

## **Development of a Backward Prediction Model Based on Limited Historical Datasets**

Jaeho Lee

*(PhD Candidate, Griffith University, Gold Coast, Queensland, Australia)*

Khoa Le

*(Lecturer, Griffith University, Gold Coast, Queensland, Australia)*

Michael Blumenstein

*(Senior Lecturer, Griffith University, Gold Coast, Queensland, Australia)*

Yew-Chaye Loo

*(Professor, Griffith University, Gold Coast, Queensland, Australia)*

### **Abstract**

In almost the last two decades, commercial Bridge Management System (BMS) packages have been remarkably developed. However, inconsistency between BMS inputs and bridge agencies' existing data is an obstacle to implement and to operate a BMS software application. A large number of bridge datasets for a BMS database is an essential requirement to analyze a bridge network. Among many requirements, historical structural datasets are vital to compute the prioritization of bridge stock for maintenance and repair activities and are mostly unavailable for bridges of more than 20 years in age.

This study focuses on the abovementioned difficulty to overcome the lacking historical data problem faced by bridge agencies to effectively use BMS applications. This paper proposes an artificial neural network (ANN) technique to predict missing components of time-series datasets to estimate historical bridge element condition ratings. Although this study only estimates historical condition ratings, the proposed concept can be used to compute other historical dataset requirements in the BMS database and hence improving the reliability of various BMS analysis modules.

### **Keywords**

BMS, Element condition ratings, ANN

## **1. Introduction**

The primary goal of a BMS is to determine and to implement the best possible strategy that ensures an adequate level of safety at the lowest possible life cycle cost (Frangopol et al., 2000). A BMS is necessary to extend the life cycle of the entire bridge network and to optimise the maintenance expenditures, because most infrastructure facilities are planned, designed, constructed, operated and modified or rehabilitated under uncertain and risky conditions (Hudson et al., 1997).

Although BMS software applications have been commercially implemented for almost two decades, many bridge agencies are still hesitated to implement such systems to manage their bridge assets. The

main reason for this is the large amounts of BMS data requirements. Many datasets are different from their BMS requirements such as format or inexistence in bridge agencies. This inconsistency is the major reason why BMSs have not been practically implemented. For that reason, many researchers and infrastructure asset practitioners have recognized that deterioration of infrastructure facilities is not deterministic (Mishalani and McCord, 2006). To overcome this fundamental issue, a bridge agency is required to collect BMS data as early as possible. To adopt analytical methods in a BMS, currently adopted analytical methods in BMS applications regularly require updates of historical datasets to ensure reliable results (Das, 1996). However, during the data collection period, insufficient datasets may yield inaccurate BMS outcomes until sufficient data are recorded in the BMS database.

Consequently, the results of periodic bridge inspection are important to keep maintain the reliability of BMSs. Bridge-condition related data are imperative resources and the most time consuming of the total requirements for proper operation of a BMS software application. Historical bridge condition ratings from periodic bridge inspection can be used directly and indirectly as input data for many significant functions in a BMS software application. More than half of the BMS outputs, about 50% of project-level and about 67% of network-level outputs, are affected by condition rating datasets(Godart and Vassie, 1999 a). Thus, it is evident that the operations of those BMS modules are very difficult without bridge condition ratings.

Time-series predictions are important resources to make decisions in many application (Weigend and Gershenfeld, 1994). In the field of bridge management, there are various available techniques to forecast and to analyze time-series datasets such as regression, Markov model, genetic algorithm and artificial neural network models. To obtain reliable prediction from conventional techniques, the size of missing patterns from a large dataset must be about 5% or less (Tabachnick and Fidell, 2001). Conventional prediction methods also cannot be applied to irregularly-sampled datasets (Karna et al., April 2006).

This fundamental limitation is discussed in the conceptual model of this study. The reliable outcomes in recent predictions can be achieved by using ANNs, but inaccurate prediction results in over longer period are also obtained. The main reason is that predictions with unreliable weighting factors are trained under rare condition rating variances using a limited amount of existing bridge condition rating datasets (Lee et al., 2005a; Lee et al., 2005b).

An enhanced version of ANN-based bridge condition rating models is presented in this paper. This model can be used to overcome the initial problems in the conceptual design. A pilot study conducted to measure its predictive performance using the National Bridge Inventory (NBI) from the Maryland Department of Transportation (DoT). This refined model called the Backward Prediction Model (BPM) predicts the entire or selected periods of historical bridge condition ratings to support the existing BMS input requirements to improve long-term prediction dependability.

## **2. Outline of the BPM**

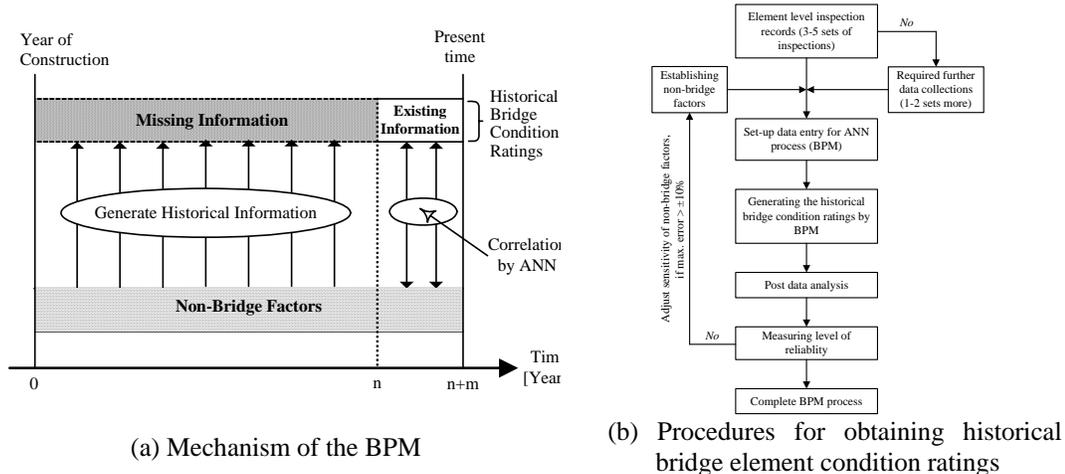
Generally, this research aims to use a small number of existing datasets collected over short periods to predict large datasets spanning over a much longer time period. In the field of bridge management, the history of commercialized element level inspection-based BMS is usually less than 15 years. A BMS requires condition rating information of structures for at least every second year. The bridge agencies that implemented the BMS during the early period still have only about 6 to 7 inspection records for their bridge assets. This can be a problem for aging bridges to recognize their historical patterns by using commonly available time series prediction methods.

Bridge condition ratings normally do not change much over short time periods. As such, it is difficult to detect data trends using an ANN-embedded condition rating prediction model. However, the existing bridge condition ratings can be strengthened using non-bridge factors such as historical information about

vehicles, population and climate data surrounding the bridge area. They directly and indirectly affect the variation of bridge conditions, thus, non-bridge factors provide trends into the existing small quantity of condition rating information throughout the ANN process.

Figure 1(a) schematically shows the mechanism of the BPM. It illustrates the establishment of the correlation between the existing condition rating data (from year  $n$  to year  $n+m$ ) and nominated non-bridge factors (from year  $n$  to year  $n+m$ ). The established relationships using neural networks are applied to the missing years' non-bridge factors to generate unrecorded condition rating data from year 0 to year  $n$ .

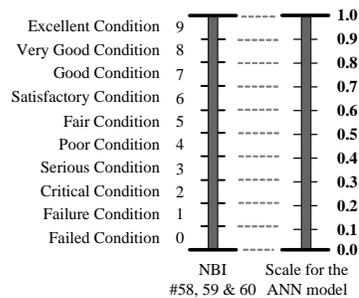
Figure 1(b) presents the procedures for obtaining historical condition ratings from the proposed model.



**Figure 1: Details of the BPM**

## 2.1 Background of the obtained sample datasets

The most widely used inspection method for a BMS operation is the element-level bridge inspection which evolves from the National Bridge Inventory (NBI). NBI information is submitted annually to FHWA by state highway agencies in America. NBI has been used for more than three and half decades to determine the needs of rehabilitation and replacement from a nation-wide perspective. The reason of using NBI instead of bridge element condition ratings in this study is the limitation of its availability. The obtained bridge condition datasets require calibration to fit into the proposed BPM model using the typical ANN input environment. The acceptable numerical scale for ANN modeling is from  $-1$  to  $1$  (or  $0$  to  $1$ ). Figure 2 illustrates the scale of NBI information for this particular study. The Condition Index (CI) in NBI is scaled between 0 and 9 for NBI #58 (deck), #59 (superstructure) and #60 (substructure). Every calibration step is assigned a different Condition State (CS) to express the bridge component's condition ratings.



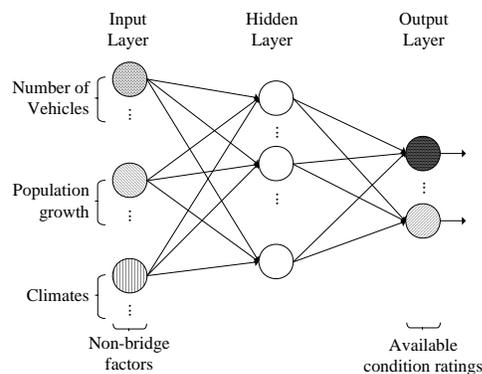
**Figure 2: Scales of NBI for the BPM**

For the non-bridge factors, based on the approximate bridge locations provided from the Maryland DoT, historical vehicle registrations, census population and climate data were collected by the Federal Highway Administration, U.S. Census Bureau, and the U.S. Department of commerce National Oceanic & Atmospheric Administration, respectively.

## 2.2 Composition of the BPM

Figure 3 illustrates the proposed single-layer feed-forward back-propagation neural-network model. It consists of an input layer, hidden layer(s) and an output layer where existing neurons in the hidden and output layers are connected by weighted connections. A neuron in the hidden layer obtains data from the input layer, which is processed by the calculation of a weighted sum and subsequently passed to another neuron in the output layer via a weighted connection.

The specifications of the inputs, outputs and functions of the proposed BPM are detailed in Table 1. The input layer has 21 variables including 4 factors for the yearly vehicle data, 2 factors for the population set and 15 factors for climate conditions. This information is used to train the ANN to determine the correlation with currently-available bridge condition rating data in the output layer.



**Figure 3: The schematic diagram of the proposed BPM**

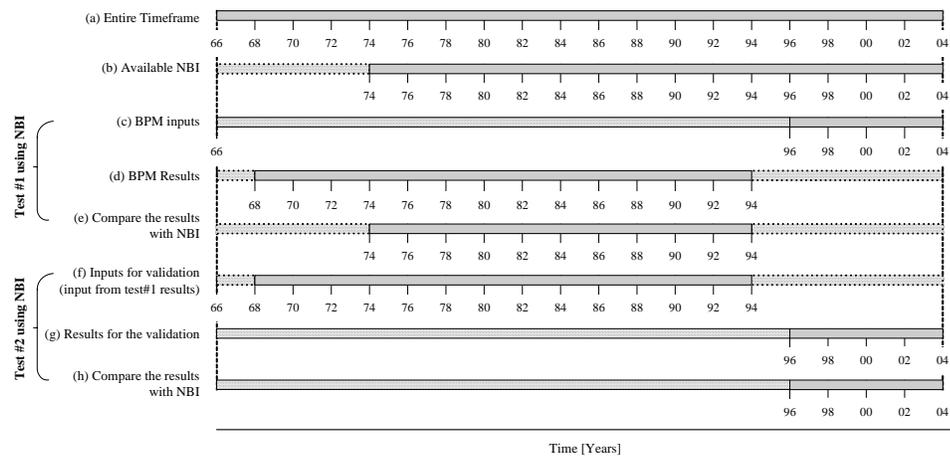
**Table 1: Components of the proposed neural-network model**

Training Algorithm	Back Propagation Algorithm
Transfer Function	Log-sigmoid Function
Inputs	Vehicles (4 factors) Populations (2 factors) Climate conditions (15 factors)
Output	Bridge Condition Ratings (1 output)

## 3. Validation of the BPM

The entire timeframe of the bridge' data used in the BPM is from 1966 to 2004. Among them, on 5 occasions, inspection results have been used as BPM trained inputs and outputs (from 1996 in 2-year increments to 2004). The assumed condition rating at year zero (1966) of the bridge has also been used. The remaining years (from 1968 to 1994 with 2-year increments) of historical condition ratings can be generated by using the proposed BPM. Generated historical condition ratings are compared with the existing information to assess its reliability.

The timeframe of Tests #1 and #2 for the proposed BPM using NBI historical datasets is described in Figure 4 which shows the timeframe of inputs (Figure 4(c) and (f)) and its results (Figure 4(d) and (g)). The two different results are compared with the existing NBI datasets (Figure 4(e) and (h)) to measure the BPM performance.

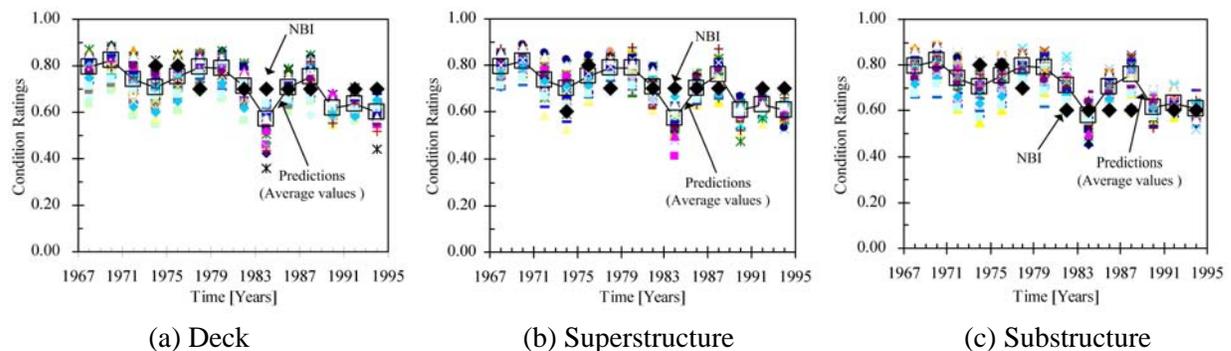


**Figure 4: BPM timeframe (Tests #1 and #2 using NBI for performance measurements)**

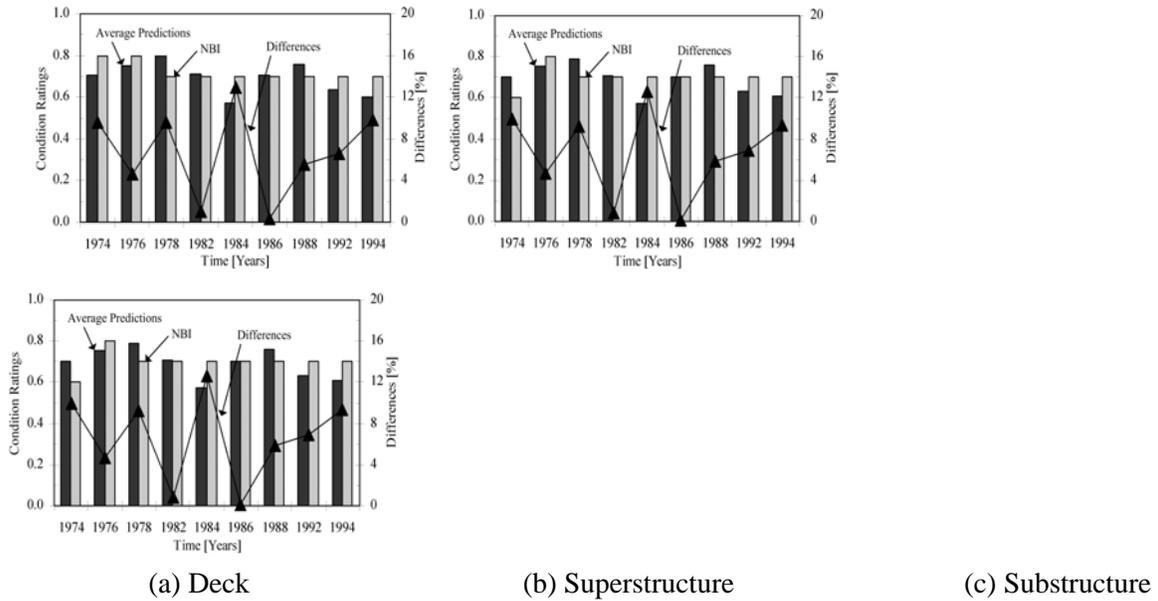
### 3.1 Backward comparisons (Test #1)

NBI data generating the historical condition ratings between 1968 and 1994 is used in the first test. Figure 5 shows the results of the generated historical condition ratings for decks, superstructures and substructures. Approximately 78.6% of the generated data from Test #1 can be directly compared with the actual NBI data to measure the prediction performance. Most artificially-generated historical condition ratings are obtained within the prediction error scale of less than  $\pm 10\%$  as detailed in Figure 6.

However, year 1982 in deck, year 1984 in superstructure, and years 1982, 1984 and 1986 in substructure possess errors larger than the allowed maximum error. In the case where the trained dataset using the existing uncorrelated condition ratings, the proposed neural-network model cannot provide acceptable condition ratings for a specific year. For example, sudden physical damages to the bridge are not reflected by the non-factors used in this proposed model, yielding inaccurate predictions. Nevertheless, the results of Test #1 can still be used as BMS input resources in further tests as will be shown in Section 3.2.



**Figure 5: BPM results (Bridge #0312xxx1)**



**Figure 6: Performance measurements (Bridge #0312xxx1)**

### 3.2 Forward comparisons (Test #2)

Another method, which can be used to validate the proposed model, is conducted in Test #2. The input of the second test uses the results of the first test between 1968 and 1994 as illustrated in Figure 4(f). It produces the future condition ratings between 1996 and 2004. The results are also compared with the existing condition ratings from 1996 to 2004 and summarized in Table 2. Each year of condition ratings estimated using the neural network model provides satisfactory results within the error bound. Therefore, input resources (the prediction results of Test #1 for this model) are validated and can be used as historical condition rating datasets for various BMS project/network level analysis modules.

**Table 2: Summary of prediction performance for test 2**

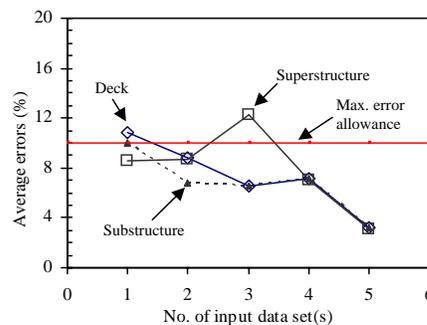
Bridge components	Years	NBI records	Average prediction	Difference	Mean Errors (%)
Deck	1996	0.6	0.568	0.032	3.20
	1998	0.6	0.590	0.010	
	2000	0.6	0.665	0.065	
	2002	0.6	0.590	0.010	
	2004	0.6	0.555	0.045	
Superstructure	1996	0.6	0.567	0.033	3.10
	1998	0.6	0.593	0.007	
	2000	0.6	0.666	0.066	
	2002	0.6	0.595	0.005	
	2004	0.6	0.556	0.044	
Substructure	1996	0.6	0.569	0.031	3.20
	1998	0.6	0.591	0.009	
	2000	0.6	0.669	0.069	

	2002	0.6	0.591	0.009	
	2004	0.6	0.557	0.043	

#### 4. Minimum datasets verification for the BPM

In this study, neural-network models are developed to predict historical bridge condition ratings. To develop the proposed neural-network model, 21 inputs (non-bridge factors) and 3 outputs (condition rating of decks, superstructures and substructures) using a single hidden layer feed-forward network are used to predict historical NBI condition ratings. Fourteen sets of historical condition ratings using 5 sets of available condition ratings from 1966 to 1994 are obtained, excluding the initial year of bridge condition. The number of neurons is denoted by the number of inspections (including those in the initial year) in which 6 neurons in the hidden layer are used.

To verify the minimum resource requirements and the response with the number of inspection records in the proposed model, additional tests are also conducted for the various amounts of inspections as input data. The proposed model assumes that the bridge agency only retains 1, 2, 3, 4 and 5 sets of historical condition ratings. Figure 7 demonstrates the average prediction errors which are gradually decreasing with a larger number of inspections. Therefore, the effective range of inputs for the number of inspections is from 2 sets of records, which partially provide predictions within the maximum acceptable error bounds. It is observed that the proposed BPM can provide satisfactory results when more than 4 sets of inspection records are used as inputs.



**Figure 7: Average errors for different numbers of trained inputs (Bridge #0312xxx1)**

#### 5. Conclusion

This research has been inspired by the lack of historical condition ratings for the BMS database. The neural-network has been employed to backward-predict the historical condition ratings for inspection items of bridge elements. The generated datasets fulfill one of the vital BMS requirements of improving the prediction reliability in future bridge condition ratings. This also affects many processes in the BMS analysis modules.

The ANN-based BPM prediction results using historical NBI datasets have been verified using two different tests with NBI raw data. The comparison results using backward predictions with NBI were shown to possess about 6.7%, 6.7% and 7.5% of the prediction errors from 1974 to 1994 with 2-year increments in decks, superstructures and substructures respectively. The other results in the future direction have also been validated using the concept of BPM. The proposed model uses the generated

historical data from Test #1 as input data to the BPM, which produces current condition ratings to compare with NBI datasets (5 sets of inspections from 1996 to 2004). The results show that the average errors in decks, superstructures and substructure are about 3.20%, 3.10% and 3.20% respectively.

It has been observed that the proposed model could be successfully used to predict the historical bridge condition ratings as long as the trained input dataset and collected non-bridge factors are sensitive to the bridge conditions.

It should be noted that this work is only valid for the obtained bridge sample dataset to provide a framework for bridge agencies. The proposed BPM also provides a cornerstone for further practical developments.

In addition, an extension of the proposed study can be applied to similar problems in many other Infrastructure Management Systems (IMS) to provide a wide range of historical resources for various analytical processes to improve the reliability of IMS outcomes.

## 6. References

- Das, P.C. (1996) "Bridge Maintenance Management Objectives and Methodologies." Bridge Management 3:inspection, maintenance, assessment and repair, J. E. Harding, G. A. R. Parke, and M. J. Ryall, eds., E&FN Spon, London, 1-7.
- Frangopol, D. M., Gharaibeh, E. S., Kong, J. S., and Miyake, M. (2000) "Optimal Network-Level Bridge Maintenance Planning Based on Minimum Expected Cost." *Transportation Research Record*, Florida, 26-33.
- Godart, B., and Vassie, P. R. (1999 a) "Review of existing BMS and definition of inputs for the proposed BMS", Deliverable D4 (BRIME Report: Project PL97-2220), *BRIME Report Project PL97-2220*.
- Hudson, W. R., Haas, R., and Uddin, W. (1997). *Infrastructure Management: Integrating Design, Construction, Maintenance, Rehabilitation and Renovation*, McGraw-Hill, New York.
- Karna, T., Rossi, F., and Lendasse, A. (April 2006) "LS-SVM functional network for time series prediction." *In Proceedings of XIVth European Symposium on Artificial Neural Networks (ESANN 2006)*, Bruges (Belgium), 473--478.
- Lee, J., Sanmugarasa, K., and Blumenstein, M. (2005a) "An Artificial Neural Network Model for the Condition Rating of Bridges." *International Conference on Construction Engineering and Management (ICCEM 2005): Globalization and Collaboration in Construction*, Seoul, S.Korea.
- Lee, J., Sanmugarasa, K., Le, K., and Blumenstein, M. (2005b) "Use of non-bridge element-level inspection data in the implementation of Bridge Management Systems (BMS)." *Collaboration and Harmonization in Creative Systems Volume II: Proceedings of the 3rd International Structural Engineering and Construction Conference (ISEC-03)*, Japan, 705-712.
- Mishalani, R. G., and McCord, M. R. (2006) "Infrastructure Condition Assessment, Deterioration Modeling, and Maintenance Decision Making: Methodological Advances and Practical Considerations." *Journal of Infrastructure Systems*, 12(3), 145-146.
- Tabachnick, B. G., and Fidell, L. S. (2001). *Using Multivariate Statistics*, Allyn and Bacon.
- Weigend, A., and Gershenfeld, N. (1994) "Time Series Prediction: Forecasting the Future and Understanding the Past." Addison-Wesley, Reading (MA).