Comparative Analysis of Machine Learning Models for Prediction of Remaining Service Life of Flexible Pavement

Narjes Nabipour 1, Nader Karballaezadeh 2, Adrienn Dineva 3, Amir Mosavi 3,4, Danial Mohammadzadeh S. 5,6 and Shahaboddin Shamshirband 7,8,*

1 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; narjesnabipour@duytan.edu.vn
2 Department of Civil Engineering, Shahrood University of Technology, Shahrood 3619995161, Iran; n.karballaezadeh@gmail.com
3 Institute of Automation, Kalman Kando Faculty of Electrical Engineering, Obuda University, 1034 Budapest, Hungary; dineva.adrienn@kvk.uni-obuda.hu (A.D.); amir.mosavi@kvk.uni-obuda.hu (A.M.)
4 Queensland University of Technology, 130 Victoria Park Road, Queensland 4059, Australia
5 Department of Civil Engineering, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran; dmvk1989@yahoo.com
6 Department of Elite Relations with Industries, Khorasan Construction Engineering Organization, Mashhad 9185816744, Iran
7 Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam
8 Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam
* Correspondence: shahaboddin.shamshirband@tdtu.edu.vn

Received: 19 October 2019; Accepted: 2 December 2019; Published: 6 December 2019

Abstract: Prediction of the remaining service life (RSL) of pavement is a challenging task for road maintenance and transportation engineering. The prediction of the RSL estimates the time that a major repair or reconstruction becomes essential. The conventional approach to predict RSL involves using non-destructive tests. These tests, in addition to being costly, interfere with traffic flow and compromise operational safety. In this paper, surface distresses of pavement are used to address the aforementioned challenges. To implement the proposed theory, 105 flexible pavement segments are considered. For each pavement segment, the type, severity, and extent of surface damage and the pavement condition index (PCI) were determined. The pavement RSL was then estimated using non-destructive tests include falling weight deflectometer (FWD) and ground-penetrating radar (GPR). After completing the dataset, the modeling was conducted to predict RSL using three techniques include support vector regression (SVR), support vector regression optimized by the fruit fly optimization algorithm (SVR-FOA), and gene expression programming (GEP). All three techniques estimated the RSL of the pavement by selecting the PCI as input. The correlation coefficient (CC), Nash–Sutcliffe efficiency (NSE), scattered index (SI), and Willmott’s index of agreement (WI) criteria were used to examine the performance of the three techniques adopted in this study. In the end, it was found that GEP with values of 0.874, 0.598, 0.601, and 0.807 for CC, SI, NSE, and WI criteria, respectively, had the highest accuracy in predicting the RSL of pavement.

Keywords: machine learning; flexible pavement; remaining service life prediction; pavement condition index; support vector regression (SVR); fruit fly optimization algorithm (FOA); gene expression programming (GEP)
1. Introduction

Pavement performance prediction models are an essential part of pavement management systems (PMSs) [1–4]. In addition to estimating the future condition of the pavement, these models help pavement engineers in the following [5,6]: (1) determining the optimal time for maintenance, rehabilitation, and reconstruction (MR&R) activities, (2) suggesting the most economical MR&R strategies, (3) estimating the required budget for MR&R activities, and (4) anticipating the results of various strategies. Until the late 1950s, little attention was paid to pavement performance and pavements were often categorized as satisfactory or unsatisfactory [7]. With the development of computational methods and the presentation of ML methods, many researchers have studied and developed pavement performance prediction models. These models differ in two respects: firstly, the index that was adopted as a criterion for pavement performance, secondly, the selected approach for modeling.

Through reviewing the pavement performance models, it was found that the Pavement Condition Index (PCI) [8], International Roughness Index (IRI) [3,9–12], Present Serviceability Index (PSI) [5,13], Remaining Service Life (RSL) [14], and Present Serviceability Ratio (PSR) [15] indices were more commonly used by the researchers. In terms of the modelling approach, methods such as artificial neural networks (ANN) [5,12], support vector machine (SVM) [1,11,14,15], radial basis function (RBF) [11], gene expression programming (GEP) [10], regression tree (RT) [13,16], Random Forest (RF) [3,17], and genetic programming (GP) [8] were used in the literature.

In general, pavement management activities are split into two categories [18]: network-level management and project-level management. Project-level management specifies roads that need to be repaired, repair process, and repair timetable. Therefore, predicting future conditions of pavement is essential for project management [18]. Forecasting pavement future conditions will require ongoing pavement assessment and inspection that will improve the operational quality of maintenance operations [19,20]. On the other hand, network-level management focuses on determining the budget required to preserve the pavement network at the standard level. Hence, it is indispensable to determine the RSL of pavement at this level of management [18]. Various factors such as traffic, characteristics of pavement materials, subgrade properties, climatic conditions, and maintenance quality have a destructive impact on the road pavement. As a result, pavement service life is surrounded by uncertainties that complicate the prediction of RSL [21].

The necessity of determining RSL in the previous paragraph has been recognized. Besides RSL, the other common indicators of pavement performance is PCI, which represents the general conditions of the pavement surface and ranges from zero for a practically unusable pavement to 100 for a flawless pavement (according to ASTM D-6433-07) [22]. PCI is determined based on pavement inspection results in terms of type, severity, and the extent of distresses [23,24]. After a thorough inspection of pavement surface distresses by an experienced inspector, PCI will be determined according to the related standard [22]. PCI is an index adopted in the project-level pavement management process. Hence, this index is specified before entering network-level pavement management. Conversely, the RSL of pavement, which is one of the pillars of network-level pavement management, is determined by applying the falling weight deflectometer (FWD) non-destructive test. In this test, a traffic lane is first blocked. Then an impulsive loading is applied to the pavement to induce pavement surface deflections. By analyzing pavement surface deflections, the RSL is calculated. In addition to the characteristics of the pavement layers materials (Young’s modulus and Poisson’s ratio) required for back-analysis in stress and strain calculations, a key point in calculating the RSL using the above method is knowing the pavement layers thickness for the analysis of deflections. The thickness of the pavement layers is determined using the ground-penetrating radar (GPR) non-destructive test. As a result, two non-destructive tests are required to determine the RSL of pavement. Traffic interference during FWD and GPR tests should also not be forgotten.

In light of the above, the current method of determining RSL is not only costly but also compromises the safety of users due to traffic interference during testing. Given the limited budget resources of transportation agencies, determining the RSL of pavement to manage a pavement
network represents one of the ongoing concerns of such agencies [20]. Considering the necessity of
RSL determination for PMSs as well as the problems mentioned in the current procedure of RSL
determination, a new pavement performance model is proposed in this paper. In this study, three
methods of SVR, SVR-FOA (support vector regression optimized by the fruit fly optimization
algorithm), and GEP have been applied to predict the RSL of road pavements. Given the importance
and widespread use of PCI in PMSs, RSL modeling was based on PCI.

The main objective of this study was to save time, reduce costs, and increase safety in the RSL
determination process. According to the PCI calculation on project-level activities, using PCI for
network-level management operations, estimation of RSL, will contribute to the overlapping of
activities and saving time. The methods presented in this paper, by excluding non-destructive tests
from the process of determining the RSL of pavement, drastically reduce costs. On the other hand,
by eliminating non-destructive tests, traffic interference during testing and the potential safety
hazards to road users are also eliminated.

The data required for developing the innovative SVR-FOA method proposed in this paper was
collected from Shahrood–Damghan highway in the Semnan province, Iran. The pavement type of
this highway is flexible. In this highway, a 100-m long pavement segment was taken from the
beginning of each kilometer, and after assessing the surface distresses of each segment, its PCI was
calculated. In the next step, FWD and GPR tests were applied to the selected segments to determine
the RSL. Finally, using three SVR, SVR-FOA, and GEP methods, the RSL modeling based on PCI was
implemented.

This paper is organized as follows: Section 2 offers a review of studies on the prediction of RSL.
Section 3 introduces the method employed in this study. This section is made of four sub-sections
titled PCI, RSL, machine learning techniques, and case study. In Section 4, the results of the analyses
are presented and discussed. Finally, the conclusions are drawn in Section 5.

2. Literature Review

PMS is an Assessment Management System (AMS) used by road network administrators to
maintain the entire network at the desired level. Predicting pavement performance is a key factor in
PMSs [25]. The models of predicting pavement’s RSL fall into the category of the pavement
performance prediction models. In this section, studies carried out by other researchers on predicting
RSL are reviewed. Table 1 lists the results of studies that have strived to estimate RSL to date.

<table>
<thead>
<tr>
<th>Category</th>
<th>Model Inputs</th>
<th>Equation</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Based on pavement responses</td>
<td>$\varepsilon_1 =$ Tensile strain at the bottom of the asphalt layer, $E_1 =$ Elastic modulus of asphalt, $f_1, f_2,$ and $f_3 =$ Regression coefficients.</td>
<td>$RSL_{fatigue} = f_3(\varepsilon_1)^{-f_2}(E_1)^{-f_3}$</td>
<td>Huang (1993)</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_2 =$ Compressive strain at the top of the subgrade, $f_4$ and $f_5 =$ Regression coefficients.</td>
<td>$RSL_{rutting} = f_4(\varepsilon_2)^{-f_5}$</td>
<td>Huang (1993)</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_3 =$ Tensile strain at the asphalt layer bottom, $E_{AC}$ = Modulus of asphalt, $a$, $b$, and $c =$ Constant coefficients of regression.</td>
<td>$\ln(RSL_{fatigue}) = a - b \ln(\varepsilon_3) - c \ln(E_{AC})$</td>
<td>Hossain &amp; Wu (2002)</td>
</tr>
<tr>
<td></td>
<td>$\varepsilon_4 =$ Horizontal tensile strain at the bottom of the asphalt layer, $K$ and $c =$ Regression coefficients.</td>
<td>$RSL_{fatigue} = K(\varepsilon_4)^{-c}$</td>
<td>Park &amp; Kim (2003)</td>
</tr>
</tbody>
</table>
2. Based on pavement quality indices

IRI = International roughness index,
\( a = \) Initial IRI (where age is zero),
\( b = \) Curvature of performance line.

\[
\text{PCI} = \ln\left(\frac{\text{IRI}_{\text{terminal}}}{a}\right) - \text{Current age} \\
\text{RSL} = \frac{\ln(\text{PCI}) - 14.728}{4.1872} \\
\text{Al-suleiman \\ & Shiyab (2003)}
\]

3. Based on the result of the non-destructive test

\( \delta = \) Pavement surface curvature,
\( \delta = D_0 - D_{30} \)
\( \alpha \) and \( \beta \) = Material constants.

\[
\text{RSL}_{\text{fatigue}} = a\left(\frac{1}{0.00236 + 0.00002}\right)^\beta \\
\text{Saleh (2016)}
\]

\[
\text{AUPP} = \frac{5D_0 - 2D_{30} - 2D_{60} - D_{90}}{2} \\
\alpha \) and \( \beta \) = Material constants.
\]

\[
\text{RSL}_{\text{fatigue}} = a\left(\frac{1}{0.0000023AUPP^{0.012}}\right)^\beta \\
\text{Saleh (2016)}
\]

In Table 1, the models of determining the RSL of pavement are divided into three categories based on the model inputs:

- **First category**: Models that predict the RSL based on the response (stress and strain) of pavement to the applied loads.
- **Second category**: Models that predict the RSL based on pavement quality indices.
- **Third category**: Models that predict the RSL based on the results of pavement non-destructive tests.

Among the above categories, models that predict the remaining pavement service based on qualitative indices appear to be more appropriate. It is because such models neither call for the analysis of pavement behavior and response, as in the first category nor require non-destructive tests, like the third category. Instead, they estimate the RSL of the pavement by assessing pavement and calculating a qualitative index in the simplest possible way. In light of the above points, in this paper, PCI has been adopted as a qualitative index for predicting RSL. Setyawan et al. conducted a similar study on East Line of South Sumatera the results of which are displayed in Table 1. They evaluated the condition of road performance and damages and determined PCI. Then they calculated the RSL of the pavement using the deflection data acquired from falling weight deflectometer measurements. Finally, the relationship between these two values was examined. The research involves five sections of the route with the different damaged condition. After calculating the PCI and RSL values, the relationship between these two parameters was determined with the help of Microsoft Excel software. The model presented in their study has two drawbacks: (1) The number of data for building a reliable model is too small, (2) instead of using simple regression with Microsoft Excel, more advanced methods, such as ML can be used to achieve better results. In contrast, the methods proposed in this paper are based on the data gathered from 105 pavement segments using machine learning methods.

In general, PCI offers a valid index accepted by all transportation agencies around the world, and it is widely used in their evaluations. Compared to the current method of estimating the RSL of pavements (using two non-destructive FWD and GPR tests), the proposed method provides a far simpler, safer, and less costly way of estimating RSL.

3. Methodology

3.1. Pavement Condition Index (PCI)

Extensively used in roads, parking lots, and airports, PCI is recognized as a standard practice by many organizations around the world, including the Federal Aviation Administration, the American Public Works Association, and the U.S. Air Force [34]. PCI is a numerical index that expresses the rate
of pavement surface distresses. PCI exhibits structural integrity and surface operational condition but is not able to measure structural capacity [22].

The first step in determining the PCI of each pavement segment is to determine the type, extent, and severity of surface distresses. The second step is to determine the deduct values (DVs) with the help of the specific curves for each type of distress. Next, reduce the number of DVs to the maximum number allowed ($m_i$) [22,34]:

$$m_i = 1 + \frac{(100 - HDV) \times 9}{98},$$

where HDV = greatest individual deduct value.

Next, specify the number of DVs greater than 2 ($q$). In the next step, the corrected deduct value (CDV) should be specified in a special curve with the help of TDV (sum of DVs) and $q$ [34]. Then among DVs that are larger than 2, decrease the smallest ones to 2. At this point, repeat the process of calculating $q$, computing CDV, and reducing DV until $q$ is equal to 1. By calculating the maximum CDV, PCI can be calculated [22,34]:

$$PCI = 100 - CDV_{\text{max}}$$

Readers can refer to [22] and [34] for reaching more details and related curves. The classification of a pavement segment based on PCI follows Table 2.

<table>
<thead>
<tr>
<th>Rating Scale</th>
<th>0–10</th>
<th>10–25</th>
<th>25–40</th>
<th>40–55</th>
<th>55–70</th>
<th>70–85</th>
<th>85–100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Failed</td>
<td>Serious</td>
<td>Very Poor</td>
<td>Poor</td>
<td>Fair</td>
<td>Satisfactory</td>
<td>Good</td>
</tr>
</tbody>
</table>

3.2. Remaining Service Life (RSL)

The RSL of pavement under operation is a key factor in implementing PMSs. It is because learning about the future conditions of the pavement network is essential for decision making, life cycle cost analysis, planning, and budget allocation [14,35]. In general, the definitions of RSL by different agencies and departments of transportation can be split into two general categories [36]:

- The remaining time to reach a level of distress when the pavement needs to be rehabilitated or reconstructed. For example, the Minnesota Department of Transportation (MnDOT) defines the RSL as the time until the next major rehabilitation.
- The time until pavement conditions reach a specific condition index limit. For example, the Michigan Department of Transportation (MDOT) defines the RSL based on the Michigan Ride Quality Index, assuming an RSL of zero when the said index is 50.

The RSL of pavement segments in this paper has been determined using the heavy falling weight deflectometer (HWD). The HWD is an FWD, the application of which is not constricted to the road and could be used for airport pavement assessment. Figure 1 shows the HWD device employed in this study.

![Figure 1. the heavy falling weight deflectometer (HWD) used in this study for determining the remaining service life (RSL).](image)
The HWD applies a tension similar to the standard axle load (8.2 t) to the pavement surface over 10 to 35 ms. A number of geophones are placed on the pavement surface at specified distances from the loading center. The task of the geophones is to record the pavement deflections induced by the load applied with the HWD device. Standard axle load simulation in HWD is generated by a series of weight drops on a loading plate placed on the pavement surface [14]. Table 3 reveals the details of the HWD test undertaken in this study.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tension (kPa)</td>
<td>600–900</td>
</tr>
<tr>
<td>Number of geophones</td>
<td>9</td>
</tr>
<tr>
<td>Geophone distance from center of loading plate (cm)</td>
<td>0, 20, 30, 45, 60, 90, 120, 150, and 180</td>
</tr>
<tr>
<td>Number of weights falling</td>
<td>Four times</td>
</tr>
<tr>
<td>Loading plate radius (mm)</td>
<td>150</td>
</tr>
</tbody>
</table>

The deflections recorded by geophones are transferred to the central computer in HWD. In this computer, the analysis of pavement surface deflections is performed using ELMOD6 software. ELMOD uses Miner’s Law to calculate critical stresses and strains. Miner’s Law is used to interact with the impact of distresses that occur during each season and by each load. With the help of back-calculation analysis, ELMOD provides important outputs including RSL, overlay thickness, and elasticity module [38].

A prerequisite of estimating RSL in accordance with the process described in this section is knowing the thickness of pavement layers. For determining the pavement layer’s thickness, GPR non-destructive testing was carried out for all pavement segments. GPR is capable of calculating pavement thickness as a continuous profile by transmitting electromagnetic waves through a transmit antenna and receiving recursive signals [39]. The receiver of GPR takes reflected waves and shows them as a plot of amplitude and time [40]. The layers thickness (\(h_i\)) of the pavement is calculated by the following equation [41]:

\[
h_i = \frac{c \Delta t_i}{\sqrt{\varepsilon_r}},
\]

where \(\Delta t_i\) = time between amplitudes \(A_i\) and \(A_{i+1}\), \(c\) = electromagnetic wave speed through the vacuum, and \(\varepsilon_r\) = the relative dielectric constant of the layer. Table 4 shows the complete details of the GPR experiment carried out in this study.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenna type</td>
<td>1000 MHz</td>
</tr>
<tr>
<td>Record speed (km/hr)</td>
<td>10</td>
</tr>
<tr>
<td>Number of scanning (per meter)</td>
<td>10</td>
</tr>
<tr>
<td>Maximum pulse penetrating depth (ns)</td>
<td>32</td>
</tr>
</tbody>
</table>

3.3. Machine Learning Techniques

In this paper, the main objective is to present a new approach for predicting the RSL of flexible pavements based on PCI using machine learning techniques. The data set consisted of the RSL and PCI of all segments under study, which were analyzed using the GEP, SVR, and SVR-FOA methods. GEP is an evolutionary algorithm that investigates the relationship between input and output variables by developing computer programs [44]. Different from both the genetic algorithm (GA) and genetic programming (GP), GEP is a combination of both introduced by Ferreira in 2001 [45]. SVR is a supervised machine learning technique employed to solve regression problems. SVR is especially popular due to its desirable management and performance in handling nonlinear issues [46]. The success of an SVR in problem-solving depends on the proper selecting of its basic parameters. SVR’s
basic parameters include $c$, $\varepsilon$, and kernel function parameters [14]. Improper values of the basic parameters in the SVR can lead to under-fitting or over-fitting. Thus, the optimum values of them must be selected during training the SVR method [47]. In other words, different optimization algorithms were developed and may be utilized for selecting the suitable values of SVR. FOA is an intelligent swarm algorithm introduced by Pan in 2012, which utilizes the food searching strategy of a fruit fly to find the optimal values for SVR basic parameters [48]. These methods are introduced in the following subsections.

3.3.1. Gene Expression Programming (GEP)

GEP is a developed GP method that solves a problem by creating expression trees (ETs). In fact, GEP is an evolutionary algorithm for creating computer programs. The designed computer programs have sophisticated tree structures that are trained similar to a living organism by changing size, shape, and composition, and adapted to the conditions. Like living organisms, GEP programs are coded as simple fixed-length linear chromosomes. Hence, GEP is a genotype-phenotype system that employs a simple genome to store and transmit genetic information and adopts a complex phenotype to explore and adapt to the environment. The genome consists of a chromosome or a fixed-length string that combines one or more genes of the same size. In fact, each chromosome contains one or more genes known as Sub-ETs. In GEP, all Sub-ETs are linked through the root with connection functions. The connection functions in GEP include division, multiplication, subtraction, and addition [49]. Figure 2 shows an instance of the genotype-phenotype structure in GEP.

![Figure 2](image)

**Figure 2.** A sample of the genotype-phenotype structure in gene expression programming (GEP) [50].

These genes, despite their fixed length, are coded for ETs of varying size and shape. It implies that the size of the coding region varies from one gene to another to allow for progressive adaptation and evolution. Each gene has a coding area called open reading frame (ORF), which, after being coded as an expression tree, provides a solution to the problem [51]. Figure 3 demonstrates the coding region (ORF), the non-coding region, and the expression tree for a gene.
Like other evolutionary approaches, GEP begins by randomly generating the initial population chromosomes. In the first population, each chromosome is assessed based on the fitness function and receives a fitness value. Various fitness functions have been used in GEP, including root relative squared error (RRSE), relative square error (RSE), root mean square error (RMSE), and mean square error (MSE) [44]. The proper chromosomes are more likely to be picked in the next generation. After being selected, chromosomes are amended by genetic operators (including transposition, inversion, mutation, recombination, and gene crossover) and then reconstructed. This process is sustained until a suitable solution or the maximum number of generations is reached [44,53].

3.3.2. Support Vector Regression (SVR)

The SVR is considered as a support vector machine (SVM) used for regression problems. SVR has been widely used for civil engineering problems [54,55]. Supervised learning techniques, including SVR, utilize structural risk minimization (SRM), while conventional neural networks use empirical risk minimization (ERM). ERM minimizes the error of training samples, but SRM is able to minimize a higher level of error. As a result, SVM is capable of overcoming the deficiencies of conventional neural networks [14,46]. The main idea in SVR is to map nonlinear information into a higher dimensional space and then solve a linear regression problem in the new space [46,56,57]. In the new space, a simple linear kernel function is adopted to solve the problem. However, in complex problems, a simple linear kernel function will be inadequate. The kernel function \( k(x, z) \) for all \( x, z \in X \) is defined as follows [46]:

\[
k(x, z) = \langle \varphi(x), \varphi(z) \rangle.
\]

Each kernel function must have two features [46]:

1. Symmetricity
   \[
   k(x, z) = \langle \varphi(x), \varphi(z) \rangle = \langle \varphi(z), \varphi(x) \rangle = k(z, x),
   \]

2. Compliance with the Cauchy-Schwartz criterion
   \[
   k(x, z)^2 = \langle \varphi(x), \varphi(z) \rangle^2 \leq \|\varphi(x)\|^2 \|\varphi(z)\|^2.
   \]

These two conditions guarantee that the new space is definable by the kernel function. The most famed kernel functions are polynomial kernel, radial basis functional kernel, linear kernel, and sigmoid kernel [57]. Given the above explanation, it is clear that SVR requires an appropriate function to explain the nonlinear relationship between input \((x)\) and output \((y)\) [57]:

\[
f(x_i) = w. \varphi(x_i) + b,
\]

where \( \varphi(x_i) \) = transformation function, \( w \) = weight, \( b \) = bias. \( w \) and \( b \) are obtained by minimizing the following function, which is known as the regularized risk function [57]:

\[
k(x, z) = \langle \varphi(x), \varphi(z) \rangle.
\]
\[ R(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} L_{\varepsilon}(y_i, f(x_i)), \]  

where \( \frac{1}{2} \|w\|^2 \) = regularization term, \( C \) = penalty coefficient, and \( L_{\varepsilon}(y_i, f(x_i)) = \varepsilon \)-insensitive loss function:

\[ L_{\varepsilon}(y_i, f(x_i)) = \max(0, |y_i - f(x_i)| - \varepsilon), \]

where \( \varepsilon \) = permitted error threshold.

To solve the optimization boundaries, two factors of \( \xi \) and \( \xi^* \) are defined \[57\]:

\[ \min f(w, \xi, \xi^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (\xi_i, \xi^i), \]

subject to

\[
\begin{align*}
& y_i - [w \cdot \varphi(x_i)] - b \leq \varepsilon + \xi_i, & \xi_i \geq 0 \\
& [w \cdot \varphi(x_i)] + b - y_i \leq \varepsilon + \xi^i, & \xi^i \geq 0
\end{align*}
\]

We now need to define a Lagrange function based on the objective function and boundary conditions \[57\]:

\[
\max H(\partial_{\xi_i}, \partial_{\xi_i^*}) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\partial_{-i}^\ast - \partial_{i}^\ast) (\partial_{j}^\ast - \partial_{j}^\ast) K(x_i, x_j) + \sum_{i=1}^{N} y_i (\partial_{i}^\ast - \partial_{i}^\ast) - \varepsilon \sum_{i=1}^{N} y_i (\partial_{i}^\ast + \partial_{i}^\ast),
\]

subjected to

\[
\sum_{i=1}^{N} y_i (\partial_{i}^\ast - \partial_{i}^\ast) = 0 , \quad \partial_{i}^\ast, \partial_{i}^\ast \in [0, C].
\]

Therefore, the regression function can be shown as follows \[57\]:

\[ f(x) = \sum_{i=1}^{N} (\partial_{i}^\ast - \partial_{i}^\ast) K(x_i, x) + b, \]

where \( K(x_i, x_j) \) = kernel function. Figure 4 summarizes the overall structure of the SVR.

Figure 4. The transformation process in the support vector regression (SVR) \[58\].

In general, SVR performance depends on its parameters. SVR parameters are \[59,60\]:

- \( \varepsilon \): this parameter supervises the width of the \( \varepsilon \)-insensitive zone, used to fit the training data. The value \( \varepsilon \) can affect the number of support vectors used to build the regression function. For the bigger \( \varepsilon \), estimates are more 'flat', and the fewer support vectors are chosen.
- \( C \): this parameter specifies the trade-off between the complexity of the model and the grade to which deviations larger than \( \varepsilon \) are bearable in optimization formulation.
- \( \gamma \): this parameter determines the relation between error minimization and smoothness of the estimated function.

These parameters are chosen by the user based on prior knowledge of SVR, so this method is not suitable for non-professional users. Various algorithms have been developed to optimize the amounts of SVR parameters. In the following subsection, one of the optimization algorithms used in this paper is introduced.
3.3.3. Fruit Fly Optimization Algorithm (FOA)

FOA is an optimization algorithm developed based on the food search behavior of the Drosophila insect [48]. The fruit fly has superior smell and vision senses, which discriminates it from other insects. Fruit flies track the smell of the food sources dispersed through the air and heads towards it. This insect is even capable of smell tracking from a distance of 40 km. When approaching the food source, the fruit fly employs a sense of vision to locate food and other fruit flies. The best information is shared among the fruit flies, and finally, the route leading to the food source is identified [61]. Figure 5 illustrates the process of food search by a fruit fly.

![Figure 5. Fruit fly swarm optimization method [62].](image)

In this paper, this optimization algorithm has been selected to find the optimal values of SVR parameters, which is known as SVR-FOA. The general steps of SVR-FOA can be summarized as follows [63,64]. The flowchart of the SVR-FOA method is illustrated in Figure 6.

**Step 1.** SVR parameters ($\varepsilon$, $c$) initialization and kernel function determination.

**Step 2.** Parameter initialization; including the maximum number of iterations, location of initial population (X-axis, Y-axis), population size, and random flight distance domain:

$$X_{\text{axis}} = \text{rands}(1,2),$$  
$$Y_{\text{axis}} = \text{rands}(1,2).$$

**Step 3.** Population initialization

A random location ($X_i$, $Y_i$) and food founding distance are assigned to each fruit fly:

$$X_i = X_{\text{axis}} + \text{Random value},$$

$$Y_i = Y_{\text{axis}} + \text{Random value},$$

where: $i$ = Population size.

**Step 4.** Population evaluation

The distance from the origin to the food source ($D$) and the smell concentration parameter ($S$) are calculated:

$$D_i = \sqrt{X_i^2 + Y_i^2},$$

$$S_i = \frac{1}{D_i}.$$

**Step 5.** Replacement

$S$ value is substituted with the fitness function or smell concentration judgment function so that the smell concentration for each fruit fly location can be attained:

$$\text{Smell}_i = \text{Function}(S_i).$$

**Step 6.** Detect the maximal smell concentration
At this point, the fruit fly with the highest $S_i$ is identified and located within the population.

\[
[bestSmell\_bestIndex] = \max(\text{Smell}).
\]  

**Step 7. Keep smell concentration**

The coordinates of the maximum smell concentration are set, and the fruit fly swarm flows in that direction,

\[
X_{\\_axis} = X(\text{bestIndex}),
\]

\[
Y_{\\_axis} = Y(\text{bestIndex}).
\]

**Step 8. Iterative optimization**

Steps 3 to 6 are repeated until the smell concentration does not show any improvement compared to the previous one or the maximum number of repetitions in Step 2 is reached.

**Step 9. Output the optimum parameter of SVR**

![Flowchart of the support vector regression optimized by the fruit fly optimization algorithm (SVR-FOA) method [62].](image)

**Figure 6.** The flowchart of the support vector regression optimized by the fruit fly optimization algorithm (SVR-FOA) method [62].

### 3.4. Case Study

To implement the theory proposed in this paper, a stretch of 105 km from Shahrood–Damghan highway in Iran was selected and inspected. Given that this highway is part of the route between Tehran (the capital of Iran) and Mashhad (the second most important city of Iran), it constitutes one of the major roads. The highway consists of two lanes in each direction and uses the flexible pavement. A prerequisite of implementing the proposed theory is to select a number of sample segments from this highway. To do so, segments 100 m in length and 7 m in width were selected from the beginning of each kilometer. By inspecting the selected segments, the surface distress data (type, severity, and extent) of each segment was recorded in the assessment forms. The PCI of each segment was calculated as described in Section 3.1. With PCI known, the RSL of the pavement segments need to be known. Hence, HWD and GPR tests were performed on all segments and the RSL of each
segment was determined. Figure 7 shows the location of the Shahrood–Damghan Highway as well as the starting and ending points of the segments understudy on this Google map.

![Figure 7. Highway No. 44 (Shahrood–Damghan) map used in this study.](image)

4. Results and Discussion

The results of the analysis are presented in this subsection. First, the statistical specifications of the input and output modeling variables are listed in Table 5. The data was extracted from IBM SPSS 23 software (version 2015, International Business Machines Corporation (IBM), Armonk, New York, U.S.).

**Table 5. Statistical characteristics of the utilized data.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard Deviation</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>Sig. in Kolmogorov–Smirnov Test</th>
<th>Correlation with RSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCI</td>
<td>59.97</td>
<td>19.00</td>
<td>100.00</td>
<td>21.51</td>
<td>-1.093</td>
<td>0.03</td>
<td>0.068</td>
<td>0.572</td>
</tr>
<tr>
<td>RSL</td>
<td>17.77</td>
<td>0.00</td>
<td>40.00</td>
<td>15.04</td>
<td>-1.393</td>
<td>0.50</td>
<td>0.000</td>
<td>1</td>
</tr>
</tbody>
</table>

As depicted in Table 5, the mean PCI of all pavement segments is 59.97, which according to Table 2, is indicative of the fair state of all segments. The mean RSL of pavement is 17.77 years. For interpreting the mean RSL, it is worth noting that ELMOD6 software does not suggest the application of an overlay layer for this RSL. Hence, the average RSL of the segments is fairly desirable. Before calculating the correlation coefficient of the modeling input and output, the normality of the data must be determined. This is determined by kurtosis and skewness coefficients, as well as the results of the Kolmogorov–Smirnov test. In probability theory and statistics, kurtosis is a measure of the “tailedness” of the probability distribution of a real-valued random variable. Moreover, skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive or negative, or undefined. A zero-skewness dataset has perfect symmetry (completely normal distribution). A positive/negative skew value indicates that the tail on the right/left side of the distribution is longer than the left/right side and the bulk of the values lie to the left/right of the mean. Kurtosis acts as a measure of the distribution peakedness. For a perfectly normal distribution, kurtosis should be zero. Positive kurtosis means a high peak distribution curve, and negative kurtosis refers to the flat-topped distribution curve [65,66]. It is very difficult to achieve zero skewness and Kurtosis, then the data distribution is normal when the
skewness and Kurtosis values are between 2 and −2 [67]. So, it is clear from Table 5 that PCI, with lower kurtosis value, had better normal distribution than RSL. Kolmogorov–Smirnov is a non-parametric test that helps to determine if the data is normal or not. If the significance level (sig.) in this test is above 0.05, the data are normal [67]. When choosing a correlation test, even if one of the variables is abnormal, then the Spearman test is used. According to Table 5, since PCI variables have a normal distribution but RSL distribution is abnormal, the Spearman correlation test must be used. The correlation between PCI and RSL is 57.2%, which represents an average value.

As noted in Section 3.3.2, SVR consists of three basic parameters (C, ε, and γ), with the quality of SVR performance depending on the values selected for these three parameters. Table 6 shows the values of these basic parameters for the SVR as well as the optimized values of these parameters by the FOA algorithm.

### Table 6. Parameters of the SVR and SVR-FOA models.

<table>
<thead>
<tr>
<th>Model</th>
<th>SVR</th>
<th>SVR-FOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1.0000</td>
<td>1.0022</td>
</tr>
<tr>
<td>SVR parameter</td>
<td>ε 0.0100</td>
<td>0.2561</td>
</tr>
<tr>
<td></td>
<td>γ 0.0010</td>
<td>0.0760</td>
</tr>
</tbody>
</table>

In Table 7, the characteristics of the GEP model used in this study, including model parameters and genetic operators, are shown. It should be noted that the default parameters of Gene Expro Tools 4.0 software (Version 2.0, Gepsoft Limited, Bartolomeu Messines, Portugal) were selected for further computations.

### Table 7. Characteristics of the GEP model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head size</td>
<td>8</td>
</tr>
<tr>
<td>Number of Genes</td>
<td>3</td>
</tr>
<tr>
<td>Chromosomes</td>
<td>30</td>
</tr>
<tr>
<td>Linking function</td>
<td>Addition (+)</td>
</tr>
<tr>
<td>One-point recombination rate</td>
<td>0.3</td>
</tr>
<tr>
<td>Two-point recombination rate</td>
<td>0.3</td>
</tr>
<tr>
<td>Inversion rate</td>
<td>0.1</td>
</tr>
<tr>
<td>Gene recombination rate</td>
<td>0.1</td>
</tr>
<tr>
<td>Mutation rate</td>
<td>0.044</td>
</tr>
<tr>
<td>Gene transposition rate</td>
<td>0.1</td>
</tr>
<tr>
<td>Used functions</td>
<td>+, −, ×, ÷, power</td>
</tr>
</tbody>
</table>

Equation (25) shows the proposed formula of the GEP method, extracted from Gene Expro Tools 4.0, for estimating pavement remaining service life in terms of PCI:

$$\text{RSL} = -6.59964 + \frac{322.568_{PCI-9.96265}}{2.61881\sqrt{\text{PCI}} + \text{PCI} - 1.18921\sqrt{\text{PCI}}^2 + \frac{0.35393(3.8325 \text{ PCI} - 9.83936)}{9.83936 + \text{PCI}}}.$$ (25)

The results of scientific research are generally assessed with indicators that exhibit the accuracy and error of the analysis. In this study, four criteria entitled Correlation Coefficient (CC), Scattered Index (SI), Nash-Sutcliffe efficiency (NSE) and Willmott’s Index of agreement (WI) were used to determine the quality of the outputs [68,69]:

$$CC = \frac{\sum_{i=1}^{n} RSL_{Oi}RSL_{pi} - \frac{1}{n}\sum_{i=1}^{n} RSL_{Oi} \left( \sum_{i=1}^{n} RSL_{pi} \right)^2}{\left( \sum_{i=1}^{n} RSL_{Oi}^2 - \frac{1}{n}\left( \sum_{i=1}^{n} RSL_{Oi} \right)^2 \right)^{1/2} \left( \sum_{i=1}^{n} RSL_{pi}^2 - \frac{1}{n}\left( \sum_{i=1}^{n} RSL_{pi} \right)^2 \right)^{1/2}}.$$ (26)

$$SI = \frac{\frac{1}{n}\sum_{i=1}^{n}(RSL_{pi} - RSL_{Oi})^2}{RSL_{Oi}}.$$ (27)
\[
NSE = 1 - \frac{\sum_i^n (RSL_{Oi} - RSL_{Pi})^2}{\sum_i^n (RSL_{Oi} - \overline{RSL_{Oi}})^2}, \tag{28}
\]

\[
WI = 1 - \left[ \frac{\sum_i^n (RSL_{Pi} - RSL_{Oi})^2}{\sum_i^n (|RSL_{Pi} - RSL_{Oi}| + |RSL_{Oi} - \overline{RSL_{Oi}}|)} \right], \tag{29}
\]

where RSL_{Oi} = observed RSL i\text{th} value, RSL_{Pi} = predicted RSL i\text{th} value, and \( \overline{RSL_{Oi}} \) = average of RSL_{Oi}.

The CC is a number in the range of \([-1, +1]\), with values of +1 and −1 indicating a complete correlation between model inputs and outputs. Positive values represent direct correlation and negative values demonstrate an inverse correlation. As the absolute value of CC approaches zero, the strength of the correlation decreases. SI represents an error and smaller values indicate lower errors in modeling. The highest NSE value is one with values close to one indicating greater modeling accuracy so that NSE = 1 represents the best modeling quality. WI is an index between 0 and 1 with values close to one suggesting higher modeling accuracy.

Table 8 reveals the four criteria introduced for all three methods employed in this study. To shed further light on the results of Table 8, a three-dimensional bar histogram of the criteria is presented in Figure 8. Based on the description in the preceding paragraph and the values in Table 8, it can be concluded that FOA has improved SVR results. Comparing the SVR-FOA and GEP modeling results, it can be contended that the GEP results are partially superior in the modeling proposed in this paper. Moreover, although the CC values of studied methods are not so high, due to lower SI values, they can be used with acceptable accuracy in RSL estimation.

Table 8. Performance evaluation indices for GEP, SVR, and SVR-FOA models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>GEP</th>
<th>SVR</th>
<th>SVR-FOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>0.874</td>
<td>0.865</td>
<td>0.879</td>
</tr>
<tr>
<td>SI</td>
<td>0.598</td>
<td>0.894</td>
<td>0.616</td>
</tr>
<tr>
<td>NSE</td>
<td>0.601</td>
<td>0.110</td>
<td>0.577</td>
</tr>
<tr>
<td>WI</td>
<td>0.807</td>
<td>0.369</td>
<td>0.786</td>
</tr>
</tbody>
</table>

Figure 8. Three-dimensional bar graphs of the statistical parameters.

There has been no standard method for splitting training and testing data. For instance, Choubin et al. [70] used a total of 63% of their data for model development, whereas Shamshirband et al. [71] utilized 67% of data, Mohammadzadeh et al. [72] 70%, and Samadarianfard et al. [73] implemented 80% of total data to develop their models. In this study, the dataset containing 105 pavement segments is used, where, approximately 70% of the data (i.e., 75 segments) are used for training, and
the remaining 30 segments are utilized for testing. Figure 9 shows the RSL predicted by the three SVR, SVR-FOA, and GEP methods as well as the RSL measured in the HWD test for the segments selected as the test. As can be seen in Figure 9, the studied models do not accurately predict the maximum and minimum RSL values. So, this may affect the scheduling of maintenance actions.

![Figure 9](image)

**Figure 9.** Observed and estimated values of RSL with SVR, SVR-FOA, and GEP models for test data.

Figure 10 displays the predicted RSL values versus the RSL values calculated by the HWD test for all three machine learning techniques adopted in this paper. In this regard, the method with the best prediction accuracy is the one that has a fit line equation of $y = x$, meaning that the line slope is equal to one and its intercept is equal to zero. Figure 10 is plotted for the test dataset. It can be comprehended from Figure 10, although the slope of trend lines of all studied methods is lower than 0.5, the accuracy of GEP is higher than SVR and SVR-FOA models.

![Figure 10](image)
Figure 10. The scatter plots of calculated RSL by HWD and estimated RSL by SVR, SVR-FOA, and GEP models for test data.

Figure 11 represents the Taylor diagram used in this study. Introduced by Taylor in 2001, the Taylor diagram is a mathematical diagram that graphically allows a comparison of several models of a system. In this diagram, there are three categories of contours [74]:

- **Blue contours**
  It shows the Pearson correlation coefficient.

- **Orange contours**
  It indicates the RMS error that is proportional to the distance from a green spot on the horizontal axis called observed.

- **Black contours**
  It indicates the standard deviation proportional to the radial distance from the center.

Figure 11. Taylor diagrams of estimated RSL for all models.
The presented Taylor Diagram shows the superior accuracy of the GEP model due to the lower distance of its correspondent point (the point with Magenta color) from the observed green point. Overall, by examining Figures 8–11, it can be concluded that the SVR technique offers an average accuracy for the purpose of this article. Using the FOA algorithm to select the basic parameters of this technique significantly enhanced the accuracy of this method. On the other hand, the GEP method provides a formula for RSL prediction. By re-examining Figures 8–11, it turned out that both SVR-FOA and GEP methods yielded desirable accuracy for RSL prediction. However, the accuracy of the GEP method was slightly higher than that of the SVR-FOA method.

5. Conclusions

Pavement management at both project and network levels are always associated with substantial costs. Due to the budget constraints inflicted on organizations in charge of PMS, optimizing pavement management costs is one of the priorities of any organization. RSL is a crucial factor for pavement management at the network level. The current procedure for determining RSL involves using FWD and GPR tests. These devices are not only costly but also interfere with the traffic flow and compromise the safety of road users. The aim subject of the study was to present a new approach for predicting the RSL of flexible pavement, which eliminated the drawbacks of current methods. Therefore, the proposed method can lead to lower costs, reduced time consumed, and also increase safety. The idea of estimating RSL has been followed by various researchers. The major differences between their studies were in the methodology and methods of analysis. After a review of previous studies on estimating the pavement RSL, we decided to use pavement surface distresses as a criterion for predicting RSL. Therefore, PCI pavement was employed as an input variable in modeling pavement RSL. Modeling was done with the help of ML techniques, which is the innovation of this research. The dataset utilized for modeling was selected from the Shahrood–Damghian highway in Iran. After selecting 105 pavement segments from the highway, PCI and RSL of all segments were determined. Modeling was conducted using GEP and SVR techniques after completing the dataset. The results of modeling with these techniques were evaluated based on four criteria include CC, SI, NSE, and WI, to determine the most appropriate technique for estimating pavement RSL. After exploring all four criteria, it was found that the GEP outcomes were far more accurate than the SVR. Then, to improve the accuracy of the SVR method, the FOA optimization algorithm was employed to add a third technique, as an innovative model (namely SVR-FOA), to the methods applied in this paper. Again, the four criteria CC, SI, NSE, and WI revealed a significant improvement in the accuracy of the SVR-FOA method compared to the SVR method, yet the GEP method still had the highest prediction precision. In sum, the findings of this paper suggested that the GEP outcomes were far more accurate than the SVR. Then, to improve the accuracy of the SVR method, the FOA optimization algorithm was employed to add a third technique, as an innovative model (namely SVR-FOA), to the methods applied in this paper. Again, the four criteria CC, SI, NSE, and WI revealed a significant improvement in the accuracy of the SVR-FOA method compared to the SVR method, yet the GEP method still had the highest prediction precision. In sum, the findings of this paper suggested that the GEP method (with values of 0.874, 0.598, 0.601, and 0.807 for the four criteria CC, SI, NSE, and WI, respectively) offered an alternative to current methods of predicting pavement RSL. Based on the desired accuracy and the cost-benefit analysis, transportation agencies can use the GEP method to optimize the model accuracy, robustness, and reliability considering either new projects or for calibrating the former models.

In general, this study highlights the usefulness of the use of artificial intelligence in transportation engineering. The results show that artificial intelligence techniques can be used as an optimization tool (at time and cost) in pavement management systems. In terms of pavement engineering, this study concluded that the RSL of pavement, as a fundamental factor in the planning of the pavement network maintenance, is a flexible factor and is not limited to current cost-intensive methods. With the help of similar methods with the proposed method in this paper, the RSL can be determined as soon as possible. This will ease the future planning of road maintenance activities.

Acknowledgments: This work has been supported by the project GINOP-2.3.4-15-2016-00003.

Funding: This work has been supported by the project GINOP-2.3.4-15-2016-00003

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

Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSL</td>
<td>Remaining Service Life</td>
</tr>
<tr>
<td>PCI</td>
<td>Pavement Condition Index</td>
</tr>
<tr>
<td>FWD</td>
<td>Falling Weight Deflectometer</td>
</tr>
<tr>
<td>GPR</td>
<td>Ground Penetrating Radar</td>
</tr>
<tr>
<td>SVR</td>
<td>Support Vector Regression</td>
</tr>
<tr>
<td>SVR-FOA</td>
<td>Support Vector Regression Optimized by Fruit Fly Optimization Algorithm</td>
</tr>
<tr>
<td>GEP</td>
<td>Gene Expression Programming</td>
</tr>
<tr>
<td>CC</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td>NSE</td>
<td>Nash–Sutcliffe Efficiency</td>
</tr>
<tr>
<td>SI</td>
<td>Scattered Index</td>
</tr>
<tr>
<td>WI</td>
<td>Willmott’s Index of agreement</td>
</tr>
<tr>
<td>PMSs</td>
<td>Pavement Management Systems</td>
</tr>
<tr>
<td>MR&amp;R</td>
<td>Maintenance, Rehabilitation, and Reconstruction</td>
</tr>
<tr>
<td>IRI</td>
<td>International Roughness Index</td>
</tr>
<tr>
<td>PSI</td>
<td>Present Serviceability Index</td>
</tr>
<tr>
<td>PSR</td>
<td>Present Serviceability Ratio</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>RBF</td>
<td>Radial Basis Function</td>
</tr>
<tr>
<td>GEP</td>
<td>Gene Expression Programming</td>
</tr>
<tr>
<td>RT</td>
<td>Regression Tree</td>
</tr>
<tr>
<td>RF</td>
<td>Random Forest</td>
</tr>
<tr>
<td>GP</td>
<td>Genetic Programming</td>
</tr>
<tr>
<td>AMS</td>
<td>Assessment Management System</td>
</tr>
<tr>
<td>DV</td>
<td>Deduct Value</td>
</tr>
<tr>
<td>CDV</td>
<td>Corrected Deduct Value</td>
</tr>
<tr>
<td>TDV</td>
<td>Sum of DVs</td>
</tr>
<tr>
<td>MnDOT</td>
<td>Minnesota Department of Transportation</td>
</tr>
<tr>
<td>MDOT</td>
<td>Michigan Department of Transportation</td>
</tr>
<tr>
<td>HWD</td>
<td>Heavy Falling Weight Deflectometer</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>ETs</td>
<td>Expression Trees</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>RRSE</td>
<td>Root Relative Squared Error</td>
</tr>
<tr>
<td>RSE</td>
<td>Relative Square Error</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>SRM</td>
<td>Structural Risk Minimization</td>
</tr>
<tr>
<td>ERM</td>
<td>Empirical Risk Minimization</td>
</tr>
<tr>
<td>FOA</td>
<td>Fruit Fly Optimization Algorithm</td>
</tr>
</tbody>
</table>

References

21. Wahyudi, W.; Sandra, P.A.; Mulyono, A.T. Analysis of Pavement Condition Index (PCI) and Solution Alternative of Pavement Damage Handling Due to Freight Transportation Overloading (Case Study: National Road Section West Sumatra Border-Jambi City). In Proceedings of the Eastern Asia Society for Transportation Studies, Lalkatta, India 2013.