ESTIMATING THE CONSTRUCTION YEAR OF ROAD BRIDGES IN LAO PDR FROM BRIDGE CHARACTERISTIC AND CONDITION DATA

Shibaura Institute of Technology PhD Student Student Member ○Silimanotham Hatthaphone
Shibaura Institute of Technology Associate Professor Regular Member Michael Henry

1. Introduction

Bridge database is necessary for the work of bridge maintenance and bridge management in developing countries, particularly in Lao PDR. Lao PDR has more than three thousand bridges on six road classifications as national road (NR), provincial road (PR), district road (DR), urban road (UR), rural road (RR) and specific road (SR). Over 600 bridges belong to the national road category, with six bridge types: concrete slab, concrete girder, steel girder, steel truss, wooden and other. However, there are issues in bridge database as 49% of the bridges in NR have missing data, especially the year of construction. Furthermore, bridge maintenance and management need to have a complete database to support on the maintenance works and the management planning in the future.

This research focuses on the data for construction year which are missing in the bridge database for concrete bridge type belonging to NR. The research objective is to estimate the missing data from the available characteristic and condition data in the bridge database.

2. Methodology

2.1. Sample

The bridge database provides the information of 651 bridges belonging to NR, wherein 330 bridges have a known construction year and 321 bridges are unknown. Furthermore, this research uses the bridge characteristics namely concrete bridge type, maximum span length and condition rating to support in the experimentation.

2.2. Analysis

This research uses the multiple regression model to predict the construction year (dependent variable) with 3 independent variables as concrete bridge type, maximum span length and condition rating. On the known construction year, the data is divided into 2 groups as 70% of 330 bridges for the model testing and 30% of 330 bridges for the accuracy testing by using random sampling method. Next, the procedure, being the trial multiple regression model, was tested many times by using 70% of 330 bridges. After getting the result of multiple regression model test, this research uses the linear equations to calculate the construction year on the 30% 330 bridges and comparing the actual construction year with the average predicted construction year. Furthermore, the linear equation by Yuan, Y.C. [1] shows below:

$$Y = B_0 + (B_1 \times X_1) + (B_2 \times X_2) + (B_3 \times X_3) \qquad (1)$$

Y: is a year of construction.

B_0: is a coefficient of intercept getting from result of multiple regression.

B_1: is a coefficient of concrete bridge type from the result of multiple regression.

B_2: is a coefficient of maximum span length from the result of multiple regression.

B_3: is a coefficient of condition ratings from the result of multiple regression.

X_1: is actual value of concrete bridge type on database.

X_2: is actual value of maximum span length on database.

X_3: is actual value of condition ratings on database.

3. Results & discussion

3.1. Multiple regression model

Multiple regression represents the relationship between independent variables and dependent variables by the observation of the coefficients value and p-value via on Table 1. Furthermore, the higher coefficient value means the stronger the relationship between the dependent and independent variable in question. This is comparative strength within the model.

Table 1 Result of multiple regression for 70% trial 1, trial 2, trial 3 and trial 4

	70% trial 1		70% trial 2		70% trial 3		70% trial 4	
	Coefficients	*P-value*	*Coefficients*	*P-value*	*Coefficients*	*P-value*	*Coefficients*	*P-value*
Intercept (B_0)	2008.65	0	2006.79	0	2008.47	0	2007.03	0
Concrete bridge type (B_1)	-3.71	0.0436	-3.06	0.0925	-3.34	0.0519	-1.37	0.4429
Max Span Length (B_2)	0.00	0.9817	0.02	0.0236	0.03	0.0359	0.02	0.0405
Condition Rating (B_3)	-3.39	0.0001	-2.57	0.0044	-3.47	0.0000	-3.06	0.0002

Keywords: bridge management, bridge database, construction year, multiple regression, national road, Lao PDR.
Address: Toyosu 3-7-5, Koto-ku, Tokyo, Japan 135-8548. TEL +81-035-859-8363

The result of four multiple regressions in Table 1 shows three independent variables are significant ($p<0.05$). First independent variable, the concrete bridge type (B_1) has the highest coefficient value in trial 1 and trial 2 and has a high coefficient value in trial 3 and trial 4 as -3.71, -3.06, -3.34 and -1.37, respectively. But it has the highest p-value on trial 2, trial 3 and trial 4 as 0.0925, 0.0519 and 0.4429. That result means the concrete bridge type is the stronger relationship and not significant variable because p-value are bigger than 0.05. Second independent variable, the maximum span length (B_2) has the lowest coefficient value in four trials as coefficient value range 0.00 to 0.03. But the p-value are smaller than 0.05 in trial 2, trial 3 and trial 4. That result means the maximum span length is the weak relationship but be significant variable. Third independent variable, the condition rating (B_3) has a high coefficient value in trial 1, trial 2 and has the highest coefficient value in trial 3 and trial 4 as -3.39, -2.57, -3.47 and -3.06, respectively. It also has the lowest p-value in four trials as 0.0011, 0.0044, 0.0000 and 0.0002. That result means the condition rating is a stronger relationship and is a significant variable. Based on the results and discussion above, three independent variables are the variety trend as first variable is significant variable in trial 1, second variable is not significant in trial 1 and variable 3 is significant in four trials. In short, the most independent variables impacted must be the condition rating because it has a high coefficient value and be significant variable in all multiple regression.

3.2. Comparing the average estimated construction year and the actual construction year

Based on the result of multiple regression, this research gets four linear equations based on the form in Eq. (1) to apply the equations to 30% of the 330 bridges for estimating the construction year. After getting the construction year from the four models, the procedure averages all predicted construction year and compares to the actual construction year of the 30% of 330 bridges.

To investigate the accuracy multiple regression model, this research uses the comparing method between the average predicted construction year and the actual construction year by plotting both construction year data on scatter (Figure 1). This illustrates the difference between the actual construction year and the estimated construction year by observing the distance between the points and red line.

The data points in Figure 1 were distributed on many trends. The trend points located closely the red line represent the similar result of multiple regression between average predicted construction year and actual year. In contrast, the trend points far from the red line illustrate the low accuracy result of multiple regression. That mean the average predicted construction year be remote from the actual construction year. Therefore, the multiple regression model on this research is shares poor performance in accuracy because the result of multiple regression is unexpected result as some older bridges in Lao PDR still have the good condition. On the other hand, some recent bridges have the bad condition rating.

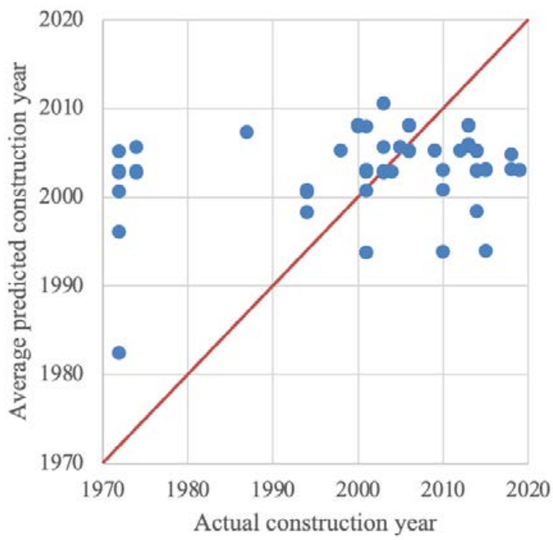

Figure 1 Relationship between average predicted construction year and actual construction year

4. Conclusion

In conclusion, this research showed the most impacted independent variable on multiple regression model be the condition rating because all regression result is the highest change probability, ranked as 99%. However, the multiple regression model might have the other impact factors like the traffic loading, environmental attractive, the limitation of data etc. In short, this research cannot accept the multiple regression model to estimate the construction year on the unknown bridge database belonging to NR.

For the future research, researcher will try to use another method for imputing the missing data and will increase the variables and amount of samples size.

Acknowledgement

This research was supported by a scholarship for road asset management from the Japan International Cooperation Agency.

References

[1] Yuan, Y.C. (2000, April). “Multiple Imputation for missing data: Concepts and new development”. In Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference (Vol. 267, No. 11)