Predicting Bridge Conditions in Ontario: A Case Study

M. Taghaddos[a] and Y. Mohamed[a]

[a] Department of Civil & Environmental Engineering, University of Alberta, Canada
E-mail: taghaddo@ualberta.ca, yasser.mohamed@ualberta.ca

Abstract –
Maintenance and repair of bridges represent significant costs in provincial and municipal government budgets. Prediction of bridge conditions can help managers in annual cost estimating and budget allocation. To assess Bridge Condition Index (BCI), each bridge component must be inspected every two years, tested if it is required, and rated. Bridge condition can be affected over time by different attributes such as material, structure, location, and use. This paper presents a study conducted to model and predict BCI based on a historical dataset of 2803 bridges in Ontario from 2000 to 2014. The paper describes the work related to data collection, cleaning and transformation. In addition, a comparison of the cross-validation performance of alternative BCI prediction models is presented and discussed.

Keywords –
Bridge maintenance; Bridge Condition Index (BCI); Prediction model; Cluster; Data analysis

1 Introduction

Bridges are essential infrastructures which are commonly used in transportation networks, and it is vital for them to function up to the acceptable level during their service life. However, to keep their functionality acceptable maintaining them and the corresponding cost is unavoidable. In the United States, the annual cost of rehabilitation is $7 billion [3]. To optimize the functionality of bridge infrastructure, continuous bridge access must be balanced with repair and maintenance costs as well as safety consideration [1]. Early identification of deterioration and following repairs enable decision makers to expedite the maintenance significantly with a lower cost and minimum disruption [4].

Continues bridge assessment can improve the maintenance of structure since it can accurately determine the structure condition, provide condition factor for load rating calculation, find the main reason of any deterioration and finally, determine appropriate rehabilitation process and corresponding budget [4]. Current tests for identifying bridge conditions are costly and time-consuming. Hence, a cost and time effective model for predicting bridge condition is beneficial from the budget allocation perspective.

The objective of this paper is to develop a framework to forecast the bridge condition for the future using historical data. Various prediction models are implemented and compared, and the best one is selected.

The advantages of the proposed approach are:

1. Enables managers to prioritize the bridges based on the level of urgent action required in a time and cost effective manner before formal assessments are conducted.
2. Helps to prioritize the expensive and invasive tests for the upcoming years.
3. Benefit ministries to proportionally allocate budget for repair and maintenance in advance.
4. Enable managers to evaluate different maintenance- scenarios over time to choose the best strategy.

2 Bridge Condition Index

One tool that can assist in the management of highway structure is Bridge Condition Index (BCI). BCI was developed based on two factors of the extent/severity factor Sf and element factor Ef, which both can have values between 1 to 10. Development of this index is entitled to High-Point Rendel and Taywood Engineering [5]. This index serves as a principle of resource allocation within a network [6] and can be helpful in various purposes naming [5]:

- It is an indication of a change in condition state over a period of time for the entire or part of the bridge
- By considering the entire bridge over a long period of time, the level of provided funding can be assessed to recognize its adequacy to keep the stock in a steady state
- By considering applied funding, BCI and type of material, the best and the most economic material of construction in the long run, can be assessed in each area
- By considering the level of funding and the BCI, it might be possible to achieve an

indication for the performance of agents who are responsible to take corrective actions.

Aside from these advantages, BCI has its own constraints in its applicability, such as [5]:

- BCI is not an indicator for the functionality of a bridge from a traffic point of view
- Safety is not a primary concern in BCI

Finally, since BCI is widely used in different countries including Canada, this paper considers BCI as a representative factor for the bridge conditions.

3 Case Description

The Ministry of Ontario adopted the BCI measurement system to evaluate the conditions of the bridge [7]. They inspected and assessed all bridges (2803 bridges) in Ontario from 2000 to 2014. In each inspection, experienced engineers and inspectors followed Ontario's Structure Inspection Manual (OSIM) which provides inspection procedures in great details. To identify maintenance procedure, they assess each bridge component comprehensively, such as barriers, sidewalks, deck asphalt, expansion joints, beams, pier cap, pier column, bearings, soffits, wingwall, and abutment. They performed a detailed visual inspection by checking the general condition of bridges, assessing components of each bridge, looking for any potential problem and reporting any safety issue (if there is any). Four common bridge test methods are used in the assessment.

1. External test using ultrasonic and magnetic particle tests, to identify hidden cracks in the structure.
2. Steel fatigue test using ultrasonic to identify cracks where steel parts are connected (it is common in older steel bridges).
3. Internal test using small samples for the test in the lab.
4. Bridge load capacity test with a driving special truck loaded with concrete blocks while instruments are attached to the bridge and record the movement to find the weight that bridge can safely carry at one time [7].

After scoring the bridge based on the BCI system, they assess the score according to Table 1 [7]. The Ministry of Ontario uses an innovative technology of rapid bridge replacement. In this method, crews lift the old bridges in a few hours and replace them with the new ones which are built nearby. Examples of this technology include Toronto 401 off-ramp bridge at Yorkdale Shopping Centre (2012), Ottawa - Island Park (2007) and Hamilton Aberdeen Bridge (2010) [7].

Table 1. The bridge condition index

BCI	condition	Maintenance
70-100	Good	Is not required in the next 5 years
60-70	Fair	Is required in the next 5 years
<60	Poor	required in the next year

4 Data mining

4.1 Explaining data set

This data set contains 2803 records for bridges in Ontario. Each record contains name, exact location, longitude, latitude, structural system (e.g. slab, beam/girder or frame), material, year built, last major and minor rehabilitation years, details of spans and measured BCI since the year 2000.

4.2 Proceeding data preparation

Preparing data is one of the most important parts of data mining problems. In this case study, data mining is a major consideration. Most of Bridge Condition Indexes are measured on a bi-yearly basis, meaning data organization is crucial. Many classification algorithms consider the gap years between measurements as "missing values," which is not true in this study because the measurements are not taken on an annual basis. To overcome this misperception, the BCIs are sorted chronologically based on the year in which they were measured and then assigned sequential numbers. One record is illustrated as an example in Table 2, which was then transformed into Table 3.

Table 2. Original dataset of one record for BCI

Year	2013	2012	2011	2010	2009	2008
BCI	88	-	89.9	-	96.2	-

Table 3. Transformed dataset of one record for BCI

Year	BCI1(BCI)	BCI2	BCI3
BCI	88	89.9	96.2

Many bridges have only minor or major rehabilitation, which affects the BCI. Although the better approach is to capture both, since the majority of bridges only have one type of rehabilitation, it will increase the missing data to a great extent. Therefore, it is decided to take the most recent rehabilitation (regardless of the type, either major or minor). Data cleaning is coded in R [8], an open-source statistical environment.

4.3 Feature Selection

Feature selection is an essential step in machine learning, particularly in big data. Many gathered variables are irrelevant to the classification, and its relevancy is unknown unless tested [9]. Most importantly, using large feature sets data, slows down algorithms [9] and decreases its accuracy when numbers of utilized attributes are more than optimal [10]. Hence, from the practical perspective selecting small and possibly minimum feature set is highly desirable [9]. This problem considered a minimal optimal problem [11], which has been studied for several years to reduce the feature set. Although using more attributes may result in better accuracy, the model is hard to interpret, and overfitting is more likely. One of the feature selection methods is called wrapper, which finds the best combination of features (subset). Boruta feature selection is a wrapper built based on random forest algorithm. Its algorithm is coded in R- an open source statistical environment in a package called Boruta [8]. It helps find the most important attributes or recognize unimportant ones. This package iteratively compares the importance of each feature with the importance of shadow attributes, that is initiated by shuffling original dataset [9]. Then, those attributes that are significantly better than shadow attributes are confirmed; otherwise, they will be dropped. This process continues until the maximum run occurs or only the confirmed attributes are left. If the former one occurs, it is called Tentative. Hence, the user can increase the number of runs or decrease p-value. To calculate the importance measure of an attribute, first the loss of accuracy of classification for all trees of the forest (that have this attribute) is calculated separately. This loss of accuracy is the result of the random disposition of attribute values between objects. Then the importance measure would be the result of dividing the average loss by its standard deviation [9]. Although it has some deficiencies, it is a useful tool since it can count the fluctuations of the mean accuracy loss among trees in the forest [9].

In order to conduct the Boruta selection algorithm, first the dataset is cleaned, similar attributes were removed, and derived attributes are mutated. The result of this algorithm is shown in Figure 1. Based on this selection feature algorithm, the previous measure BCI (CI2) has the most crucial role in the prediction of bridge condition. Furthermore, different feature selection methods such as subset selection, forward and backward methods as well as filtering method (single factor analysis to evaluate the prediction power of each attribute) are also conducted. All these feature selection methods confirm that the most effective attribute is the most recent BCI.

Hence, the features that have importance value greater than 15 are selected for prediction model which are BCI2, BCI3 (the two recent BCI), the year of most recent rehab and age. Adding more attributes such as longitude and latitude (which implicitly is in the country), material or category, not significantly change the accuracy of the models. It is important to note that different models are conducted based on the result of other feature selection algorithms, namely subset selection, forward, backward, filter and Boruta selection (as each feature selection algorithm suggests a different set of attributes). Although all of them are common in BCI2, the Borouta model is the best based on its performance, hence, proposed.

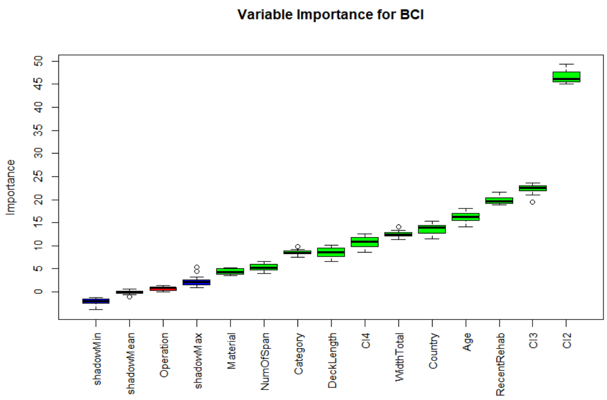

Figure 1. Boruta selection algorithm

4.4 Prediction Model

In this case study, in order to predict the BCI, different numerical prediction algorithms are conducted. Each model is tested in a cross-validation test. Then, the results are compared to choose the best one.

4.4.1 Model Development

The first step in developing the framework is to export the cleaned data from R [8] in a spreadsheet, which is the input for RapidMiner Studio as the Data-Mining Software [12] (Its graphical environment eases its usage for data mining users). As Figure 2 shows, linear regression, KNN, neural network, support vector machine, and random forest are modelled.

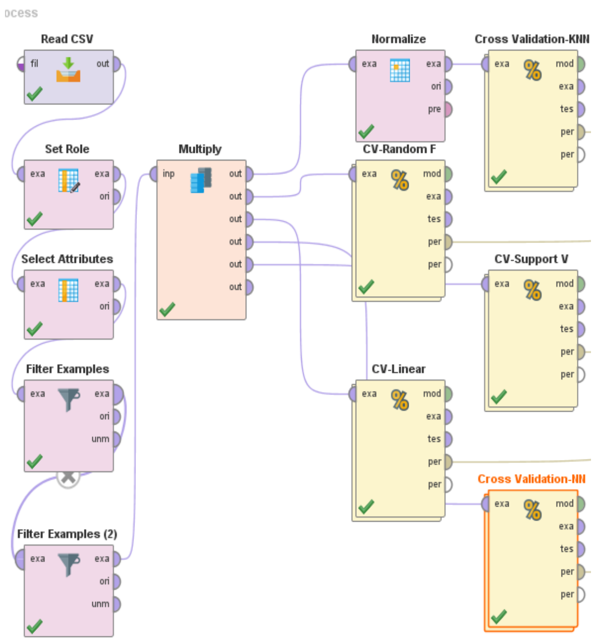

Figure 2. Classification model development

4.4.2 Linear Regression (LR)

In this classification, the relationship between attributes and target attribute is defined with a linear predictor function and their unknown parameters of this function are calculated from training dataset. Table 2 shows the statistical results of this model.

The formula to calculate the upcoming BCI for the next two years will be as equation 1.

$$\begin{aligned} &Next\ BCI \\ &= 0.823 \times BCI2 - 0.075 \times BCI3 + 0.08 \\ &\times Year\ of\ Most\ Recent\ Rehab - 0.024 \\ &\times Age - 139.854 \end{aligned} \tag{1}$$

In this equation, BCI is the predicted Bridge Condition Index for the next year, BCI2 and BCI3 are two formers measured BCIs for the past years. Year of the most recent rehab is the previous year of rehabilitation (either minor or major) and Age is the current year minus the year built of the bridges.

4.5 K-Nearest Neighborhood (KNN)

K-Nearest Neighbour (KNN) is a lazy classification algorithm based on the instances. The mechanism of this method is by calculating the distance for each instance from other instances and assigning them to a class based on k-nearest point. There are different equations for calculation distances such as Manhattan or Euclidian distance. In this study, Euclidian distance is used to measure. David William Aha studied KNN in his doctoral dissertation (1990) and suggested normalizing the values of each feature before applying the model [13]. That is because KNN is sensitive to the scale of data and if they have a different scale, the effect of bigger attributes may neutralize the value of small attributes. Finding the best K in the KNN method can be achieved by try and error as there is no precise rule for this purpose.

4.6 Neural Network (NN)

Artificial Neural Network (ANN or NN) is a data-driven model which is trained iteratively from a random state to estimate target value. This algorithm tries to mimic brain behaviour to learn through neurons. The neurons receive signals from other neurons through the links from the previous layer, which may strength or weaken through weights. When the signal excitation reaches to a certain extent, the neurons will react and fire. However, it is not clear what happens inside this algorithm completely as it is called a black box engine. NN is not depended on the distribution or probability; hence, is considered as a universal approximator [14]. It can be used for both classification (which is well known for it) and regression. It is important to note that by the iterative process, the algorithm finds the weights in a way that can minimize the error.

4.7 Support Vector Machine (SVM)

It is a supervised learning classification model that analyses data for classification or regression analysis. SVM finds the algorithm that had the best and well-separated categories with the clear gap (as wide as possible) between the classes, which increase the probability of a new sample to belong to only one group. Generally, in this algorithm, the wider the margin, the lower the generalization error [15].

4.8 Random Forest (RF)

Random forest is based on the idea of a random selection of features to generate trees. It is an ensemble learning method based on numerous decision trees in training. This algorithm can be used for classification or regression, and the predicted output will be the mode of the classes or mean of prediction, respectively [16, 17].

4.9 Model Selection and Comparison

To select the best model, the first step is to test the accuracy of the model. If there is an extremely big data set, the best practice is to use half of the data for training and the rest for testing. However, usually, data-limitation is a common problem. Thus, it is important to find the best approach for dividing the dataset into training and testing sets. That is crucial because if all data is used in training, although the model is highly accurate, the performance of the model facing new unseen input is highly unknown (validity). On the other hand, if a small

portion of data is utilized for training, the model will not be precise enough. Cross-validation is a tool to overcome this difficulty by dividing data to n folds." The technique of cross-validation usually is recommended as a better test of the model because of the well-known bias induced by testing the predictive validity of a model on the same data that were used to estimate its parameters" [18]. Hence, by cross-validation, we can use all of the data-set in the model. Table 4 shows the error comparison of classifiers by cross-validation test.

Based on Table 5, the relative error of support vector machine is the lowest, root mean squared error of the random forest is the lowest, and correlation of random forest is the highest among all prediction models. However, statistically, there is not a meaningful difference in the accuracy of these models based on statistical t-test with alpha 0.05. Therefore, it is better to find the model which has the smallest range for the error. For this purpose, the box plot of accuracies, based on root mean square error is plotted in Figure 3.

As Figure 3 shows, although the ranges do not vary too much, support vector machine, KNN (K=4) and neural network have a relatively wider range. Based on Figure 3, the author recommends using random forest prediction model, which has a relatively denser box-plot, or linear regression because of its simplicity.

Table 4. Statistics of Linear Regression Model

attribute	coefficient	std. Error	std. Coefficient	Tolerance	t-stat	p-value	code
BCI2	0.823	0.031	0.818	0.288	26.847	0.000	****
BCI3	−0.075	0.031	−0.074	0.286	−2.414	0.016	**
Most Recent Rehab	0.080	0.010	0.126	0.995	8.028	0.000	****
age	−0.024	0.010	−0.042	0.868	−2.484	0.013	**
(Intercept)	−139.854	20.059	NaN	NaN	−6.972	0.000	****

Table 5. Error comparison of classifiers

Classifier	Root Mean Squared Error	Relative Error	Correlation
KNN(K=2)	4.511 +/- 0.758	3.29% +/- 0.35%	0.749 +/- 0.090
KNN(K=3)	4.292 +/- 0.614	3.26% +/- 0.30%	0.772 +/- 0.069
KNN(K=4)	4.178 +/- 0.705	3.22% +/- 0.39%	0.783 +/- 0.079
KNN(K=5)	4.169 +/- 0.688	3.21% +/- 0.36%	0.784 +/- 0.076
LR	4.092 +/- 0.821	3.24% +/- 0.79%	0.807 +/- 0.083
RF	3.947 +/- 0.806	2.91% +/- 0.36%	0.810 +/- 0.091
NN	4.092 +/- 0.821	3.24% +/- 0.79%	0.807 +/- 0.083
SVM	4.203 +/- 0.878	2.19% +/- 0.32%	0.799 +/- 0.086

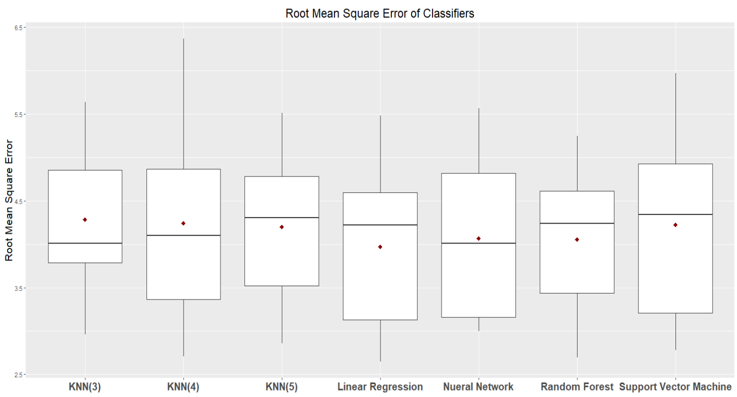

Figure 3. Root Mean Square Error comparison of classifiers

5 Conclusion

Annually, Maintenance and repair of bridges impose significant costs to municipalities. Having a good prediction of bridge conditions can help to recognize the bridges that need an urgent repair, or require repair in a short interval of time. This can be highly beneficial for governments for budget allocation and sequencing the maintenance operation. Furthermore, it can be helpful for planners in order to have a targeted bridge inspection.

This study is based on the historical data of the Municipality of Ontario to find a predictive model for bridge condition index. After feature selection, four prediction models are assessed, namely, linear regression, random forest, neural network, KNN and support vector machine. Then, based on the corresponding accuracy between these models, random forest and support vector machine are suggested. As the study shows, this model can predict the next BCI with 97% accuracy having two former BCIs, age and the year of most recent rehabilitation. By annual update of the database, the suggested framework can enhance further to achieve better accuracy. If the historical data is available, this model can be conducted for a different location.

6 Limitation

The scope of this study is limited to develop a prediction model for BCI and excludes improvement of measurement in BCI experiment or incorporating safety in BCI. This model is not also for selecting the type of material or any design specification of the bridges.

7 References

[1] NEWMAN J. Concrete Repair Strategies, Essentials of Concrete Investigations. Institution of Engineers of Ireland, Dublin, 2003.

[2] Federal Highway Administration. Delivering Infrastructure for America's Future. US Department of Transportation, Washington DC 2002.

[3] Sánchez-Silva M, Frangopol DM, Padgett J, Soliman M. Maintenance and operation of infrastructure systems. Journal of Structural Engineering. 2016 Jun 2;142(9): F4016004.

[4] Pearson-Kirk D. The benefits of bridge condition monitoring. In Proceedings of the Institution of Civil Engineers-Bridge Engineering 2008 Sep (Vol. 161, No. 3, pp. 151-158). Thomas Telford Ltd.

[5] Das PC, editor. Management of highway structures. Thomas Telford; 1999.

[6] Chase SB, Adu-Gyamfi Y, Aktan AE, Minaie E. Synthesis of national and international methodologies used for bridge health indices. United States. Federal Highway Administration; 2016 May 1.

[7] Ontario Ministry of Transportation, Bridge repairs, Online:http://www.mto.gov.on.ca/english/ highway-bridges/ontario-bridges.shtml, Accessed: 12/04/2018

[8] Team RC. R: A language and environment for statistical computing, 2013.

[9] Kursa MB, Rudnicki WR. Feature selection with the Boruta package. Journal of Statistical Software. 2010 Sep 16;36(11):1-13.

[10] Kohavi R, John GH. Wrappers for feature subset selection. Artificial intelligence. 1997 Dec 1;97(1-2):273-324.

[11] Nilsson R, Peña JM, Björkegren J, Tegnér J. Consistent feature selection for pattern recognition in polynomial time. Journal of Machine Learning Research. intelligence. 1998 Aug;20(8).

[12] Cooil, B., Winer, R., & Rados, D. Cross-Validation for Prediction. Journal of Marketing Research, 24(3), 271-279, 1987.