Engineering Structures 80 (2014) 211–221

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

Methodologies for predicting natural frequency variation of a suspension bridge

Irwanda Laory [a,*], Thanh N. Trinh [b], Ian F.C. Smith [b], James M.W. Brownjohn [c]

[a] School of Engineering, University of Warwick, Coventry, United Kingdom
[b] Applied Applied Computing and Mechanic Laboratory, Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland
[c] Vibration Engineering Section, College of Engineering, Mathematics and Physical Sciences, The University of Exeter, Exeter, United Kingdom

ARTICLE INFO

Article history:
Received 15 November 2011
Revised 29 August 2014
Accepted 1 September 2014
Available online 22 September 2014

Keywords:
Environmental effect
Artificial neural network
Support vector regression
Regression tree
Random forest
Variable importance
Suspension bridge

ABSTRACT

In vibration-based structural health monitoring, changes in the natural frequency of a structure are used to identify changes in the structural conditions due to damage and deterioration. However, natural frequency values also vary with changes in environmental factors such as temperature and wind. Therefore, it is important to differentiate between the effects due to environmental variations and those resulting from structural damage. In this paper, this task is accomplished by predicting the natural frequency of a structure using measurements of environmental conditions. Five methodologies – multiple linear regression, artificial neural networks, support vector regression, regression tree and random forest – are implemented to predict the natural frequencies of the Tamar Suspension Bridge (UK) using measurements taken from 3 years of continuous monitoring. The effects of environmental factors and traffic loading on natural frequencies are also evaluated by measuring the relative importance of input variables in regression analysis. Results show that support vector regression and random forest are the most suitable methods for predicting variations in natural frequencies. In addition, traffic loading and temperature are found to be two important parameters that need to be measured. Results show potential for application to continuously monitored structures that have complex relationships between natural frequencies and parameters such as loading and environmental factors.

1. Introduction

Many vibration-based approaches in structural health monitoring have been designed to identify changes in natural frequency values for the purpose of detecting changes in structural conditions that may indicate structural damage and degradation. In reality, however, civil engineering structures are subject to environment and operating effects caused by changes in temperature, traffic, wind, humidity and solar-radiation [1–5]. Such environmental effects also change natural frequency values, hence concealing changes due to structural damage [6–10]. Therefore, it is important to distinguish between changes due to structural damage and changes resulting from environmental effects. This task is managed observing then modeling dependencies of natural frequencies on environmental parameters [11]. The prediction of natural frequencies of structures under environmental changes has been studied using methods such as linear regression analysis, artificial neural networks and support vector regression.

Multiple linear regression (MLR) was employed to predict changes in natural frequencies of the Alamosa Canyon Bridge (USA) due to environmental temperature variation [9] with natural frequencies formulated as a linear function of temperature data. It was found that the changes in the frequencies were linearly correlated with temperature taken from different locations on the bridge. Peeters et al. [12] conducted a 1-year monitoring study for the Z24-Bridge (Switzerland) before it was deliberately damaged, applying a linear regression analysis to distinguish normal frequency changes from abnormal changes due to damage. Also, for this concrete box girder bridge, Peeters and Roeck [13] applied an autoregressive method with exogeneous inputs (ARX) to predict the bridge natural frequencies, where no relationship was found between natural frequencies and wind, rainfall and humidity. Liu and Dewolf [3] simulated the varying natural frequencies under temperature changes using a linear regression analysis, concluding that the long-term variations of natural frequencies are closely related to the variation in in-situ concrete temperature for the three frequencies they measured. The MLR method has also been used to predict natural frequencies of suspension bridges and a footbridge using long-term monitoring data [11,14].

* Corresponding author.
E-mail address: I.Laory@warwick.ac.uk (I. Laory).

http://dx.doi.org/10.1016/j.engstruct.2014.09.001
0141-0296/

Artificial neural networks (ANNs) have been successfully applied in fields such as pattern recognition [15], artificial intelligence [16] and civil engineering [17–20]. For long-term monitoring of structures, ANNs have been employed to predict time-dependent natural frequencies of a structure in order to eliminate the environmental effects on vibration-based damage detection procedures. For example, Ni et al. [21] applied an ANN to formulate the correlation between the natural frequencies and environmental temperatures taken from the cable-stayed Ting Kau Bridge (Hong Kong). Zhou et al. [22] further investigated the performance of the ANNs formulated using the early stopping technique by constructing three different kinds of input, including mean temperatures, effective temperatures and principle components (PCs) of temperatures. The results indicated that when a sufficient number of PCs were taken into account, the ANN using temperature PCs as inputs predicted natural frequencies more accurately than that when using the mean temperatures. More studies on ANNs for the prediction of structural responses are found in Refs. [22–25].

Support vector regression (SVR) is an application form of support vector machines that is a learning system using a high dimension feature space [26,27]. An attractive characteristic of SVR is that instead of minimizing the observed training error such as with MLR and ANNs, SVR involves minimizing the generalized error bound in order to achieve good performance. The generalized error bound is the combination of the training error and a regularization term that controls the complexity of prediction functions. A good overview of SVR is given in [28,29]. SVR has been successfully employed in fields such as text categorization and pattern recognition as well as structural health monitoring [27,30]. Ni et al. [31] applied SVR to predict natural frequencies of the cable-stayed Ting Kau Bridge (Hong Kong) subjected to temperature variations taken from 1-year measurement data, the method exhibiting better prediction capability than the MLR method. Also using measurement data of this bridge, Hua et al. [32] combined principle component analysis (PCA) and SVR to simulate temperature–frequency correlations. It was found that the SVR method trained using the PCs of measured temperature data outperformed that trained using measured temperature data directly.

The methodologies used above are based on parametric functions that specify the form of the relationship between inputs and a response (output) but in many cases, the form of the relationship is unknown. Regression tree (R_Tree) methods offer a non-parametric alternative [33] that has been used extensively in a variety of fields. The method has been found to be especially useful in biomedical and genetic research, speech recognition and other applied sciences [34]. Recent studies in the machine-learning field found that significant improvements in prediction accuracy have resulted from growing an ensemble of trees in a random way, a methodology called *random forest* (RF) [35]. It has been demonstrated that RF has improved prediction accuracy in comparison to other regression methods [36] but additionally provides measures of variable importance for each input variable [37,38]. This method has not been evaluated for its applicability to structural health monitoring, so this paper investigates the performance of RF on predicting natural frequencies through a case study of a suspension bridge.

The studies mentioned above have proposed methodologies for predicting the dynamic responses of bridges, but none has compared methodologies for prediction accuracy. This paper compares five methodologies – multiple linear regression, artificial neural networks, support vector regression, regression tree and random forest – in terms of their ability to predict natural frequencies of a suspension bridge. Confidence intervals are then defined for the best method to differentiate the effects due to environmental changes from those caused by structural damage. Furthermore, the individual effects of temperature, wind and traffic loading on the natural frequency responses of the bridge are evaluated using the variable importance metric in regression analysis.

2. Methodologies for predicting natural frequencies of the bridge

2.1. Multiple linear regression (MLR)

Assuming that a response variable y (for example natural frequency) is linearly related to the p input variables (for example temperature, wind and traffic loading) $x_1, \ldots, x_p$ so that

$$y = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + e. \tag{1}$$

This relationship is known as a linear regression analysis, where β_i is the regression coefficient associated with the ith input variable x_i and e the random error with mean zero and variance σ^2. Using the dataset of n observations in measurement time series, the unknown coefficients β_i are determined using the least-squares method.

2.2. Artificial neural networks (ANNs)

Artificial neural networks can be used as a nonlinear regression method to predict the natural frequency of a bridge. ANN is a two-stage regression in which the first stage is to create derived features Z_m, represented by hidden layer, from linear combinations of the inputs and the second stage is to model the output Y_m as a function of linear combinations of the Z_m. Z_m could be considered as a basis expansion of the original input X.

$$\begin{aligned} Z_m &= \phi(\alpha_{0m} + \alpha_m^T X), \quad m = 1, \ldots, M, \\ T_k &= \beta_{0k} + \beta_k^T Z, \quad k = 1, \ldots, K, \\ f_k(X) &= T_k + e, \quad k = 1, \ldots, K, \end{aligned} \tag{2}$$

where $Z = (Z_1, Z_2, \ldots, Z_M)$, $\phi(v)$ is the activation function which is usually chosen to be the sigmoid $\phi(v) = 1/(1 + e^{-v})$, e the random error, α_i and β_i are unknown parameters. Given a training set $\{x_i, y_i\}$ ($i = 1, \ldots, N$), the ANN regression model is formulated by searching these unknowns so that the sum-of-squared errors as a measure of fit reaches a minimum value.

$$R(\alpha, \beta) = \sum_{k=1}^{K} \sum_{i=1}^{N} (y_{ik} - f_k(x_i))^2 \tag{3}$$

The generic approach to minimizing, $R(\alpha, \beta)$, is by gradient descent, called back-propagation. A two-layer back-propagation neural network (BPNN) is employed to predict the natural frequencies of a structure. BPNN is first trained using the training set in order to formulate the relationship between the natural frequencies and environmental factors including direct loading such as traffic. BPNN is composed of one hidden layer and one output layer with a tan-sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The tan-sigmoid transfer function is capable of capturing the nonlinear relationship between input variables (in our example three of them) and output variables (in our example individual natural frequencies).

An important parameter to be determined when using BPNN for prediction tasks is the optimal number of hidden nodes in the hidden layer. A network with too few hidden nodes might not have enough flexibility to capture the nonlinearities in the relationship while a network with too many hidden nodes may have a tendency to overfit the training data.