Engineering Structures 33 (2011) 903–910

Contents lists available at ScienceDirect

Engineering Structures

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

Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models

J. Mata *

Monitoring Division, Concrete Dams Department, National Laboratory for Civil Engineering, Av. do Brasil 101, 1700-066, Lisbon, Portugal

ARTICLE INFO

Article history:
Received 15 July 2010
Received in revised form
9 November 2010
Accepted 1 December 2010
Available online 11 January 2011

Keywords:
Concrete dam
Dam behaviour
Ceteris paribus
Artificial neural network
Multiple linear regression

ABSTRACT

The safety control of large dams is based on the measurement of some important quantities that characterize their behaviour (like absolute and relative displacements, strains and stresses in the concrete, discharges through the foundation, etc.) and on visual inspections of the structures. In the more important dams, the analysis of the measured data and their comparison with results of mathematical or physical models is determinant in the structural safety assessment.

In its lifetime, a dam can be exposed to significant water level variations and seasonal environmental temperature changes. The use of statistical models, such as multiple linear regression (MLR) models, in the analysis of a structural dam's behaviour has been well known in dam engineering since the 1950s. Nowadays, artificial neural network (NN) models can also contribute in characterizing the normal structural behaviour for the actions to which the structure is subject using the past history of the structural behaviour. In this work, one important aspect of NN models is discussed: the parallel processing of the information.

This study shows a comparison between MLR and NN models for the characterization of dam behaviour under environment loads. As an example, the horizontal displacement recorded by a pendulum is studied in a large Portuguese arch dam. The results of this study show that NN models can be a powerful tool to be included in assessments of existing concrete dam behaviour.

1. Introduction

The main objective of the safety control of a concrete dam is to guarantee the functions for which it was built by maintaining its functionality and its structural integrity. The safety control is supported by monitoring activities and is based on models.

The ultimate purpose of the models is to predict the behaviour of a concrete dam and to identify whether the behaviour of the structure is still similar to past behaviour under the same loads or if there is any difference. If indeed the evolution is divergent between the model prediction and actual behaviour, then the assumptions of the models have changed and the reason for the change should be identified to assess the consequences.

Models based on mechanical principles are often difficult to construct and it is necessary to deal with the uncertainty in the parameters. In general, it is interesting to find out how changes in the input variables affect the values of the response variables. An empirical formulation for structural response is usually obtained as the sum of three terms: the temperature variation, the hydrostatic pressure variation and other unexpected unknown causes such as the result of time effects. The uncertainty of the model is represented by the residual term of the model. Some structural identification techniques have been successfully obtained by De Sortis and Paoliani [1] and Léger and Leclerc [2], although using a very complex procedure. On the other hand, with a large amount of observation data it is possible to define the characterization of a normal dam's behaviour by using statistical models without the knowledge of mechanical principles [3]. Nowadays, there is great experience in using MLR model methods for the characterization of a concrete dam's behaviour.

The NN models have been applied in different areas, including dam engineering. Some works related to this subject can be mentioned such as Perner et al. [4], Gomes and Awruch [5], Fedele et al. [6], Feng and Zhou [7], Bakhary et al. [8], Wang and He [9], Wen et al. [10], Liu et al. [11], Joghataie and Dizaji [12] and Yi et al. [13].

Both MLR and NN approaches have potential value for assessing the behaviour of the control variables that support the safety assessment of the concrete dam as is shown with a *ceteris paribus*[1] analysis in this study. In the period of normal operation of a

* Tel.: +351 218443372; fax: +351 218443026.
E-mail address: jmata@lnec.pt.

[1] *Ceteris paribus* is a Latin phrase, that can be translated as "all other things being equal".

0141-0296/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engstruct.2010.12.011

concrete dam, the main actions and the structural response are well characterized and there is a strong correlation between these two. The study of a structural response of horizontal displacements with *ceteris paribus* analysis for the temperature effect is presented for different water levels. In the same way, a similar study for the hydrostatic pressure effect is carried out for different levels of temperature.

2. Statistical models

The variations of hydrostatic pressure and temperature are the main actions to be taken into account when analysing the results of the concrete dam observations.

The simultaneous effects of hydrostatic pressure and temperature variations create an observed effect which is the result of both loads. It is also important to take in account any effects of the loads separately in interpreting a concrete dam's behaviour. Such isolation can be promptly obtained when the variation of only one load appears between seasons. However, the most general way to separate their effects is the use of statistical methods. These methods use simultaneous consideration of a large number of observations, allowing the establishment of the correlations between the observed behaviour and corresponding loads.

The MLR model is one of the statistical techniques most widely used for analysing multifactor effects. A MLR model is a statistical technique for investigating and modeling the relationship between variables [14,15]. In dam engineering, MLR models have a long history and were initially known as quantitative analysis models [16]. A MLR model does not imply a cause effect relationship between the variables and in almost all applications of regression, the regression equation is only an approximation to the true relationship between the variables.

In recent years, the field of Artificial Intelligence has introduced some tools that are able to perform cognitive tasks such as pattern recognition and function approximation [17]. This is the case of the Multilayer Perceptron (MLP) models that were used in this work.

Generally, MLR and NN models are valid only within the region of the observed data.

2.1. Multiple linear regression

A MLR model is a method used to model the linear relationship between a dependent variable and one or more independent variables. The dependent variable is sometimes also called the predictand or response, and the independent variables the predictors.

These models consider that the effects associated with a limited time period at a specific point can be approximated by Eq. (1).

$$E(h,\theta,t) = E_h + E_\theta + E_t \tag{1}$$

where $E(h,\theta,t)$ is the observed effect; E_h is the elastic effect of hydrostatic pressure; E_θ is the elastic effect of temperature, depending on the thermal conditions; E_t is the effect function of time, considered irreversible.

Eq. (1) is based on several simplifying assumptions concerning the behaviour of materials, such as: (i) the analysed effects refer to a period in the life of a concrete dam, for which there is no relevant structural changes; (ii) the effects of the normal structural behaviour for normal operating conditions can be represented by two parts. A part of the elastic nature (reversible and instantaneous, resulting from the variations of the hydrostatic pressure and the temperature) and another part of inelastic nature (irreversible) such as a time function; (iii) the effects of the hydrostatic pressure and temperature changes can be studied separately.

The effects of hydrostatic pressure variation, $E_h(h)$, are usually represented by polynomials, depending on the height of water in the reservoir h, Eq. (2).

$$E_h(h) = \beta_1 h + \beta_1 h^2 + \beta_2 h^3 + \beta_4 h^4. \tag{2}$$

The effect of the temperature changes can be considered as a proportional attenuation of the air temperature changes with a phase shift with depth along section.

Very simple MLR models usually do not use temperature measurements because it is assumed that the thermal effect $E_\theta(d)$ can be represented by the sum of sinusoidal functions with one-year period and six-month period [18,2]. Thus, the effect of temperature variations is defined by a linear combination of sinusoidal functions, which only depends on the day of the year, Eq. (3).

$$E_\theta(d) = \beta_5 \sin(d) + \beta_6 \cos(d) + \beta_7 \sin^2(d) + \beta_8 \sin(d)\cos(d) \tag{3}$$

where $d = \frac{2\pi\cdot j}{365}$ and j represents the number of days between the beginning of the year (January 1) until the date of observation ($0 \le j \le 365$).

To represent the time effects, $E_t(t)$, it is usual to consider the functions presented in Eq. (4), where t is the number of days since the beginning of the analysis.

$$E_t(t) = \beta_9 t + \beta_{10} e^{-t}. \tag{4}$$

Suppose that there are p independent variables and n observations, $(X_1, \ldots, X_p, Y)$ where Y represents the observed effect and X are the functions of water level height, temperature and time, respectively.

The model relating the independent variable to the dependent variable is $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_j X_j + \cdots + \beta_p X_p + \epsilon$.

The model is obtained by a system of n equations that can be expressed in a matrix notation as $\mathbf{Y} = \mathbf{X}\beta + \epsilon$, where $\mathbf{Y}$ is a $(n \times 1)$ vector of the dependent variable or response, $\mathbf{X}$ is a $(n \times (p+1))$ matrix of the levels of the p independent variables, β is a $((p+1) \times 1)$ vector of the regression coefficients, and ϵ is a $(n \times 1)$ vector of random errors. This method assumes that the expected value of the error term is zero, which is $E(\epsilon) = 0$; the variance $V(\epsilon) = \sigma^2$ and that the errors are uncorrelated [14,19,20].

MLR models are based on least squares: the model is fit such that the sum-of-squares of differences of observed and predicted values, $L = \sum_{i=1}^{n} \epsilon_i^2 = \epsilon^T \epsilon = (\mathbf{Y} - \mathbf{X}\beta)^T (\mathbf{Y} - \mathbf{X}\beta)$, is minimized. The regression coefficient estimator, $\hat{\beta}$, is the solution for β in the equation $\frac{\partial L}{\partial \beta} = 0$.

In matrix notation, the least squares estimator of β is $\hat{\beta} = (\mathbf{X^T X})^{-1} \mathbf{X}^T \mathbf{Y}$, the fitted model is $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$ and the vector of the residuals is denoted by $\hat{\epsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$.

2.2. Multilayer Perceptron

The NN model is a simplified mathematical model of a natural neural network. NN models are inspired on the efficiency of the brain process. Many of the important issues concerning the application of artificial neural networks can be introduced in the simpler context of polynomial curve fitting [17,21].

NN models have been employed successfully to solve complex problems in various fields of application including classification, pattern recognition, prediction, optimization, function approximation and control systems [22,23]. The increasing interest for this area derives from the learning ability of these models, which relate the variables without imposing relationships between them.

A neuron is the main element of an artificial neural network. It is an operator with inputs and outputs, associated with a transfer function, f, interconnected by synaptic connections or weights, w. Fig. 1 illustrates how information is processed through a single neuron.