Reliability Engineering and System Safety 169 (2018) 40–50

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Hydrostatic-season-time model updating using Bayesian model class selection

Sonja Gamse[a], Wan-Huan Zhou[b,*], Fang Tan[b], Ka-Veng Yuen[b], Michael Oberguggenberger[c]

[a] Unit for Surveying and Geoinformation, Faculty of Engineering Science, University of Innsbruck, Technikerstr. 13, Innsbruck 6020, Austria
[b] Department of Civil and Environmental Engineering, Faculty of Science and Technology, University of Macau, Avenida da Universidade, Taipa, Macau, China
[c] Unit for Engineering Mathematics, Faculty of Engineering Science, University of Innsbruck, Technikerstr. 13, Innsbruck 6020, Austria

ARTICLE INFO

MSC:
00-01
99-00

Keywords:
Bayesian model class selection
Geodetic observations
Hydrostatic-season-time model
Model class selection
Multiple linear regression
Rock-fill embankment dam

ABSTRACT

The aim of this paper is to present a novel attempt for parametric estimation in the hydrostatic-season-time (HST) model. The empirical HST-model has been widely used for the analysis of different measurement data types on dams. The significance of individual parameters or their sub-groups for modelling the influence of the water level, air and water temperature, and irreversible deformations due to the ageing of the dam, depends on the structure itself. The process of finding an accurate HST-model for a given data set, which remains robust to outliers, cannot only be demanding but also time consuming. The Bayesian model class selection approach imposes a penalisation against overly complex model candidates and admits a selection of the most plausible HST-model according to the maximum value of model evidence provided by the data or relative plausibility within a set of model class candidates. The potential of Bayes interference and its efficiency in an HST-model are presented on geodetic time series as a result of a permanent monitoring system on a rock-fill embankment dam. The method offers high potential for engineers in the decision making process, whilst the HST-model can be promptly adapted to new information given by new measurements and can enhance the safety and reliability of dams.

1. Introduction

The safety control of constructed facilities, including large dams, is based on measuring activities, modelling of registered observations and engineering judgement of the results. In the process of modelling registered observations, a reliable mathematical, i.e. functional and stochastic model, should be developed according to the nature of the observations and the observed physical process. The evaluation model is an essential part for a quantitative safety and reliability assessment of engineering structures. An overview of basic models used for the analysis of monitoring data on dams is described in [3]. The main intent of the mathematical modelling of measurements is to compare the actual response of the dam, measured by different sensors, with the predictions of the model, with the aim of detecting anomalies early and the prevention of failures [19]. The prediction models should enable the separation of effects due to different factors [13]. The reversible influences of water level and temperature and the irreversible time influence as the main three external forces on dams can be decomposed using a well-known empirical hydrostatic-season-time (HST) model. The parameters of a reliable HST-model used should not be treated as constant and should be re-evaluated according to the nature of measurements, frequency of observations, and changeable or even additional influences on the structure. Multiple linear regression (MLR) is a widely used method for estimating unknown parameters in any functional model and can also be implemented for the HST-model. The most difficult task or challenge of the MLR is the selection of significant parameters. In the process of defining a reliable model, it is important to keep the number of variables as low as possible and to use the simplest possible model that is still consistent with the data [14].

The Bayesian probabilistic approach provides a framework for parametric identification in a prescribed functional model and uncertainty quantification [22,28]. While the accuracy measures such as R^2 and *RMSE*, which are based on minimising a norm of the fitting error between the output/measured data and the corresponding estimated prediction, select an optimal model class with often many parameters included, the crux of the Bayesian probabilistic approach is to select the most plausible model within a set of model class candidates. The implementation of the Bayesian probabilistic modelling for the structural reliability analysis of dams has been discussed in several research works with the main focus on the implementation of Bayesian networks. A comprehensive human risk analysis model using Bayesian networks for

* Corresponding author.
E-mail address: hannahzhou@umac.mo (W.-H. Zhou).
URL: http://www.fst.umac.mo/en/staff/fstwhz.html (W.-H. Zhou)

http://dx.doi.org/10.1016/j.ress.2017.07.018
Received 8 February 2017; Received in revised form 17 July 2017; Accepted 28 July 2017
Available online 1 August 2017
0951-8320/

estimating human risks due to dam-break floods is presented in [17], with the application to the landslide dam failure in [18]. In the analysis of earth-fill embankment dams [9], a non-parametric Bayesian belief network was implemented in order to assess the interaction of several influences in the risk assessment models. In [11], a Bayesian approach was employed to develop a dynamic probability description, which can evaluate and predict the dam safety levels. To update the reliability analysis of a steel miter gate, the Bayesian updating was applied to the Condition Index visual inspections for locks and dams [10]. Further, the potential of Local Sensitivity Analysis for analysis of uncertainty with respect to river flooding and dam failure risks is assessed in [8]. The probabilistic approaches for life-cycle maintenance of structural systems are discussed in [4].

In this work, the Bayesian probabilistic approach is used to define a reliable HST-model for displacement time series analysis. Since there is a fixed parameter set, with measured input and output values, the authors suggest a parametric modelling of influences. The proposed Bayesian model class selection can present an objective and prompt tool for real-time estimation and updating of the influences on the dam. The method is computationally inexpensive and can be implemented straightforward into the monitoring systems.

The paper is organised as follows. After the introduction, the mathematical background of the HST-model and MLR are described in Section 2. The theoretical foundations of the Bayesian probabilistic approach and Bayesian model class selection are given in Section 3. In Section 4, the numerical example and analysis of implementation of the Bayesian probabilistic approach in the HST-model is given. In the numerical example, the geodetic data of a permanent monitoring system on a rock-fill embankment dam are used. The implications of this work are discussed in Section 5.

2. Hydrostatic-season-time model and multiple linear regression

2.1. Hydrostatic-season-time model

Statistical (empirical) models, which are derived from the knowledge about a dam's behaviour and are based on the previous measurements, have been used for decades in dam safety monitoring. In particular, the hydrostatic-season-time (HST) model has been fully implemented and widely used in the assessment of different types of observations on dams. It is based on the assumption that the dam response is a consequence of three main influences on the dam: the influence of water level in an impounding reservoir (a), influence of water and air temperature (b) and time influence (c). Deformations under (a) and (b) have a reversible nature, whereas the time influence causes the irreversible deformations. We can outline the main features of the HST-model as follows [3,20]:

– The HST-model is simple to be used by engineers for the prediction of dam performance.
– The HST-model can detrend and estimate different influences on the dam, and above all it can separate the irreversible deformations from the reversible responses.
– The HST-model admits the dam response simulation under particular loads.

A detailed review on dam behaviour models, based on monitoring data, and references on further practical implementations, are given in a comprehensive paper by Salazar et al. [20]. In the paper, the advantages and also the limitations of the HST-model are discussed. The three effects are considered as independent, although it is known from the experience of engineers that a certain correlation exists and it is difficult to separate the individual effects unambiguously. The water level in an impounding reservoir affects the thermal response of the dam along the vertical due to a changeable surface influenced by the air or water temperature. This fact features especially the behaviour of concrete arch dams, where the temperature influence is more significant. In some cases, i.e. if the dam is constructed for electricity production, the strong correlation between water level fluctuations and the annual cycle of the air temperature is present. Consequently, the multicollinearity between the variables may impact the usefulness of the HST-model. It may lead to high sensitivity of unknown coefficients to the data and poor prediction [14].

The mathematical formulation of the HST-model is as follows. The effect of the hydrostatic load is considered with a fourth-order polynomial regression [3,6]:

$$H_i = a_1 \cdot h_i + a_2 \cdot h_i^2 + a_3 \cdot h_i^3 + a_4 \cdot h_i^4. \tag{1}$$

Here, h_i presents a relative water level for a time point t_i [7]:

$$h_i = \frac{h_{\max} - h(t_i)}{h_{\max} - h_{\min}}, \tag{2}$$

where

$h(t_i)$ ... measured water level for a time point t_i, [m],

$h_{\max}$, $h_{\min}$ ... maximal and minimal water level, [m], for the time period of modelling,

$t_i : i = 0, 1, 2, \ldots, N$ ... consecutive time stamp of observations.

In the HST-model, the thermal effects of the water and air temperature are approximated with a linear combination of sinusoidal functions, S_i, that depend only on the day of the year [3,6]:

$$\begin{aligned} S_i = {} & b_1 \cdot \sin(\omega_a \cdot t_i) + b_2 \cdot \cos(\omega_a \cdot t_i) + b_3 \cdot \sin^2(\omega_a \cdot t_i) \\ & + b_4 \cdot \sin(\omega_a \cdot t_i) \cdot \cos(\omega_a \cdot t_i), \end{aligned} \tag{3}$$

with annual frequency ω_a.

For the modelling of long term irreversible deformations of a time influence, different, strictly monotone functions are proposed [3]. The influence of time or irreversible deformations can be modelled, as suggested by Bonelli et al. [6], by the sum of a linear term, a positive exponential and a negative exponential of reduced time during the analysed period:

$$T_i = c_1 \cdot \tau_i + c_2 \cdot e^{\tau_i} + c_3 \cdot e^{-\tau_i}, \tag{4}$$

with $\tau_i = \frac{t_i - t_0}{t_N - t_0}$ as reduced time during the analysed time period $[t_0, t_N]$.

The HST-model, with a constant a_0 and error term ϵ_i can be written for a time step t_i as

$$y_i = a_0 + H_i + S_i + T_i + \epsilon_i, \tag{5}$$

with the model prediction $\hat{y}_i = a_0 + H_i + S_i + T_i$ corresponding to the measurement y_i with an error ϵ_i: $y_i = \hat{y}_i + \epsilon_i$. In a case study (Section 4), y_i represents the measured relative coordinate at the time step t_i in one direction of a predefined three-dimensional coordinate system.

2.2. Multiple linear regression

The unknown parameters of the HST-model (12 parameters in Eq. (5)) can be estimated by multiple linear regression (MLR), which is a statistical technique for investigating and modelling a linear relationship between the measurements and observations. The estimated model gives the predicted/estimated value(s) of dependent variable(s) in terms of several explanatory or measured variables. The process of fitting the model to the data can be done according to the parameter estimation techniques. The most widely used technique is the least square method, which minimises the sum of residual squares, $\sum_{i=1}^{N} \epsilon_i^2 \rightarrow \min$, where N is the size of measured data. A functional regression model is given in matrix form as $\mathbf{y} = \mathbf{X} \cdot \mathbf{b} + \epsilon$, with the vector of N_b-unknown regression coefficients, $\mathbf{b}_{[N_b \times 1]}$, and the vector of dependent variables or measurements, $\mathbf{y}_{[N \times 1]}$. The residuals, given in the vector $\epsilon_{[N \times 1]}$, are assumed to have a Gaussian distribution with zero mean and variance σ_ϵ^2, $\epsilon_i \sim N(0, \sigma_\epsilon^2)$. The errors are uncorrelated, $\sigma(\epsilon_i, \epsilon_j) = 0 \ \forall i, j \ni i \neq j$. The matrix $\mathbf{X}_{[N \times N_b]}$ is the design matrix or matrix of regressor variables [14].

After computing the regression parameters, an assessment of the model prediction accuracy should be performed, since it provides a measure of model accuracy, it allows comparison of different models, and it