Article

Evaluation and Prediction of Pavement Deflection Parameters Based on Machine Learning Methods

Xueqin Chen 1, Qiao Dong 2,3 and Shi Dong 4,5,*

1 Department of Civil Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
2 Department of Roadway Engineering, School of Transportation, Southeast University, Nanjing 211189, China
3 National Demonstration Center for Experimental Road and Traffic Engineering Education, Southeast University, Xi’an 710064, China
4 Engineering Research Center of Highway Infrastructure Digitalization, Ministry of Education of PRC, Chang’an University, Xi’an 710064, China
5 College of Transportation Engineering, Chang’an University, Xi’an 710064, China

* Correspondence: dongshi@chd.edu.cn

Abstract: The deflection measurements made using Falling Weight Deflectometers (FWDs) are widely used in the back-calculation of pavement layer moduli. Pavement structural characteristics, changes in temperature, and other related factors exert a significant effect on the deflection measurements. Therefore, three machine learning methods—Classification and Regression Tree (CART), Random Forest (RF), and Gradient Boosting Decision Tree (GBDT)—were used to evaluate the importance of influencing factors including FWD test conditions, pavement structural parameters, climatic factors, traffic level, rehabilitation level, and service age, on the FWD measurements of deflection basin in this study. The results indicated that structural number was an important feature for all FWD measurements but its importance on \( \lg(D_0 - D_{20}) \) and \( \lg(D_0 - D_{30}) \) was smaller than other FWD measurements. The relative feature importance of the asphalt layer, base, and subbase on \( \lg(D_0 - D_{20}) \) and \( \lg(D_0 - D_{30}) \) was asphalt layer > subbase > base; their relative importance on \( \lg(D_{20} - D_{60}) \), \( \lg(D_{30} - D_{60}) \), and \( \lg(D_{30} - D_{90}) \) was asphalt layer > base > subbase; and their relative importance on \( \lg(D_{90} - D_{120}) \) and \( \lg(D_{60} - D_{120}) \) was base > subbase > asphalt layer. Among the FWD test condition variables, drop load was the most significant factor influencing deflection measurements. The second-layer temperature was also important for \( \lg(D_0 - D_{20}) \), \( \lg(D_0 - D_{30}) \), and \( \lg(D_0 - D_{45}) \). The importance of precipitation was greater than the freeze index. The prediction results shown that the accuracy of GBDT was as high as 99%. Besides, GBDT outperformed RF, and RF outperformed CART. The analyses between FWD deflection parameters and influencing factors, especially the structural characteristics of the pavement, provide theoretical evidence for the evaluation of pavement layer strength on the basis of FWD data.

Keywords: falling weight deflectometer; feature importance; GBDT; CART; Random Forest; LTPP

1. Introduction

1.1. Background

The falling weight deflectometer (FWD) is a non-destructive testing device used to evaluate the physical properties of pavements. During the FWD test, the pavement surface is subjected to a load pulse produced by dropping a large weight onto a buffered load plate, as shown in Figure 1. The load level, load duration, and load area are adjusted to simulate the actual loading caused by a rolling vehicle wheel traveling on an in-service pavement. A series of deflection sensors are used to measure the pavement surface deformation in response to the load pulse. The sensors are installed at various distances from the load plate. Typically, the offsets of sensors are 0, 200, 300, 450, 600, 900, 1200, and 1500 mm from the load plate center, respectively. The acquired deflection data can be used to evaluate the compaction quality and uniformity of the subgrade, and estimate the...
pavement’s structural capacity. Currently, the inverse analysis or back-calculation process is widely used to calculate the pavement layer moduli and stiffnesses based on FWD measurements [1,2]. Mechanical models [3], mechanistic-empirical methods [4], computer programs such as EVERCALC, ELMOD, MODCOMP, MODULUS, MICHBACK, etc. [5], among which EVERCALC and MODCOMP can be primarily used to back-calculate the LTPP pavement sections, and some novel approaches like Artificial Neural Network [6,7], and Recurrent Neural Network and Wide & Deep structure [8] can be used to back-calculate pavement layer moduli. The results have been useful in evaluating the structural condition of in-service pavements, predicting the pavement performance, and determining treatment strategies.

Figure 1. Diagram of FWD test [9].

1.2. Influencing Factors of Deflections

Since deflections have been extensively used in the assessment of pavement structural conditions, the data quality of FWD measurements may influence the back-calculation accuracy of pavement structural parameters. Many studies have shown the factors influencing deflection measurements. Mehta and Roque [10] indicated that the pavement structural characteristics, including milling operations, damage layers, variation in layer thickness, and temperature, had a significant effect on the deflection data. Nobakht et al. [11] developed a rehabilitation selection methodology based on a damage ratio calculated from FWD data and two rehabilitation categories including different overlay thicknesses were considered as alternatives for pavement sections with damage ratios between 0.3 and 0.7. Besides, the crack type and crack width affected the deflection basin. Ma et al. [12] found that pavement cracks would make the deflection basin steeper. The drop load of FWD also affected the deflection measurements [1,13]. The effects of moisture content of subgrade soil and pavement temperature on the pavement structural response, and the effects of pavement surface temperature, air temperature, and precipitation on FWD measurements were analyzed and quantified [14,15]. The results showed that temperature had the most significant effect on the measured deflections. Wang et al. [16] analyzed the influence of material property, layer thickness, loading magnitude, and pavement temperature on the pavement surface deflections through finite element simulations.

According to China’s code “Specifications for Design of Highway Asphalt Pavement” [17], deflection is no longer a pavement design index, but it is still an important construction acceptance index. In the previous version of Specifications for Design of Highway Asphalt Pavement (JTG D50-2006), design deflection was determined on the basis of highway classification, cumulative Equivalent Single Axle Load (ESAL), layer type, and base type. The temperature correction of FWD deflection value for the construction acceptance of asphalt pavement depends on whether the asphalt layer thickness is greater than 5 cm or not. The temperature correction coefficient is related with the pavement surface temperature and asphalt layer thickness. The calculated deflection value is also associated with the temperature correction coefficient, seasonal influence coefficient, and
moisture influence coefficient. Therefore, the influence of the above factors on the deflection should be analyzed and quantified.

1.3. Machine Learnings

With the rapid development and wide use of data science, an increasing number of machine learning methods are being used in the analysis of pavement engineering, including pavement performance evaluation, performance prediction, distress recognition, automatic driving, etc. Machine learning has exhibited powerful and excellent prediction capability. For the use of deflection data, ANNs can be established to predict the pavement layer moduli, interlayer condition, tensile strains at the bottom of asphalt layer, compressive strains on the top of subgrade, deviator stresses on the top of subgrade, Poisson’s ratio, and layer thickness [7,18,19]. Wang et al. [16] employed the data obtained from finite element simulations to establish and train an ANN model and found that the prediction accuracy was better than that of traditional methods regarding FWD field measurements. Li and Wang [20] used the ANN combined with genetic algorithm (GA) optimization to verify the database of pavement surface deflections and strains of different pavement structures, material properties, loadings, and temperatures. ANN has also been used to predict the two deflection basin parameters based on the input of pavement structural and functional characteristics, environmental factors, and subgrade soil attributes [21]. Han et al. [8] proposed a hybrid neural network structure, combined with Residual Neural Network, Recurrent Neural Network, and Wide & Deep (ResRNN–W&D) structure to analyze the layer moduli, and the ResRNN–W&D structure presented a stronger generalization ability than ANNs. Rabbi and Mishra [22] determined the Deflection Basin Parameters (DBPs), such as base layer index, middle layer index, and lower layer index, using the measured deflection data. Haridas et al. [23] developed a Deep Neural Network (DNN)-based approach to predict the deflection parameters through the data of roughness, traffic, pavement age, temperature, and climatic factors. The prediction accuracy of the DNN approach was about 82% for the test dataset.

More machine learning methods can be used effectively in other aspects of pavement engineering [24]. Dong et al. [25] adopted the classification and regression tree (CART) to quantify the effects of weather, traffic, materials, and construction practice factors on the effectiveness of pavement treatments. Gujar and Vakharia [26] used a Support Vector Machine (SVM) to predict and validate the composition of mineral filler in the microsurfacing mix design. Random Forest can be used to estimate pavement international roughness index (IRI) considering the factors including distress measurements, traffic, climate, structural parameters, and maintenance data [27], and to evaluate the importance of aggregate gradation, mixture volumetric parameters, asphalt binder properties, physical properties of hot-mix asphalt (HMA), age, and pavement thickness on the pavement performance [28]. Zhang et al. [29] proposed a Random Forest to predict the Pavement Condition Index (PCI) using the performance indices, pavement structure, traffic parameters, and meteorological data. The gradient boosted tree can be used to predict the IRI and PCI [30], and determine the key factors affecting asphalt overlay performance including IRI, rutting, fatigue cracking, transverse cracking, and longitudinal cracking [31]. Compared with traditional methods, machine learning methods have been demonstrated to be more effective approaches offering better performance.

1.4. Long-Term Pavement Performance (LTPP) Program

The Long-Term Pavement Performance (LTPP) program was started as a part of the Strategic Highway Research Program (SHRP) in 1987, and has been managed by the Federal Highway Administration (FHWA) since 1992. It aims to study the pavement performance and understand how and why pavements perform as they do. The performance data of more than 2000 pavement sections throughout the United States and Canada are collected and stored in the LTPP database. Over the past thirty years, the LTPP data have been demonstrated to be valuable in advancing pavement engineering technology. LTPP data
can be used for the back-calculation of layer modulus [5,8], prediction of structural number [32], investigation on pavement performance [28], effectiveness evaluation of pavement maintenance [25], and optimization of preventive maintenance strategy [33].

1.5. Objectives and Scope
The objective of the present study is to evaluate and predict the different deflection measurements of asphalt pavements applied rehabilitation strategies through three machine learning methods. Historical asphalt pavement rehabilitation projects were extracted from the LTPP InfoPave™ system. The rehabilitation types included Asphalt Concrete Overlay, Hot-Mix Recycled Asphalt Concrete Overlay, Surface Recycled Asphalt Concrete, Mill Off AC and Overlay with AC, and Mill Existing Pavement and Overlay with Cold-Mix Recycled AC. The importance of features including pavement rehabilitation, traffic level, climate, structural parameters, pavement temperature, as well as service age on deflections was analyzed. The important features of pavement deflection basin parameters were determined and analyzed. The analyses can provide theoretical evidence for the pavement layer strength evaluation through FWD data.

2. LTPP Data
2.1. FWD Measurements
The FWD data were extracted from the table MON_DEFＬ_DROP_DATA. Seven deflection measurements—PEAK_DEFＬ_1, PEAK_DEFＬ_2, PEAK_DEFＬ_3, PEAK_DEFＬ_4, PEAK_DEFＬ_5, PEAK_DEFＬ_6, and PEAK_DEFＬ_7—were recorded in the table. The corresponding sensor locations can be found in column CENTER_OFFSET of table MON_DEFＬ_DEV_SENSORS. As shown in Table 1, the center offsets of the seven deflection measurements were 0, 203, 305, 457, 610, 914, and 1219 mm, respectively. To simplify the variable notations, the seven deflections are denoted as D₀, D₂₀, D₃₀, D₄₅, D₆₀, D₉₀, and D₁₂₀ in this study. The deflection basin is composed of the deflections measured at discrete locations along the pavement cross-section, as shown in Figure 2. The deflection decreased with increasing center offset.

<table>
<thead>
<tr>
<th>Deflection</th>
<th>CENTER_OFFSET (mm)</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEAK_DEFＬ_1</td>
<td>0</td>
<td>D₀</td>
</tr>
<tr>
<td>PEAK_DEFＬ_2</td>
<td>203</td>
<td>D₂₀</td>
</tr>
<tr>
<td>PEAK_DEFＬ_3</td>
<td>305</td>
<td>D₃₀</td>
</tr>
<tr>
<td>PEAK_DEFＬ_4</td>
<td>457</td>
<td>D₄₅</td>
</tr>
<tr>
<td>PEAK_DEFＬ_5</td>
<td>610</td>
<td>D₆₀</td>
</tr>
<tr>
<td>PEAK_DEFＬ_6</td>
<td>914</td>
<td>D₉₀</td>
</tr>
<tr>
<td>PEAK_DEFＬ_7</td>
<td>1219</td>
<td>D₁₂₀</td>
</tr>
</tbody>
</table>

The study of Cong et al. [34] showed that the subgrade modulus can be evaluated by the difference between the seventh and eighth deflections, and the difference between the third and fourth deflections can be used to evaluate the base layer structural condition. The relationships between other deflection parameters and the asphalt layer condition, base condition, and subbase condition were also investigated through the sensitivity analysis. The deflection parameters can be determined from the seven measured deflection data values. Therefore, D₀–D₂₀, D₀–D₃₀, D₀–D₄₅, D₂₀–D₆₀, D₃₀–D₆₀, D₃₀–D₉₀, D₆₀–D₉₀, D₆₀–D₁₂₀, D₉₀–D₁₂₀ were further calculated to analyze the relationships between the deflection basin parameters and influencing factors.
2.2. Influencing Factors

Data on the influencing factors, including rehabilitation level, climate factors, traffic volume, pavement structural parameters, FWD test conditions, and service age, were also collected. The rehabilitation types of Asphalt Concrete Overlay, Hot-Mix Recycled Asphalt Concrete Overlay, Surface Recycled Asphalt Concrete, Mill Off AC and Overlay with AC, and Mill Existing Pavement and Overlay with Cold-Mix Recycled AC, with IMP_TYPE = 19, 43, 45, 51, and 56, respectively, were extracted from the LTPP database. All the rehabilitation projects were in states/provinces in North America, as shown in Figure 3.

![Figure 2. Deflection basin along the pavement cross-section.](image1)

![Figure 3. Distribution map of pavement rehabilitation projects.](image2)

The variable IMP_THICKNESS is a variable that can represent the rehabilitation level, and refers to the increase in pavement thickness due to the rehabilitation activity. The climate factors include precipitation AVG_ANN_PRECIP and freeze index AVG_FREEZE_INDEX. The traffic level was characterized by the variable ANNUAL_kESAL. The structural parameters of the pavement include the structural number SN_VALUE, the thickness of the asphalt layer LAYER_THICKNESS, the type of base BASE_TYPE (granular base and treated base), the thickness of the base BASE_THICKNESS, and the thickness of the subbase SUBBASE_THICKNESS. It was noted that the variable BASE_TYPE was a categorical variable, including granular base (GB) and treated base (TB). The data for the FWD test conditions included the variables DROP_LOAD, LAYER_TEMPERATURE_1, LAYER_TEMPERATURE_2, LAYER_TEMPERATURE_3, and their corresponding measures.
measurement depth LAYER_TEMP_DEPTH_1, LAYER_TEMP_DEPTH_2, and LAYER_TEMP_DEPTH_3. The variable AGE, indicating the service age of the pavement, was extracted as well. A description of all of the variables is provided in Table 2. All imperial units were converted into metric units. In this study, 11,075 samples consisting of 635 pavement rehabilitation projects were investigated.

Table 2. Variable description.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMP_THICKNESS</td>
<td>Increase in pavement thickness due to the rehabilitation activity (cm).</td>
<td>0.2</td>
<td>21.1</td>
<td>7.3</td>
<td>4.2</td>
</tr>
<tr>
<td>AVG_ANN_PRECIP</td>
<td>Average annual precipitation (mm).</td>
<td>1753</td>
<td>63,322</td>
<td>22,746</td>
<td>12,058</td>
</tr>
<tr>
<td>AVG_FREEZE_INDEX</td>
<td>Average freeze index.</td>
<td>0</td>
<td>1746</td>
<td>224</td>
<td>379</td>
</tr>
<tr>
<td>ANNUAL_kESAL</td>
<td>Estimated annual kilo ESAL (Equivalent standard axle load).</td>
<td>2</td>
<td>3409</td>
<td>391</td>
<td>530</td>
</tr>
<tr>
<td>SN_VALUE</td>
<td>Structural number used for AC pavements.</td>
<td>2</td>
<td>14</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>LAYER_THICKNESS</td>
<td>Thickness of the asphalt layer (cm).</td>
<td>3</td>
<td>56</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>BASE_THICKNESS</td>
<td>Thickness of the base (cm).</td>
<td>0</td>
<td>71</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>SUBBASE_THICKNESS</td>
<td>Thickness of the subbase (cm).</td>
<td>0</td>
<td>525</td>
<td>29</td>
<td>38</td>
</tr>
<tr>
<td>DROP_LOAD</td>
<td>Peak drop load (plate pressure) (kPa).</td>
<td>187</td>
<td>1171</td>
<td>688</td>
<td>236</td>
</tr>
<tr>
<td>LAYER_TEMPERATURE_1</td>
<td>Measured subsurface layer temperature at depth LAYER_TEMP_DEPTH_1 (deg C).</td>
<td>−7</td>
<td>58</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>LAYER_TEMPERATURE_2</td>
<td>Measured subsurface layer temperature at depth LAYER_TEMP_DEPTH_2 (deg C).</td>
<td>−8</td>
<td>56</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>LAYER_TEMPERATURE_3</td>
<td>Measured subsurface layer temperature at depth LAYER_TEMP_DEPTH_3 (deg C).</td>
<td>−9</td>
<td>53</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>LAYER_TEMP_DEPTH_1</td>
<td>Layer temperature hole depth for temperature measurements LAYER_TEMPERATURE_1 (mm).</td>
<td>1.3</td>
<td>6.4</td>
<td>2.5</td>
<td>0.2</td>
</tr>
<tr>
<td>LAYER_TEMP_DEPTH_2</td>
<td>Layer temperature hole depth for temperature measurements LAYER_TEMPERATURE_2 (mm).</td>
<td>2.5</td>
<td>27.9</td>
<td>9.4</td>
<td>4.5</td>
</tr>
<tr>
<td>LAYER_TEMP_DEPTH_3</td>
<td>Layer temperature hole depth for temperature measurements LAYER_TEMPERATURE_3 (mm).</td>
<td>3.8</td>
<td>50.8</td>
<td>17</td>
<td>8.5</td>
</tr>
<tr>
<td>AGE</td>
<td>Service age of pavements.</td>
<td>0</td>
<td>23</td>
<td>5</td>
<td>4.6</td>
</tr>
</tbody>
</table>

2.3. Data Preprocessing

Previous research indicated that there was a linear relationship between the logarithm of deflection basin parameter and logarithm of layer modulus. The shape of the deflection basin of semi-rigid base asphalt pavement at a center offset of 30–120 cm was very close to the exponential curve. Therefore, logarithmic transformation was conducted for FWD measurements. It can be seen from Figure 4 that the distribution of lg(D₀–Dₚ) is close to a normal distribution after the logarithmic transformation. We found that the prediction accuracy of the machine learning methods improved from 0.908 to 0.932 after performing the logarithmic transformation in this study. Feng et al. [35] concluded that there was a linear relationship between the design deflection and ESAL in the double logarithmic coordinate. Thus, variable ANNUAL_kESAL was also transformed to lg(ANNUAL_kESAL) for further analysis.
3. Machine Learning Methods

Three machine learning methods—Classification and Regression Tree (CART), Random Forest (RF), and Gradient Boosting Decision Tree (GBDT)—were used to evaluate the feature importance on deflection parameters and evaluate the prediction accuracy.

3.1. Classification and Regression Tree (CART)

Classification and Regression Tree, known as CART, was proposed by Breiman et al. [36]. CART is a supervised machine learning algorithm, composed of feature selection, tree generation, and pruning. The essence of CART is constructing a binary decision tree recursively. It can split nominal and continuous attributes. CART can be used for classification and regression. For the regression tree, the least-squared deviation criterion is used to split attributes. For the classification tree, the Gini index minimization criterion is used for feature selection to generate a binary tree. In this study, the target variable was pavement deflection parameter which was a continuous variable. Therefore, the regression tree was used.

For the regression tree, there are two critical problems to be solved: one is to select splitting points; the other is to determine the output value of nodes in the tree. The following procedure is used to establish the regression tree.

1. Select \( j \) variable and its value \( s \) as the splitting variable and splitting point to split the input space. Traverse variable \( j \), find its corresponding splitting point \( s \), and select the optimal pair \((j, s)\) that minimizes Equation (1).

\[
\min_{j,s} \left[ \min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \tag{1}
\]

where \( R_1 \) and \( R_2 \) are the split input spaces. \( R_1(j, s) = \{x | x^{(i)} \leq s\} \) and \( R_2(j, s) = \{x | x^{(i)} > s\} \). \( c_1 \) and \( c_2 \) are corresponding output values of \( R_1 \) and \( R_2 \) space.

2. Split the input space based on the optimal pair \((j, s)\) and determine the output value using Equation (2).

\[
\hat{c}_m = \frac{1}{N_m} \sum_{x_i \in R_m(j,s)} y_i, \; x \in R_m, \; m = 1, 2 \tag{2}
\]

3. Continue to perform step 1 and 2 on the two subspaces until the stop condition is satisfied.
The input space is split into \( M \) subsets \( R_1, R_2, \ldots, R_M \) and the output value of each subset is \( \hat{c}_m \). The decision tree is generated as shown in Equation (3).

\[
f(x) = \sum_{m=1}^{M} \hat{c}_m I(x \in R_m)
\]  

(3)

Pre-pruning and post-pruning can be used to cut back the tree to avoid overfitting. The tree is pruned back to the point where the cross-validated error is minimized in our study.

3.2. Random Forest (RF)

Random Forest (RF) is an ensemble learning method based on the bagging algorithm developed by Breiman and Cutler [37]. RF builds multiple decision trees during the training stage to improve accuracy. It can be used for both classification and regression. For classification, the output of RF is the majority outcome that most trees vote. For regression, the average prediction of the individual trees is taken as the final prediction. RF overcomes the decision tree’s habit of overfitting, and it generally outperforms decision trees.

To generate a Random Forest, the training algorithm known as bootstrap aggregating or bagging is employed whereby the training data for each tree making up the forest is selected randomly with replacement. Therefore, each tree will be trained using a different subset of training data. Given a training dataset consisting of \( n \) samples and \( p \) features, a decision tree in RF is grown and propagated by the following steps.

1. Create a bootstrap sample of equivalent size \( n \) by randomly sampling with replacement from the pool of \( n \) samples.
2. Select designated \( m_{\text{try}} \) (<<\( p \)) features, sampling without replacement from the available \( p \) feature pool for each tree, with one feature deciding the split at each node of a decision tree.
3. Grow the tree to maximum depth without any pruning. For each tree in a forest, bootstrap samples and composition of \( m_{\text{try}} \) nodal features vary. The number of trees and the maximum number of features are two critical hyperparameters to be tuned to improve RF accuracy.

3.3. Gradient Boosting Decision Tree (GBDT)

Gradient Boosting Decision Tree (GBDT) is an ensemble learning method developed by Friedman [38] that can also be used in classification and regression. Unlike bagging in RF, boosting does not involve bootstrap sampling. Figure 5 shows the difference between bagging and boosting. Boosting algorithm is used to sequentially combine individual weak learners in a way that each new learner fits the residual from the previous step to achieve a strong learner. For GBDT, the weak learners are decision trees. Each tree attempts to minimize the residual of the previous tree. It uses a loss function to minimize the residuals and converge to a final output value. For instance, mean squared error can be used for regression, while logarithmic loss can be used for classification. The loss function is optimized using the gradient descent. It is worth noting that existing trees in the model do not change when a new tree is added. Every time a new tree is added, it fits on a modified version of the initial dataset. Generally, the accuracy of GBDT is higher than RF.
minimize the residual of the previous tree. It uses a loss function to minimize the residuals and converge to a final output value. For instance, mean squared error can be used for regression, while logarithmic loss can be used for classification. The loss function is optimized using the gradient descent. It is worth noting that existing trees in the model do not change when a new tree is added. Every time a new tree is added, it fits on a modified version of the initial dataset. Generally, the accuracy of GBDT is higher than RF.

Figure 5. Diagram of bagging and boosting: (a) bagging; (b) boosting.

For a training set \( \{(x_i, y_i)\}_{i=1}^n \), a differentiable loss function \( L(y, f(x)) \), and number of iteration \( M \), the steps as follows can be used to conduct a GBDT.

1. Initialize the weak learner using Equation (4).

\[
f_0(x) = \arg\min_c \sum_{i=1}^n L(y_i, c)
\]  

(4)

2. For \( m = 1 \) to \( M \), conduct the following steps (a) to (d).

(a) For sample \( i = 1, 2, \cdots, n \), compute the residual as shown in Equation (5).

\[
\gamma_{im} = \left[ \frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x_i) = f_{m-1}(x)}
\]  

(5)

(b) Take the residual obtained in the previous step as the new true value, and regard the dataset \( (x_i, \gamma_{im}) \), \( i = 1, 2, \cdots, n \) as the training dataset of next decision tree to obtain a new decision tree \( f_m(x) \) whose leaf node region is \( R_{jm} \), \( j = 1, 2, \cdots, J \). \( J \) is the number of its leaves.

(c) For leaf region \( j = 1, 2, \cdots, J \), solve the following optimization problem as shown in Equation (6).

\[
\gamma_{jm} = \arg\min_\gamma \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma)
\]  

(6)

(d) Update the strong learner as shown in Equation (7).

\[
f_m(x) = f_{m-1}(x) + \sum_{j=1}^J \gamma_{jm} I(x \in R_{jm})
\]  

(7)
3. Output the final learner as shown in Equation (8).

\[ f(x) = f_M(x) = f_0(x) + \sum_{m=1}^{M} \sum_{j=1}^{J} \gamma_{jm} I(x \in R_{jm}) \] (8)

Generally, the number of trees and learning rate are two key hyperparameters that affect the performance and accuracy of a GBDT model.

4. Discussions of Results

The three machine learning methods elaborated above were used to evaluate and predict the relationship between the influencing factors and deflection measurements \( \lg(D_0-D_{20}), \lg(D_0-D_{30}), \lg(D_0-D_{45}), \lg(D_{20}-D_{60}), \lg(D_{30}-D_{60}), \lg(D_{60}-D_{90}), \lg(D_{60}-D_{120}), \) and \( \lg(D_{90}-D_{120}) \). The dataset was randomly divided into 70% training set and 30% validation set. For CART, R-square was used to determine the number of split. For RF, the number of trees and maximum number of features were set as 100 and 13, respectively. For GBDT, the number of trees and learning rate were set as 300 and 0.3, respectively.

4.1. Feature Importance

Figure 6 shows the feature importance of all the influencing factors on deflection parameters evaluated by GBDT, RF, and CART, respectively. The overall tendency of feature importance obtained by the three methods was similar except some minor differences. Among the variables of pavement structural parameters including SN VALUE, LAYER_THICKNESS, BASE_TYPE, BASE_THICKNESS, and SUBBASE_THICKNESS, SN_VALUE was a very important feature of all FWD deflection parameters, which was consistent with the theoretical explanations. Feature importance of SN_VALUE on \( \lg(D_0-D_{20}) \) and \( \lg(D_0-D_{30}) \) was smaller than other FWD measurements, indicating that the effect of pavement structural number on the deflection near the basin center was smaller. The relative feature importance of asphalt layer, base, and subbase on \( \lg(D_0-D_{20}) \) and \( \lg(D_0-D_{30}) \) was asphalt layer > subbase > base; their relative importance on \( \lg(D_{20}-D_{60}), \lg(D_{30}-D_{60}), \) and \( \lg(D_{60}-D_{90}) \) was asphalt layer > base > subbase; and their relative importance on \( \lg(D_{90}-D_{120}) \) and \( \lg(D_{60}-D_{120}) \) was base > subbase > asphalt layer. This indicates that when the center offset of the deflection sensor location was small, the correlation between FWD measurements and the thickness of asphalt layer was greater than the correlation with the thickness of base and subbase. With increasing center offset, the correlation between FWD measurements and thickness of base was greater than the correlation with the thickness of asphalt layer. For the importance of different base types, the deflection of TB was lower than GB.

Among the FWD test condition variables including DROP_LOAD, LAYER_TEMPERATURE_1, LAYER_TEMPERATURE_2, LAYER_TEMPERATURE_3, LAYER_TEMP_DEPTH_1, LAYER_TEMP_DEPTH_2, and LAYER_TEMP_DEPTH_3, DROP_LOAD was the most significant factor influencing deflection measurements. LAYER_TEMPERATURE_2 was also important for \( \lg(D_0-D_{20}), \lg(D_0-D_{30}), \) and \( \lg(D_0-D_{45}) \). This indicates that the second layer temperature, whose average depth was 9.4 mm, was correlated with the deflections near the basin center. The importance of other temperatures and temperature depths was small.

For the climate factors, the importance of AVG_ann_precip was greater than AVG_FREEZE_INDEX. When the sensor locations were far away from the basin center, the importance of precipitation on deflections became greater. The traffic level ANNUAL_kESAL was an important feature for all the measured deflections. The importance of rehabilitation level IMP_THICKNESS on deflections was not great and the values of feature importance at various deflection sensor locations were similar. The importance of service age on \( \lg(D_{90}-D_{120}) \) was greater than the importance on other deflection parameters.
Figure 6. Cont.
Figure 6. Feature importance of three machine learning methods: (a) GBDT; (b) RF; (c) CART.

4.2. Prediction Accuracy

Table 3 shows the R-squares of GBDT, RF, and CART used to predict the different deflection parameters. It can be seen that all three methods exhibit good prediction performance. Generally, GBDT outperformed RF, and RF outperformed CART when analyzing the same dataset. The prediction accuracy of GBDT was as high as 99%, indicating that GBDT was an effective machine learning method. Figure 7 shows the plot of actual and predicted $\lg(D_{0-D20})$ using the three different methods. It also can be concluded that the prediction accuracy ranking is GBDT > RF > CART.

Table 3. R-square of three machine learning methods.

<table>
<thead>
<tr>
<th>Target</th>
<th>Method</th>
<th>GBDT</th>
<th>RF</th>
<th>CART</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lg(D_{0-D20})$</td>
<td>0.991</td>
<td>0.904</td>
<td>0.729</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{0-D30})$</td>
<td>0.993</td>
<td>0.911</td>
<td>0.796</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{0-D45})$</td>
<td>0.992</td>
<td>0.917</td>
<td>0.803</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{20-D60})$</td>
<td>0.993</td>
<td>0.933</td>
<td>0.799</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{30-D60})$</td>
<td>0.993</td>
<td>0.935</td>
<td>0.849</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{30-D90})$</td>
<td>0.994</td>
<td>0.933</td>
<td>0.827</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{60-D120})$</td>
<td>0.990</td>
<td>0.896</td>
<td>0.695</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{60-D90})$</td>
<td>0.992</td>
<td>0.918</td>
<td>0.787</td>
<td></td>
</tr>
<tr>
<td>$\lg(D_{60-D120})$</td>
<td>0.992</td>
<td>0.911</td>
<td>0.812</td>
<td></td>
</tr>
</tbody>
</table>
Figure 7. Actual vs. predicted of $\lg(D_0-D_{20})$ using three machine learning methods: (a) GBDT; (b) RF; (c) CART.

5. Conclusions

In this study, three machine learning methods—CART, RF, and GBDT—were used to evaluate and predict the FWD measurements of deflection basin based on LTPP data. The feature importance of influencing factors including FWD test conditions, pavement structural parameters, climatic factors, traffic level, rehabilitation level, etc. on deflections were analyzed. The main conclusions that can be drawn from this study are summarized below:

1. Among the six variables of pavement structural parameters, structural number was an important feature for all FWD measurements, but its importance on $\lg(D_0-D_{20})$ and $\lg(D_0-D_{30})$ was smaller than other FWD measurements. The relative feature importance of asphalt layer, base, and subbase on $\lg(D_0-D_{20})$ and $\lg(D_0-D_{30})$ was asphalt layer > subbase > base; their relative importance on $\lg(D_{20}-D_{60})$, $\lg(D_{30}-D_{60})$, and $\lg(D_{30}-D_{90})$ was asphalt layer > base > subbase; and their relative importance on $\lg(D_{90}-D_{120})$ and $\lg(D_{60}-D_{120})$ was base > subbase > asphalt layer.

2. Among the FWD test condition variables, drop load was the most important factor influencing deflections. The second layer temperature was also important for $\lg(D_0-D_{20})$,
lg(D₀–D₃₀), and lg(D₀–D₄₅). The importance of temperature depths and other temperatures was small.

(3) The importance of precipitation on deflection was greater than freeze index. The traffic level was also an important feature for all measured deflection.

(4) All three methods—GBDT, RF, and CART—were demonstrated to be effective and efficient approaches analyzing the deflection. Additionally, the prediction accuracy of GBDT was as high as 99%. Generally, GBDT outperformed RF, and RF outperformed CART.

FWD deflection basin can comprehensively represent the pavement bearing capacity. The analyses between FWD deflection basin parameter and influencing factors especially the pavement structural characteristics can provide theoretical evidence for the pavement layer strength evaluation through FWD data. Moreover, deflection is widely used in the back-calculation of pavement layer modulus. Therefore, the precise prediction of deflection can improve the back-calculation accuracy. In our further study, the machine learning methods will be used to back-calculate the pavement layer modulus based on the FWD data.

Author Contributions: Conceptualization, X.C.; Data curation, X.C.; Funding acquisition, X.C., Q.D. and S.D.; Writing—original draft, X.C.; Writing—review and editing, Q.D. and S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities, CHD grant number 300102341508, the Natural Science Foundation of Jiangsu Province grant number BK20200468, the National Natural Science Foundation of China grant number 52208439 and 51978163.

Data Availability Statement: Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

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

References
22. Rabbi, M.F.; Mishra, D. Using FWD deﬂection basin parameters for network-level assessment of ﬂexible pavements. *Int. J. Pavement Eng.* 2021, 22, 147–161. [CrossRef]