackova@fsv.cvut.cz (P.V.)

<sup>3</sup> Department of Civil Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; emanthos@civil.auth.gr

\* Correspondence: nicola.baldo@uniud.it; Tel.: +39-0432-558-745

**Abstract:** This paper presents a study about a Machine Learning approach for modeling the stiffness of different high-modulus asphalt concretes (HMAC) prepared in the laboratory with harder paving grades or polymer-modified bitumen which were designed with or without reclaimed asphalt (RA) content. Notably, the mixtures considered in this study are not part of purposeful experimentation in support of modeling, but practical solutions developed in actual mix design processes. Since Machine Learning models require a careful definition of the network hyperparameters, a Bayesian optimization process was used to identify the neural topology, as well as the transfer function, optimal for the type of modeling needed. By employing different performance metrics, it was possible to compare the optimal models obtained by diversifying the type of inputs. Using variables related to the mix composition, namely bitumen content, air voids, maximum and average bulk density, along with a categorical variable that distinguishes the bitumen type and RAP percentages, successful predictions of the Stiffness have been obtained, with a determination coefficient ( $R^2$ ) value equal to 0.9909. Nevertheless, the use of additional input, namely the Marshall stability or quotient, allows the Stiffness prediction to be further improved, with  $R^2$  values equal to 0.9938 or 0.9922, respectively. However, the cost and time involved in the Marshall test may not justify such a slight prediction improvement.

**Keywords:** asphalt concretes; road pavements; polymer modified bitumen; recycled asphalt pavement; stiffness modulus; Marshall stability; machine learning modeling; shallow neural networks; Bayesian optimization; data augmentation



**Citation:** Baldo, N.; Miani, M.; Rondinella, F.; Valentin, J.; Vackcová, P.; Manthos, E. Stiffness Data of High-Modulus Asphalt Concretes for Road Pavements: Predictive Modeling by Machine-Learning. *Coatings* **2022**, *12*, 54. <https://doi.org/10.3390/coatings12010054>

Academic Editor: Qiao Dong

Received: 16 December 2021

Accepted: 31 December 2021

Published: 3 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

High modulus asphalt concretes (HMAC) were first designed and used in France nearly 40 years ago [1]. HMAC or interchangeable term EME (*Enrobé a Module Élevé*) is a special type of asphalt mixture (asphalt concrete) with a strong aggregate structure, slightly higher amount of bituminous binder and high stiffness. This type of mixture is used in both heavy-duty and structural rehabilitation projects where it is desirable to minimize the impact of grade change, yet still, ensure pavement longevity.

Apart from high stiffness modulus, HMAC is defined by high resistance to fatigue mainly if stiffness is well balanced with fatigue parameters, high strength characteristics, minimized permanent deformation occurrence, and good resistance to traffic load and climate conditions. On the other hand, if very hard bituminous binders or higher reclaimed asphalt content is applied there is a potentially higher risk of low-temperature cracking as compared with conventional asphalt concretes (AC). Nevertheless, in general, it is assumed that due to slightly higher bitumen content the fracture toughness and the fracture energy

reaches usually high values, and the only risk can be represented by brittle type of thermal-induced cracking. In HMAC usually low gradation bituminous binders (e.g., 10/20, 15/25, 20/30) or polymer-modified bitumen are used—in this case again rather binders with lower penetration grades (PMB 10/40-65, PMB 25/55-60, etc.) are preferred. Additive for stiffening can be used as well like for example fatty amide acids.

Use of HMAC in base (or potentially binder) layer can potentially lead to a reduction of the thickness of asphalt layers in pavement structure in comparison to pavement structure with conventional asphalt concretes, while the service life of such a construction remains unchanged. The effort to reduce the thickness of the asphalt layer is related to the reduction of construction costs and later life-cycle costs related to maintenance (LCC optimization) [2–4]. In addition, the material resources can be saved. Some of the published papers or research outputs present a reduction in thickness between 25 and 30 percent in pavement structure [5,6]. With regard to road pavement bases, such HMAC technology is somehow comparable, from the performance point of view, to cement-bound mixtures. In this case, the stiffness modulus is represented by the dynamic elastic modulus of the cement-bound mixtures, investigated through ultrasonic tests [7,8]. Other innovative equipment for advanced testing on cementitious materials, even derived from asphalt machine testing, can be found in proper literature [9].

In the case of long-life pavement, the overall costs have to be assessed not only from the perspective of construction costs, but mainly from the viewpoint of life cycle costs. Primary cost can be higher, but the pavement shows less demand for repairs and rehabilitation actions, and therefore the life cycle costs are significantly lower than for other types of asphalt mixtures. For these reasons, it is necessary to focus on the life cycle cost assessment during the selection of the right pavement design and not only on the lowest construction price as currently often happens.

Espersoon [1] showed the results of the experimental research that has been done to calculate the reduction in thickness of the base layer with HMAC compared to a base layer with conventional paving grade bitumen for runway pavements based on measuring and evaluation of dynamic modulus of different mixtures at different temperatures. Rys et al. [3] presented an analysis of 80 selected road sections in Poland, of a total length of about 1300 km, and compared low-temperature cracking properties of pavements with HMAC mix type and conventional asphalt concrete base. It was revealed that pavements with high modulus asphalt bases have a 2.45 times higher occurrence of odd cracks than pavements with conventional asphalt concrete bases. H.J. Lee et al. [10] designed HMAC mixtures as well as high modulus asphalt binders (HMAB). First, the binders were tested and compared with the conventional unmodified and SBS modified binders. Then HMABs were used in HMAC mixes. The results showed that dynamic modulus is 50% higher for HMAC than for conventional AC at high temperatures and rutting is on the other hand twice smaller. The fatigue test showed 5 to 10 times better results. Moghaddam et al. [11] investigated the rheological properties of different types of highly modified asphalt binders and EME mixes. A good correlation between the rheological properties and the binders' microstructure was determined and from their results, it was stated that the performance of asphalt mix can, to some extent, be predicted by the rheological properties of bituminous binders used. Si et al. [12] investigated the micro-mechanical characteristics of HMAC pavement structure under loads. The application of high modulus materials decreased the vertical stress in all asphalt layers except the surface layer. Zaumanis et al. [13] designed HMAC containing 100% reclaimed asphalt (RA). Through multiple design iterations, it was found that it was not possible to fully fulfill fatigue, stiffness modulus, and rutting requirements for most of the recycled HMAC mix variants. Even an increase of the binder content did not help to fulfill the fatigue requirements and therefore use of additives was recommended in mixtures containing high content of RAP.

The Czech Republic started to use HMAC mixtures from 2001 based on collected experience (mainly following the French knowledge and good practice) and formulated, after an initial period of practical experience from diverse job sites, own technical specifi-

cations: they were adopted in 2009 by the Ministry of Transportation and are known as TP 151. After additional 11 years, the technical specifications will be upgraded in 2021 to a national technical standard and the HMAC mixtures will be treated similarly to regular asphalt concretes in terms of type testing, factory production control, and site control testing requirements. Like the French early-stage approach the HMAC mixtures were mainly understood as a technical solution of an asphalt layer providing high stiffness. The aspect of fatigue characteristics is still underestimated which results in mixed designs with stiffness modulus at 15 °C > 14,000 MPa (the minimum required limit is 9000 MPa). Since usually hard paving grades like 20/30 are used this can result, from the perspective of longer lifetimes, in potential problems in the future especially if these mixtures are used for heavy-duty pavements. A similar problem can arise if such mixtures are used in binder courses, especially in the case of sudden changes in the temperatures and repeated freezing-thawing cycles, which might result in temperature-induced cracks. Reflecting these potential risks, the new technical standard assumes already that for an HMAC mixture where the stiffness modulus at 15 °C is higher than 13,500 MPa a fatigue test needs to be done to reduce the probability of an early failure. For the stiffness testing, either repeated indirect tensile strain test on cylindrical specimens (IT-CY) or 2-point beam test on trapezoidal test specimens can be done following the standard EN 12697-26. In terms of HMAC mix design, comparing the requirements with traditional asphalt concrete for base or binder courses, the grading curve limits are closer, and more stringent requirements are defined for voids content with a range between 3%-vol. and 5%-vol. for the initial mix design. The minimum bitumen content shall be 4.2% for HMAC 22 and 4.4% for HMAC 16, usually, the mix designs used for real job sites vary in bitumen content between 4.6 and 5.1%. Further details related to requirements and restrictions for HMAC mixtures used in the Czech Republic can be found in [14].

In recent years, mathematical approaches have been applied to evaluate the characteristics and performance of asphalt pavements with the aim of implementing a numerical mix design. Although advanced constitutive models of the asphalt concrete's mechanical behavior exist nowadays [15–18], Machine Learning represents a numerical framework of application interest for the development of predictive models of the ACs' mechanical response [19–24]. Among Machine learnings approaches, Artificial Neural Networks (ANNs) are nonlinear fitting systems that are imitating the brain's biological learning process to correlate information or data. A detailed description of the mathematical framework has already been widely discussed e.g., by Baldo et al. [25].

Among the literature studies, Ozsahin and Oruc [26] have employed a Shallow Neural Network (SNN), i.e., a three-layer perceptron neural network, to establish the relation between the emulsified asphalt mixtures' resilient modulus and some relevant predictors (such as cement addition level, curing time and residual asphalt content), demonstrating that Artificial Intelligence can reduce the time consumed at the design stage to define the proper mixture. Tapkın et al. [27] have applied ANNs to model the creep behavior under repeated loading of polypropylene modified asphalt concretes. Similarly, Saoudi and Haddadi [28] predicted the creep deformation of asphalt concrete modified with different contents of industrial rubber waste to identify the percentage of additive beyond which the creep rate stops decreasing. Mirzahosseini et al. [20] have analyzed the rutting potential of dense ACs, implementing an NN that maps the pattern between the flow number and the main features of standard Marshall specimens, such as coarse aggregate to fine aggregate ratio, filler and bitumen contents, percentage of voids in mineral aggregate and Marshall quotient. Artificial Intelligence has also enabled us to accurately model the fatigue performance [29,30] and dynamic modulus [31,32] of ACs under different environmental and loading conditions. Given that the mechanical response of ACs also depends on the volumetric properties (which have to meet the limits set by standards or local specifications), Zavrtnik et al. [33] have relied on ANNs to predict air void content in seven types of asphalt mixture. Recently, Hussain et al. [34] have developed an ANN model that can be used to characterize the phase angle behavior of wearing and base course AC mixtures.

Despite the positive results reported in the documented literature, few researchers have bothered to employ an objective and effective strategy to find the neural network “structure” with the best score on a certain performance metric, also called optimal NN. Such structure derives from the so-called hyper-parameters that define both the network topology and the algorithmic functioning. These “model settings”, which concur to define the optimal functioning of the neural network, are commonly set up by means of a trial-and-error procedure [20,21], such as grid or random search, that employs a certain performance metric (e.g., the mean square error). These approaches, which involve evaluating the selected performance metric over and over (for all desired hyper-parameters combinations), are extremely time-consuming and do not guarantee that the optimal model will be identified [35]. In this context, the Bayesian Optimization offers an efficient and semi-autonomous process for fine-tuning the hyperparameters of the optimal NN model [36]. By keeping a record of past evaluations, the Bayesian approach builds a probabilistic model of the performance function, which is used to make decisions on the next set of hyperparameters to be evaluated so that the expected error is minimized [37,38].

The purpose of this study is to develop suitable Machine Learning strategies that allow to design and specify an optimal ANN model for the ACs’ stiffness prediction task to be accurately identified, even when the available experimental dataset is relatively small. In particular, Bayesian Optimization was employed to properly set some relevant topology-related hyperparameters. The motivation for adopting Machine Learning for assessing or predicting HMAC mix behavior is based on a recommendation provided by the Czech technical specifications TP 151. The case study involved a set of 38 variants of high modulus asphalt concretes, with a total of 115 Marshall test specimens, prepared in the laboratory as part of real-case mix design processes and characterized by different binder types, bitumen contents, aggregate gradations, and Reclaimed Asphalt Pavement (RAP) percentages. The SNN approach was employed to identify a reliable correlation between the Stiffness and the main features related to the mix composition, such as bitumen content (% by mass of mix), air voids content (%), maximum and average bulk density ( $\text{g}/\text{cm}^3$ ), along with a categorical variable that distinguishes the bitumen type and RAP percentages. Moreover, SNN models have also been developed that include the Marshall tests results among the input feature aforementioned: in fact, the Marshall Stability and the Marshall Quotient can be correlated with the target variable and help to improve the prediction accuracy.

## 2. Materials and Methods

In the experimental study presented in this article and for the sake of using neural networks a set of 38 variants of HMAC mixtures was used. These mixtures were either experimentally designed or sampled from a real asphalt mix production, whereas in both options the mixtures had to fulfill requirements set in Czech technical specifications TP 151. The intention was to collect a wide range of possible mix designs which differ in used aggregate type, bitumen type, and gradation or partial substitution of virgin aggregates by reclaimed asphalt. The only basic condition was that all mixtures have to be of gradation 0–22 mm. With respect to used bituminous binders conventional hard paving grade bitumen of lower penetration values as well as polymer-modified binders were used in the HMAC mixtures, whereas hard paving grade 20/30 or PMB 25/55-60 dominated as being mostly recommended and used for practical application. Characteristics of the binders used in particular HMAC mixtures are not presented by this paper since it was not the intention to check and validate the required characteristics like penetration or softening point. Since most of the tested mixtures originated from regular production on mixing plants and each mixing plant has to follow the factory production control rules according to EN 13108-21, it is assumed that all the binders were within the specifications. Similarly, it was not the intention to collect and validate the characteristics of used aggregates, and in this stage of neural networks modeling the possible impact of aggregate type (in terms of its mineralogy) was not expected to be included in the machine learning process. Generally,

the aggregates used in HMAC in the Czech Republic are coming from various regions of the country and are rich in different minerals like basalt, hornblende, spilite, diorite, gneiss, greywacke, etc. Nevertheless, the scope of this study was not to investigate the impact of input materials properties on the resulting asphalt mix stiffness, but to estimate asphalt stiffness out of asphalt characteristic properties values by applying neural networks.

#### High Modulus Asphalt Concrete Characterization

As can be seen from Table 1 some mix variants contained between 10% and 30% reclaimed asphalt. The limitation of 30% is given by the technical specifications TP 151. In some of these cases, a 50/70 paving grade bitumen was used to soften slightly the resulting binder in the asphalt mix due to the elevated content of reclaimed asphalt with an aged bituminous binder. This technical solution so far presents a preferred option besides the possibility to use rejuvenators. In two cases a hard bitumen 15/25 was used since being required for some pavement structures. Finally, in one case (marked VMT 22 NT, VMT stands for Vysokým Modulem Tuhosti, i.e., HMAC in Czech) a warm mix asphalt concept was applied using an amide fatty acid additive for reduction of the mix production temperatures.

**Table 1.** Summary of experimental test results and specimens' characteristics.

Mix	Bitumen Type	ID	Bulk Density	Max Bulk Density	Binder Content	Voids Content	Maximum Strength	Marshall Stability	Marshall Flow	IT-CY 15 °C
			(g/cm <sup>3</sup> )	(g/cm <sup>3</sup> )	(%)	(%)	kN	kN	(0.1 mm)	(MPa)
VMT 22 with 30% RA (Froněk-A)	20/30	M1	2.455	2.640	4.9	7.0	20.6	20.0	33	16,062
			2.429		4.9	8.0	22.4	22.7	35	14,283
			2.456		4.9	7.0	21.6	23.1	28	16,078
VMT 22 with 30% RA (Froněk-B)	20/30	M1	2.459	2.647	4.6	7.1	20.7	20.9	51	14,867
			2.453		4.6	7.3	19.6	20.5	41	15,616
			2.456		4.6	7.2	21.0	21.4	43	14,350
VMT 22 with 30% RA (Froněk-C)	20/30	M1	2.473	2.663	4.3	7.2	22.8	22.4	22	15,974
			2.475		4.3	7.0	24.7	25.9	24	15,535
			2.485		4.3	6.7	24.1	24.6	27	15,452
VMT 22 with 20% RA (Froněk-1)	20/30	M2	2.467	2.676	4.3	7.8	20.1	20.5	58	12,049
			2.463		4.3	8.0	19.0	19.8	42	14,419
			2.461		4.3	8.0	20.2	21.0	30	13,003
VMT 22 with 20% RA (Froněk-2)	20/30	M2	2.486	2.682	4.6	7.3	20.7	21.7	59	13,792
			2.462		4.6	8.2	18.9	20.0	42	11,559
			2.480		4.6	7.5	19.6	19.8	47	12,452
VMT 22 with 20% RA (Froněk-4)	20/30	M2	2.460	2.678	4.9	8.1	23.5	23.5	53	14,441
			2.460		4.9	8.1	24.3	23.9	45	15,113
			2.443		4.9	8.8	23.9	24.6	30	16,558
VMT 22 with 20% RA (Froněk-6)	20/30	M2	2.422	2.667	5.2	9.2	18.6	21.8	35	13,116
			2.411		5.2	9.6	19.1	22.2	27	11,548
			2.422		5.2	9.2	22.3	25.4	34	12,370
VMT 22 with 30% RA var. 5.1	50/70	M3	2.547	2.617	5.1	2.7	17.1	19.7	71	13,171
			2.554		5.1	2.4	17.2	20.0	55	11,659
			2.538		5.1	3.0	19.6	21.9	45	13,242
VMT 22 with 30% RA. var. 4.8	50/70	M3	2.538	2.607	4.8	2.6	17.4	19.9	58	12,739
			2.535		4.8	2.8	14.8	16.9	47	13,287
			2.539		4.8	2.6	22.7	25.5	61	13,217
VMT 22 with 30% RA (Froněk)	50/70	M3	2.549	2.602	4.8	2.0	17.4	20.2	53	13,025
			2.539		4.8	2.4	15.3	17.9	63	14,267
			2.548		4.8	2.1	16.8	19.0	66	13,325
VMT 22 with 30% RA (Froněk)	50/70	M3	2.553	2.626	4.6	2.8	20.6	20.7	51	15,871
			2.548		4.6	3.0	18.6	21.0	54	15,666
			2.548		4.6	3.0	20.2	23.4	50	16,707

Table 1. Cont.

Mix	Bitumen Type	ID	Bulk Density	Max Bulk Density	Binder Content	Voids Content	Maximum Strength	Marshall Stability	Marshall Flow	IT-CY 15 °C
			(g/cm <sup>3</sup> )	(g/cm <sup>3</sup> )	(%)	(%)	kN	kN	(0.1 mm)	(MPa)
VMT 22 with 20% RA (Froněk-3)	50/70	M4	2.473	2.639	4.8	6.3	18.1	19.0	34	12,729
			2.495		4.8	5.4	20.2	21.6	34	12,282
			2.477		4.8	6.1	21.5	22.3	46	14,101
VMT 22 with 20% RA (PKB-A)	50/70	M4	2.397	2.496	4.4	4.0	14.2	13.6	48	8666
			2.421		4.4	3.0	13.4	13.4	50	9064
			2.412		4.4	3.4	12.2	12.4	51	8135
VMT 22 with 10% RA (PKB-101)	50/70	M5	2.358	2.559	4.6	7.9	12.1	11.4	35	8950
			2.351		4.6	8.1	15.3	14.1	37	9339
			2.355		4.6	8.0	12.8	14.5	34	9311
VMT 22 with 10% RA (PKB-102)	50/70	M5	2.341	2.559	4.5	8.5	17.1	16.2	90	9203
			2.343		4.5	8.4	17.1	16.1	80	9142
			2.323		4.5	9.2	15.1	14.2	96	9361
VMT 22 NT	20/30	M6	2.362	2.490	4.7	5.1	18.9	17.1	46	14,357
			2.409		4.7	3.2	21.4	20.3	56	14,601
			2.409		4.7	3.3	20.8	19.7	52	14,784
VMT 22 NT	20/30	M6	2.296	2.490	4.7	7.8	19.4	16.3	84	13,653
			2.313		4.7	7.1	20.8	18.8	84	15,529
			2.296		4.7	7.8	19.7	16.5	59	15,345
VMT 22 (SK-1)	20/30	M6	2.330	2.449	4.6	4.9	25.4	24.1	26	12,102
			2.324		4.6	5.1	24.5	23.2	20	12,027
			2.305		4.6	5.9	22.5	20.5	27	10,528
VMT 22 (VIA-1)	20/30	M6	2.702	2.789	4.7	3.1	21.1	22.8	42	17,417
			2.691		4.7	3.5	19.2	21.3	34	17,262
			2.680		4.7	3.9	20.5	23.0	29	17,478
VMT 22 (SK-2)	30/45	M7	2.414	2.490	4.6	3.1	22.0	20.6	49	12,483
			2.416		4.6	3.0	20.8	19.6	54	12,129
			2.396		4.6	3.8	21.7	20.6	59	11,734
VMT 22 (VHS)	30/45	M7	2.641	2.747	4.7	3.9	11.5	14.7	38	12,136
			2.648		4.7	3.6	14.4	15.8	55	11,478
			2.650		4.7	3.5	14.2	15.6	42	12,566
VMT 22 (VIA-2)	TSA 15/25	M8	2.709	2.818	4.8	3.9	19.7	22.8	39	16,182
			2.724		4.8	3.3	18.0	21.6	31	17,571
			2.712		4.8	3.8	18.4	22.3	31	17,227
VMT 22 (TPA-1)	TSA 15/25	M8	2.458	2.566	4.6	4.2	21.5	21.5	45	12,629
			2.454		4.6	4.4	22.5	21.8	40	12,412
			2.460		4.6	4.1	23.5	23.0	52	13,627
VMT 22 (EV)	PMB 25/55–60	M9	2.574	2.655	4.9	3.0	24.8	25.1	81	13,203
			2.567		4.9	3.3	23.2	23.6	71	11,688
			2.576		4.9	3.0	27.2	27.7	65	13,772
VMT 22 (SK-3)	PMB 25/55–60	M9	2.357	2.436	4.7	3.2	19.7	19.9	67	10,581
			2.362		4.7	3.0	20.4	20.6	62	10,940
			2.363		4.7	3.0	20.9	21.1	76	10,505
VMT 22 (SK-4)	PMB 25/55–60	M9	2.366	2.436	4.9	2.9	11.8	18.5	34	6632
			2.367		4.9	2.8	10.8	17.3	28	6001
			2.358		4.9	3.2	11.3	17.8	28	6699
VMT 22 (TPA-2)	PMB 25/55–60	M9	2.338	2.457	4.8	4.8	16.3	16.1	30	9024
			2.329		4.8	5.2	20.2	20.4	31	9134
			2.334		4.8	5.0	17.6	17.4	39	9097
VMT 22 (ESLAB)	PMB 25/55–60	M9	2.476	2.558	4.7	3.2	18.4	19.5	31	9585
			2.485		4.7	2.9	17.9	18.4	47	9322
			2.481		4.7	3.0	17.2	17.4	46	10,656
VMT 22 (TPA-3)	PMB 25/55–60	M9	2.415	2.632	4.9	8.3	20.8	19.7	53	7102
			2.427		4.9	7.8	18.8	17.8	51	8203
			2.430		4.9	7.7	19.4	18.5	47	7986

Table 1. Cont.

Mix	Bitumen Type	ID	Bulk Density	Max Bulk Density	Binder Content	Voids Content	Maximum Strength	Marshall Stability	Marshall Flow	IT-CY 15 °C
			(g/cm <sup>3</sup> )	(g/cm <sup>3</sup> )	(%)	(%)	kN	kN	(0.1 mm)	(MPa)
VMT 22 (TPA-4)	PMB 25/55–60	M9	2.416	2.485	4.9	2.8	21.1	21.7	42	9174
			2.408		4.9	3.1	19.4	19.6	39	10,245
			2.404		4.9	3.3	19.6	19.8	50	9421
VMT 22 (VIA-3)	PMB 25/55–60	M9	2.415	2.510	4.8	3.8	18.4	17.5	57	9731
			2.416		4.8	3.7	18.1	16.9	45	10,044
			2.415		4.8	3.8	19.2	17.7	48	9446
VMT 22 (TPA-5)	PMB 25/55–60	M9	2.338	2.457	5.0	4.8	18.7	18.9	46	8691
			2.329		5.0	5.2	20.7	20.9	36	8759
			2.334		5.0	5.0	17.4	17.6	34	8611
VMT 22 (TPA-6)	PMB 25/55–60	M9	2.335	2.467	4.8	5.3	18.2	18.8	36	9953
			2.332		4.8	5.5	16.5	18.0	35	9018
			2.339		4.8	5.2	17.5	17.9	28	9884
			2.330		4.8	5.5	16.3	16.4	26	9521
VMT 22 (Chvaletice)	PMB 25/55–60	M9	2.392	2.532	4.7	5.5	28.3	26.3	47	16,134
			2.397		4.7	5.3	25.6	23.8	41	15,808
			2.386		4.7	5.8	28.3	26.8	33	15,855
VMT 22 (SK-5)	PMB 25/55–60	M9	2.373	2.460	4.9	3.5	14.5	12.9	37	5685
			2.362		4.9	4.0	14.2	12.0	34	5636
			2.378		4.9	3.3	15.2	13.7	52	5991
VMT 22 (TPA-7)	PMB 25/55–65	M10	2.439	2.610	5.1	6.6	15.1	14.8	85	6686
			2.441		5.1	6.5	15.8	14.9	90	6223
			2.433		5.1	6.8	14.1	13.6	76	6848
VMT 22 (TPA-8)	PMB 25/55–65	M10	2.553	2.648	4.8	3.6	19.8	19.2	79	11,989
			2.556		4.8	3.5	21.8	22.0	61	12,075
			2.545		4.8	3.9	18.9	19.1	58	11,958
VMT 22 (TPA-9)	PMB 25/55–65	M10	2.554	2.629	5.0	2.8	21.8	22.7	54	11,849
			2.548		5.0	3.1	19.8	20.2	59	11,603
			2.543		5.0	3.3	19.5	19.3	56	12,071

For all mix variants bulk density and maximum density were determined and void content was calculated (EN 12697-8), whereas the void content requirement in TP 151 is set with a range of 2.5%–6.0%-vol. for control testing. The test specimens were compacted by 2 × 75 blows using an impact compactor (EN 12697-30). From all tested mixtures 12 variants did not fulfill the voids content requirement, whereas most of them were variants with reclaimed asphalt and in one case it was the warm mix variant of HMAC, where the test specimens were compacted at a lower temperature. Soluble bitumen content was determined as another empirical characteristic that could be used for neuronal network analysis and machine learning approach. The TP 151 prescribes a bitumen content of 4.2% to 5.4%. All analyzed mixtures fulfill this criterion and none of them was neither at the lower limit nor approaching the upper limit. On average the bitumen content was around 4.8%–4.9% which has provided the asphalt mixtures with sufficient bitumen film.

Test specimens of all mixtures were then conditioned at 15 °C and tested according to EN 12697-26, annex C by repeated indirect tensile strain test on cylindrical specimens (IT-CY) for stiffness. The technical specifications TP 151 require a minimum stiffness value at 15 °C 9000 MPa. In Table 1 values that did not meet this criterion are marked in red. On the other hand, there is no upper limit for the stiffness which is potentially dangerous with respect to fatigue life. This might be true especially for hard paving grade binders which can lead often to stiffness values > 14 GPa. In these cases, the new technical standard, which is under final review, recommends running fatigue test according to EN 12697-24, which shall provide sufficient confidence that too high stiffness values will not result in early loss of fatigue life.

After stiffness testing specimens were conditioned at 60 °C for 45–60 min in a water bath prior to running the Marshall test, which was performed according to EN 12697-34.

Such test is commonly used in asphalt concretes mix-design, including epoxy asphalt concretes for wearing surface on orthotropic steel bridge decks [39]. Generally, there is no requirement according to the existing specifications to execute this test. The TP 151 just indicates that the mix designer can prior to stiffness testing run the Marshall test and if the Marshall Stability reaches min. 14 kN and/or the Marshall Flow is in the range of 20–60 dmm such mix might reach the required stiffness as well. The reason for this optional validation was in the past defined in the technical specifications because there is a limited number of laboratories able to run stiffness tests (actually, in the Czech Republic only two university laboratories are able to perform such tests). This recommended procedure was originally one of the incentives why to apply the neural network approach. If it could be proven that based on characteristic properties the neural network model can predict with good accuracy the stiffness modulus, the Marshall test could be used, e.g., for control testing during asphalt mix production and paving.

The results of bulk densities vary with aggregate types which are used in the Czech Republic. Values approaching bulk densities of 2.7 g/cm<sup>3</sup> are usually related to mixing designs where basalt is used. On the other hand, values between 2.3 g/cm<sup>3</sup> and 2.4 g/cm<sup>3</sup> are mostly related to minerals like spilite or hornblende.

Since the Marshall test and its characteristics are not required as a standard parameter for HMAC, it is even not possible to compare the results with some limiting threshold values. Most of the mixtures reached Marshall Stability values in the range of 17–21 kN; only two mix variants had a value < 14 kN. Additionally, with respect to the recommendation provided by the TP 151 and described earlier in the text, related to the limit of 14 kN as an indication for reaching stiffness > 9000 MPa, it has been verified that for some mixes (Table 1: VMT 22 (SK-4), VMT 22 (TPA-3), VMT 22 (TPA-5), VMT 22 (TPA-7)) with Marshall Stability values over 14 kN, the stiffness modulus resulted lower than 9000 MPa. Similarly, it is true for the Marshall Flow values and recommendation provided by TP 151.

In Section 3.3 a detailed analysis of the correlation strength between the Marshall mechanical parameters and stiffness is presented. However, the scatter plot suggesting the type of correlation between stiffness modulus and Marshall stability (or Marshall flow) is presented in Figure 1 (and Figure 2, respectively): the stiffness seems to have some dependence on the Marshall stability, whereas no dependence was found for the flow (please refer to Table 1 for Legend understanding).

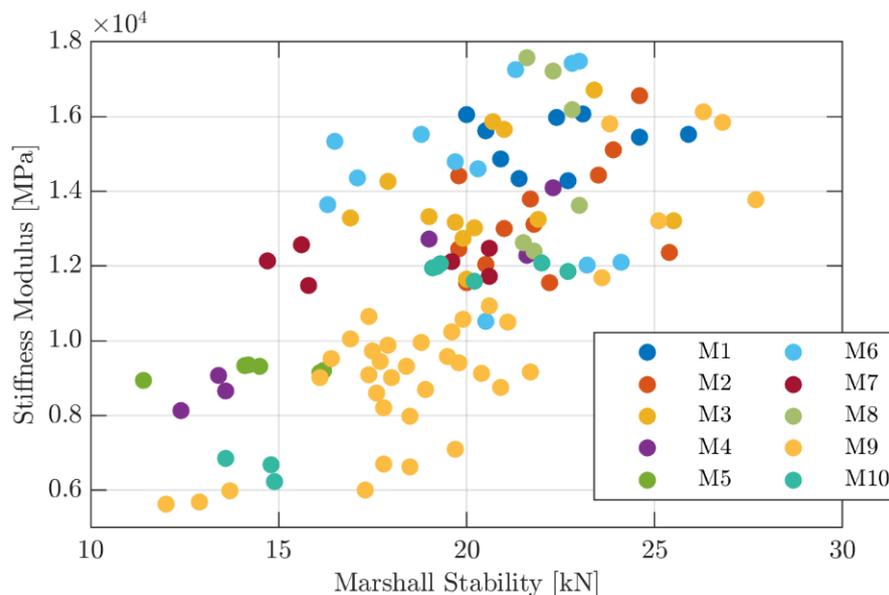
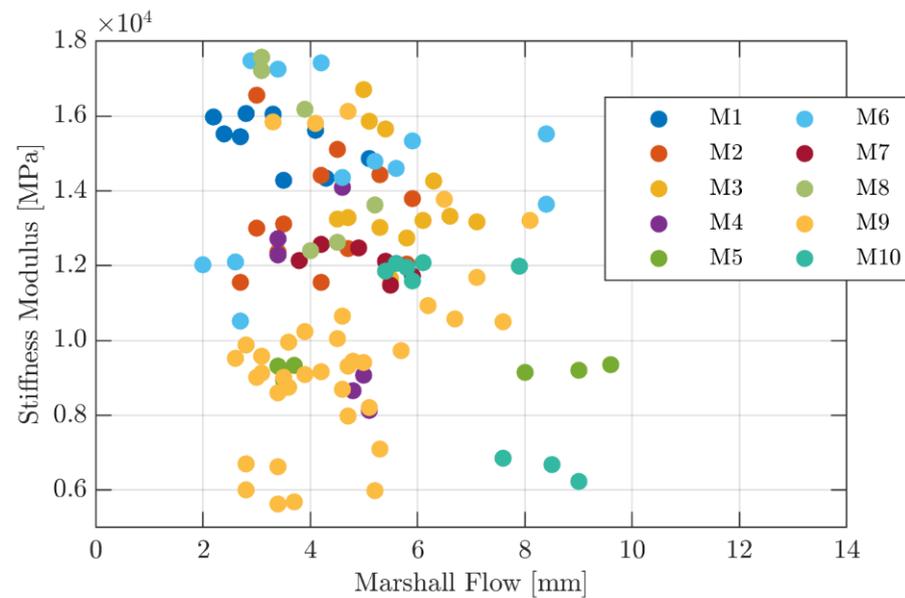


Figure 1. Scatter plot between Stiffness Modulus and Marshall Stability.



**Figure 2.** Scatter plot between Stiffness Modulus and Marshall Flow.

Table 1 provides data about HMAC stiffness at 15 °C. Six mixtures showed lower values than the minimum requirement from the specification TP 151—variant VMT 22 with 20% RA (PKB-A), VMT 22 (SK-4), VMT 22 (TPA-3), VMT 22 (TP-5), VMT 22 (SK-5), and VMT 22 (TP-7). For three of them, the values are very low. Additionally, except the VMT 22 with 20% RA (PKB-A) where the stiffness could be influenced by the used 50/70 paving grade, the remaining 5 mixtures with low stiffness were all designed and produced with PMB. Often for HMACs with PMB 25/55-60, it is challenging to exceed the limit of 9000 MPa due to the elastic behavior of this polymer-modified binder, and several optimizations and modifications of the grading curve are needed. This might explain why most variants with lower stiffness are at the same time containing this type of PMB.

### 3. Methodology

#### 3.1. Shallow Neural Networks with Backpropagation Algorithm

Shallow Neural Networks (SNNs), also called Multilayer Perceptron Neural Networks (MLPs), is a type of Artificial-Intelligence-based on the feedforward network architecture, which can approximate essentially any relation between input and output patterns to an arbitrary degree of accuracy [40,41]. To fully understand the topology and algorithmic functioning of an SNN, the reader should be reminded that the biological nervous system, the brain, acquires knowledge and abilities by a learning process, and that specific synaptic weights of interneuron's connections store the acquired know-how to replicate it. Similarly, to perform a particular task or function of interest, a Shallow Neural Network has to be trained using a process, typically supervised (i.e., based on input pattern  $x \in R^F$  and ground-truth target  $y \in R$  known beforehand), that iteratively updates connections' weights, according to a learning rule, until the network is able to replicate the desired task with the required accuracy. In fact, a Multilayer Perceptron Neural Networks consists of an input layer, at least one hidden layer, and an output layer. Each of these layers has several processing units, the artificial neurons, and each of them is interconnected with weighted connections to any other unit in the subsequent layer: thereby, the information to be processed travels in just one direction, from input to output layer. Figure 3 shows a typical SNN topology.

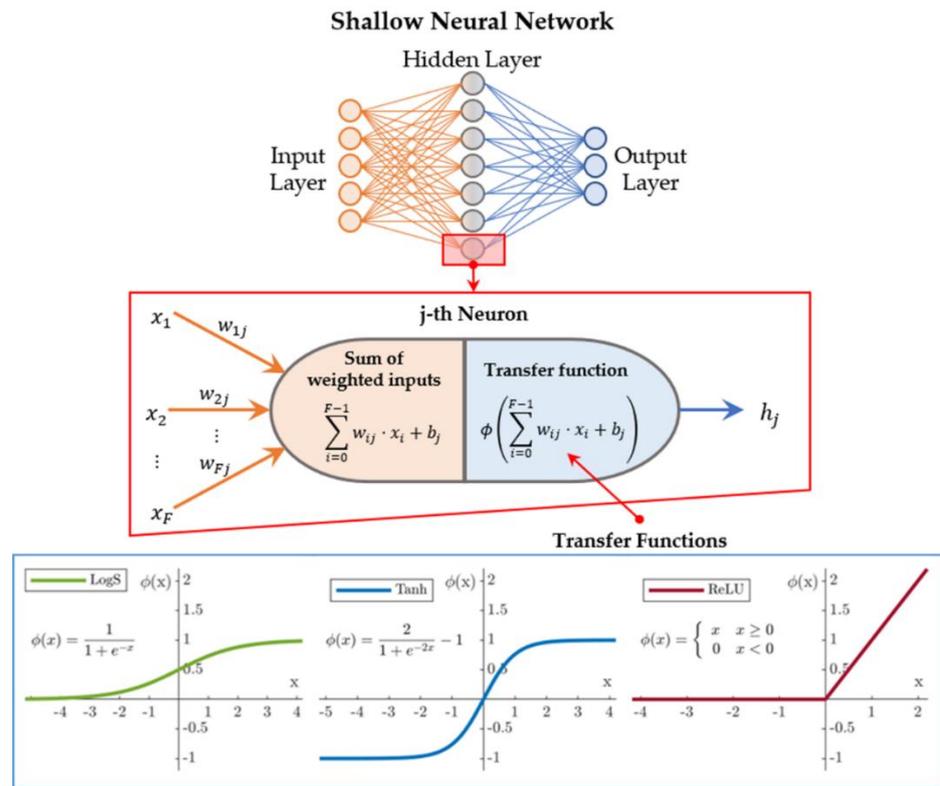


Figure 3. The SNN topology, the artificial neuron unit, and the main used transfer functions.

Artificial neurons perform the network computations: these are logistic regression models with a nonlinear activation function. This means that each neuron in the first hidden layer implements a dot product between the input feature vector  $x \in R^F$  and the vector  $w_j^{(1)} \in R^F$  that represents the weights of the connections it has made with each component of the input feature vector (or the neurons  $x_i$ ) and adds a scalar bias  $b_j^{(1)}$ . The resulting scalar is passed to a function  $\phi^{(1)}$ , called transfer or activation unit, which returns the neuron’s output:  $h_j^{(1)} = \phi^{(1)}(\sum_{i=0}^{F-1} w_{ij}^{(1)} \cdot x_i + b_j^{(1)})$ . Some activation functions, which are also employed in this study, are shown in Figure 3: the rectified linear (ReLU), the hyperbolic tangent (TanH), and the logistic sigmoid (LogS) units. The first hidden layer output is the vector  $h^{(1)} = [h_0^{(1)}, \dots, h_{N-1}^{(1)}]$ , with  $N$  the number of neurons that form the layer itself, but it is also the input vector of the subsequent hidden layer. Therefore, the outlined process is repeated until the output of the network is computed:  $\hat{y} = \phi^{(out)}(h^{(L)})$ . The matrix  $W = \{w_{ij}^{(l)}, b_j^{(l)} \mid l \in \{1, 2, \dots, L\}\}$ , with  $L$  the number of network layers, forms the set of parameters to be defined in the supervised training process, in which a back-propagation algorithm is typically used [42]. These algorithms compute the first-order approximation of the gradients of a loss metric  $L(\hat{y}, y)$  (or error function), between the network output  $\hat{y}$  and the expected target vector  $y$ , with respect to the parameters  $W$ ; then, a learning rule uses the derivatives of  $L(\cdot)$  to correct the weights and biases of the network, performing several iterations until the loss value is minimized; the resulting parameters are kept fixed to process new feature patterns. In this study, the loss optimization objective  $L(\cdot)$  is the Mean Squared Error (MSE):

$$L(\hat{y}(W^e), y, W^e) = \frac{1}{M} \sum_{m=1}^M (\hat{y}_m(W^e) - y_m)^2 \tag{1}$$

where  $W^e$  is the parameters’ matrix at iteration  $e$  of the training process and  $\hat{y}(W^e) - y$  the errors of the corresponding network, whereas  $M$  is the number of observations in the

training set. In fact, since a single experimental data set is usually available, it is randomly partitioned, according to fixed percentages, into a training, a validation, and a test subset: this practice is known as the hold-out method.

The training set contains the data used to train the SNN, while the validation and test sets hold the data intended to evaluate the model's effectiveness. Evaluation on the validation set occurs at regular intervals during the training process, while the test set evaluation is typically performed on the fully trained model when it is desired to estimate the model's generalization capability (i.e., the out-of-sample performance) and then to compare the accuracy of different models. In fact, generalization refers to how well the concepts learned by a Machine Learning model apply to specific observations not used by the neural network during the training process since the model, in order to minimize the MSE, might fit too much of the training data. Such a situation, known as Overfitting, is caused by large values of the connections' weights and results in an excessively fluctuating and physically unjustified interpolating surface/curve of the experimental data. A good generalization, stated on the basis of a given performance metric's value, allows reliable predictions to be made in the future about data out of the training sample, since the overfitting problem has been successfully prevented. In particular, among the methods to avoid overfitting and improve generalization, Early Stopping is the most widely used and simplest approach [41]. This method exploits the iterative nature of the training process and the validation data set to decide when to stop the training, according to a procedure called cross-validation: in fact, at each iteration of the training algorithm, the performance of the training process can be measured by monitoring the model error on the validation set. In general, it is observed that up to a certain number of iterations, new iterations improve the model. However, after that point, the model's capability to generalize can weaken as the neural network begins to overfit the training data and compromise the accuracy of the validation set. When the error on the validation set goes up for several iterations (i.e., in this study, for 7 iterations), the training is stopped, and the weights and biases that produced the minimum MSE on the validation data set are employed as the final trained network parameters.

In the current study, the Levenberg-Marquardt backpropagation algorithm [43], jointly with the Early Stopping method, was employed due to its efficiency in training Shallow Neural Networks among those implemented in MATLAB® [41]. In order to ensure convergence to the minimum of the error surface (and thus to the optimum values of the connections' weights), the training process was repeated 10 times and the best performing model was selected. With regard to the hold-out method, it is worth pointing out that such a practice has two major drawbacks when the number of observations is small: first, some relevant patterns may be excluded from the training set; second, the training-test splitting makes the model sensitive to the randomness of data in the training set. Therefore, since the purpose of the study is to predict the representative mechanical response of each asphalt concrete considered here, the original data set was augmented by  $Q = 38$  data by averaging the experimental results of specimen replicas of each HMAC variant and associating the main mix composition features of the original variant with each new synthetic target. The obtained stiffness modulus values form a representative set of all situations for which the network will be used. For this reason and to avoid the aforementioned problems, the cross-validating (20%) and the training (80%) sets for the SNN training procedure were selected randomly from the experimental data set (115 Marshall test specimens), while the 38 synthetic data were used for testing the trained SNN model.

### 3.2. Bayesian Optimization

The definition of the neural network topology and of the training algorithm parameters (i.e., the model's hyper-parameters) is an essential step in the application of any Machine Learning approach to a specific input-output fitting problem. This process, known as hyper-parameters tuning, represents a challenge for any research engineer and is closely related to model selection, since the hyper-parameters determine the response of the neural

network and therefore the performance of the model of which they are part. However, this tuning process is simplified when dealing with Shallow Neural Networks: Hagan et al. [41] assert that an SNN, as defined in the previous Paragraph, can arbitrarily well solve any multidimensional regression problem between given data by providing a sufficient number of neurons and an adequate transfer function to its single hidden layer. It means that the SNN hyper-parameters, when the parameters of the Levenberg-Marquardt training algorithm are set to their default values (please refer to the LM algorithm implementation proposed by the MathWorks® within the well-known MATLAB® ANN Toolbox), are related only to the network architecture, namely the number  $N$  of neurons that form the hidden layer and the activation unit  $\phi$  that characterizes them. Therefore, it becomes clear that the choice to exploit SNNs stems from the possibility of simplifying the hyper-parameters tuning problem as well as from the need to identify an easily interpretable and usable modeling tool for the research engineer or laboratory technician interested in the development of predictive models of the asphalt concrete's mechanical response. Figure 4 shows a flowchart of the optimization procedure followed in this study.

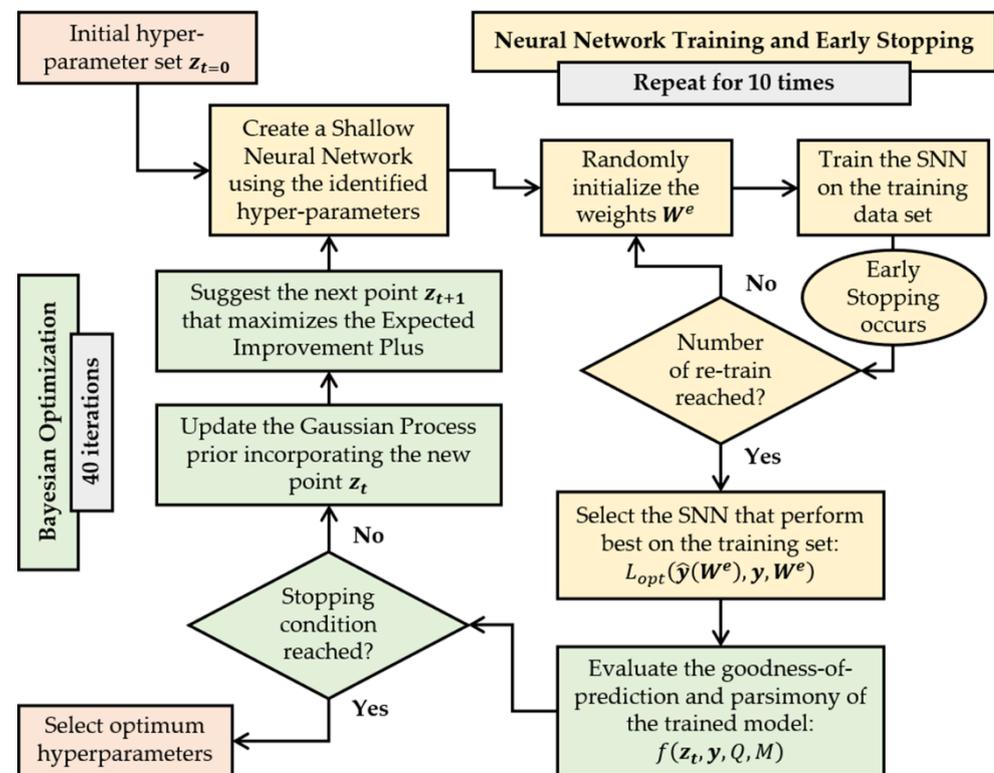


Figure 4. Flow chart of the optimization procedure.

Regarding strategies for carefully defining model hyper-parameters, standard approaches involve a random or grid search for the optimal set within pre-defined ranges on the basis of a sampling method. However, there are many optimization algorithms available nowadays that can reduce the time spent in searching for the best model hyperparameters. Among them, the Bayesian optimization (BO) algorithm [44] has found considerable success mainly due to the work of Snoek et al. [45]. The goal of the optimization process is to minimize a given objective function  $f(z)$  for  $z = [z_p]$ ,  $p \in \{0, \dots, P\}$ ,  $P \in N$  in a bounded domain  $Z_p \subset R$ . For this purpose, the BO algorithm adopts a Gaussian Process regression model (GPR) for  $f(z)$  [46], i.e., builds a non-parametric probabilistic model of the objective function, and updates it based on Bayesian statistics for each new evaluation of  $f(z)$ . Then, an acquisition function  $a(z)$  based on the prior distribution is maximized to determine the next point  $z_{next}$  for the evaluation:  $z_{next} = argmax_z a(z)$ . By repeating this procedure iteratively (until the set maximum number of iterations is reached), the BO algorithm

progressively improves the underlying probabilistic model for the objective function and thereby allows the acquisition function to sample points that improve the metric  $f(z)$ . For optimal results, a rule of thumb asserts to set a maximum number of runs greater than or equal to 20 times the number of hyper-parameters to be optimized. In particular, the acquisition function best known for its capability to efficiently reduce the number of evaluations required to find the global minimum of many black-box functions [47,48] is the Expected Improvement (EI) algorithm [49]. Such EI function evaluates the goodness of a point  $z$  by the expected amount of improvement in  $f(z)$ , ignoring the solutions that cause an increase in the objective metric. To escape a local minimum, the correction proposed by Bull [50] allows the EI acquisition function to modify its behavior when it estimates the over-exploitation of an area of the bounded domain  $Z_p$ : the resulting acquisition function is called Expected-Improvement-Plus (EIP) and it has been used in the current study setup.

Focusing on the current problem, namely the modeling of high-modulus asphalt concrete stiffness by means of Shallow Neural Networks, the objective function has to consider both out-of-sample performances, i.e., generalization capability, and model parsimony. In fact, as the complexity of the network increases (i.e., in this study, as the number of hidden neurons increases), the SNN does not use all of its weights to minimize the mean squared error because early stopping restricts the number of effective parameters, i.e., limits the weights value, to avoid overfitting. This means that if a “large” network was selected for the given problem, it would not necessarily overfit the data (due to early stopping) but it would definitely take longer to compute the network response, which might not be very different from that of a “smaller” network. If this happened, the model would be over-parameterized for the task, with direct consequences on efficiency. Therefore, the function  $f(\cdot)$  used in this study adds a penalty term, reliant on the number  $K = 3N + 1$  of network parameters (i.e., weights and biases) and the number  $M = 92$  of observations in the training set, to a term measuring the goodness-of-prediction of the model to the  $Q = 38$  test data [51]:

$$f(N, \phi, \mathbf{y}, Q, M) = \log \left( \frac{1}{Q} \sum_{q=1}^Q (\hat{y}_q(N, \phi) - y_q)^2 \right) + \frac{2(3N + 1)}{M} \quad (2)$$

Given the  $P = 2$  hyperparameters,  $N$  and  $\phi$ , and fixed their ranges  $Z_N = \{1, \dots, 20\}$ ,  $Z_\phi = \{\text{TanH}, \text{ReLU}, \text{LogS}\}$ ,  $f(\cdot)$  is a function that implements an SNN with  $N$  neurons in its hidden layer and  $\phi$  as activation function; such SNN is trained on a  $M$ -observation training set for 10 times and early stopped each time to avoid overfitting; the best performing SNN is then selected and performs on a  $Q$  test data set;  $f(\cdot)$  returns a single scalar that expresses the balance between model generalization and model parsimony. The Bayesian optimization algorithm is run for 40 iterations. The optimal hyper-parameters  $N$  and  $\phi$ , and then the optimal model, are selected when  $f(\cdot)$  is minimized by the BO algorithm.

All codes required for this study were implemented in MATLAB<sup>®</sup> using the ANN Toolbox framework and they were run on a machine provided with an Intel<sup>®</sup> Core<sup>™</sup> i7-6700HQ 2.60 GHz CPU and 16GB of RAM running Windows 10 Home. Each experiment lasted about 2 h.

### 3.3. Input Features Selection and Models Evaluation

Recalling that the aim of the study is to develop a model for predicting the stiffness modulus of different asphalt concretes, the problem arises of selecting the input variables or features that will characterize the model itself. In fact, a model that engages irrelevant features, poorly correlated with the target variable, may have learning difficulties, higher memory requirements, and increased complexity, as well as worse prediction accuracy and then poor generalization capability [52,53]. Due to the dependency of mechanical characteristics on both aggregate and bitumen properties, it is important to include the fundamental parameters of asphalt concretes related to their composition in the input feature vector, i.e., in the current study, the bitumen content (% by mass of mix), the air voids content (%), maximum and average bulk density ( $\text{g}/\text{cm}^3$ ), along with a categorical variable that distin-

guishes the bitumen type and RAP percentages (values from 0 to 10). Nevertheless, the use of the Marshall test results, in terms of Stability or Quotient, as predictors could help to improve the accuracy of the predictions, even though these parameters are often considered as target variables due to their frequent use in the mix design process or simply due to the implied costs of the test [54–56]. In particular, the relationship between Marshall Stability (MS), as well Marshall Quotient (MQ), and Stiffness Modulus was investigated using the non-parametric Spearman Rank Order Correlation ( $\rho$ ) [57]. There was a strong correlation between MS and Stiffness ( $\rho = 0.67$ ,  $n = 115$ ,  $p < 0.0005$ ), with high stiffness levels associated with high stability levels, and the same applies to the relationship between MQ and Stiffness but with a medium correlation strength ( $\rho = 0.35$ ,  $n = 115$ ,  $p < 0.0005$ ).

Therefore, three SNNs characterized by a different number of predictors were developed in this study and referred to in the following with the acronyms  $MIX_{SNN}$ ,  $MS_{SNN}$ ,  $MQ_{SNN}$ . More specifically, the first of these employs the aforementioned five mix composition variables to predict the Stiffness Modulus, while the remaining ones add MS and MQ to the input feature vector, respectively. Each variable belonging to both input and target vector has been standardized before being processed by the SNN, i.e., all variables have been rescaled to have zero mean and unity standard deviation: such an adjustment served to improve the efficiency of the Levenberg-Marquardt algorithm. In addition, all three models were optimized according to the procedure reported in Section 3.2, and then, evaluated and compared on the test data set using statistical performance metrics to identify the features combination that yields the best generalization capability. The employed model evaluation functions are the mean absolute error (MAE), the root mean square error (RMSE), the coefficient of determination ( $R^2$ ), and adjusted R-squared ( $R^2_{adj}$ ). For a detailed description of each performance metric please refer to Biecek and Burzykowski [58], as well Legates and McCabe [59].

In summary, the investigated SNNs is realized with a 5- or 6-neurons input layer (one neuron for each input feature), a  $N$ -neurons hidden layer whose processed output is passed to an activation unit  $\phi(\cdot)$  being either a ReLU, TanH or LogS function, and a 1-neuron output layer associated with the identity activation function. Figure 5 shows a summary flowchart of the full procedure followed.

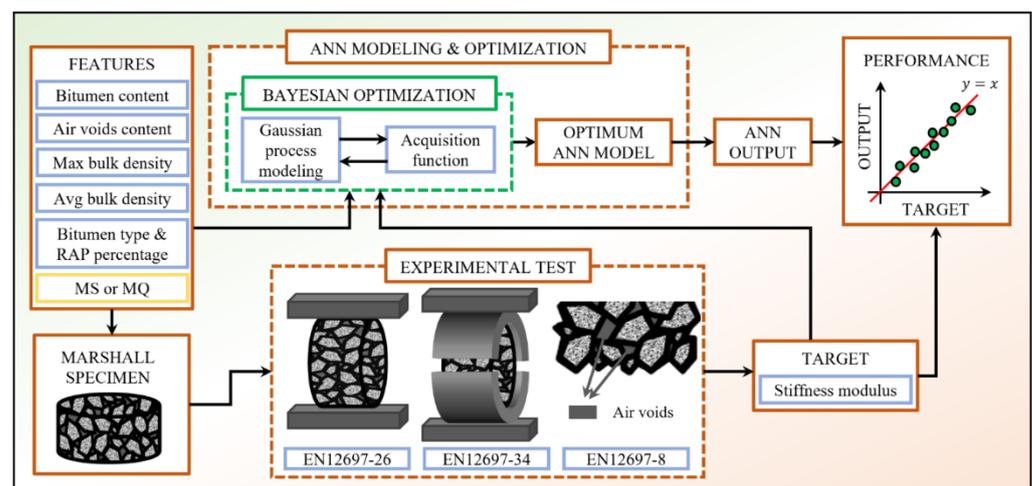


Figure 5. Step-by-step methodology applied in the current study setup.

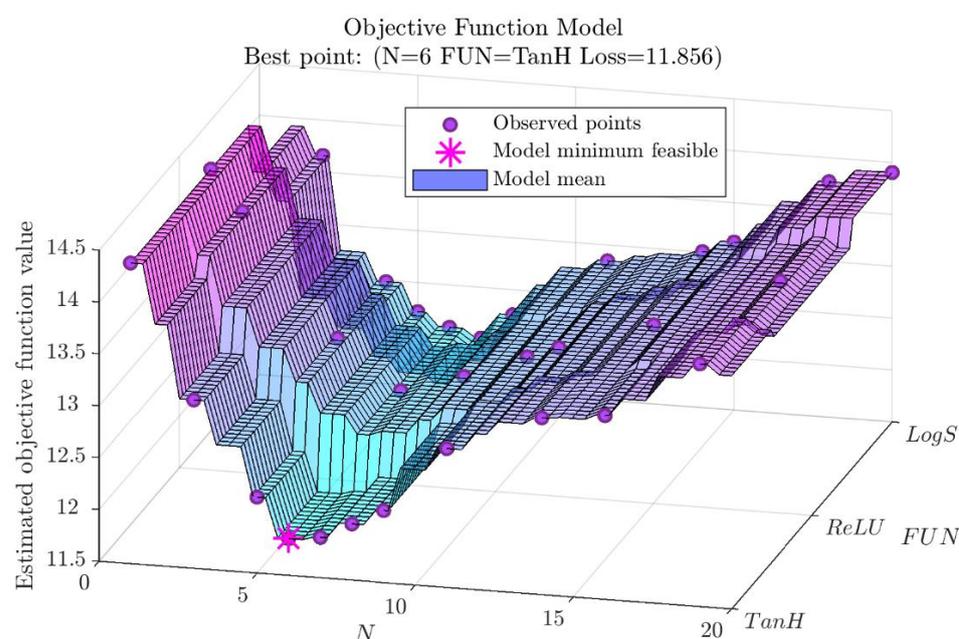
#### 4. Discussion

Table 2 summarizes the results of the hyperparameter optimization process, according to the procedure described in Section 3.2, for each group of input features and the value of the statistical metrics assessing generalization capability for each model. In particular, columns 3 and 4 show the number of neurons and the activation function of the hidden layer, identified by the Bayesian process based on the objective function  $f(\cdot)$ , for the three SNNs (i.e.,  $MIX_{SNN}$ ,  $MS_{SNN}$ ,  $MQ_{SNN}$ ): these shallow solutions ( $N < 8$  neurons) represent the most

efficient neural configuration for each modeling problem considered (Section 3.3). The best result in terms of  $f(\cdot)$  (fifth column of Table 2) was achieved by the  $MS_{SNN}$  ( $f(\cdot) = 11.856$ , Table 2), i.e., the network that receives as input features the Marshall Stability along with the mix composition parameters of the HMAC: Figure 6 shows the Gaussian Process regression model (overlapped with the objective function points that were sampled) that the Bayesian optimizer has progressively developed to find the minimum of the function  $f(\cdot)$  for the  $MS_{SNN}$ . This network, which is the best in terms of generalization performance (last four columns of Table 2), receives as input a 6-component features vector, which is processed by a 6-neurons hidden layer characterized by the TanH activation function.

**Table 2.** Comparison between SNN models.

ID	Features	$N$	$\phi$	$f(\cdot)$	MAE	RMSE	$R^2$	$R^2_{adj}$
$MIX_{SNN}$	5	6	TanH	12.093	209.12	293.56	0.9909	0.9894
$MS_{SNN}$	6	6	TanH	11.856	160.17	241.54	0.9938	0.9923
$MQ_{SNN}$	6	8	LogS	12.373	174.91	272.61	0.9922	0.9902



**Figure 6.** Gaussian Process regression model for the  $MS_{SNN}$ .

In general, the addition of the empirical mechanical parameters among the predictors of the stiffness modulus improved the prediction accuracy compared to the use of mixes' composition parameters alone, as shown by the model evaluation functions: in particular, the values of the  $R^2_{adj}$  parameter (last column of Table 2), a modified version of  $R^2$  which assesses the effect of adding predictors to a model, increase with the use of MS or MQ, showing that the new independent term improves the model more than would be expected by chance, but the percentage gain in model accuracy is really paltry. In fact, although the percentage variation in MAE between  $MIX_{SNN}$  and  $MS_{SNN}$  is +23.4%, in terms of  $R^2_{adj}$  the gain is only +0.29% and therefore such that it may not justify the use of additional data, such as any results of the Marshall test. Therefore, even though in the following the results of the  $MS_{SNN}$  will be presented as it represents the best model among those evaluated, it is advised to use the Marshall parameters when the prediction of the IT-CY asphalt mix stiffness from the main composition parameters of the HMACs should not be satisfactory.

Figure 7 shows the linear regression model between the network outputs and the experimental targets in the test data set: the high value of the coefficient of determination  $R^2 = 0.99384$ , very close to unity, suggests that the  $MS_{SNN}$  is capable of accurately predicting

the average mechanical response of HMAC specimen replicas in the training data set and that the problem of overfitting has been avoided despite the small data set available. The same is true for the other SNNs, which show slightly lower performance (please refer to the second-to-last column of Table 2). Therefore, the data augmentation approach designed and implemented in the present study has been particularly suitable for the modeling problem in question and represents an effective approach whenever multiple replicates of different mixtures were available and it was desired to model the average mechanical response resulting from different aggregate compositions, binder percentages, and so on. The Stiffness-Stability comparison between network outputs and experimental targets in the training data set is presented in Figure 8: the absence of an exact matching between the blue and orange points is further evidence of the MS<sub>SNN</sub>'s capability to disregard the behavior of the single replica in order to capture the average mechanical response resulting from the mixtures' composition parameters.

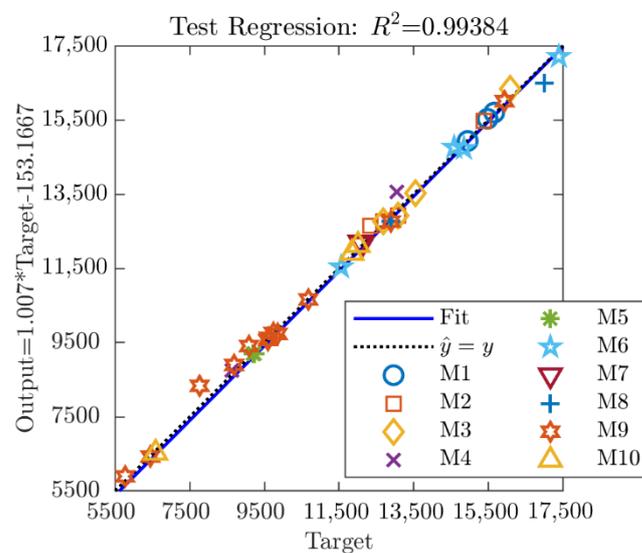


Figure 7. Linear regression analysis on the test set for the MS<sub>SNN</sub>.

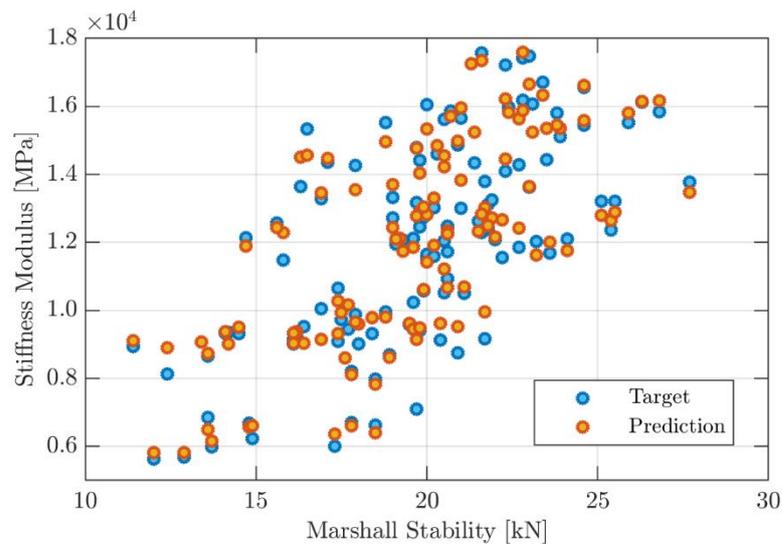


Figure 8. Comparison between experimental target and MS<sub>SNN</sub> output on the training data set.

Finally, it is necessary to point out that, generally, the data sets used for modeling the mechanical behavior of ACs are relatively small, due to the time and cost that the execution of a multitude of laboratory tests on multiple replicas of the same mixture and/or different

variants implicitly requires. Therefore, even though the experimental data set exploited in this paper may seem limited for neural network modeling, there is evidence that the application of shallow neural networks, coupled with an appropriate data augmentation technique, avoids model overfitting and, consequently, an effective procedure for small HMACs data sets has been validated. Accordingly, there is the possibility of extending the current study (using the same techniques) to other mixtures, characterized by binders and aggregates different from those considered.

## 5. Conclusions

The following conclusions can be drawn from this study:

1. The IT-CY Stiffness Modulus of 115 Marshall test specimens of high-modulus asphalt mixtures prepared in the laboratory with reclaimed asphalt pavement or polymer-modified bitumen has been investigated, according to EN 12697-26 Annex C, as part of real case-mix design processes.
2. There were good correlation strengths between the Stiffness Modulus and the Marshall test results, with high stiffness levels associated with high stability or quotient levels. Therefore, one of these empirical parameters could be used as an input feature, along with some parameters related to the HMAC composition, to improve the performance of a predictive model.
3. Machine Learning approaches have been employed for the development of a predictive model of the HMACs' stiffness modulus: the focus was particularly on Shallow Neural Networks, given their simple structure and good computational power even with respect to small data sets.
4. A Bayesian optimization process was used to identify the neural topology, as well as the transfer function, optimal for the required modeling. In addition, a data augmentation strategy was designed for the case of the IT-CY test.
5. By employing different performance metrics, it was possible to compare the optimal models obtained by varying the input feature related to the empirical Marshall test results. The SNN, which showed the best prediction accuracy of the average mechanical response of HMAC variants, receives as input a 6-component features vector, i.e., the Marshall stability (kN), the bitumen content (% by mass of mix), the air voids content (%), maximum and average bulk density ( $\text{g}/\text{cm}^3$ ), along with a categorical variable that distinguishes the bitumen type and RAP percentages (values from 0 to 10); such input features vector is processed by 6 neurons in the hidden layer characterized by a hyperbolic tangent activation unit.
6. A worthwhile future development could be an in-depth investigation of aggregate grading curves' influence on stiffness predictions by including additional inputs connected with mixture proportion. Another valuable alternative would be to replace during the modeling phase, variables referring to empirical properties (i.e., Marshall Stability) with those referring to pavement performance. In this way, it would be possible to predict by machine learning approaches fatigue life and/or permanent deformation resistance. Such an attempt would represent a significant step toward performance-based mixture design.

**Author Contributions:** Conceptualization, N.B., M.M. and J.V.; methodology, N.B., M.M., F.R., J.V. and P.V.; software, M.M.; validation, N.B., M.M., F.R., J.V. and E.M.; formal analysis, N.B., M.M., F.R., J.V., P.V. and E.M.; investigation, J.V. and P.V.; resources, N.B. and J.V.; data curation, N.B., M.M., F.R., J.V., P.V. and E.M.; writing—original draft preparation, N.B., M.M., F.R., J.V., P.V. and E.M.; writing—review and editing, N.B., M.M., F.R., J.V., P.V. and E.M.; visualization, E.M.; supervision, N.B.; project administration, N.B. and J.V.; funding acquisition, N.B. and J.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in Table 1 of the current paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Espersson, M. Effect in the high modulus asphalt concrete with the temperature. *Constr. Build. Mater.* **2014**, *71*, 638–643. [[CrossRef](#)]
2. Miró, R.; Valdés, G.; Martínez, A.; Segura, P.; Rodríguez, C. Evaluation of high modulus mixture behaviour with high reclaimed asphalt pavement (RAP) percentages for sustainable road construction. *Constr. Build. Mater.* **2011**, *25*, 3854–3862. [[CrossRef](#)]
3. Rys, D.; Judycki, J.; Pszczola, M.; Jaczewski, M.; Mejlun, L. Comparison of low-temperature cracks intensity on pavements with high modulus asphalt concrete and conventional asphalt concrete bases. *Constr. Build. Mater.* **2017**, *147*, 478–487. [[CrossRef](#)]
4. Corté, J.F. Development and uses of hard-grade asphalt and of high-modulus asphalt mixes in France. *Transp. Res. Circ.* **2001**, *503*, 12–31.
5. Newcomb, D.E.; Willis, R.; Timm, D.H. *Perpetual Asphalt Pavements—A Synthesis*; Asphalt Pavement Association of Michigan: Okemos, MI, USA, 2002.
6. Geng, H.; Clopotel, C.S.; Bahia, H.U. Effects of high modulus asphalt binders on performance of typical asphalt pavement structures. *Constr. Build. Mater.* **2013**, *44*, 207–213. [[CrossRef](#)]
7. Marvila, M.T.; Azevedo, A.R.G.; Alexandre, J.; Zanelato, E.B.; Azeredo, N.G.; Simonassi, N.T.; Monteiro, S.N. Correlation between the properties of structural clay blocks obtained by destructive tests and Ultrasonic Pulse Tests. *J. Build. Eng.* **2019**, *26*, 100869. [[CrossRef](#)]
8. Pasetto, M.; Baldo, N. Re-use of industrial wastes in cement bound mixtures for road construction. *Environ. Eng. Manag. J.* **2018**, *17*, 417–426. [[CrossRef](#)]
9. Wang, L.; Song, Z.; Zeng, S.; Liu, J.; Ma, K. Exploration of the load fatigue test method for cement concrete wheels. *Case Stud. Constr. Mater.* **2022**, *16*, e00793. [[CrossRef](#)]
10. Lee, H.J.; Lee, J.H.; Park, H.M. Performance evaluation of high modulus asphalt mixtures for long life asphalt pavements. *Constr. Build. Mater.* **2007**, *21*, 1079–1087. [[CrossRef](#)]
11. Moghaddam, T.; Baaj, H. Rheological characterization of high-modulus asphalt mix with modified asphalt binders. *Constr. Build. Mater.* **2018**, *193*, 142–152. [[CrossRef](#)]
12. Si, C.; Zhou, Y.; You, Z.; He, Y.; Chen, E.; Zhang, R. Micro-mechanical analysis of high modulus asphalt concrete pavement. *Constr. Build. Mater.* **2019**, *220*, 128–141. [[CrossRef](#)]
13. Zaumanis, M.; Arraigada, M.; Poulikakos, L.D. 100% recycled high-modulus asphalt concrete mixture design and validation using vehicle simulator. *Constr. Build. Mater.* **2020**, *260*, 119891. [[CrossRef](#)]
14. Baldo, N.; Valentin, J.; Manthos, E.; Miani, M. Numerical Characterization of High Modulus Asphalt Concrete Containing RAP: A Comparison among Optimized Shallow Neural Models. In *IOP Conference Series: Materials Science and Engineering, Proceedings of the 5th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium, Prague, Czech Republic, 15–19 June 2020*; IOP Publishing: Bristol, UK, 2020. [[CrossRef](#)]
15. Yun, T.; Kim, Y.R. Viscoelastoplastic modeling of the behavior of hot mix asphalt in compression. *KSCE J. Civ. Eng.* **2013**, *17*, 1323–1332. [[CrossRef](#)]
16. Di Benedetto, H.; Sauzéat, C.; Clec’h, P. Anisotropy of bituminous mixture in the linear viscoelastic domain. *Mech. Time-Depend. Mater.* **2016**, *20*, 281–297. [[CrossRef](#)]
17. Rahmani, E.; Darabi, M.K.; Little, D.N.; Masad, E.A. Constitutive modeling of coupled aging-viscoelastic response of asphalt concrete. *Constr. Build. Mater.* **2017**, *131*, 1–15. [[CrossRef](#)]
18. Darabi, M.K.; Huang, C.W.; Bazzaz, M.; Masad, E.A.; Little, D.N. Characterization and validation of the nonlinear viscoelastic-viscoplastic with hardening-relaxation constitutive relationship for asphalt mixtures. *Constr. Build. Mater.* **2019**, *216*, 648–660. [[CrossRef](#)]
19. Specht, L.P.; Khatchaturian, O.; Brito, L.A.T.; Ceratti, J.A.P. Modeling of asphalt-rubber rotational viscosity by statistical analysis and neural networks. *Mater. Res.* **2007**, *10*, 69–74. [[CrossRef](#)]
20. Mirzahosseini, M.R.; Aghaeifar, A.; Alavi, A.H.; Gandomi, A.H.; Seyednour, R. Permanent deformation analysis of asphalt mixtures using soft computing techniques. *Expert Syst. Appl.* **2011**, *38*, 6081–6100. [[CrossRef](#)]
21. Androjić, I.; Marović, I. Development of artificial neural network and multiple linear regression models in the prediction process of the hot mix asphalt properties. *Can. J. Civ. Eng.* **2017**, *44*, 994–1004. [[CrossRef](#)]
22. Alrashydah, E.I.; Abo-Qudais, S.A. Modeling of creep compliance behavior in asphalt mixes using multiple regression and artificial neural networks. *Constr. Build. Mater.* **2018**, *159*, 635–641. [[CrossRef](#)]
23. Ziari, H.; Amini, A.; Goli, A.; Mirzaiyan, D. Predicting rutting performance of carbon nano tube (CNT) asphalt binders using regression models and neural networks. *Constr. Build. Mater.* **2018**, *160*, 415–426. [[CrossRef](#)]
24. Montoya, M.A.; Haddock, J.E. Estimating asphalt mixture volumetric properties using seemingly unrelated regression equations approaches. *Constr. Build. Mater.* **2019**, *225*, 829–837. [[CrossRef](#)]
25. Baldo, N.; Manthos, E.; Pasetto, M. Analysis of the mechanical behaviour of asphalt concretes using artificial neural networks. *Adv. Civ. Eng.* **2018**, *2018*, 1650945. [[CrossRef](#)]

26. Ozsahin, T.S.; Oruc, S. Neural network model for resilient modulus of emulsified asphalt mixtures. *Constr. Build. Mater.* **2008**, *22*, 1436–1445. [[CrossRef](#)]
27. Tapkın, S.; Çevik, A.; Uşar, Ü. Accumulated strain prediction of polypropylene modified marshall specimens in repeated creep test using artificial neural networks. *Expert Syst. Appl.* **2009**, *36*, 11186–11197. [[CrossRef](#)]
28. Saoudi, B.S.; Haddadi, S.H. Predicting creep deformation of asphalts modified with polymer using artificial neural networks. *Mag. Civ. Eng.* **2021**, *101*, 10106. [[CrossRef](#)]
29. Xiao, F.; Amirkhania, S.; Juang, C.H. Prediction of fatigue life of rubberized asphalt concrete mixtures containing reclaimed asphalt pavement using artificial neural networks. *J. Mater. Civ. Eng.* **2009**, *21*, 253–261. [[CrossRef](#)]
30. Ahmed, T.M.; Green, P.L.; Khalid, H.A. Predicting fatigue performance of hot mix asphalt using artificial neural networks. *Road Mater. Pavement Des.* **2017**, *18*, 141–154. [[CrossRef](#)]
31. Ceylan, H.; Schwartz, C.W.; Kim, S.; Gopalakrishnan, K. Accuracy of predictive models for dynamic modulus of hot-mix asphalt. *J. Mater. Civ. Eng.* **2009**, *21*, 286–293. [[CrossRef](#)]
32. Gong, H.; Sun, Y.; Dong, Y.; Han, B.; Polaczyk, P.; Hu, W.; Huang, B. Improved estimation of dynamic modulus for hot mix asphalt using deep learning. *Constr. Build. Mater.* **2020**, *263*, 119912. [[CrossRef](#)]
33. Zavrtnik, N.; Prosen, J.; Tušar, M.; Turk, G. The use of artificial neural networks for modeling air void content in aggregate mixture. *Autom. Constr.* **2016**, *63*, 155–161. [[CrossRef](#)]
34. Hussain, F.; Ali, Y.; Irfan, M. Quantifying the Differential Phase Angle Behaviour of Asphalt Concrete Mixtures Using Artificial Neural Networks. *Int. J. Pavement Res. Technol.* **2021**, *121235*, 1–19. [[CrossRef](#)]
35. Baldo, N.; Manthos, E.; Miani, M. Stiffness modulus and marshall parameters of hot mix asphalts: Laboratory data modeling by artificial neural networks characterized by cross-validation. *Appl. Sci.* **2019**, *9*, 3502. [[CrossRef](#)]
36. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **2015**, *104*, 148–175. [[CrossRef](#)]
37. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for hyper-parameter optimization. In Proceedings of the 25th Annual Conference on Neural Information Processing Systems (NIPS 2011), Granada, Spain, 12–14 December 2011; Neural Information Processing Systems Foundation: La Jolla, CA, USA, 2011; Volume 24.
38. Bergstra, J.; Yamins, D.; Cox, D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 115–123.
39. Hoang, V.H.; Nguyen, Q.T.; Tran, A.T.; Tran, T.C.H.; Do, T.A. Mechanical behavior of the asphalt wearing surface on an orthotropic steel bridge deck under cyclic loading. *Case Stud. Constr. Mater.* **2022**, *16*, e00836. [[CrossRef](#)]
40. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control. Signal Syst.* **1989**, *2*, 303–314. [[CrossRef](#)]
41. Demuth, H.B.; Beale, M.H.; De Jess, O.; Hagan, M.T. *Neural Network Design*; Martin Hagan: Stillwater, OK, USA, 2014.
42. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. *Neurocomputing: Foundations of Research*; MIT Press: Cambridge, MA, USA, 1988; pp. 696–699.
43. Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **1994**, *5*, 989–993. [[CrossRef](#)] [[PubMed](#)]
44. Miani, M.; Dunnhofer, M.; Rondinella, F.; Manthos, E.; Valentin, J.; Micheloni, C.; Baldo, N. Bituminous Mixtures Experimental Data Modeling Using a Hyperparameters-Optimized Machine Learning Approach. *Appl. Sci.* **2021**, *11*, 11710. [[CrossRef](#)]
45. Snoek, J.; Larochelle, H.; Adams, R.P. Practical Bayesian Optimization of Machine Learning Algorithms. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 13–16 December 2012; pp. 2951–2959.
46. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.
47. Kushner, H.J. A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise. *J. Basic Eng.* **1964**, *86*, 97–106. [[CrossRef](#)]
48. Srinivas, N.; Krause, A.; Kakade, S.M.; Seeger, M. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. *arXiv* **2009**, arXiv:0912.3995. Available online: <https://arxiv.org/abs/0912.3995> (accessed on 11 June 2021).
49. Mockus, J.; Tiešis, V.; Zilinskas, A. The application of Bayesian methods for seeking the extremum. In *Towards Global Optimization*, 2nd ed.; Dixon, L.C.W., Szego, G.P., Eds.; North Holland Publishing Co.: Amsterdam, The Netherlands, 1978; pp. 117–129.
50. Bull, A.D. Convergence Rates of Efficient Global Optimization Algorithms. *J. Mach. Learn. Res.* **2011**, *12*, 2879–2904. Available online: <https://arxiv.org/abs/1101.3501> (accessed on 11 June 2021).
51. Qi, M.; Zhang, G.P. An investigation of model selection criteria for neural network time series forecasting. *Eur. J. Oper. Res.* **2001**, *132*, 666–680. [[CrossRef](#)]
52. Bowden, G.J.; Maier, H.R.; Dandy, G.C. Input determination for neural network models in water resources applications. Part 2. Case study: Forecasting salinity in a river. *J. Hydrol.* **2005**, *301*, 93–107. [[CrossRef](#)]
53. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.
54. Tapkın, S.; Çevik, A.; Uşar, Ü. Prediction of Marshall test results for polypropylene modified dense bituminous mixtures using neural networks. *Expert Syst. Appl.* **2010**, *37*, 4660–4670. [[CrossRef](#)]
55. Ozgan, E. Artificial neural network based modelling of the Marshall Stability of asphalt concrete. *Expert Syst. Appl.* **2011**, *38*, 6025–6030. [[CrossRef](#)]

- 
56. Zumrawi, M.M.; Khalil, F.O. Experimental study of steel slag used as aggregate in asphalt mixture. *Am. J. Constr. Build. Mater.* **2017**, *2*, 26–32. [[CrossRef](#)]
  57. Spearman, C. The proof and measurement of association between two things. *Am. J. Psychol.* **1904**, *15*, 72–101. [[CrossRef](#)]
  58. Biecek, P.; Burzykowski, T. *Explanatory Model Analysis. Explore, Explain, and Examine Predictive Models*; Chapman and Hall/CRC: New York, NY, USA, 2021.
  59. Legates, D.R.; McCabe, G.J. A refined index of model performance: A rejoinder. *Int. J. Climatol.* **2013**, *33*, 1053–1056. [[CrossRef](#)]