Backpropagation Neural Network to estimate pavement performance: Dealing with Measurement Errors

Amin, S & Amador, L

Original citation & hyperlink:
Amin, S & Amador, L 2016, 'Backpropagation Neural Network to estimate pavement performance: Dealing with Measurement Errors' *Road Materials and Pavement Design*, vol 18, no. 5, pp. 1218-1238
https://dx.doi.org/10.1080/14680629.2016.1202129

DOI 10.1080/14680629.2016.1202129
ISSN 1468-0629
ESSN 1468-0629

Publisher: Taylor and Francis

This is an Accepted Manuscript of an article published by Taylor & Francis in Road Materials and Pavement Design on 1st July 2016, available online: http://www.tandfonline.com/10.1080/14680629.2016.1202129

Backpropagation Neural Network to estimate pavement performance: dealing with measurement errors

Md. Shohel Reza Amin

Research Assistant, Dept. of Building, Civil and Environmental Engineering,

Concordia University, 1455 de Maisonneuve Blvd. West, EV-6.139

Montreal, Quebec, H3G 1M8, Canada

Tel: +14389364119. Fax: +1-514-848-7965.

Email: md_amin@encs.concordia.ca

(Corresponding Author)

Luis E. Amador-Jiménez

Associate Professor, Dept. of Building, Civil and Environmental Engineering,

Concordia University, 1455 de Maisonneuve Blvd. West, EV-6.164

Montreal, Quebec, H3G 1M8, Canada

Tel: +1-514-848-2424 (ext. 5783). Fax: +1-514-848-7965.

Email: luis.amador@concordia.ca

Backpropagation Neural Network to estimate pavement performance: dealing with measurement errors

Abstract

The objective of this study is to apply the Backpropagation Neural Network (BPN) with Generalized Delta Rule (GDR) learning algorithm for reducing the measurement errors of pavement performance modeling. The Multi-Layer Perceptron (MLP) network and sigmoid activation function are applied to build the BPN network of Pavement Condition Index (*PCI*). Collector and arterial roads of both flexible and rigid pavements in Montreal City are taken as a case study. The input variables of PCI are Average Annual Daily Traffic (*AADT*), Equivalent Single Axle Loads (*ESALs*), Structural Number (*SN*), pavement's age, slab thickness and difference of *PCI* between current and preceding year (ΔPCI). The BPN networks estimates that the *PCI* has inverse relationships with *AADT*, *ESALs* and pavement's age. The *PCI* has positive relationships with these variables for roads that have recent treatment operations. The *PCI* has positive relationships with *SN* and slab thickness that imply the increase of pavement condition with increasing structural strength and slab thickness. The ΔPCI significantly influences the estimation of *PCI* values. The *AADT* and *ESALs* have considerable importance, however, pavement's age and structural characteristics of pavement have insignificant influence in determining the *PCI* values except in the case of flexible arterial roads.

Keywords

Pavement condition index; Backpropagation Neural network; measurement errors; annual average daily traffic; equivalent single axle loads; pavement's age; structural strength.

1. Introduction

An appropriate pavement performance curve is the fundamental component of pavement management system (PMS) and ensures the accuracy of pavement maintenance and rehabilitation (M&R) operations (Jansen and Schmidt, 1994; Johnson and Cation, 1992; Attoh-Okine, 1999). The pavement performance models help PMS to optimize M&R operations and to estimate the consequences of M&R operations on the future pavement condition during the life span of pavement (George, Rajagopal and Lim, 1989; Li, Haas and Xie, 1997). Early PMSs did not have performance curves rather they evaluated only the current pavement condition. The simplified performance curves were later introduced based on engineering opinions on the expected design life of different M&R operations (Kulkarni and Miller, 2002). The only predictive variable of these performance curves was pavement's age. The development of performance curve is explicitly complicated since the pavement performance depends on a large number of dynamic and static attributes (Amin, 2015).

There are two streams of pavement performance modeling such as deterministic and stochastic. The major differences between deterministic and stochastic performance models are model development concepts, modeling process or formulation and output format of the models (Li, Xie, and Haas, 1996; Amin, 2015). Deterministic models include primary response, structural performance, function performance and damage models for pavements (George et al., 1989; Amin, 2015). Different methods of deterministic models are mechanistic, mechanistic-empirical and regression models (Saleh, Mamlouk, and Owusu-Antwi, 2000; AASHTO, 1985; George et al., 1989; De Melo e Siva, Van Dam, Bulleit, and Ylitalo, 2000). Mechanistic models draw the relationship between response parameters such as stress, strain, and deflection (Li et al., 1996).

Mechanistic-empirical models draw the relationship between roughness, cracking, and traffic loading. Regression models draw the relationship between a performance (e.g. riding comfort index) and predictive parameters (e.g. pavement thickness, pavement material properties, traffic loading, and age) (Li et al., 1996). A large number of deterministic models are developed for regional or local PMSs such as traffic related, time related, interactive-time related and generalised models (Attoh-Okine, 1999).

Deterministic models cannot address some important issues such as (a) randomness of traffic loads and environmental conditions, (b) difficulties in quantifying the factors or parameters that substantially affect pavement deterioration, (c) measurement errors associated with pavement condition and (d) bias from subjective evaluations of pavement condition (Li et al., 1997; Amin, 2015). These constraints of deterministic models open the application of stochastic modeling.

Stochastic models recently have received considerable attentions from pavement engineers and researchers (Wang, Zaniewski, and Way, 1994; Karan, 1977). Typically, the Markov Decision Process (MDP) defines a stochastic model (Li et al., 1997). The Markov process predicts the 'after' condition of pavement knowing the 'before' condition (George et al., 1989). The main challenges of these stochastic models are to develop the Transition Probability Matrices (TPMs) and to obtain and process a large amount of measured performance data for all pavement categories in a road network (Li et al., 1997). However, the main drawbacks of MDP approach are (a) it does not accommodate budget constraints along with condition state and (b) pavement sections are grouped into a small number of roughly homogeneous families based on pavement or road or traffic characteristics (Liebman, 1985; Li, Cheetham, Zaghloul, Helali, and Bekheet, 2006). The MDP suggests that pavement sections should be categorized into small numbers of families to avoid dealing with large number of pavement families. Similarly, the optimization programming of M&R strategies are estimates for a group of pavement sections rather than for each road section under a given budget. The optimization programming of M&R strategies are calculated using the steady state probabilities of pavement condition. In reality, pavements under a given maintenance policy usually take many years to reach the steady state and the pavement proportion under a particular state is changing every year. The application of steady state probabilities in the optimization objective function does not fully reflect reality (Li et al., 2006).

2. Pavement Performance Models dealing with measurement errors

Pavement performance models are associated with data collection and measurement errors. Ben-Akiva et al. (1993) developed the latent performance approach dealing with forecasting uncertainties during condition data collection. A latent variable captures the ambiguity in measuring infrastructure condition (Durango-Cohen, 2007). This latent model suffers from computational limitations. Finding an optimal action for a given period requires estimating and assigning a probability to every possible outcome of data-collection process. Number of outcomes, probabilities and computational effort to obtain M&R policies increases exponentially with the number of distresses being measured (Durango-Cohen, 2007; Amin, 2015).

Durango-Cohen (2007) applied the Polynomial Linear Regression (PLR) model to define the dynamic system of infrastructure deterioration process. The PLR model includes condition data and a set of exogenous (deterministic and stochastic) inputs. Durango-Cohen's PLR model cannot define the proportion of errors contributed by each of the factors to the distress outcome (Amin, 2015).

Attoh-Okine (1994) proposed the Artificial Neural Network (ANN) for predicting the roughness progression in flexible pavements. However, some built-in functions of ANN such as learning rate and momentum term of ANN algorithm were not investigated properly. Inaccurate application of these built-in functions may affect the aptness of ANN (Attoh-Okine, 1999). Attoh-Okine (1999) analyzed the contribution of learning rate and momentum term in Back Propagation Neural (BPN) algorithm for the pavement performance prediction of Kansas pavement condition data during 1993. The BPN model estimates International Roughness Index (IRI) as a function of rutting, faulting distress, transverse cracking distress, block cracking and Equivalent Single Axle Loads (*ESALs*) (Attoh-Okine, 1999). Shekharan (1999) applied the partitioning of connection weights in ANN to estimate the relative contribution of structural number, age of pavement, and cumulative *ESALs* to the present serviceability rating (PSR) of pavement. The weights of output layer connection are partitioned into input node shares. The weights, along the paths from input to output nodes, indicate the relative predictive importance of input variables. These weights are used to partition the sum of effects on the output layer (Shekharan, 1999). However, Attoh-Okine (1999) and Shekharan (1999) models have not yet overcome the functional limitations of neural network algorithms (Amin, 2015).

3. Objective

This study applies the Backpropagation Neural Network (BPN) method with Generalized Delta Rule (GDR) learning algorithm to reduce the measurement error of the pavement performance modeling. Collector and arterial roads of Montreal City are taken as a case study.

4. Methodology

4.1. Data Collection

Data on pavement condition, age, traffic volume and road characteristics are collected from the Ville de Montréal. Pavement condition data in 2010 and 2009 are used in this study since the Ville de Montréal has complete pavement condition data only for these years.

The AASHTO Design Guide (AASHTO, 1993) terms Average Annual Daily Traffic (*AADT*) as the 80-KN Equivalent Single Axle Loads (*ESALs*) that are the total damage of road pavement caused by commercial vehicles. The *ESALs* are calculated based on number, type and distribution of commercial vehicles, road characteristics and truck growth factor on the road network of Montreal City. Data on type and distribution of trucks on the road network of Montreal City and annual truck growth rate (2 percent) are adopted from the report prepared by the Cement Association of Canada (Cement Association of Canada, 2012). Truck distribution and truckloads on the collector and arterial roads are shown in Table 1.

[Table 1]

Since data on the thickness of pavement's layers are not available from the Ville de Montréal, thickness data for different layers of Portland Cement Concrete (PCC) and Hot Mix Asphalt (HMA) pavements in Montreal City are also adopted from the report prepared by the Cement Association of Canada (2012). The Structural Number (*SN*) of the flexible pavements is calculated from the thickness of pavement layers and climate condition of Montreal City.

This study categorizes the road segments into four categories based on pavement types (e.g. flexible and rigid) and road hierarchies (e.g. arterial and collector). The predictive variable for all types of pavement is Pavement Condition Index (*PCI*). The input variables for the flexible pavements are *AADT*, *ESALs*, *SN*, pavement's age (*N*) and difference of *PCI* between current and preceding year ($\Delta PCI = PCI_{2009} - PCI_{2010}$). The ΔPCI helps to track the condition deterioration or application of treatment operations at the preceding year. The input variables for the rigid pavements are *AADT*, *ESALs*, slab thickness (*T*), *N* and ΔPCI. Since *AADT* and *ESALs* are log-linearly related to *PCI*, Log_{10} (*AADT*) and Log_{10} (*ESALs*) are taken as input variables of *PCI*.

4.2. Learning Process in the Backpropagation neural network

The fundamental concept of BPN network for a two-phase propagate-adapt cycle is that input variables are applied as a stimulus to the input layer of network units that are propagated through each upper layer until an output is generated. This estimated output are compared with the desired output to estimate the error for each output unit. These errors are transferred backward from the output layer to each unit in the intermediate layer that contributes directly to the output. Each unit in the intermediate layer receives only a portion of the total error signal based roughly on the relative contribution to the original output. This process repeats layer-by-layer until each node receives an error representing its relative contribution to the total error. Based on the error received, connection weights are updated by each unit to cause the network to converge toward a state allowing all the training patterns to be encoded (Freeman and Skapura, 1991). Diagrams of BPN networks for flexible and rigid pavements are shown in Figure 1 and 2 respectively.

[Figure 1]

[Figure 2]

This study applies a GDR learning algorithm of BPN network. The learning process of BPN network for pavement performance modeling is described in this section. Let assume that we have a set of P vector-pairs in the training set $\{(x_1, y_1), (x_2, y_2)\ldots (x_p, y_p)\}$ and the functional mapping is $y = \phi(x): x \in R^N, y \in R^M$. The processing function is $\{(x_1, d_1), (x_2, d_2)\ldots (x_p, d_p)\}$ with input vectors ($\mathbf{x}_k$) and desired output value (d_k). The mean square error (ε_k^2) is defined by Equation 1 (Freeman and Skapura, 1991).

$$\varepsilon_k^2 = \theta_k = (d_k - y_k)^2 = (d_k - \mathbf{w}^t\mathbf{X}_\mathrm{N})^2 \qquad \text{where } y = \mathbf{w}^t\mathbf{X} \tag{1}$$

The weight vector at time t is $\mathbf{w}^t$. Since the weight vector is an explicit function of iteration (R), the initial weight vector is denoted $\mathbf{w}(0)$ and the weight vector at iteration R is $\mathbf{w}(R)$. At each step, the next weight vector is calculated following Equation 2 (Freeman and Skapura, 1991).

$$\mathbf{w}(R+1) = \mathbf{w}(R) + \Delta\mathbf{w}(R) = \mathbf{w}(R) - \mu\nabla\theta_k\big(\mathbf{w}(R)\big) = \mathbf{w}(R) + 2\mu\varepsilon_N\mathbf{X}_N \;\; \forall\, \nabla\theta(\mathbf{w}(R)) \approx \nabla\theta(\mathbf{w}) \tag{2}$$

Equation 2 is the Least Mean Square (LMS) algorithm, where $\Delta\mathbf{w}(R)$ is the change in weight vector **(w)** at the R^{th} iteration, and μ is the constant of negative gradient of the error surface. The error surface is either hyperbolic tangent or sigmoid learning function. The constant variable (μ) determines the stability and speed of convergence of the weight vector toward the minimum error value (Freeman and Skapura, 1991).

The input layer distributes the values to the hidden or intermediate layer units. Equation 3 defines the output (net_{pj}) of input node (I_{pj}) assuming that the activation of input node is equal to the net input. Similarly, Equation 4 defines the output (net_{pk}) of output node (O_{pk}) (Freeman and Skapura, 1991).

$$I_{pj} = f_j\left(net_{pj}\right) \quad net_{pj} = \sum_{i=1}^{N} w_{ji}x_{pi} + \theta_j \tag{3}$$

$$O_{pk} = f_k\left(net_{pk}\right) \quad \forall\, net_{pk} = \sum_{j=1}^{L} w_{kj}I_{pj} + \theta_k \tag{4}$$

Where w_{ji} is the weight on the connection from i^{th} input unit to j^{th} hidden unit, w_{kj} is the weight on the connection from j^{th} hidden unit to p^{th} output unit, and θ_j and θ_k are errors at intermediate and output layers respectively. The weight is determined by taking an initial set of weight values representing a first guess as the proper weight for the problem. The output values are calculated applying the input vector and initial weights. The calculated output is compared with the correct output and a measure of the error is determined. The amount of change in each weight is determined. The iterations with all training vectors are repeated until the error in all vectors of training set is reduced to an acceptable value (Freeman and Skapura, 1991).

Equations 3 and 4 define the output of input and output nodes, respectively. In reality, there are multiple units in a layer. A single error value (θ_k) is not suffice for BPN network. The sum of error squares for all output units is shown in Equation 5 (Freeman and Skapura, 1991).

$$\theta_{pk} = \frac{1}{2}\sum_{k=1}^{M} \varepsilon_{pk}^2 = \frac{1}{2}\sum_{k=1}^{M}(y_{pk} - O_{pk})^2$$

$$\Delta_p\theta_p(\mathbf{w}) = \frac{\partial(\theta_p)}{\partial w_{kj}} = -(y_{pk} - O_{pk})\frac{\partial}{\partial w_{kj}}(O_{pk}) = -(y_{pk} - O_{pk})\frac{\partial f_k}{\partial(net_{pk})}\frac{\partial(net_{pk})}{\partial w_{kj}} \tag{5}$$

Change in weight of output layer is expressed in Equation 6 by combining Equations 3, 4 and 5 (Freeman and Skapura, 1991).

$$\frac{\partial(\theta_p)}{\partial w_{kj}} = -(y_{pk} - O_{pk})\frac{\partial f_k}{\partial(net_{pk})}\frac{\partial}{\partial w_{kj}}(\sum_{j=1}^{L} w_{kj}I_{pj} + \theta_k) = -(y_{pk} - O_{pk})f_k'(net_{pk})I_{pj} \tag{6}$$

Where $f_k'(net_{pk})$ is the differentiation of Equation 4. This differentiation eliminates the possibility of using a linear threshold unit, since the output function for such a unit is not differentiable at the threshold value. Equation 7 estimates the weights on the output layer following Equations 2 and 6 (Freeman and Skapura, 1991).

$$w_{kj}(R+1) = w_{kj}(R) + \tau(y_{pk} - O_{pk})f_k'(net_{pk})I_{pj} \tag{7}$$

Where τ is a constant and learning-rate parameter. There are two forms of activation functions such as hyperbolic tangent $\left[f_k\left(net_{jk}\right) = \tan(net_{jk}) = (e^{net_{jk}} - e^{-net_{jk}})/(e^{net_{jk}} + e^{-net_{jk}})\right]$ and sigmoid or logistic function$\left[f_k\left(net_{jk}\right) = (1 + e^{-net_{jk}})^{-1}\right]$. The sigmoid or logistic function is for output units in a range of (0, 1) and the hyperbolic tangent function is for output units in a range of (-1, 1). Since the output of this model (e.g. pavement condition index) is positive value, sigmoid or logistic function is applied and can be expressed by Equation 8 (Freeman and Skapura, 1991).

$$w_{kj}(t+1) = w_{kj}(t) + \tau\left(y_{pk} - O_{pk}\right)O_{pk}\left(1 - O_{pk}\right)I_{pj} = w_{kj}(t) + \tau\delta_{pk}I_{pj} \tag{8}$$

The errors, estimated from the difference between calculated and desired output, are transferred backward from the output layer to each unit in the intermediate layer. Each unit in the intermediate layer receives only a portion of the total error based roughly on the relative contribution the unit made to the original output. This process repeats layer-by-layer until each node in the network has received an error that represents its relative contribution to the total error. The connection weights are updated based on the error received by each unit. Reconsidering Equations 4, 5, and 8 for Backpropagation algorithm, Equation 9 expresses the change of weights in hidden layer (Freeman and Skapura, 1991).

$$\theta_p = \frac{1}{2}\sum_{k=1}^{M}(y_{pk} - O_{pk})^2 = \frac{1}{2}\sum_{k=1}^{M}(y_{pk} - f_k(net_{pk}))^2 = \frac{1}{2}\sum_{k=1}^{M}(y_{pk} - f_k(\sum_{j=1}^{L} w_{kj}I_{pj} + \theta_k))^2$$

$$\frac{\partial \theta_p}{\partial w_{ji}} = -\sum_{k=1}^{M} (y_{pk} - O_{pk}) \frac{\partial O_{pk}}{\partial w_{ji}} = -\sum_{k=1}^{M} (y_{pk} - O_{pk}) \frac{\partial O_{pk}}{\partial (net_{pk})} \frac{\partial (net_{pk})}{\partial I_{pj}} \frac{\partial I_{pj}}{\partial (net_{pj})} \frac{\partial (net_{pj})}{\partial w_{ji}}$$

$$\frac{\partial \theta_p}{\partial w_{ji}} = -\sum_{k=1}^{M} (y_{pk} - O_{pk})\, f_k'(net_{pk}) w_{kj} f_j'(net_{pj}) x_{pi}$$

$$\Delta_p w_{ji} = \frac{\partial \theta_p}{\partial w_{ji}} = \tau f_j'(net_{pj}) x_{pi} \sum_{k=1}^{M} (y_{pk} - O_{pk})\, f_k'(net_{pk}) w_{kj} = \tau f_j'(net_{pj}) x_{pi} \sum_{k=1}^{M} \partial_{pk}\, w_{kj} \tag{9}$$

Equation 9 explains that each weight update in hidden layer depends on the error terms (∂_{pk}) in the output layer. The BPN network defines hidden layer error as $\delta_{pj} = f_j'(net_{pj}) \sum_{k=1}^{M} \partial_{pk}\, w_{kj}$ to update weight equations analogous to those for the output layer (Equation 10). Equations 8 and 10 have the same form of delta rule (Freeman and Skapura, 1991).

$$w_{ji}(t+1) = w_{ji}(t) + \tau \delta x_{pi} \tag{10}$$

5. Data analysis

This study partitions the dataset into training (60 percent), testing (30 percent), and validation (10 percent) data to estimate the BPN models for all road categories. The BPN network uses the training and testing data to train the network and to track errors during training in order to prevent overtraining respectively. The BPN algorithm finally estimates the predictive ability of the BPN network by using the validation data.

5.1 Back Propagation Neural Network Performance

This study evaluates the performance of BPN models to determine the statistically significance of BPN models. The Sum of Squares Error (SSE) and Relative Error (RE) defines the fitness of BPN models. The SSE is the cross-entropy error when the sigmoid activation function is applied to the output layer. The BPN model minimizes the SSE function during training. The RE is the percentage of incorrect predictions and is associated with dependent variable. In other words, the RE is the ratio of SSE for dependent variable and 'null model'.

Estimation of BPN models has insignificant difference between values implied by estimators and the true values of the output particularly for training data (Table 2). Testing data, used to track errors during training, also contain minor expected value of squared error loss (Table 2). Insignificant errors for validation data explain the accurate prediction ability of the constructed BPN networks (Table 2).

[Table 2]

The predicted-by-observed and residual-by-observed scatterplot are plotted to understand the relationship between predicted and observed data and residual and observed data respectively. The predicted and observed data of PCI for the combined training and testing samples are plotted on the y-axis and x-axis of the predicted-by-observed scatterplot respectively (Figure 3). Ideally, values should lie roughly along a 45-degree line starting at the origin. The scatterplots for flexible and rigid pavements of arterial and collector roads show that the BPN models do a reasonably good job of predicting PCI (Figure 3).

[Figure 3]

The residual and predicted values of *PCI* are also plotted on the y-axis and x-axis of the residual-by-observed scatterplot respectively. Figure 4(c) and 4(d) show that the residual-by-observed scatterplot for flexible and rigid collector roads are well-behaved and fit scatterplots. In case of rigid arterial roads, the residuals roughly form horizontal band and bounce randomly around the '0' line, however, there are few outliers (Figure 4(b)). These outliers do not have significant influence to estimate the BPN network for *PCI* values. The scattered distribution of residuals vs. predicted values of *PCI* questions the statistical significance or fitness of BPN network for flexible arterial roads (Figure 4(b)).

[Figure 4]

5.2 Parameter Estimation of Input Variables

The predictive variables are initially applied as stimulus to the input layer of network units that is propagated to the hidden (intermediate) layers in the BPN network. This study applies the Multi-Layer Perceptron (MLP) network that is a function of predictors minimizing the prediction error of outputs. The MLP procedure computes the minimum and maximum values of the range and find the best number of hidden layers within the range (IBM, 2010). The MLP estimates the number of hidden layers based on the minimum error in the testing data and the smallest Bayesian information criterion (BIC) in the training data (IBM, 2010). The MLP estimates that the best number of hidden layers is two. In the first hidden layer of network, the training and testing data are distributed into three sub-layers H (1:1), H (1:2) and H (1:3). The sigmoid activation function is used for the hidden layers so that the activation of the hidden unit is a Gaussian ‘bump’ as a function of input units (IBM, 2010).

In reality, the *PCI* has inverse relationships with AADT, *ESALs* and pavement’s age for both flexible and rigid pavements. Pavement condition deteriorates with increasing traffic volume, axle loads and pavement’s age. For the training and testing data of flexible arterial roads in sub-layers H (1:1) and H (1:2), the *PCI* has inverse relationships with *AADT*, *ESALs* and pavement’s age. For example, a one-unit increase in $\log_{10}$ (*AADT*) will produce an expected decrease in PCI of 0.086 and 0.249 in the hidden sub-layers H (1:1) and H (1:2) respectively (Table 3). Similarly, in the same sub-layers, a one-unit increase in $\log_{10}$ (*ESALs*) and pavement’s age will produce an expected decrease in *PCI* of 0.077 and 0.325, and 2.765 and 1.207 respectively (Table 3). In contrary, the *PCI* has positive relationships with *AADT*, *ESALs* and pavement’s age in the H (1:3) sub-layer of BPN network for the flexible arterial roads. A one-unit increase in $\log_{10}$ (*AADT*), $\log_{10}$ (*ESALs*) and pavement’s age will produce an expected increase in *PCI* of 0.069, 0.005 and 0.415 respectively (Table 3). This may be because of the inclusion of training and testing data in this sub-layer that have recent treatment operations. The *PCI* has increased for treatment operations instead of high *AADT*, *ESALs* and pavement’s age (Figure 5). This assumption is strongly supported by the negative value of ΔPCI in the H (1:3) sub-layer of BPN network for flexible arterial roads (Table 3). A one-unit increase in ΔPCI will produce an expected decrease in *PCI* of 1.031 in H (1:3) sub-layer, however, will increase 3.877 and 1.576 unit of *PCI* in sub-layers H (1:1) and H (1:2) respectively (Table 3). The positive relationship between *SN* and *PCI* explains that better structural strength of pavement increases the pavement condition. A one-unit increase in *SN* will produce an expected increase in *PCI* of 0.020, 0.052 and 0.622 in H (1:1), H (1:2) and H (1:3) sub-layers respectively (Table 3).

For flexible pavement of collector roads, a one-unit increase in ΔPCI will produce an expected decrease in *PCI* of 0.889 in H (1:2) sub-layer, however, will increase 1.025 and 0.838 unit of *PCI* in sub-layers H (1:1) and H (1:3) respectively (Table 3). A one-unit increase in $\log_{10}$ (*AADT*) will increase 0.253 unit of *PCI* in H (1:2) sub-layer and decrease 0.423 and 0.265 unit of *PCI* in sub-layers H (1:1) and H (1:3) respectively (Table 3). Similarly, a one-unit increase in $\log_{10}$ (*ESALs*) will increase 0.209 unit of *PCI* in H (1:2) sub-layer and decrease 0.176 and 0.201 unit of *PCI* in sub-layers H (1:1) and H (1:3) respectively (Table 3). The relationship between *PCI* and pavement's age shows that a one-unit increase in pavement's age will produce an expected decrease in *PCI* of 0.092, 0.021 and 0.017 in the sub-layers H (1:1), H (1:2) and H (1:3) respectively (Table 3). The *SN* has positive relationship with *PCI* for flexible collector roads. A one-unit increase in the *SN* will produce an expected increase in *PCI* of 0.111, 0.368 and 0.946 in sub-layers H (1:1), H (1:2) and H (1:3) respectively (Table 3).

[Table 3]

[Figure 5]

For rigid pavements of arterial roads, a one-unit increase in ΔPCI will produce an expected decrease in *PCI* of 0.34 in H (1:1) sub-layer, however, will increase 1.288 and 0.971 unit of *PCI* in sub-layers H (1:2) and H (1:3) respectively (Table 4). A one-unit increase in $\log_{10}$ (*AADT*) will increase 0.661 unit of *PCI* in H (1:1) sub-layer and decrease 0.059 and 0.097 unit of *PCI* in sub-layers H (1:2) and H (1:3) respectively (Table 4). Similarly, a one-unit increase in $\log_{10}$ (*ESALs*) will increase 0.348 unit of *PCI* in H (1:1) sub-layer and decrease 0.121 and 0.059 unit of *PCI* in sub-layers H (1:2) and H (1:3) respectively (Table 4). The relationship between *PCI* and pavement's age shows that a one-unit increase in pavement's age will produce an expected decrease in *PCI* of 0.286, 1.85 and 1.268 in sub-layers H (1:1), H (1:2) and H (1:3) respectively (Table 4). The slab thickness (mm) of rigid pavement has positive relationship with the *PCI*. A one-unit increase in slab thickness will produce an expected increase in *PCI* of 0.44, 0.282 and 0.745 in sub-layers H (1:1), H (1:2) and H (1:3) respectively (Table 4).

For rigid pavements of arterial roads, a one-unit increase in ΔPCI will produce an expected decrease in *PCI* of 0.496 in H (1:1) sub-layer, however, will increase 0.511 and 0.522 unit of *PCI* in sub-layers H (1:2) and H (1:3) respectively (Table 4). A one-unit increase in $\log_{10}$ (*AADT*) will increase 0.058 unit of *PCI* in H (1:1) sub-layer and decrease 0.046 and 0.241 unit of *PCI* in sub-layers H (1:2) and H (1:3) respectively (Table 4). Similarly, a one-unit increase in $\log_{10}$ (*ESALs*) will increase 0.296 unit of *PCI* in H (1:1) sub-layer and decrease 0.06 and 0.546 unit of *PCI* in sub-layers H (1:2) and H (1:3) respectively (Table 4). The relationship between *PCI* and pavement's age shows that a one-unit increase in pavement's age will produce an expected decrease in *PCI* of 0.431, 0.266 and 0.323 in sub-layers H (1:1), H (1:2) and H (1:3) respectively (Table 4). Similar to the rigid pavements of arterial roads, the slab thickness (mm) has positive relationship with *PCI* in the rigid pavements of collector roads. A one-unit increase in slab thickness will produce an expected increase in *PCI* of 0.327, 0.157 and 0.23 in sub-layers H (1:1), H (1:2) and H (1:3) respectively (Table 4).

[Table 4]

Each unit of the second hidden layer is a function of the units in the first hidden layer, and each response is a function of the units in the second hidden layer. For example, H (1:1), H (1:2) and H (1:3) sub-layers of hidden layer 1 contribute -3.553, -1.72 and -0.712 to H (2:1) sub-layer of hidden layer 2 for the training and testing data of flexible arterial roads respectively (Table 3). The H (1:1), H (1:2) and H (1:3) sub-layers of hidden layer 1 contribute -2.520, -1.303 and -0.017 to the H (2:2) sub-layer of hidden layer 2 respectively (Table 3). The H (1:1), H (1:2) and H (1:3) sub-layers of hidden layer 1 contribute 1.367, 1.043 and -2.341 to the H (2:1) sub-layer; and contribute -0.676, -.879 and 1.579 to the H (2:2) sub-layer of hidden layer 2 for the training and testing data of flexible collector roads respectively (Table 3).

For the training and testing data of arterial rigid roads, the H (1:1), H (1:2) and H (1:3) sub-layers of hidden layer 1 contribute 0.710, -0.930 and 1.380 to the H (2:1) sub-layer; and contribute 0.685, 1.442 and -0.565 to the H (2:2) sub-layer of hidden layer 2 respectively (Table 4). The H (1:1), H (1:2) and H (1:3) sub-layers of hidden layer 1 contribute 1.686, -1.685 and -1.346 to the H (2:1) sub-layer; and contribute 2.079, -1.880 and -1.210 to the H (2:2) sub-layer of hidden layer 2 for the training and testing data of rigid collector roads respectively (Table 4).

For the output layer, the activation function is the sigmoid function. The H (2:1) and H (2:2) sub-layers have almost equal weight to output unit in the flexible arterial roads (e.g. 4.151 and 4.034) and rigid collector roads (e.g. 3.292 and 3.621) (Table 3 and 4). However, H (2:1) sub-layer has approximately double weight to output units comparing to H (2:2) layer in the flexible collector roads (Table 3). The H (2:2) layer has approximately triple weight to output units comparing to H (2:1) layer in the rigid arterial roads (Table 4).

The BPN network performs the sensitivity analyses to compute the importance of input variables in determining the *PCI* based on the combined training and testing samples. The importance of an input variable is a measure of how much the *PCI* value changes for different values of an input variable. The *PCI* values for flexible arterial roads are predominantly determined by ΔPCI (36.4 percent) and pavement's age (36.3 percent) (Table 5). Other input variable such as $\log_{10}$ (*AADT*), $\log_{10}$ (*ESALs*) and *SN* have 13.8 percent, 12 percent and 1.5 percent contributions in determining the *PCI* value (Table 5). The ΔPCI also significantly influence the *PCI* values of rigid arterial, flexible collector and rigid collector roads by 33.1 percent, 33 percent and 32.9 percent respectively (Table 5). However, pavement's age does not significantly influence the *PCI* values of rigid arterial (16.2 percent), flexible collector (12.3 percent) and rigid collector (21.1 percent) roads (Table 5).

[Table 5]

The $\log_{10}$ (*AADT*) and $\log_{10}$ (*ESALs*) have considerable importance to estimate the *PCI* values in BPN models for rigid arterial, flexible collector and rigid collector roads. For example, the $\log_{10}$ (*AADT*) has 23 percent, 22.6 percent and 20.1 percent importance to estimate *PCI* values of rigid arterial, flexible collector and rigid collector roads respectively (Table 5). The $\log_{10}$ (*ESALs*) variable contributes 19.4 percent, 22.1 percent and 24.8 percent of *PCI* values for rigid arterial, collector flexible and collector rigid roads respectively (Table 5). The structural characteristics of pavement, *SN* and slab thickness, for flexible and rigid pavements do not have significant influence in determining the *PCI* values respectively (Table 5). The reason is that the categorical values of thickness of pavement's layers for broader categories of *AADT* are applied in this study both for flexible and rigid pavements from the report prepared by the Cement Association of Canada (2012). There is a strong potential that the BPN models might estimate the significant or considerable influences of *SN* and slab thickness on the *PCI* for flexible and rigid pavements respectively, if the actual data on thickness of pavement's layers for each road segment can be accommodated into the BPN network.

The *PCI* values for each segment of different road categories of Montreal are estimated by applying the developed BPN models during the period of 2009-2058 (Figure 6-9).

[Figure 6]

[Figure 7]

[Figure 8]

[Figure 9]

The BPN methods with GDR learning algorithm overcomes the prevailing functional errors of pavement performance modeling such as stability and speed of convergence of the weight vector toward the minimum error value. The uncertainty is not only associated with the statistical analysis but also with uncertainty of the traffic data collection process. Future studies should analyze the reliability of the traffic data (e.g. *AADT* and *ESALs*) to overcome these uncertainties. In addition, a complete historic record on the pavement condition, pavements' structural attributes, pavement age, traffic volume, and road characteristics will enable to estimate more accurate pavement performance model by applying BPN networks.

6. Conclusion

The pavement performance models optimize treatment operations and estimate the consequences of treatment operations on the future pavement condition during the life span of pavement. The prevailing deterministic and stochastic models cannot overcome some key drawbacks such as measurement errors. The objective of this study is to apply the Backpropagation Neural Network (BPN) method with Generalized Delta Rule (GDR) learning algorithm to reduce the measurement errors of the pavement performance modeling. The Multi-Layer Perceptron (MLP) network and sigmoid activation function are applied to build the BPN networks of pavement condition index (*PCI*). Collector and arterial roads of both flexible and rigid pavements in Montreal City are taken as a case study.

The BPN networks estimates that the *PCI* has inverse relationships with *AADT*, *ESALs* and pavement's age for flexible and rigid pavements of arterial and collector roads. However, the positive relationships are observed for roads that have recent treatments. The *PCI* has positive relationships with *SN* and slab thickness that imply that the increase of pavement condition depends on the increase of structural strength and slab thickness.

Difference between the *PCI* of consecutive years (ΔPCI) significantly influence the PCI values of all category roads. The *AADT* and *ESALs* have considerable importance to estimate the *PCI* values in BPN models. However, pavement's age does not significantly influence the *PCI* values except in the case of flexible arterial roads (36.3 percent). The structural characteristics of flexible and rigid pavements do not have significant influence in determining the *PCI* values. This is because of considering the categorical values of thickness of pavement layers. There is a strong potential that the BPN models might estimate the significant or considerable influences of *SN* and slab thickness on the *PCI* for flexible and rigid pavements respectively, if the actual data on layer thickness can be included into the BPN network. Future studies should analyze the reliability of the traffic data (e.g. *AADT* and *ESALs*) to overcome these uncertainties. In addition, a complete historic record on the pavement condition, pavements' structural attributes, pavement age, traffic volume, and road characteristics will enable to estimate more accurate pavement performance model by applying BPN networks.

References

AASHTO, (1985). *AASHTO Guide for the Design of Pavement Structures (Appendix EE)*. Washington, DC: American Association of State Highway and Transportation Officials.

AASHTO, (1993). *Guide for the Design of Pavement Structures*. Washington, DC: American Association of State Highway and Transportation Officials.

Amin, M.S.R. (2015). The Pavement Performance Modeling: Deterministic vs Stochastic Approaches. In S. Kadry and A. El Hami (eds.) *Numerical Methods for Multiscale and Multiphysics in Reliability and Safety* (pp. 179-196). Switzerland: Springer.

Attoh-Okine, N. O. (1994). *Predicting Roughness Progression in Flexible Pavements Using Artificial Neural Networks*. Paper presented at the Third International Conference on Managing Pavements, Transportation Research Board, National Research Council, San Antonio, Texas, 55-62.

Attoh-Okine, N. O. (1999). Analysis of Learning Rate and Momentum Term in Backpropagation Neural network Algorithm Trained to Predict Pavement Performance. *Advances in Engineering Software*, 30, 291-302.

Ben-Akiva, M., Humplick, F., Madanat, S., and Ramaswamy, R. (1993). Infrastructure Management under Uncertainty: The Latent Performance Approach. *Journal of Transportation Engineering*, 119(1), 43-58.

Cement Association of Canada, (2012). *Methodology for the Development of Equivalent Pavement Structural Design Matrix for Municipal Roadways- Montréal and Québec City: Including Maintenance & Rehabilitation Schedules and Life Cycle Cost Analysis*. Final Report. Ottawa: Applied Research Associates Inc.

de Melo e Siva, F., Van Dam, T. J., Bulleit, W. M., and Ylitalo, R. (2000). Proposed Pavement Performance Models for Local Government Agencies in Michigan. *Transportation Research Record: Journal of the Transportation Research Board*, 1699, 81-86.

Durango-Cohen, P. L. (2007). An Optimal Estimation and Control Framework for the Management of Infrastructure Facilities. In C. Karlsson, W. P. Anderson, B. Johansson, K. Kobayashi (eds) *The Management and Measurement of Infrastructure: Performance, Efficiency and Innovation* (pp. 177-197). Cheltenham, UK: Edward Elgar.

Freeman, J. A., and Skapura, D. M. (1991). *Neural Networks. Algorithms, Applications, and Programming Techniques*. Menlo Park, California: Addison-Wesley Publishing Company.

George, K. P, Rajagopal, A. S., and Lim, L. K. (1989). Models for Predicting Pavement Deterioration. *Transportation Research Record: Journal of the Transportation Research Board*, 1215, 1-7.

IBM, (2010). *IBM SPSS Neural Networks 19*. <http://www.sussex.ac.uk/its/pdfs/SPSS_Neural_Network_19.pdf> (January 10, 2012).

Jansen, J. M., and Schmidt, B. (1994). *Performance Models and Prediction of Increase Overlay need in Danish state highway pavement management system, BELMA*. Paper presented at the Third International Conference on Managing Pavements, San Antonio, TX, 74-84.

Johnson, K. D., and Cation, K. A. (1992). Performance Prediction Development Using Three Indexes for North Dakota Pavement Management System. *Transportation Research Record: Journal of the Transportation Research Board*, 1344, 22-30.

Karan, M. (1977). *Municipal Pavement Management System.* Dissertation, Department of Civil Engineering, University of Waterloo.

Kulkarni, R. B., and Miller, R. W. (2002). Pavement Management Systems Past, Present, and Future. *Transportation Research Record: Journal of the Transportation Research Board*, 1853, 65-71.

Liebman, J. (1985). *Optimization Tools for Pavement Management*. Paper presented at the North American Pavement Management Conference, Ontario Ministry of Transportation and Communication, U.S. Federal Highway Administration, Washington D.C., 6.6-6.15.

Li, N., Xie, W.-C., and Haas, R. (1996). Reliability-Based Processing of Markov Chains for Modeling Pavement Network Deterioration. *Transportation Research Record: Journal of the Transportation Research Board*, 1524, 203–213.

Li, N., Haas, R., and Xie, W.-C. (1997). Investigation of Relationship between Deterministic and Probabilistic Prediction Models in Pavement Management. *Transportation Research Record: Journal of the Transportation Research Board*, 1592, 70-79.

Li, Y., Cheetham, A., Zaghloul, S., Helali, K., and Bekheet, W. (2006). Enhancement of Arizona Pavement Management System for Construction and Maintenance Activities. *Transportation Research Record: Journal of the Transportation Research Board*, 1974, 26–36.

Saleh, M. F., Mamlouk, M. S., and Owusu-Antwi, E. B. (2000). Mechanistic Roughness Model based on Vehicle-Pavement Interaction. *Transportation Research Record: Journal of the Transportation Research Board*, 1699, 114-120.

Shekharan, A. R. (1999). Assessment of Relative Contribution of Input Variables to Pavement Performance Prediction by Artificial Neural Networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1655, 35-41.

Wang, K. C. P., Zaniewski. J., and Way, G. (1994). Probabilistic Behavior of Pavements. *Journal of Transportation Engineering*, 120(3), 358–375.