Identification of Structural Performance of A Steel-Box Girder Bridge Using Machine Learning Technique

Hyun-Joong KIM
Doctoral Candidate
Seoul National University
Seoul, Korea
hjkim@sel.snu.ac.kr

Wonsuk PARK
Chief Researcher
Seoul National University
Seoul, Korea
wpark@sel.snu.ac.kr

Hyun-Moo KOH
Professor
Seoul National University
Seoul, Korea
hmkoh@snu.ac.kr

Jinkyo F. CHOO
Assistant Professor
Konkuk University
Seoul, Korea
jfchoo@konkuk.ac.kr

Summary
A new procedure for multiple model identification and representative model selection is proposed. Multiple models are generated from the weighted aggregation formulation for multi-objective optimization problem which deals with multiple target data including static displacements, natural frequencies and mode shapes. By applying principal component analysis on identified models, target values and structural model parameters can be grouped. In the identified principal component space, then, a representative model is selected via *K*-means clustering method. The proposed method is applied to the Yondae Bridge in Korea, a 4×45 m continuous steel-box girder bridge. The proposed method successfully produces a representative finite element model of which static and dynamic properties are well within the given error bounds of field measured data. It also shows that structural performance of the bridge can be identified by representative finite element models in terms of deflection, moment and shear force. Rating factor will be calculated to assess structural performance more definitely corresponding updated models' status.

Keywords: Structural identification; Machine learning; Principal Component Analysis; K-means clustering; Bridge assessment.

1. Identification of multiple models set

Yondae Bridge which is open-toped two-cell steel box girder with RC slab type is taken as a case study. The nominal values of material and sectional properties of baseline FE model are derived from design drawing as well as other available information. From those values, the best fitting parameters to real structure's behaviour are identified based on the measured deflections, natural frequencies and mode shapes. In the updating procedure, 104 structural parameters were initially considered including weight density, elastic modulus, moment of inertia and section area which are a crucial factor in determining structure's behaviour. Among them, 33 parameters showing little sensitivity on static/dynamic behaviour are neglected, and finally 71 parameters are selected to be updated.

Discrepancy between response of baseline model and measured response determine the adjustment direction of the parameters. Error of each term indicating such discrepancy must be defined, from which objective function to be optimized is formulated. These error terms are aggregated into one objective function by multiplying weighting term w_k as shown in equation (1).

$$J = \sum_{i=1}^{M} w_i e_i \tag{1}$$

In this research, 5 values are assigned independently to the each of 3 weightings; 0, 1, 10, 50, 100. Larger values imply higher importance and reliability of the target data, smaller value means vice versa. Pair of weighting combination are normalized and employed in the objective function by turn, and then it ultimately leads to the differently updated models for each weighting pair. In this way, total 67 candidate models are obtained which fit given data the most.

https://doi.org/10.2749/222137813806501939

2. Clustering and representative model selection

Applying PCA to the obtained parameters of candidate models reveal principal components of structural parameters and candidate model's coordinate in the transformed space. Each axis implies the linear combination of original structural parameter spaces, as they become orthogonal to each other. PCA enables to look into the structure of the data and classify them roughly. For more rigorous classification, *K*-means clustering algorithm is applied to data points in all principal component space. As a result, 5 appropriate clusters are identified for the example problem. When we look them in the first three principal components' space, we can find out they're well divided into independent cluster.

Parameters for representative model is obtained by transforming 5 cluster centres' coordinate in principal component space into original space. Three other extreme models are also picked from the candidate models expected as the best fitting models to deflection, natural frequency and mode shape respectively, each of which is featured by high weightings to specific responses while other responses are underrated with the low weightings. Three errors for the candidate FE models are summarized in table 1. M1 ~ M5 means the representative model of cluster 1 to 5, and E1, E2 and E3 are models considerably fitted on the deflection, natural frequency, and mode shape respectively.

Table 1: Error of representative and extreme models

Model / Error	M1	M2	M3	M4	M5	E1	E2	E3
e_1	0.2075	0.2501	0.2311	**0.2027**	0.3239	**0.1940**	0.3562	0.2778
e_2	0.0539	0.0351	0.0338	**0.0235**	0.0406	0.0606	**0.0421**	0.0489
e_3	**0.0240**	0.0384	0.0404	0.0378	0.0410	0.0427	0.0448	**0.0220**

It can be easily found that the representative models do not always guarantee the smallest error for all responses, but errors of all models are located within the extreme error bound. In other words, the representative models compromises all errors in relatively acceptable range compared with extreme cases.

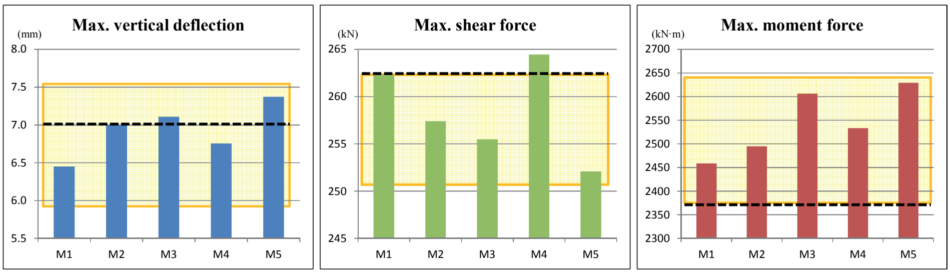

Fig. 1: Max. deflection, shear force and moment of representative, extreme and initial models

In addition, the performances of the representative models are investigated through deflection, shear force and moment. They are also compared with responses of the three extreme models as well as the value of the initial baseline model. In the comparative plotting of figure 1, yellow areas show the response bound of extreme models. In addition, black dotted line indicates the response of initial model without update. Whereas the representative models, denoted as M1 to M5, are observed as still showing moderate performance between extreme models, the initial model shows much biased response for moment and shear force. It implies that the risk of overestimation or underestimation is mitigated through the use of updated model. Especially, by using representative model than a specific model, performances can be investigated in various views.

These observations support the validity of the representative models according to model use purpose, rather than using biased model. In most situations that uncertainties exist in FE model as well as measured data itself, these types of representative model is possibly a good answer by considering relative reliability or importance in the measurement data.