sustainability

MDPI

Article

Prediction of Fracture Toughness of Intermediate Layer of Asphalt Pavements Using Artificial Neural Network

Dong-Hyuk Kim [1], Ha-Yeong Kim [1], Ki-Hoon Moon [2] and Jin-Hoon Jeong [1,*]

[1] Department of Civil Engineering, Inha University, 100 Inha-ro, Michuhol-gu, Incheon 22212, Korea; my91kim@gmail.com (D.-H.K.); ha00303@naver.com (H.-Y.K.)
[2] Korea Expressway Corporation Research Institute, Korea Expressway Corporation, 24, Dongtansunhwan-daero 17-gil, Hwaseong-si 18489, Korea; zetamkh@ex.co.kr
* Correspondence: jhj@inha.ac.kr; Tel.: +82-32-860-7574; Fax: +82-32-873-7560

Abstract: For the sustainable management of pavements, predicting the condition of the pavement structure using a consistent and accurate method is necessary. The intermediate layer situated immediately below the surface layer has the greatest effect on the condition of the pavement structure. As a result, to accurately predict the condition of a pavement structure, the mechanistic properties of the intermediate layer are very important. Fracture toughness (FT)—a mechanistic property of the intermediate layer—is an important factor in predicting the distress that develops on the pavement surface. However, measuring FT consistently is practically impossible by coring asphalt pavement over the entire pavement section. Therefore, an artificial neural network (ANN) model—developed by using pavement surface conditions and traffic volume—can predict the FT of the intermediate layer of expressway asphalt pavements. Several ANN models have been developed by applying various optimizers, ANN structures, and preprocessing methods. The optimal model was selected by analyzing the predictive performance, error stability, and distribution of the developed models. The final selected model showed small stable errors, and the distributions of the predicted and measured FTs were similar. Furthermore, FT limits were set with outliers. Future asphalt pavement conditions can be predicted throughout the expressway network by using the proposed model without coring samples.

Keywords: asphalt pavement; intermediate layer; fracture toughness; artificial neural network (ANN)

Citation: Kim, D.-H.; Kim, H.-Y.; Moon, K.-H.; Jeong, J.-H. Prediction of Fracture Toughness of Intermediate Layer of Asphalt Pavements Using Artificial Neural Network. *Sustainability* **2022**, *14*, 7927. https://doi.org/10.3390/su14137927

Academic Editor: Adelino Jorge Lopes Ferreira

Received: 24 May 2022
Accepted: 24 June 2022
Published: 29 June 2022

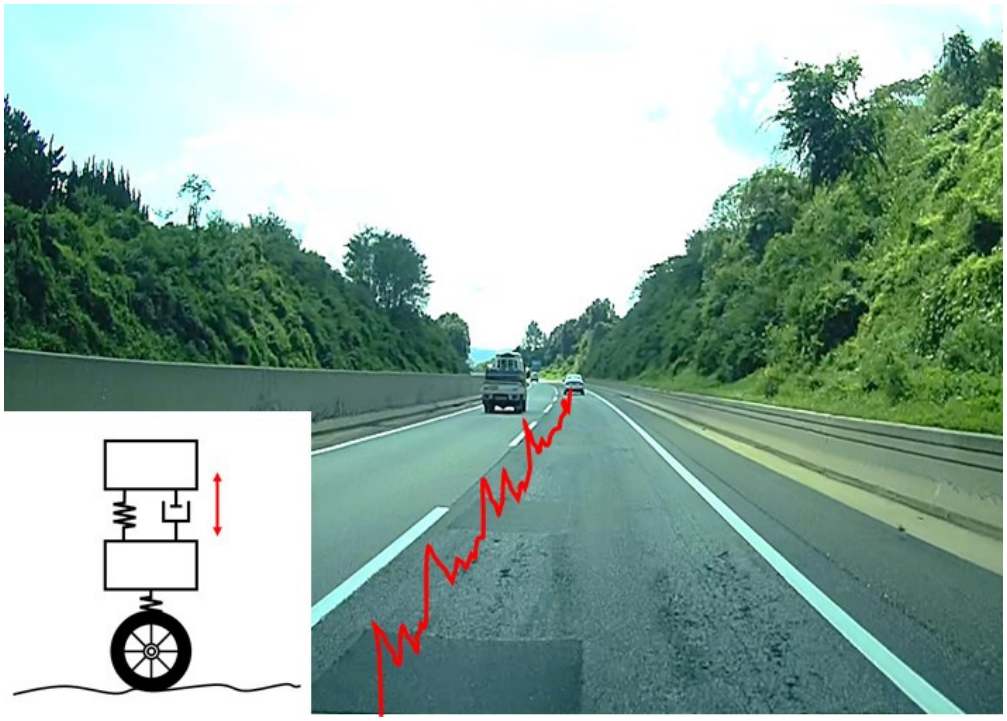
1. Introduction

Since 1996, the Korea Expressway Corporation has implemented an expressway pavement management system (HPMS) to provide sustainable expressway pavements [1]. The HPMS is used to plan pavement management, investigate pavement conditions, organize collected data, and analyze organized data for Korean expressways. Figure 1 shows how to measure the international roughness index (IRI), rut depth (RD), and surface distress (SD) as asphalt pavement conditions for HPMS.

Because IRI, RD, and SD describe only the conditions of the pavement surface, they cannot be used to estimate the conditions for the full thickness of asphalt pavement. The conditions of the surface, intermediate, and base layers of asphalt pavements differ from those of concrete pavements, which have similar conditions throughout the slab thickness. The state of the intermediate layer directly under the surface layer, in particular, influences the conditions visible on the pavement surface [2]. Thus, long-term management plans based on conditions of the intermediate layer must be employed to ensure high usability and economic efficiency of asphalt pavements.

The most common method of investigating the conditions of the intermediate layer involves examining the samples cored from asphalt pavements or measuring mechanistic

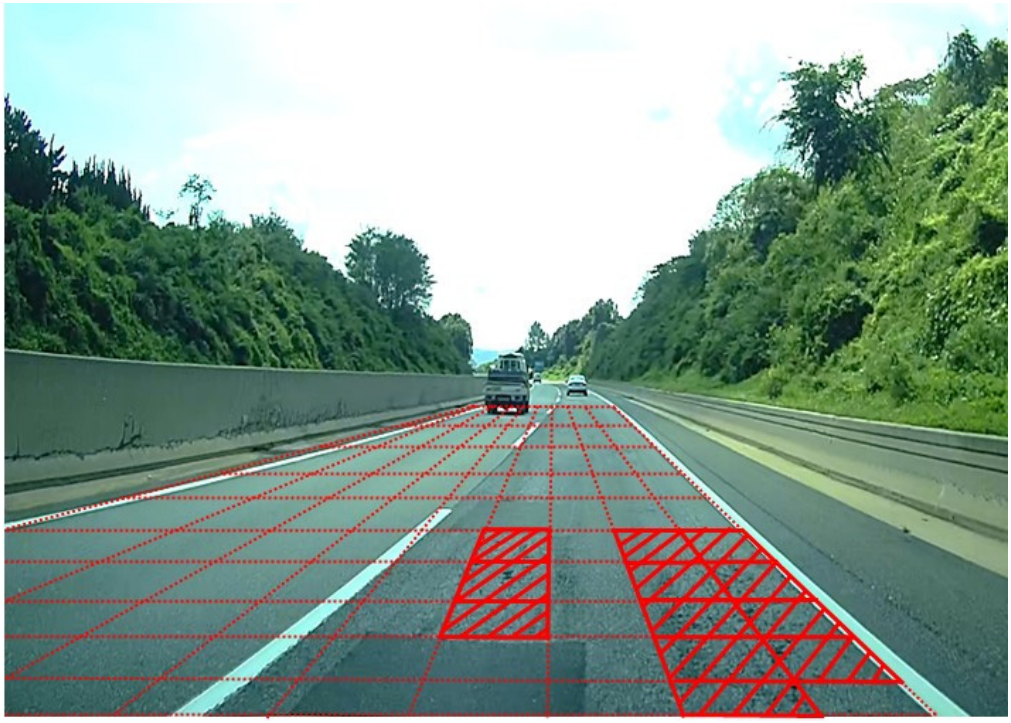
properties such as fracture toughness (FT) or indirect tensile strength. The indirect tensile strength test is an advanced method to evaluate the cracking resistance of asphalt pavement that replaces the Marshall stability test [3]. However, the indirect tensile strength has the disadvantage of not taking the deformation of the sample into account because it is calculated from the load applied at the time of destruction and the sectional area of the sample.



(**a**)

(**b**)

Figure 1. *Cont.*



(**c**)

Figure 1. Investigation method for (**a**) IRI, (**b**) RD, and (**c**) SD.

We calculated the FT by using the load–deformation curve obtained from the indirect tensile strength test. As a material property primarily, FT is used to explain the development of fatigue and low-temperature cracks, in addition to crack propagation [4]. For asphalt pavement, cracks at the nanometer and micrometer levels in a sublayer propagate to the surface layer and grow to more significant levels. The FT and the information before the destruction of the sample can adequately explain the mechanism for the development and propagation of cracks [5]. Thus, the distress propagation from the intermediate layer to the surface layer can be indirectly predicted by examining the FT of the intermediate layer.

The FT of the intermediate layer should be included in the HPMS database using a simple method for long-term sustainable pavement management. However, examining the FT of the intermediate layer by coring samples from entire expressway asphalt pavement sections is practically impossible. Thus, the HPMS database should include FT predictions that use data already collected from all expressway asphalt pavement sections. For example, an artificial neural network (ANN) model was developed to predict the indirect tensile strength of the intermediate layer of asphalt pavements by using IRI, RD, SD, and the equivalent single-axle load (ESAL) as the independent variables collected from the entire expressway asphalt pavement sections for the HPMS [6].

Using a linear regression model to predict the material properties of the intermediate asphalt pavement layers can be difficult because of the complex correlation between the material properties and influencing factors [7]. Several models predicting the material properties of the intermediate layer of asphalt pavements have been developed using ANNs. Gandhi et al. developed an ANN-based indirect tensile strength prediction model for asphalt pavement by using the asphalt binder content, conditioning duration, anti-stripping agents, aggregate source, and asphalt binder sources as variables [8]. The model performed well enough to be used in indirect tensile strength predictions for other projects. Josipa et al. developed a model for predicting pavement performance using ANN [9]. The developed model performed well enough to be used for pavement evaluation and maintenance strategy establishment at both project and network levels. Furthermore, the ANN models consider all the complex linear and nonlinear correlations between the

variables and have low standards for normality and independence [10–12]. Therefore, we developed an FT prediction model of the asphalt pavement intermediate layer by using ANN.

In this study, we developed an ANN model predicting the FT of the intermediate layer of asphalt pavements by using IRI, RD, SD, and ESAL as independent variables. Because ESAL quantifies the effect of traffic load on pavement, it can affect IRI, RD, and SD. However, we judged IRI, RD, and SD to be independent of ESAL because they were influenced not only by ESAL but also by construction quality, material condition, and environmental load.

We applied several pre-treatment methods to the variables to improve the model's prediction performance. We fabricated 100 ANN models by using combinations of different structures of hidden layers, nodes, and optimizers. To select three candidate models with the best prediction performance, we used the training dataset to train the models. We compared the errors between the FT predictions from the training and validation datasets as well as the measured FTs to examine the generality of the candidate model learning. We also compared the error rates between the FT predictions from the training and test datasets as well as the measured FTs to examine the applicability of the candidate models to the new input data. We selected the final model by comparing the averages and standard deviations of the FT predictions for the entire expressway asphalt pavement with the measured FTs. To complement the final model, we set lower and upper limits for FT predictions.

2. Methods

2.1. Measurement, Collection, and Pretreatment of Data

Figure 2 shows cylindrical samples of both the surface and intermediate layers of the asphalt pavement that were cored from 66 random locations on the No. 25 expressway route (Nonsan-Cheonan Expressway). We conducted an indirect tensile strength test according to KS F 2382 for the 50 mm thick upper part of the intermediate layer cut from the core samples. We measured the deformation of the core samples at 25 °C based on the load at a loading velocity of 50 mm/min. Figure 3 shows how the FT—defined as the area surrounded by the load–deformation curve and deformation axis—was calculated for each core sample.

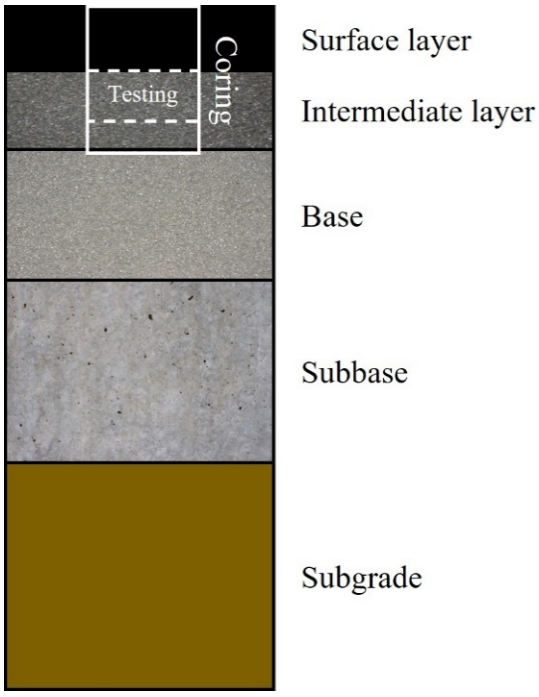


Figure 2. Coring and testing locations.

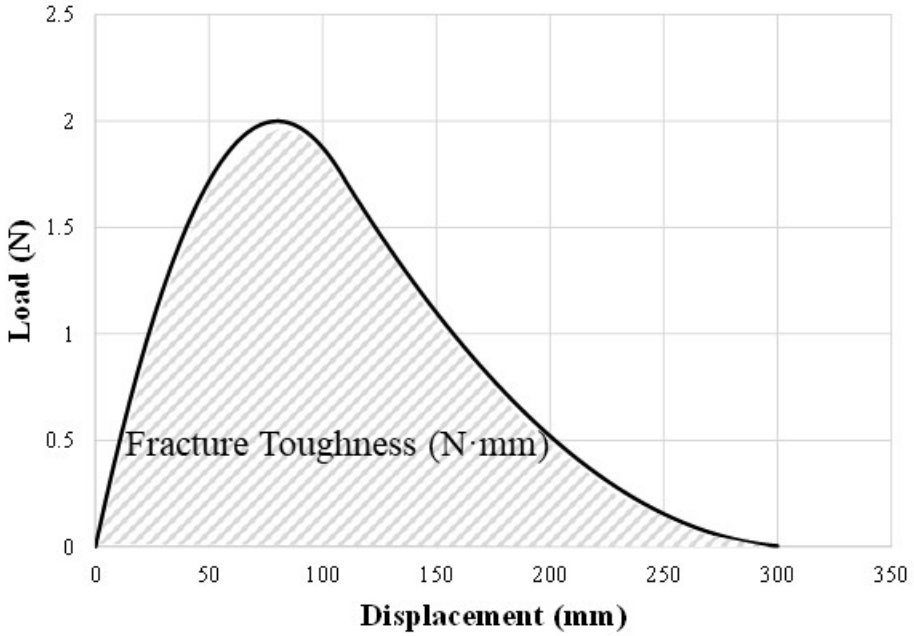


Figure 3. Concept of fracture toughness.

We collected HPMS data for 66 unit sections with lengths of 100 m, where the samples were cored. The HPMS database includes general information—such as the route, direction, lane, distance from the origin, and management branch—as well as pavement conditions, including IRI, RD, SD, and annual average daily traffic (AADT) by vehicle type. In this study, we assumed that the IRI, RD, SD, and AADT—commonly measured for entire expressway asphalt pavement sections—influence the FT and collected them to develop an ANN model for predicting the FT of entire expressway asphalt pavement sections.

Distress, such as cracks that develop in the intermediate layer, reduces the FT at the location and can propagate up to the surface layer [13]. When the FT of the intermediate layer is reduced due to distress, the unrecoverable deformation at the pavement surface caused by the repetitive passing of heavy vehicles might worsen [13]. The repetitive passing of heavy vehicles also reduces the FT of the intermediate layer [14]. In this study, we used the ESAL, which sums up the multiplication of the ESAL factor and traffic volume by vehicle type and quantifies the effect of heavy vehicles on pavement conditions as an independent variable in the FT prediction model.

The regression models, including the ANN model, should use variables whose normality is verified [15]. Therefore, in this study, we verified the normality of the dependent variable, FT, and independent variables (IRI, RD, SD, and ESAL) of the ANN model by calculating their skewness and kurtosis. We used Equation (1) to calculate the skewness, representing the asymmetry of the data of a variable. The distribution of data leans to the right if the skewness is negative and to the left if the skewness is positive [16].

$$\gamma_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left(\frac{1}{n}\sum_{i}^{n}(x_i - \overline{x})^2\right)^{\frac{3}{2}}} \quad (1)$$

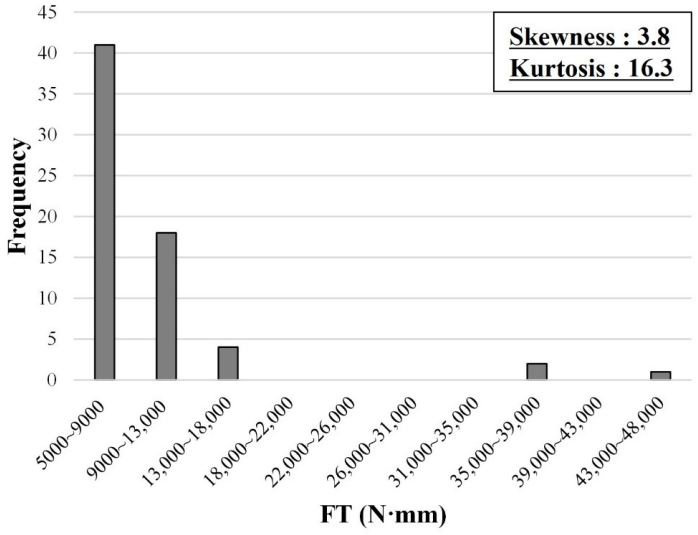
We calculated the kurtosis, which represents the degree of concentration of the data of a variable, using Equation (2). The distribution of the data approaches a normal distribution if the kurtosis approaches an absolute value of 3. The degree of concentration increases if the kurtosis becomes smaller than the absolute value of 3 and decreases if the kurtosis becomes larger than the absolute value of 3 [17]. Although data with skewness close to zero and kurtosis close to the absolute value of three shows high normality, ordinary data cannot satisfy the criteria. Therefore, data with skewness less than the absolute value of

3 and kurtosis less than the absolute value of 10 were set as the criteria for normality, as suggested by Kline [18].

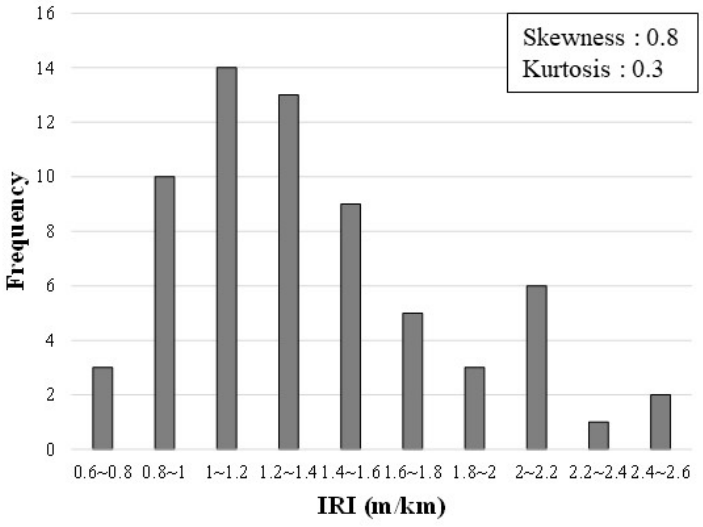
$$g_2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^2} - 3 \quad (2)$$

In Equations (1) and (2), γ_1 represents skewness, g_2 represents kurtosis, n is the number of data elements, x_i is the data element, and $\overline{x}$ is the average of the data elements.

Figure 4 shows the distribution of the dependent variable (FT) and independent variables (IRI, RD, SD, and ESAL), and their respective skewness and kurtosis. The distribution of IRI (skewness: 0.8; kurtosis: 0.3) and RD (skewness: 0.5; kurtosis: −0.5) satisfied the normality visually, and their skewness and kurtosis satisfied the normality criteria. The distribution of ESAL dissatisfied the normality visually because there was relatively little data in the middle part. However, the skewness and kurtosis of ESAL were 0.6 and −1.3, respectively, satisfying both criteria of normality. Meanwhile, FT and SD showed a distribution with the data concentrated on small numbers. Their skewness was 3.8 and 4.4, and their kurtosis values were 16.3 and 22.6, respectively; thus, they did not meet the normality criteria.

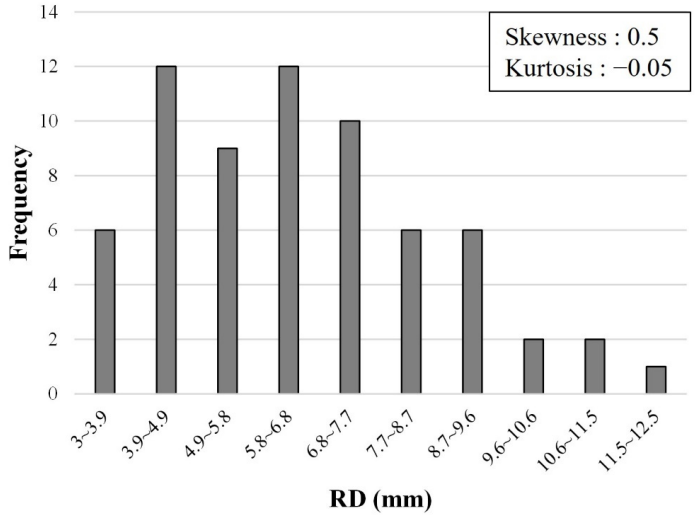


(**a**)

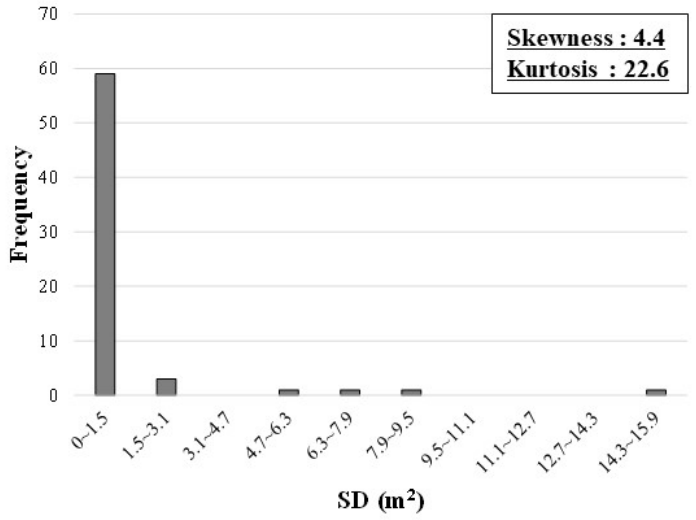


(**b**)

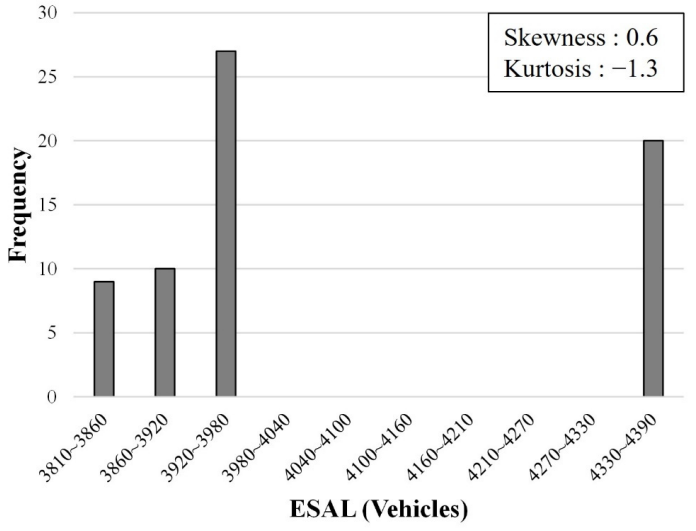
Figure 4. *Cont.*



(**c**)



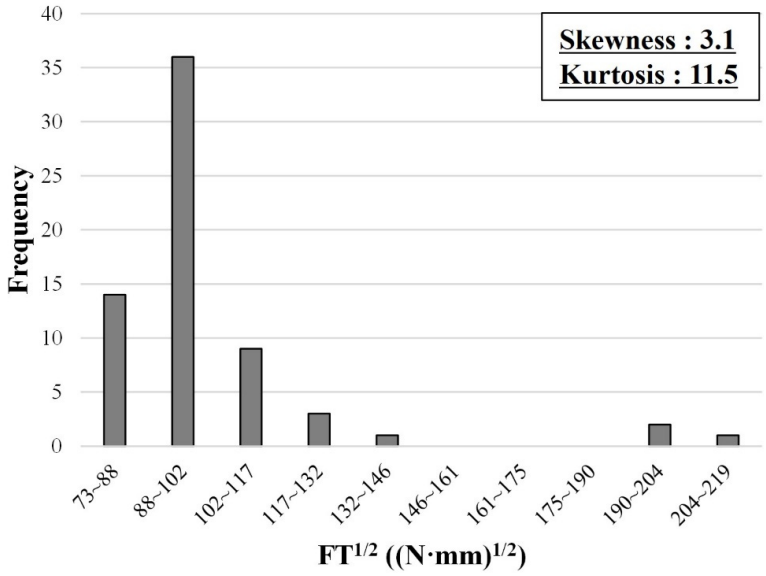
(**d**)



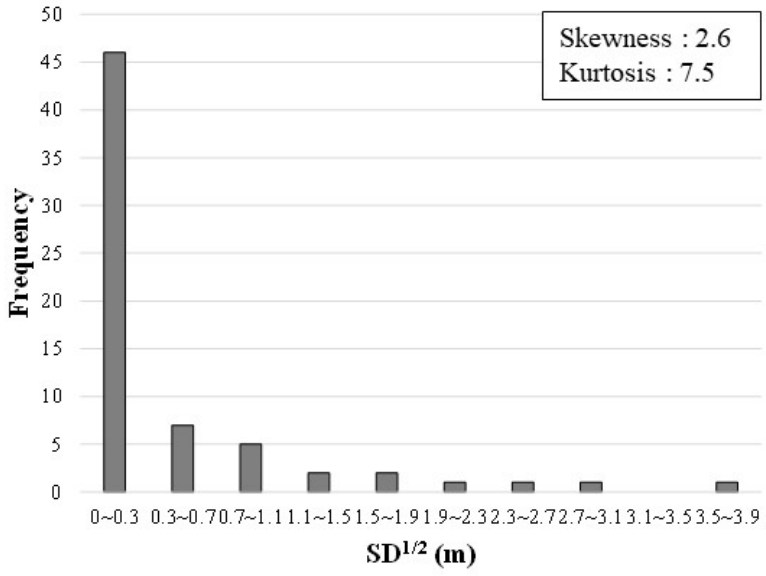
(**e**)

Figure 4. Frequency distribution of raw data of variables. (**a**) FT, (**b**) IRI, (**c**) RD, (**d**) SD, (**e**) ESAL.

If the data do not satisfy the criteria for normality, they can be adjusted to achieve normality by applying square roots, cube roots, or logarithms [19,20]. We visually verified the distributions of FT and SD after applying the square root, cube root, and logarithm, as shown in Figure 5, and examined their skewness and kurtosis. The distribution of the square root of FT in Figure 5a shows skewness and kurtosis of 3.1 and 11.5, respectively; thus, the criteria for normality were not met. However, the square root of SD in Figure 5b shows skewness and kurtosis of 2.6 and 7.5, respectively, both of which satisfy the normality criteria. The cube root of the FT (in Figure 5c) and SD (in Figure 5d) and the logarithm of the FT (in Figure 5e) and SD (in Figure 5f) satisfied the criteria for normality. Consequently, we developed the FT prediction models by using dataset 1, comprising the actual values of IRI, RD, and ESAL, the cube root of FT, and the cube root of SD. We also used dataset 2, comprising the actual values of IRI, RD, and ESAL, the logarithm of FT, and the logarithm of SD, as shown in Table 1.

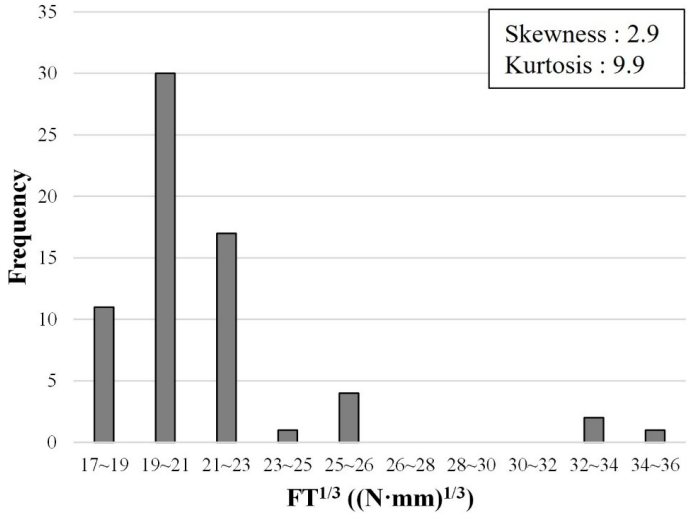


(**a**)

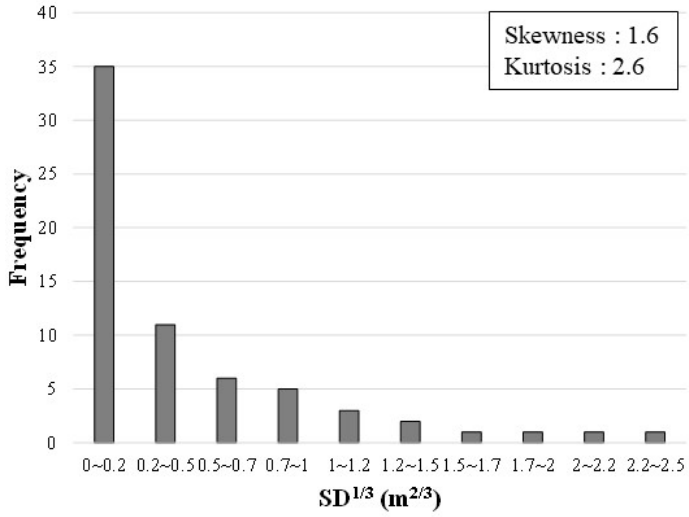


(**b**)

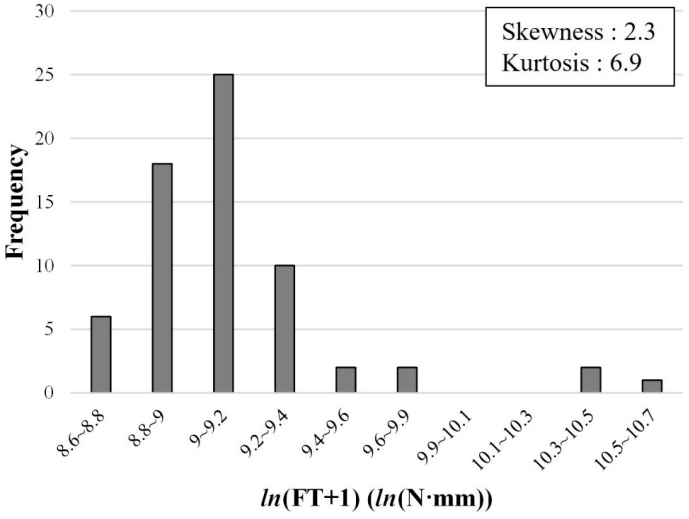
Figure 5. *Cont.*



(**c**)

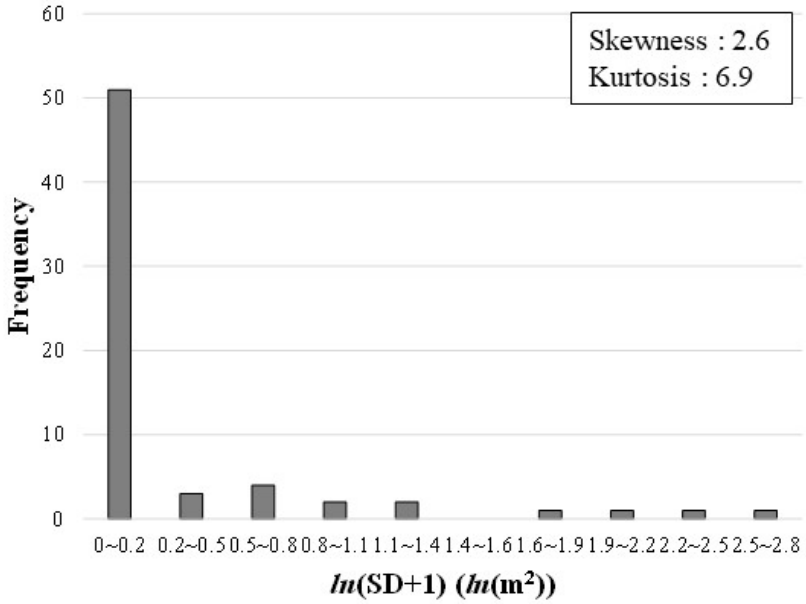


(**d**)



(**e**)

Figure 5. *Cont*.



(**f**)

Figure 5. Frequency distribution of data of FT and SD with different forms. (**a**) $FT^{1/2}$, (**b**) $SD^{1/2}$, (**c**) $FT^{1/3}$, (**d**) $SD^{1/3}$, (**e**) ln(FT + 1), (**f**) ln(SD + 1).

Table 1. Data type of variables in each dataset.

Variables	Dataset 1	Dataset 2
FT	Cube Root	Logarithm
IRI	Raw	Raw
RD	Raw	Raw
SD	Cube Root	Logarithm
ESAL	Raw	Raw

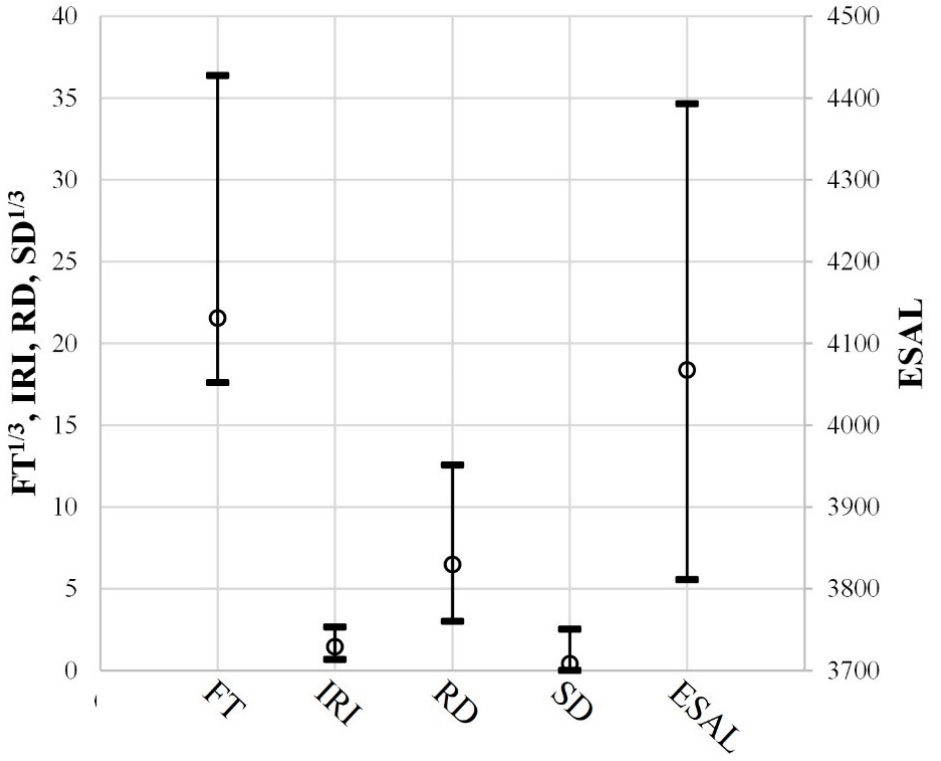
Accurately reflecting the weight of each variable in the regression models, including the ANN model, requires standardization matching of the average and standard deviation of the data for each variable [21]. Figure 6 shows the maximum, minimum, and average values of the data for each variable included in datasets 1 and 2. All variables excluding IRI and SD showed different ranges of data distribution. In particular, ESAL was distributed across a much wider range, with tens to thousands of larger values than those of the other variables. Therefore, we used Equations (3)–(5) to standardize the data elements for each variable to obtain the average and standard deviation of 0 and 1, respectively. Consequently, the distribution of data for all variables included in datasets 1 and 2 became more similar, as shown in Figure 7.

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n} \tag{3}$$

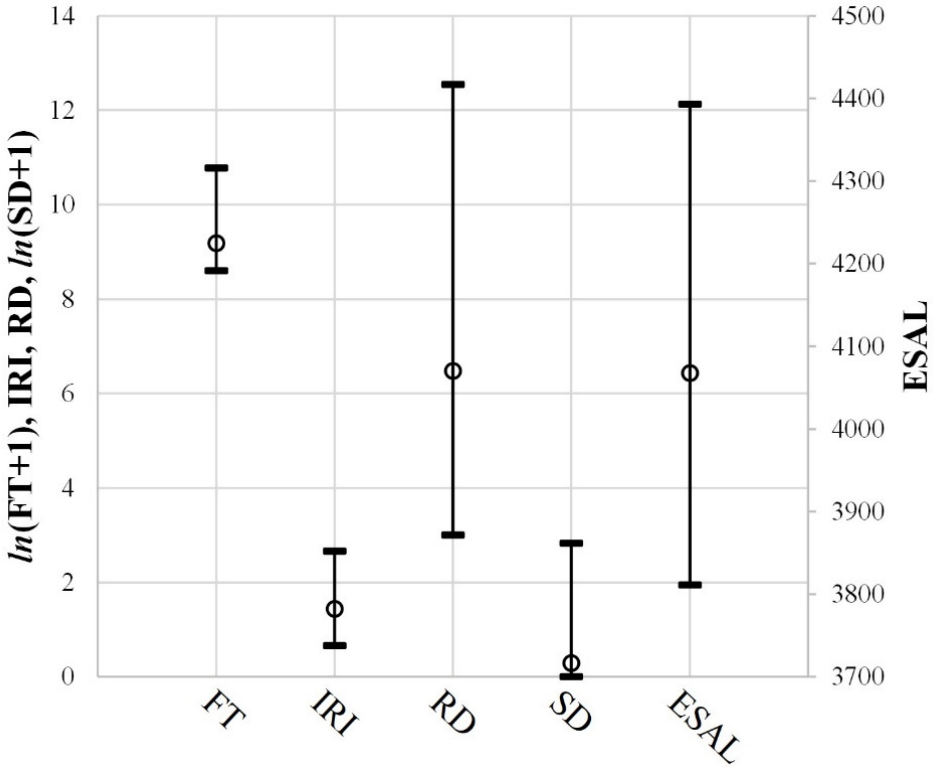
$$\sigma = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n - 1} \tag{4}$$

$$z_i = \frac{x_i - \mu}{\sigma} \tag{5}$$

In Equations (3)–(5), z_i represents the standardized data element, x_i is the data element, μ is the average of the data elements, and σ is the standard deviation of the data element.

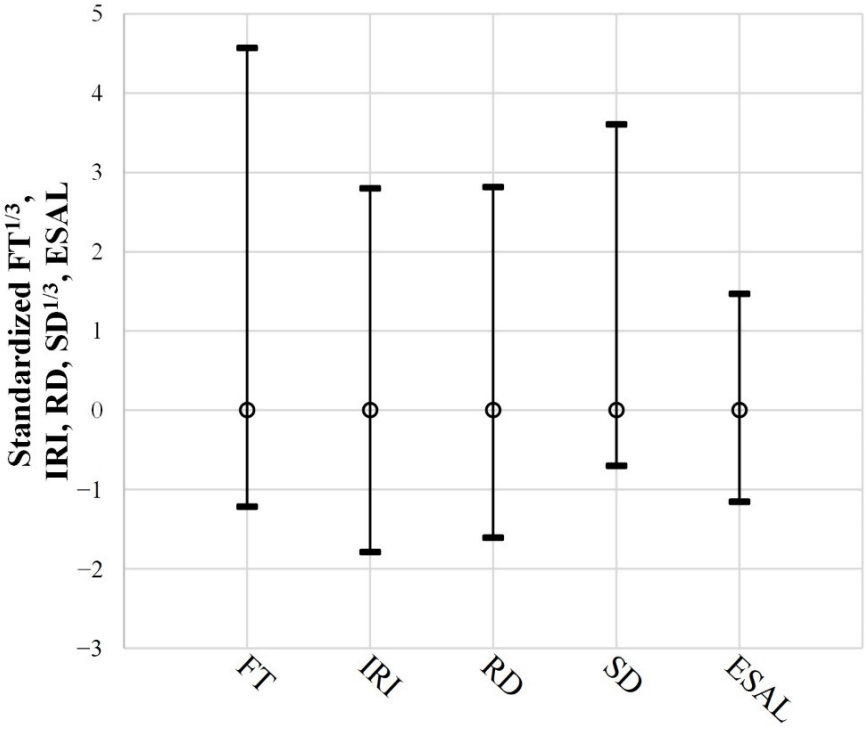


(**a**)

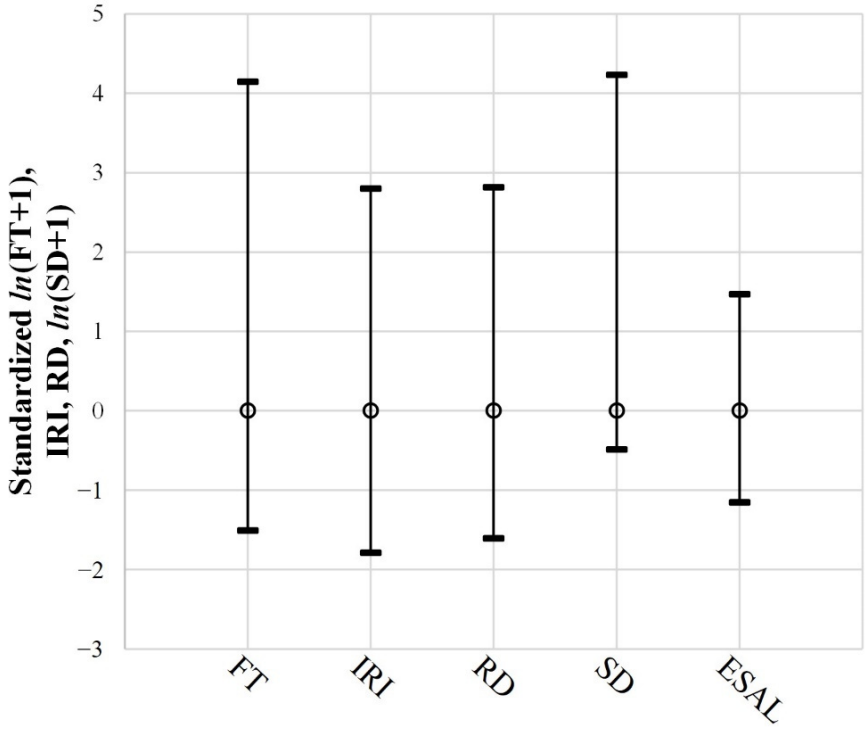


(**b**)

Figure 6. Distribution of the maximum, minimum, and average adjusted variables. (**a**) Dataset 1, (**b**) Dataset 2.



(**a**)



(**b**)

Figure 7. Distribution of the maximum, minimum, and average standardized variables. (**a**) Dataset 1, (**b**) Dataset 2.

2.2. ANN Modeling for FT Prediction

The ANN model is a prediction method that uses machine learning to imitate the function of a human neural network [22]. In the human neural network, a signal is carried from the input neuron, where the stimulus is input, to the hidden neuron based on the coupling strength (line thickness) between the neurons. The response is obtained from the output neuron when the signal is transferred again from the hidden neuron to the output neuron according to the coupling strength.

Figure 8 shows the procedure for constructing the ANN model with the optimal weight and bias, minimizing the error. First, we conducted pretreatment normalizing and standardizing of the data for each variable and determined input layers with input nodes, hidden layers with hidden nodes, output layers with output nodes, and activation functions at each hidden node to construct the ANN structure. Subsequently, we randomly set the weights and biases connecting the nodes in different layers. Using feed-forward processing, the input values are fed into the input nodes to generate results at the output nodes after passing through hidden nodes, weights, and biases. If the errors between the predicted and measured values are larger than the criterion, weights and biases are optimized using an appropriate optimizer and feed-forward processing is repeated until the error became smaller than the criterion.

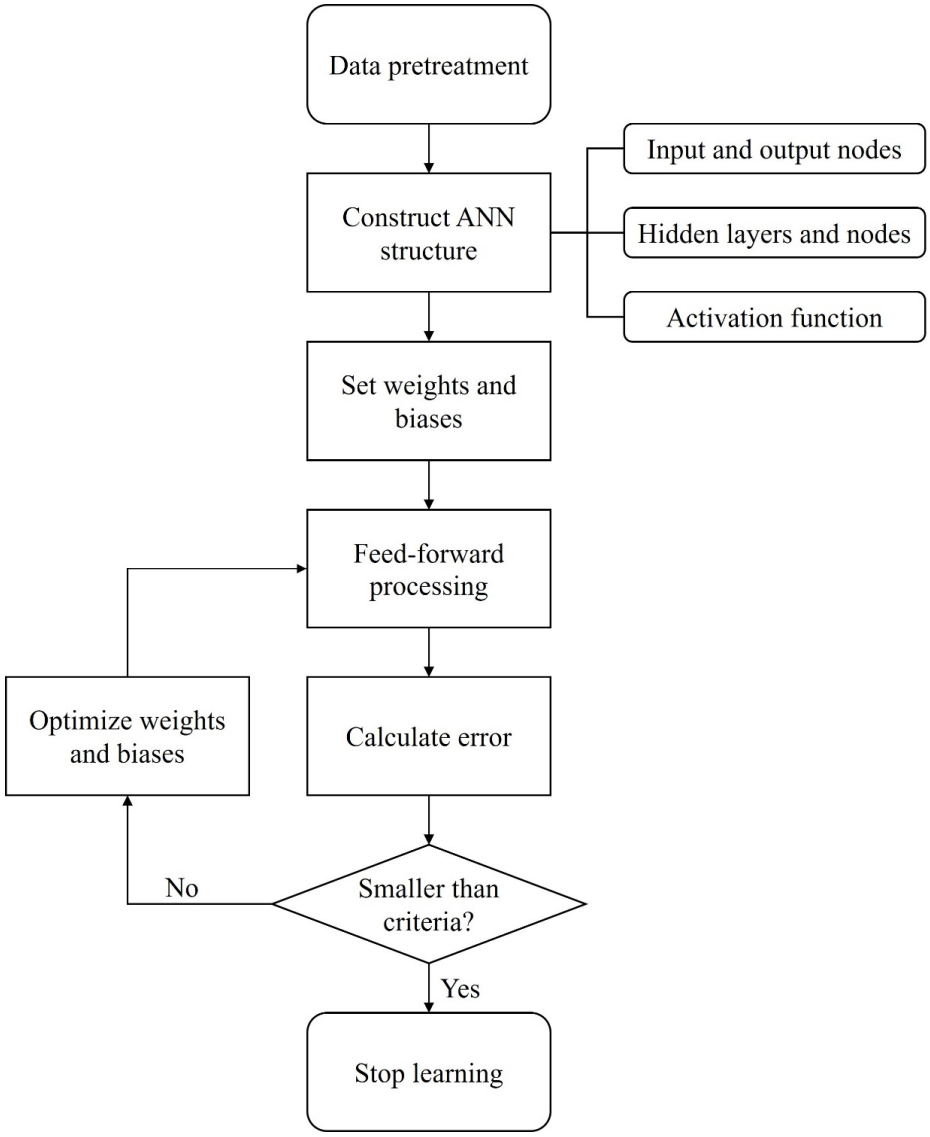


Figure 8. Learning process using ANN.

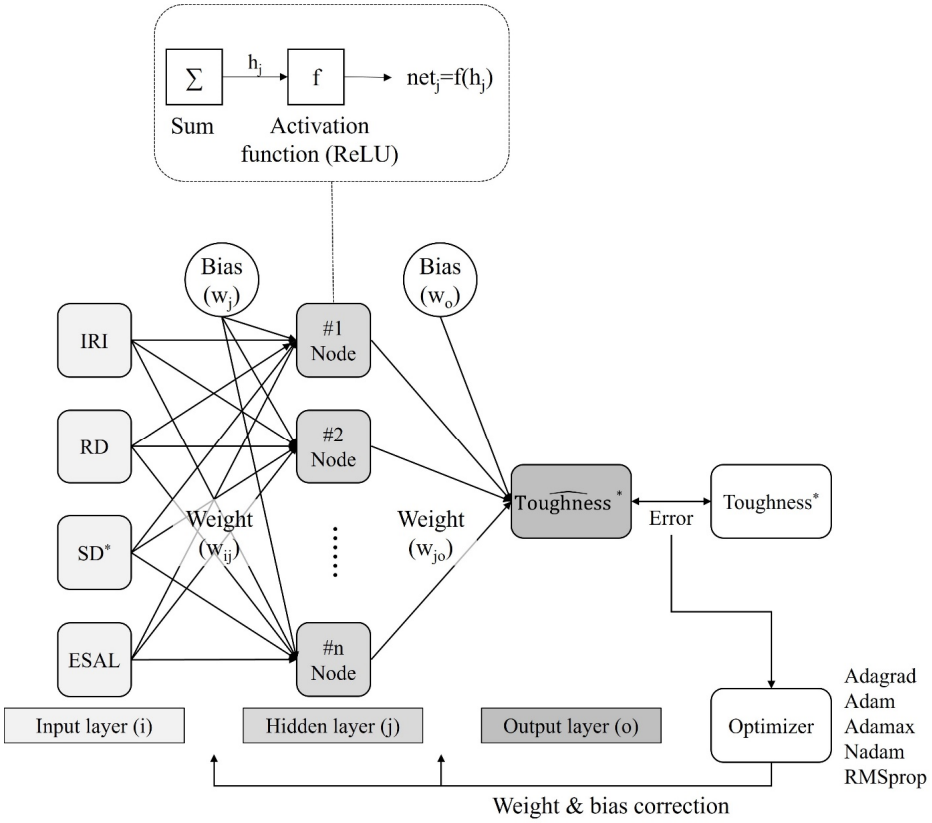
We fabricated ANN models for predicting the FT of the intermediate layer of expressway asphalt pavements by using the independent variables IRI, RD, SD, and ESAL, as shown in Table 2. We fabricated 100 ANN models by combining five types of optimizers, two types of data, and ten combinations of hidden layers and hidden nodes to compare their prediction performance. To determine the label of each model, we combined the abbreviations of the optimizer type (Adagrad: Ag; Adam: Ad; Adamax: Am; Nadam: Nd; and RMSprop: Rp), data type (dataset 1: 1 and dataset 2: 2), and the number of nodes in the first and second hidden layers (first and second numbers are the number of nodes in the first and second hidden layers, respectively). For example, if we determined the label of the model as Ad245, it means that Adam is used as the optimizer with dataset 2, and the numbers of the nodes in the first and second hidden layers are 4 and 5, respectively.

Table 2. ANN model structures.

Dataset	Number of Hidden Layers	Number of Nodes in 1st Hidden Layer	Number of Nodes in 2nd Hidden Layer	Number of Hyper-Parameters	Optimizer
Dataset 1	1	4	-	25	Adagrad (Ag) Adam (Ad) Adamax (Am) Nadam (Nd) RMSprop (Rp)
		5	-	31	
		6	-	37	
		7	-	43	
		8	-	49	
	2	3	6	46	
		3	7	51	
		4	4	45	
		4	5	51	
		5	3	47	
Dataset 2	1	4	-	25	
		5	-	31	
		6	-	37	
		7	-	43	
		8	-	49	
	2	3	6	46	
		3	7	51	
		4	4	45	
		4	5	51	
		5	3	47	

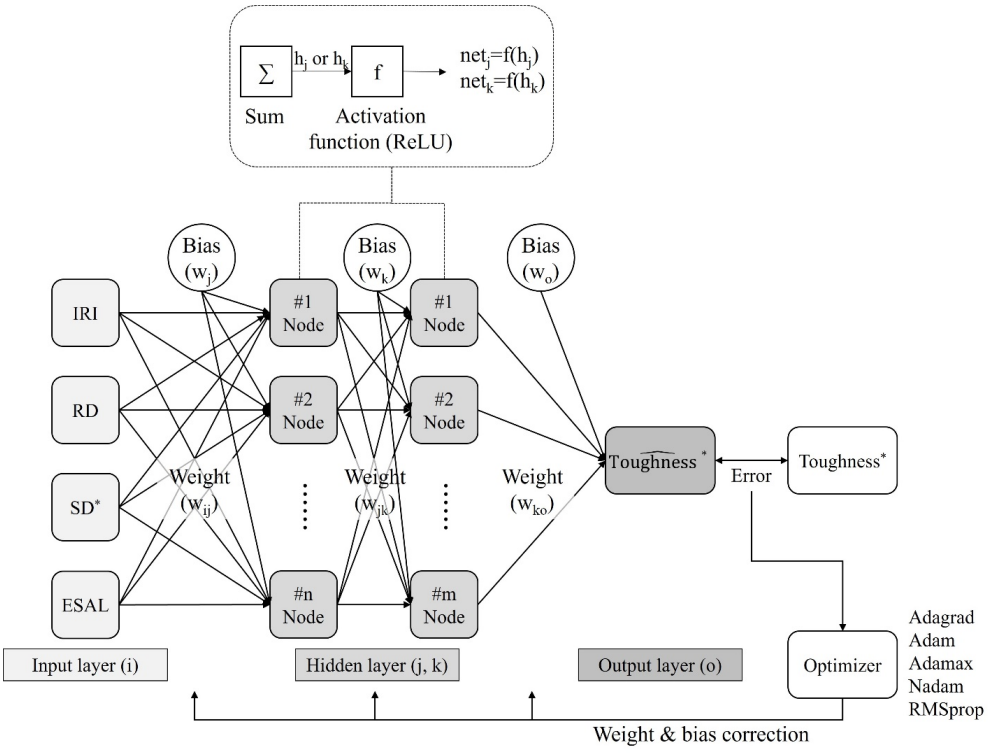
Figure 9 shows the constructed ANN model. The input layer comprises four nodes for IRI, RD, SD, and ESAL, and the hidden layer comprises one or two layers, as shown in Figure 9a,b). The activation function at each node in the hidden layers assigns a nonlinearity to the model. In the ANN model with multiple layers, the sigmoid function—the most basic activation function—can cause a gradient vanishing problem, in which the adjusted values of the weights and biases converge to zero [23]. The proposed ANN model comprises multiple layers. Therefore, we used a rectified linear unit (ReLU) based on Equation (6) as the activation function because it does not cause the vanishing gradient problem and has been used in the most extensive scope of research [24]. The output layer consists of one node that outputs the predicted FT. The model could correct the weights and biases by using an appropriate optimizer when the error between the predicted and measured values is larger than the criterion.

$$f(x) = \begin{cases} 0 \ (x < 0) \\ x \ (x \geq 0) \end{cases} \quad (6)$$



*Data #1: SD1/3, Toughness1/3; Data #2: ln(SD+1), ln(Toughness+1)

(**a**)



*Data #1: SD1/3, Toughness1/3; Data #2: ln(SD+1), ln(Toughness+1)

(**b**)

Figure 9. Structure of ANN model. (**a**) One hidden layer, (**b**) two hidden layers.

We categorized datasets 1 and 2—normalized and standardized by pretreatment—into training and test datasets at an 8:2 ratio. We also assigned 20% of the training dataset to the validation dataset to verify the errors resulting from the training dataset for each training session. Equation (7) determines the number of hyperparameters according to the number of hidden layers and nodes.

$$H_n = \sum_{k=1}^{n-1}((L_k + 1) \times L_{k+1}) + L_n + 1 \quad (7)$$

In Equation (7), H_n is the number of hyperparameters between two layers from $n-1$ to n and L_k is the number of nodes in the hidden layer k.

The hyperparameter, adjusted by training, is the weight and bias between each layer in the ANN model [25]. The prediction performance of the ANN model using the training dataset improves, whereas the one using the validation dataset deteriorates, causing overfitting when the hyperparameter is too large. However, errors caused by both the training and validation datasets increase when the hyperparameter is too small. Therefore, an appropriate number of hyperparameters should be used to limit the number of hidden layers and nodes [26].

In this study, we limited the number of hyperparameters to 53, corresponding to the number of training datasets—that is, 80% of the total data (66). Accordingly, we constructed 100 models by using five combinations of nodes for one hidden layer and five combinations of nodes for two hidden layers. First, we constructed five ANN models with four, five, six, seven, and eight nodes, respectively, with each model having one hidden layer. We also constructed five ANN models with two hidden layers, in which the layers had combinations of three and six nodes, three and seven nodes, four and four nodes, four and five nodes, and five and three nodes, respectively.

We used each of the five optimizers—Adagrad (adaptive gradient), Adam (adaptive moment estimation), Adamax (Adam max), Nadam (Nesterov momentum to Adam), and RMSprop (root mean square propagation)—to optimize the weights and biases of the ANN model:

Adagrad, based on Equation (8), optimizes the model less when the adjustment amount for the weights and biases is large, whereas it optimizes the model more when the adjustment amount is small [27].

$$\theta_t = \theta_{t-1} - \frac{\eta \nabla_\theta J(\theta)}{\sqrt{\sum_{i=1}^{t}(\nabla_\theta J(\theta))^2}} \quad (8)$$

Adam, based on Equations (9) and (10), optimizes the model by using an extremely small operational quantity that simplifies the gradient, which determines the adjusted amounts of weights and biases [28].

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla_\theta J(\theta) \quad (9)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)(\nabla_\theta J(\theta))^2 \quad (10)$$

Adamax, based on Equations (9) and (11), has a similar function to that of Adam while restricting the maximum adjustment amount [28].

$$v_t = \beta_2^p v_{t-1} + \left(1 - \beta_2^p\right)(\nabla_\theta J(\theta))^p = \max(\beta_2 v_{t-1}, |\nabla_\theta J(\theta)|) \quad (11)$$

Nadam, based on Equation (12), is also similar to Adam, but it expands the range of previous errors involved in the determination of the adjustment amount [29].

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon}(\beta_1 \hat{m}_t + \frac{(1 - \beta_1)\nabla_\theta J(\theta)}{1 - \beta_1^t}) \quad (12)$$

RMSprop, based on Equations (13) and (14), is similar to Adagrad, but it reflects more recent learning information than past information [30]:

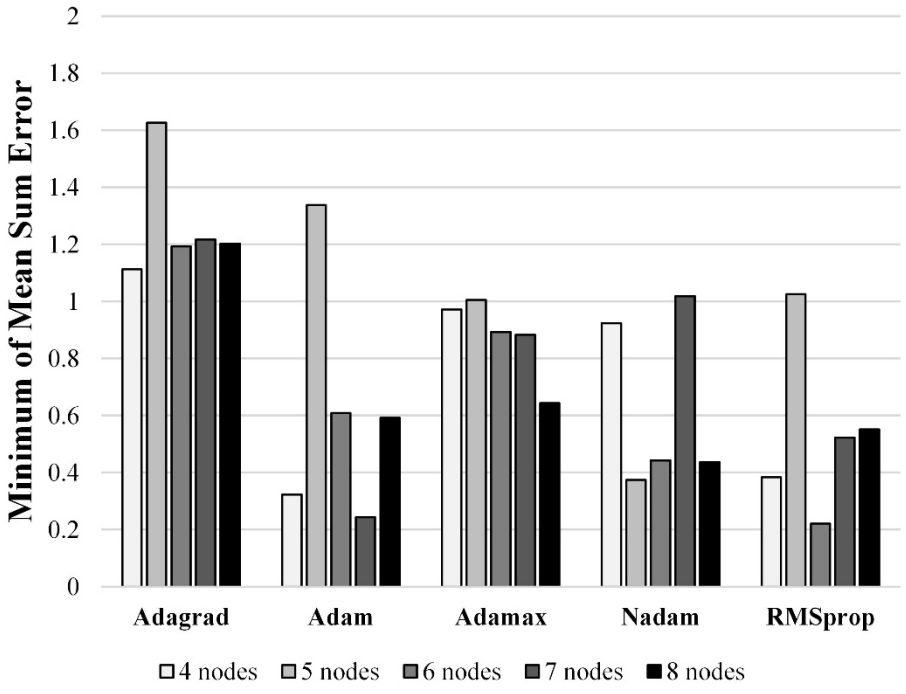
$$G_t = \gamma G_{t-1} + (1-\gamma)(\nabla_\theta J(\theta_t))^2 \quad (13)$$

$$\theta_t = \theta_{t-1} - \frac{\eta}{\sqrt{G_t + \varepsilon}} \nabla_\theta J(\theta_t) \quad (14)$$

In Equations (8)–(14), θ_t is the parameter vector, η is the step size, m_t is the first-moment vector, ∇_θ is the gradient, $J(\theta)$ is the loss function, v_t is the second-moment vector, G_t is the exponential mean of the square of the loss function gradient, and $\beta_1, \beta_2, \gamma,\ p,$ and ε serve as float values.

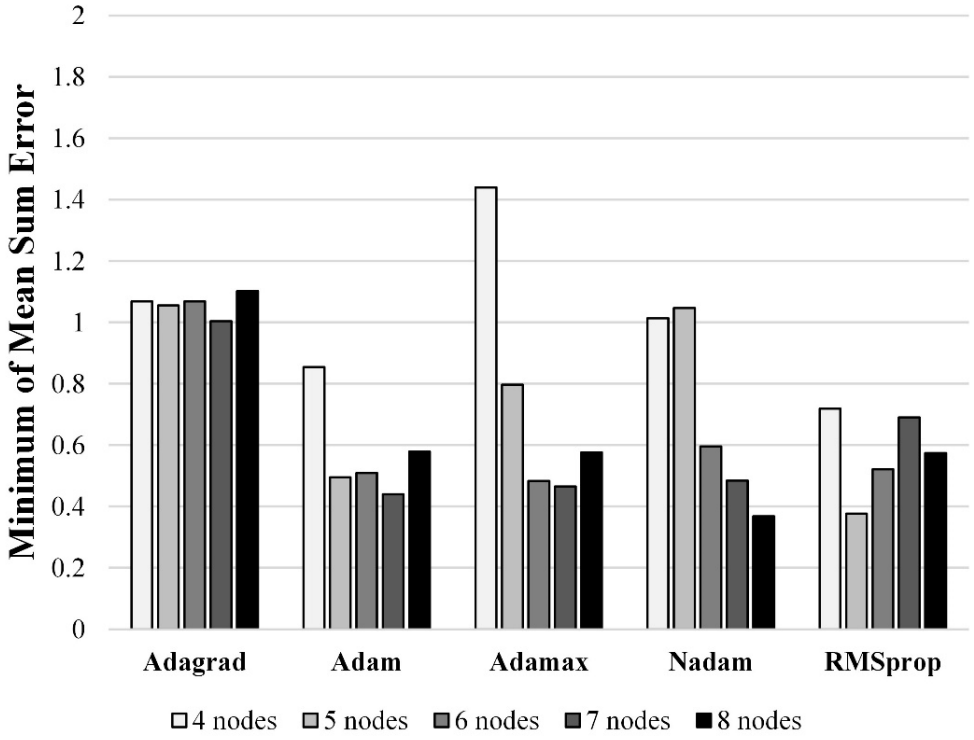
2.3. Selection and Evaluation of Candidate ANN Models

To train the models, we optimized the weights and biases a maximum of 2000 times for each model, and the models predicted the FT as the output. Figure 10 shows the minimum errors recorded between the FT training (dataset) predictions and the measured FTs as a function of the composition of the hidden layers, type of data (dataset), and type of optimizer. Figure 10a shows the errors for dataset 1 and one hidden layer. The errors for the Adam, Nadam, and RMSprop optimizers were the lowest; in particular, the Rp160 model was the best, with an error of 0.22. Figure 10b shows the errors for dataset 2 and one hidden layer. The errors for the Adam, Nadam, and RMSprop optimizers were the lowest. In particular, the Nd280 model showed superior performance with an error of 0.37. The errors for dataset 1 and two hidden layers are shown in Figure 10c. The errors for the Nadam and RMSprop optimizers were the lowest; specifically, the Rp136 model showed superior performance with the lowest error of 0.13, compared to that of all the models. Furthermore, the error for the Nd144 model was 0.18, which was the third-lowest overall error. Figure 10d shows the errors for dataset 2 and two hidden layers. The errors for the Adamax optimizer were the lowest; specifically, the Am253 model was the best, with an error of 0.15, which was the second lowest, overall.

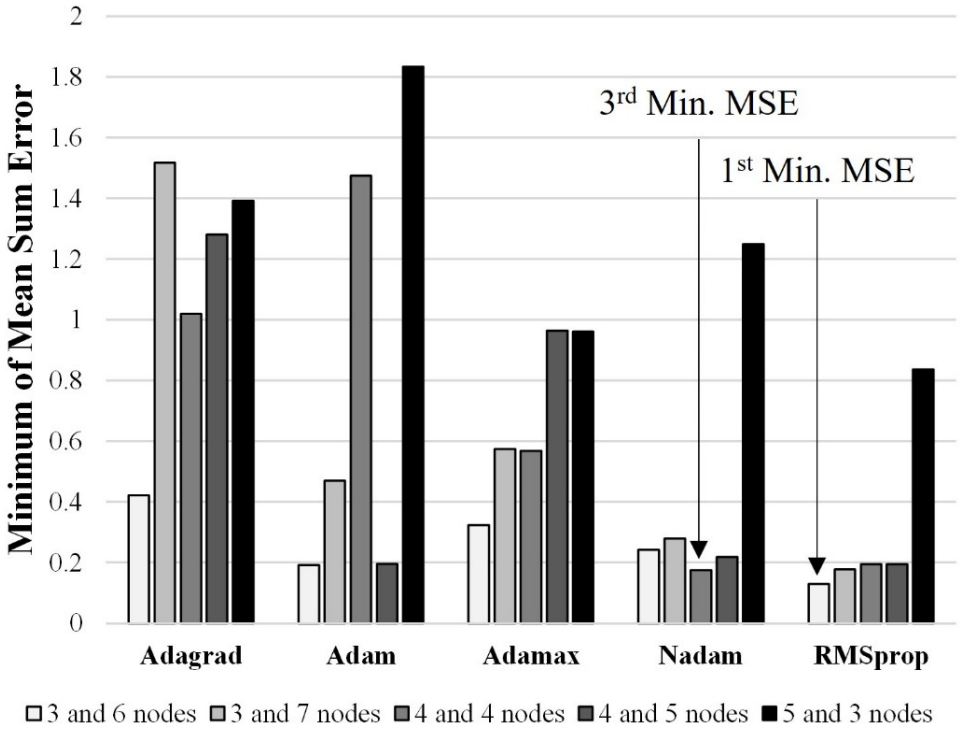


(**a**)

Figure 10. *Cont.*

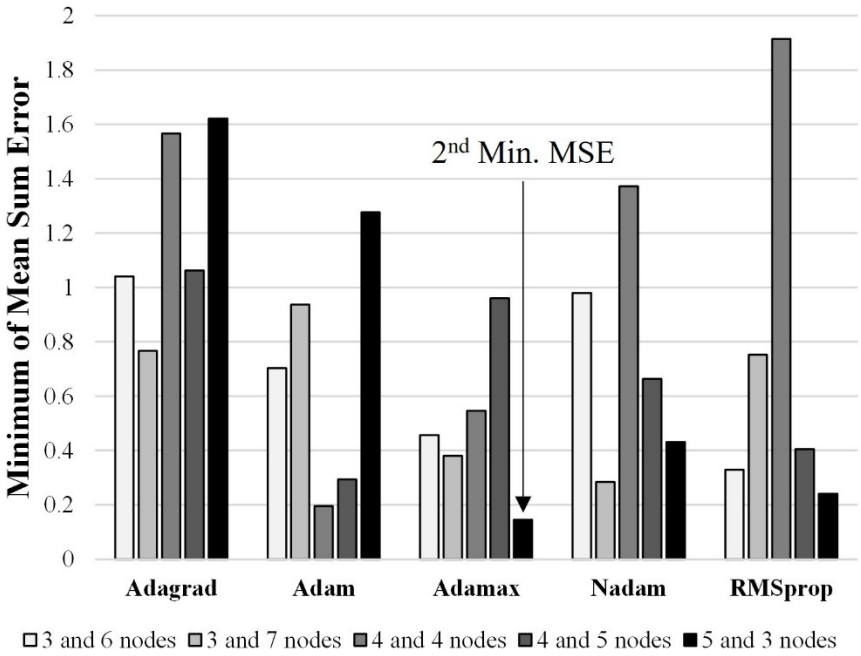


(**b**)



(**c**)

Figure 10. *Cont*.



(**d**)

Figure 10. Minimum MSE calculated by trained model according to ANN structures and optimizers. (**a**) Dataset 1 and one hidden layer, (**b**) Dataset 2 and one hidden layer, (**c**) Dataset 1 and two hidden layers, (**d**) Dataset 2 and two hidden layers.

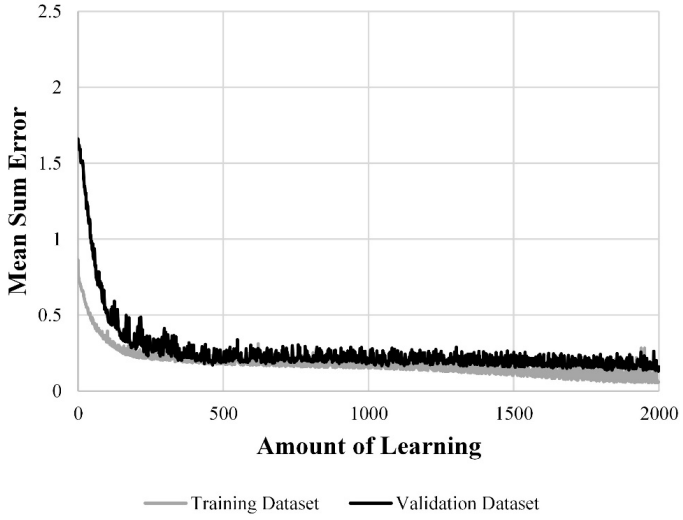
The models with one hidden layer showed higher errors than those with two hidden layers because they had fewer hyperparameters, resulting in poorer prediction performance. Therefore, we selected the Rp136, Am253, and Nd144 models—all of which contained two hidden layers—as the candidate models with the lowest errors.

We compared the errors recorded using the training dataset with those recorded using the validation dataset as a function of the amount of learning, as shown in Figure 11, to investigate the learning generality of the candidate models. We excluded the validation dataset, which occupied 20% of the training dataset, from feed-forward processing to examine the overfitting of the candidate models. The errors of all candidate models decreased with the amount of learning for both the training and validation datasets, reducing the possibility of overfitting. Figure 11a shows that for the Rp136 model, both datasets have unstable errors with large variations, even with increased amounts of learning. The Am253 model shows stable errors for the validation dataset in Figure 11b because the errors from the training dataset are sufficiently stable, regardless of the learning amount. The Nd144 model in Figure 11c shows errors with irregular variation for the validation dataset because the errors of the model for the training dataset are slightly unstable regardless of the learning amount.

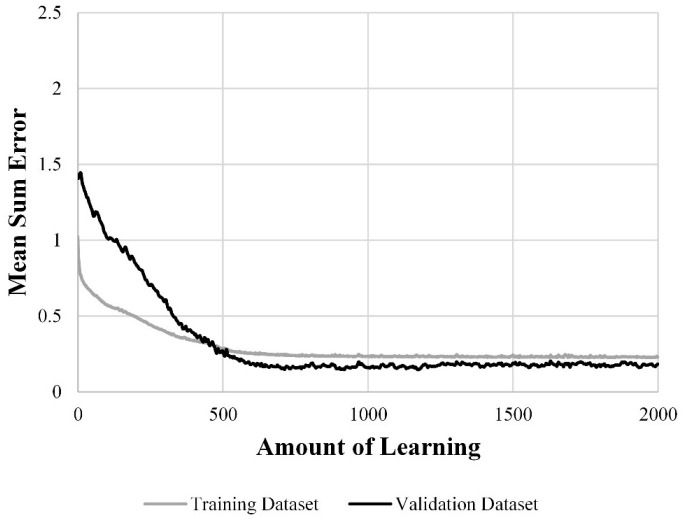
We compared the error rates from the training dataset with those from the test dataset. We calculated the error rates by dividing the error between the predicted and measured FTs by the measured FT, as expressed in Equation (15). The average of the entire error rate calculated for each FT is used to represent the error rate of each candidate model.

$$E = \frac{\sum \frac{|\hat{x}_i - x_i|}{x_i}}{n} \times 100 \quad (15)$$

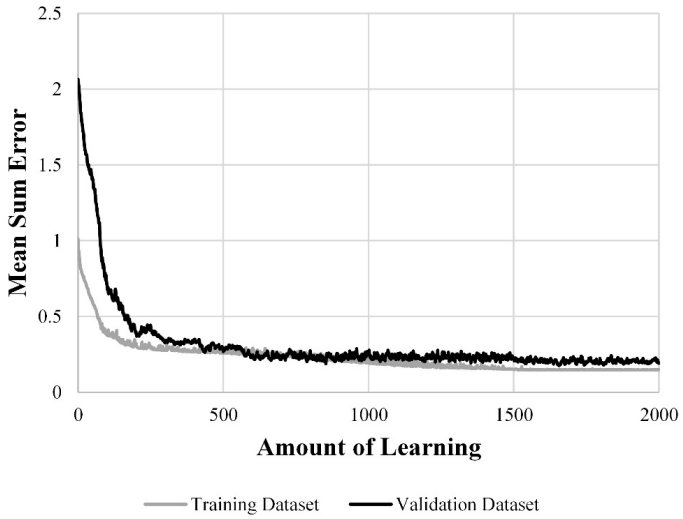
In Equation (15), E is the error rate (%), $\hat{x}_i$ is the prediction value, x_i is the actual value, and n represents the number of data elements.



(**a**)



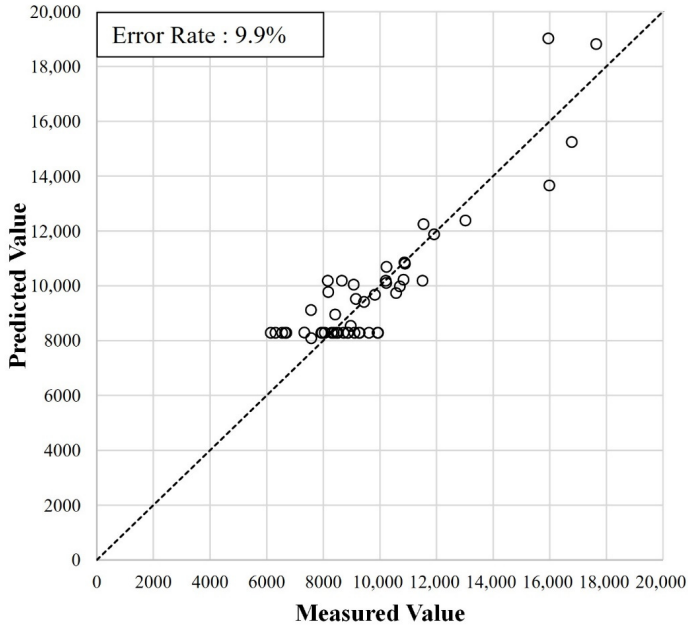
(**b**)



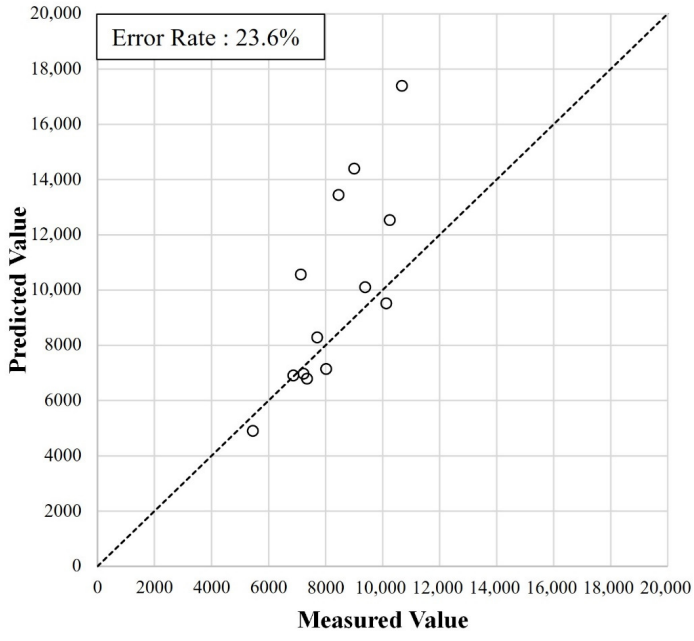
(**c**)

Figure 11. MSE according to the amount of learning for training and validation dataset. (**a**) Rp136 model, (**b**) Am253 model, (**c**) Nd144 model.

Figures 12–14 show the FT predictions using both the training and test datasets, as well as the measured FTs, and error rate of the Rp136, Am253, and Nd144 models, respectively. The error rate for the Rp136 model for the training dataset was low (9.9%), whereas the error rate for the test dataset was relatively high (23.6%). This significant difference was most likely caused by the unstable errors observed for both the training and validation datasets, regardless of the learning amount, as shown in Figure 11a. Similar error rates for the Am253 model for the training and test datasets occurred at 12.2% and 12.4%, respectively, because of the error stability for both datasets, as shown in Figure 11b. For the Nd144 model, the error rate of the test dataset was slightly higher (18.0%) than that of the training dataset (11.4%). Figure 11c shows the slightly increased error rate of the test dataset because of the slightly unstable errors observed for both datasets.

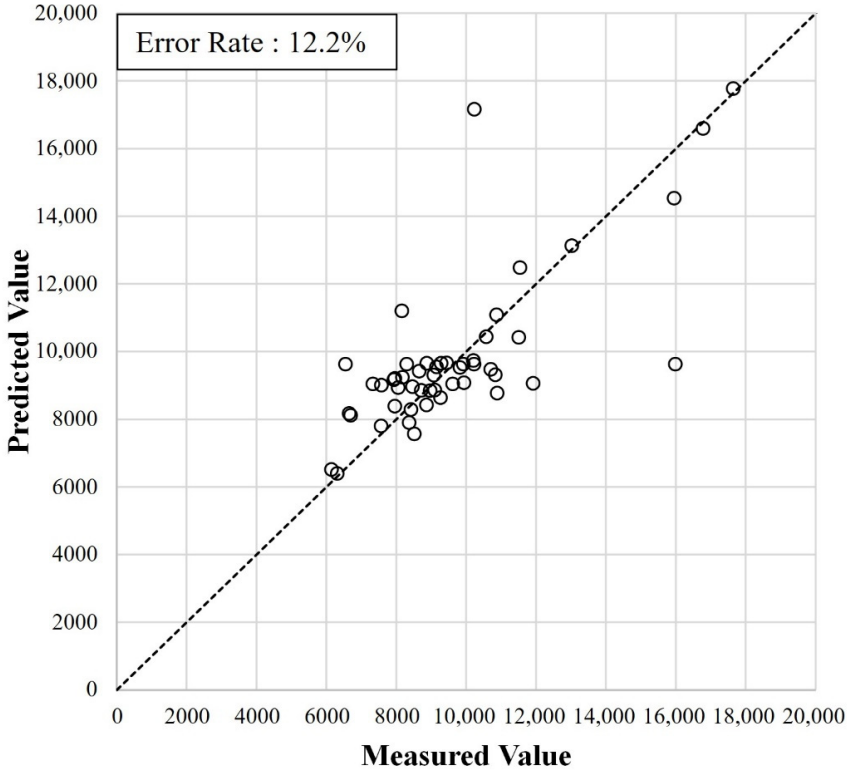


(**a**)

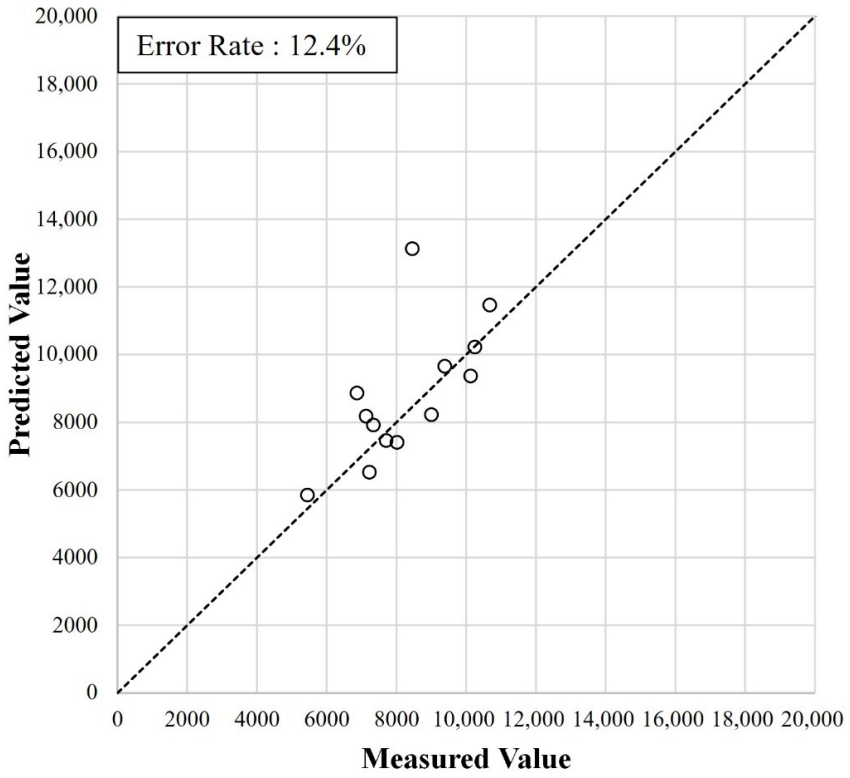


(**b**)

Figure 12. Scatter plot of measured and predicted FT for Rp136 model. (**a**) Training dataset, (**b**) test dataset.

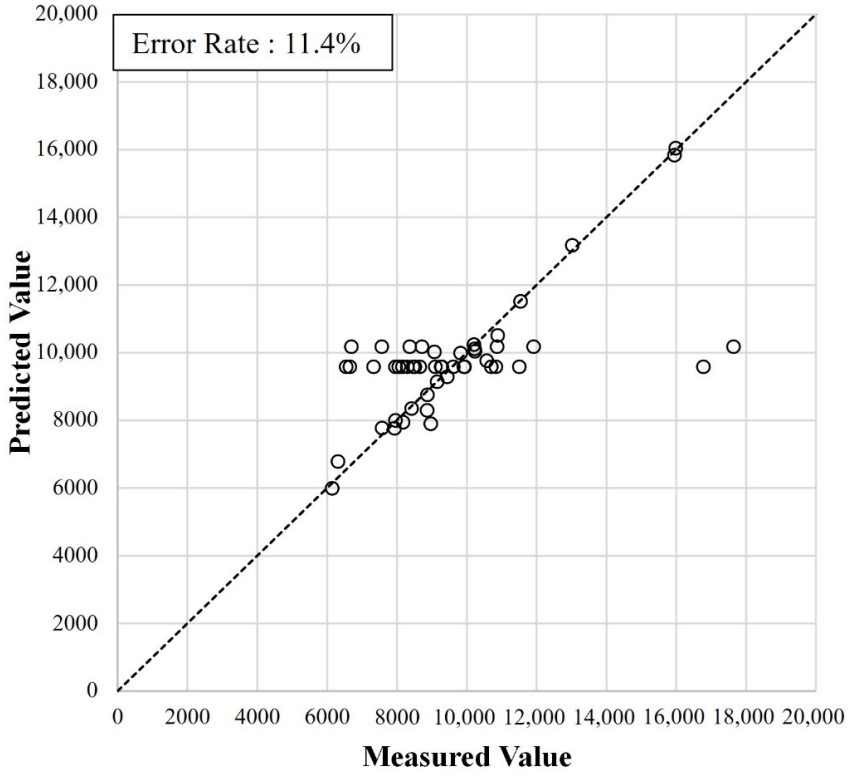


(**a**)

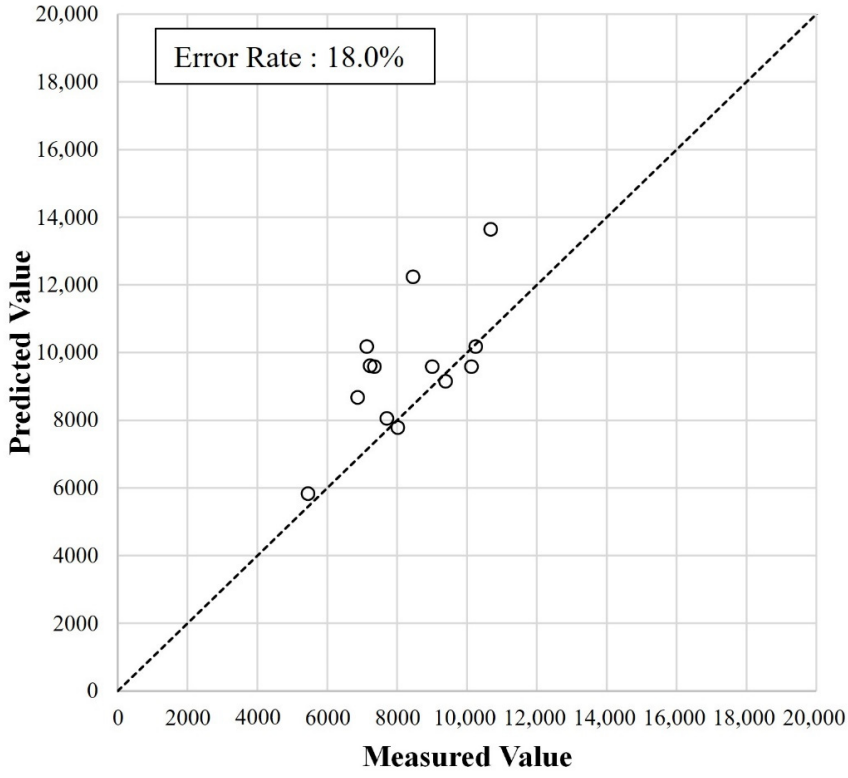


(**b**)

Figure 13. Scatter plot of measured and predicted FT for Am253 model. (**a**) Training dataset, (**b**) test dataset.



(**a**)

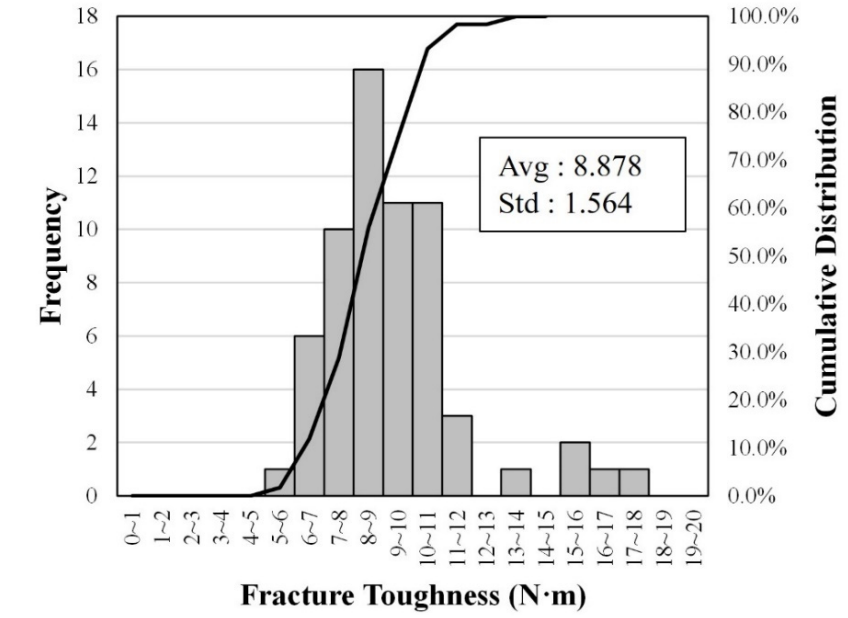


(**b**)

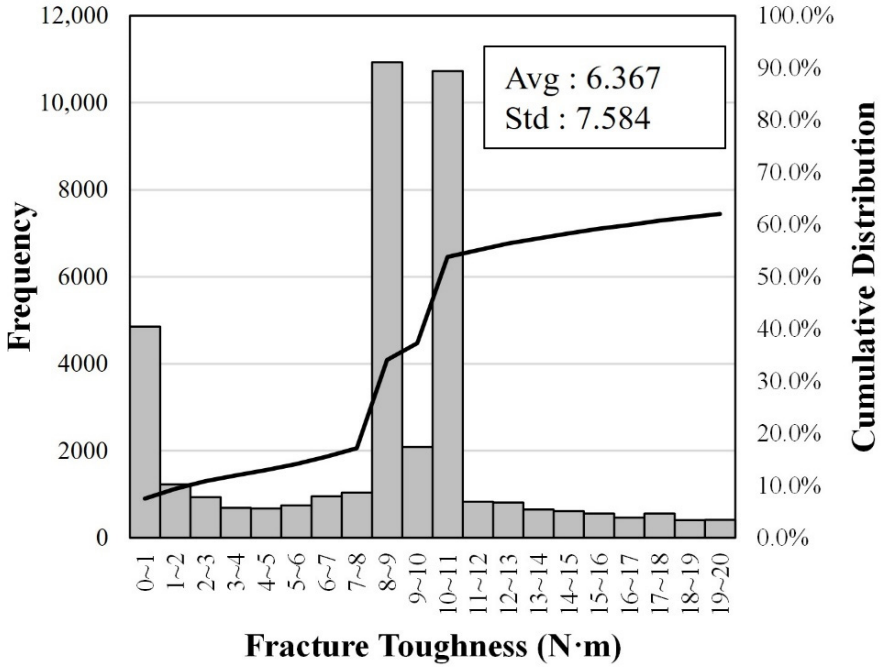
Figure 14. Scatter plot of measured and predicted FT for Nd144 model. (**a**) Training dataset, (**b**) test dataset.

3. Results and Discussion

We compared the distributions of the FT predictions of the intermediate layer of the entire expressway asphalt pavement sections by the candidate models trained with HPMS data collected in 2019 with those of 66 samples cored in the field. Figure 15a shows the frequency and cumulative probability distributions of the FTs measured for the 66 core samples. The average and standard deviation of the measured FTs—after eliminating outliers—were 8.878 and 1.564 N·m, respectively.

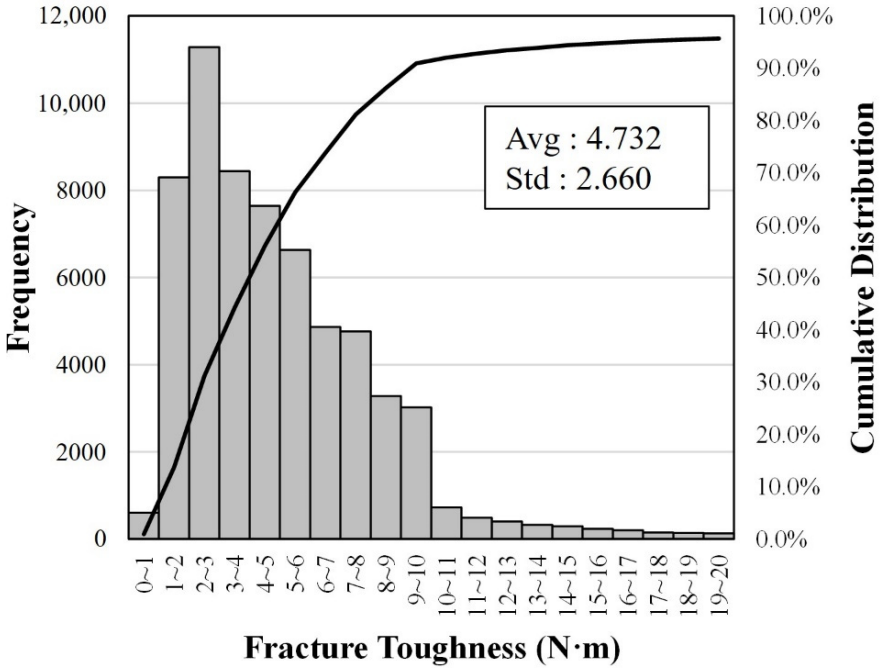


(**a**)

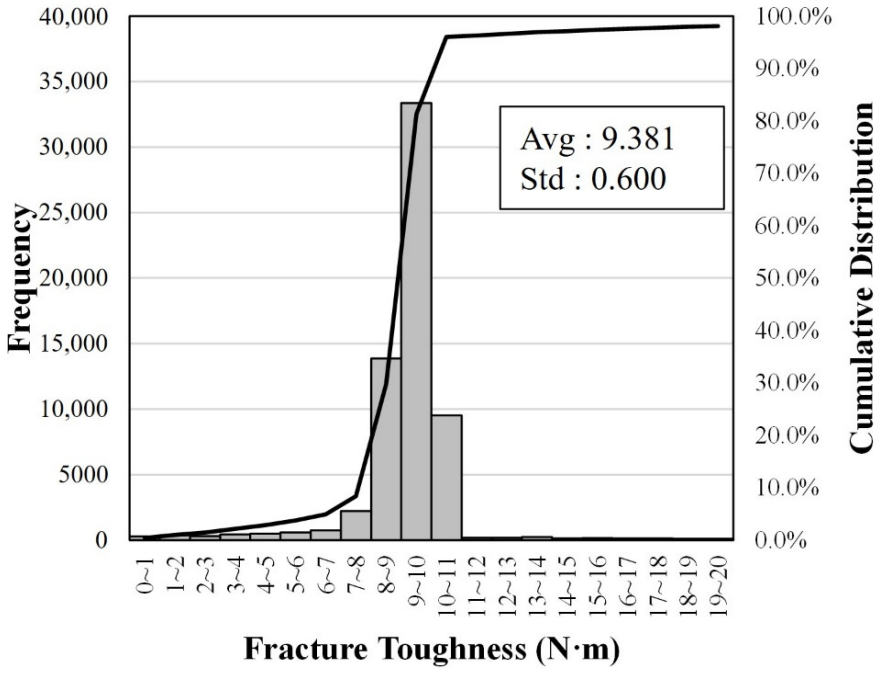


(**b**)

Figure 15. *Cont*.



(**c**)



(**d**)

Figure 15. Frequency distribution of actual and predicted FT. (**a**) Measured data, (**b**) prediction of Rp136 model, (**c**) prediction of Am253 model, (**d**) prediction for Nd144 model.

Figure 15b shows the frequency and cumulative probability distributions of the FT predictions that use the Rp136 model. The average and standard deviation of the FT predictions after eliminating outliers were 6.367 N·m and 7.584 N·m, respectively. The standard deviation of the FT predictions using the model was much larger than that of the measured FTs because of the low stability of the output values, as shown earlier in Figures 11a and 12. Therefore, Rp136 was not selected as the final model.

Figure 15c shows the frequency and cumulative probability distributions of FT predictions that use the Am253 model. The average and standard deviation of the FT predictions—after eliminating outliers—were 4.732 and 2.660 N·m, respectively. The standard deviation of the FT predictions of the model was slightly larger than that of the measured FTs. Furthermore, the average, frequency distribution, and cumulative probability distribution of the FT predictions are completely different from those of the measured FTs. Therefore, we did not select the Am253 model as the final model, although the error for the training and validation datasets and the error rate for the training and test datasets were superior to those of the other candidate models.

Figure 15d shows the frequency and cumulative probability distributions of FT predictions that use the Nd144 model. The average and standard deviation of the FT predictions—after eliminating outliers—were 9.381 and 0.600 N·m, respectively. The standard deviation of the FT predictions of the model was less than half that of the measured FTs because of the stable variation in errors and relatively small error rates, as shown earlier in Figures 11c and 14. Moreover, the average, frequency distribution, and cumulative probability distribution of the FT predictions are similar to those of the measured FTs. Accordingly, we selected the Nd144 model as the final model among the three candidate models.

The Nd144 model contained two hidden layers, with four nodes in each layer, and we used Nadam as its optimizer. We applied the cube root to the variables FT and SD to ensure their normality. Because we calculated FT predictions using a mathematical method, they can be predicted beyond the normal range of FT. The condition of the pavement may be evaluated as excessively good or poor because the FT exceeds the normal range. Therefore, we complemented the model by setting the lower and upper limits to 7.689 N·m and 10.804 N·m, respectively, and replacing the outliers outside this range with these limits, as shown in Equation (16). The upper and lower limits were defined as FT values of the cumulative distribution of 95% and 5%, respectively, as shown in Figure 15a. The range limits set in this study were reasonable because 8.0 N·m was designated as the criterion for the FT of the intermediate layer of asphalt pavement by KASCON [31].

$$FT_c = \begin{cases} 7.689\ (FT < 7.689) \\ FT(7.689 \leq FT \leq 10.804) \\ 10.804\ (FT > 10.804) \end{cases} \quad (16)$$

Here, *FTc* represents corrected fracture toughness (N·m) and *FT* represents fracture toughness (N·m).

4. Conclusions

We developed a total of 100 FT models that use ANN and selected the best model by examining the prediction performance, error stability, and prediction distribution. The primary conclusions of this study are as follows:

We used the IRI, RD, SD, and ESAL as the independent variables of the FT prediction model because they had a mechanistic relationship with the FT; we measured them every year during HPMS surveys. The distributions of FT and SD did not satisfy normality; therefore, they were normalized. The mean and standard deviation of all variables were standardized.

We developed a total of 100 ANN models with different normalization methods, number of hidden layers, number of nodes, and optimizer methods. We examined the errors between the measured and predicted FTs for all ANN models. Among the models, Rp136, Am253, and Nd144 had the lowest error rates; we selected them as the candidate models.

We compared the errors between the FT training predictions and measured FTs with the errors between the FT validation predictions and measured FTs as a function of the amount of learning to examine the generality of the learning of the candidate models. By using both the training and validation datasets, the errors of all candidate models decreased with increasing learning amounts, thereby reducing the possibility of overfitting. When comparing the stability of the errors, the Rp136 model was the most unstable, the Nd144

model was moderately stable, and the Am253 model was the most stable. If the error is stable, it is expected to show a low error even for new input data.

We compared the error rates between the FT training predictions and measured FTs with those between the FT test predictions and measured FTs to examine the applicability of the candidate models to new input data. The Rp136 model exhibited a high error rate in the test dataset because the error was unstable. The Nd144 model exhibited a slightly higher error rate on the test dataset than on the training dataset. Because the Am253 model had stable errors, the error rates of the test and training datasets were similar.

When the FT predictions from the Nd144 model were compared to those of the measured FTs at 66 randomly selected locations on expressway asphalt pavement, they exhibited the most comparable frequency distribution, cumulative probability distribution, and average. Because of the stable variation in the error and the relatively low error rate, we selected this model as the final ANN model. The model was complemented by setting the lower and upper limits to 7.689 N·m and 10.804 N·m, respectively, and by replacing the outliers with these range limits. The proposed model contributes to the sustainability of the pavement structure by simply predicting asphalt pavement conditions.

Author Contributions: Conceptualization, J.-H.J. and D.-H.K.; methodology, D.-H.K.; software, H.-Y.K.; validation, K.-H.M.; formal analysis, D.-H.K.; investigation, H.-Y.K.; resources, K.-H.M.; data curation, H.-Y.K.; writing—original draft preparation, D.-H.K.; writing—review and editing, J.-H.J.; visualization, D.-H.K.; supervision, J.-H.J.; project administration, K.-H.M.; funding acquisition, K.-H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Korea Expressway Corporation and the APC was funded by Inha University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: Korea Expressway Corporation (KEC) and Inha University financially sponsored this study.

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

References

1. KEC. *Highway Pavement Maintenance System Operation Manual*; Korea Expressway Corporation: Gimcheon-si, Korea, 1996.
2. Lee, G.H.; Gang, M.S.; Jo, M.J. Damage status and reduction method of asphalt pavement. *Korean Soc. Road Eng.* **2012**, *14*, 5–11.
3. Harold, L.; Von, Q.; Thomas, W.K. Mixture properties related to pavement performance. *Proc. Assoc. Asph. Asph. Paving Technol.* **1989**, *58*, 553–570.
4. Zafrul, K.; Hasan, M.F. Fracture toughness measurement of asphalt concrete by nanoindentation. In Proceedings of the 2017 International Mechanical Congress and Exposition, Tampa, FL, USA, 3–9 November 2017.
5. Kruzic, J.J.; Kim, D.K.; Koester, K.J.; Ritchie, R.O. Indentation techniques for evaluating the fracture toughness of biomaterials and hard tissues. *J. Mech. Behav. Biomed. Mater.* **2009**, *2*, 384–395. [CrossRef] [PubMed]
6. Kim, D.H.; Lee, S.J.; Moon, K.H.; Jeong, J.H. Prediction of indirect tensile strength of intermediate layer of asphalt pavement using artificial neural network model. *Arab. J. Sci. Eng.* **2021**, *46*, 4911–4922. [CrossRef]
7. Choi, J.H.; Adams, T.M.; Bahia, H.U. Pavement roughness modeling using back-propagation neural networks. *Comput.-Aided Civ. Infrastruct. Eng.* **2004**, *19*, 295–303. [CrossRef]
8. Gandhi, T.; Xiao, F.; Amirkhanian, S.N. Estimating indirect tensile strength of mixtures containing anti-stripping agents using an artificial neural network approach. *Int. J. Pavement Res. Technol.* **2009**, *2*, 1–12.
9. Josipa, D.; Hrvoje, D.; Tatjana, R.; Sanja, D. Application of an artificial neural network in pavement management system. *Teh. Vjesn.* **2018**, *25*, 466–473. [CrossRef]
10. Graves, A.; Liwicki, M.; Fernandez, S.; Bertolami, R.; Bunke, H.; Schmidhuber, J.; Nobel, A. Connectionist system for improved unconstrained handwriting recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 855–868. [CrossRef]
11. Achanta, A.S.; Kowalski, J.G.; Rhodes, C.T. Artificial neural networks: Implications for pharmaceutical sciences. *Drug Dev. Ind. Pharm.* **1995**, *21*, 119–155. [CrossRef]
12. Kim, S.; Gopalakrishnan, K.; Ceylan, H. Neural networks application in pavement infrastructure materials. *Intell. Soft Comput. Infrastruct. Syst. Eng.* **2009**, *259*, 47–66. [CrossRef]
13. Lushinga, N.; Cao, L.; Dong, Z. Effect of silicone oil on dispersion and low-temperature fracture performance of crumb rubber asphalt. *Adv. Mater. Sci. Eng.* **2019**, *2019*, 8602562. [CrossRef]

14. McDaniel, R.; Shah, A. *Asphalt Additives to Control Rutting and Cracking*; Federal Highway Administration: IN/JTRP-2002/29; Federal Highway Administration: Washington, DC, USA, 2002.
15. Melesse, A.M.; Ahmad, S.; McClain, M.E.; Wang, X.; Lim, Y.H. Suspended sediment load prediction of river systems: An artificial neural network approach. *Agric. Water Manag.* **2011**, *98*, 855–866. [CrossRef]
16. Abramowitz, M.; Stegun, I.A. *Hanbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed.; Dover: New York, NY, USA, 1972; p. 928. ISBN 0-486-61272-4.
17. Kenney, J.F.; Keeping, E.S. *Mathematics of Statistics*, 3rd ed.; D. Van Nostrand Company: New York, NY, USA, 1962; pp. 102–103. ISBN 978-1114612600.
18. Kline, R.B. *Principles and Practice of Structural Equation Modeling*, 2nd ed.; Guilford Press: New York, NY, USA, 2005; ISBN 978-1606238769.
19. Wilson, E.B.; Hilferty, M.M. The distribution of chi-square. *Proc. Natl. Acad. Sci. USA* **1931**, *17*, 684–688. [CrossRef] [PubMed]
20. Maclean, C.J.; Morton, N.E.; Elston, R.C.; Yee, S. Skewness in commingled distributions. *Biometrics* **1976**, *32*, 695–699. [CrossRef] [PubMed]
21. William, H.K.; Judith, M.T. *International Encyclopedia of Statistics*, 1st ed.; Free Press: New York, NY, USA, 1978; pp. 523–541, ISBN 002917970X.
22. McCulloch, W.; Walter, P. A logical calculus of ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [CrossRef]
23. Lagaris, I.E.; Likas, A.; Fotiadis, D.I. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans. Neural Netw.* **1998**, *9*, 987–1000. [CrossRef] [PubMed]
24. Ramachandran, P.; Barret, Z.; Quoc, V.L. Searching for activation functions. *Comput. Sci. Neural Evol. Comput.* **2017**. Available online: https://openreview.net/forum?id=SkBYYyZRZ (accessed on 23 May 2022).
25. Claesen, M.; Bart, D.M. Hyperparameter search in machine learning. In Proceedings of the XI Metaheuristics International Conference, Agadir, Morocco, 7–10 June 2015.
26. Probst, P.; Boulesteix, A.L.; Bischl, B. Tunability: Importance of hyperparameters of machine learning algorithms. *J. Mach. Learn. Res.* **2018**, *20*, 1–32.
27. Lydia, A.A.; Francis, F.S. Adagrad—An optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci.* **2019**, *6*, 566–568.
28. Diederik, P.K.; Jimmy, L.B. Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Represent, 2015 Workshop Track, San Diego, CA, USA, 7–9 May 2015.
29. Dozat, T. Incorporating Nesterov momentum into Adam. In Proceedings of the International Conference on Learning Represent, 2016 Workshop Track, San Juan, Puerto Rico, 2–4 May 2016.
30. Ruder, S. An overview of gradient descent optimization algorithms. *Comput. Sci. Mach. Learn.* **2016**. [CrossRef]
31. KASCON. *Hot Mix Asphalt*; SPS-KAI0002-F2349-5687; Korea Asphalt Concrete Industrial Cooperative Association: Seoul, Korea, 2018.