

The prognostic modelling of piezometric levels based on seepage monitoring in earthen dams

Alla Demianiuk^{1,*}, and Dmytro Stefanyshyn^{1,2}

¹National University of Water and Environmental Engineering, Hydraulic construction and Hydraulics Department, 33028 Rivne, Ukraine

²Institute of Telecommunications and Global Information Space, 03186 Kyiv, Ukraine

Abstract. This paper presents an innovative approach to the prognostic modelling of piezometric levels in earthen dams equipped with automated monitoring systems. The main idea of the approach and the expected prognostic results are illustrated with the example of prediction of piezometric levels in the earthen dam of the Kyiv hydropower plant. This usually complex prediction task is simplified in this approach by means of simple regression models and combined situational and inductive modelling which enables overcoming the excessive uncertainty in time series. To calibrate the interpretation and prognostic models, daily monitoring data of the piezometric levels over a period of eight years was used. To verify the prediction results, monitoring data collected in the three years following this eight-year period was used. The goodness of fit of interpretation models was performed by R^2 testing. To assess the goodness of the prediction fit, mean absolute and relative error estimators, as well as the Nash-Sutcliffe efficiency coefficient, were employed.

1 Introduction

Dams, which are widely used in various economic sectors and spheres of human life, are among the most typical objects relating to the hazard potential structures. Today, there are approximately 850,000 dams located around the world; more than 40,000 are categorised as large dams [1]. Earthen dams are the most prevalent type of water retaining structure in the world. In Ukraine, the share of earth dams is estimated to be more than 95% of the total number of dams of various types [2].

The widespread use of earthen dams considering the devastating accidents that have occurred in the past [3, 4], makes the problem of the reliable operation of dams particularly important. Yet, there remain a lot of issues regarding the reliability of earthen dams because of their vulnerability to such impacts as flood overtopping (30%) and seepage (60%) [3]. Moreover, 45% of accidents were directly caused by seepage, and in 15% of cases, seepage was recognised as an indirect cause. Relevant monitoring is identified as being the main factor in facilitating the survivability of earth dams during accidents [5-7]. In particular, in [5] it is stated that the majority of damaged dams had no or dysfunctional monitoring and warning systems. The availability of advanced monitoring facilities makes it possible to

* Corresponding author: a.v.demianiuk@nuwm.edu.ua

expand the class of tasks of reliability assurance of earthen dams with the addition of modelling and the prediction of the behaviour of dams under long-term operation.

The paper summarises the main results of the prognostic modelling of piezometric levels in one of the earthen dams, where regular piezometric observations with automated monitoring system (AMS) have been provided for a long time.

2 Overview of the problem

Since earthen dams are prone to be damaged by seepage, special attention must be paid to the monitoring of seepage in dams [2, 5-7]. Piezometric monitoring is the most common and effective type of instrumental control of seepage in earthen dams [2, 4, 5].

Until recently, the shortage of monitoring data has often been denoted as one of the main shortcomings in the implementation of the basic principles of technical diagnostics [8], which include regular control of the diagnostic parameters and the detection of changes in their behaviour. Currently, AMSs provide the acquisition and reliable storage of large arrays of different data. As a consequence, another issue has arisen regarding the proper handling of so-called big data. As the amount of data expands, data homogeneity reduces but heteroscedasticity and multicollinearity increase.

The main diagnostic models based on instrumental data that are used nowadays for the modelling of earthen dam behaviour are regression models [2, 5, 9]. The classical approach to model construction is focused on developing sophisticated models of a multiple regressions type, which usually include a large number of explanatory variables [10].

As experience shows, the construction of adequate regression models for prediction purposes, especially for extrapolation, can become challenging even in simple cases [2, 10-13]. Good interpretation models may have a tendency to over-predict low values and under-predict high values or vice versa [10]. When there is a large amount of statistical data it often results in that the more sophisticated regression model is developed based on this data, the more complicated to achieve the goodness of prediction fit. In this regard sophisticated regression models may be quite successful for data interpretation, but not for prediction.

Regression models are highly adaptable – they can be modified in many ways depending on the particularity of the forecast task [10]. Thus, regression models are still the most convenient type of model to solve practical prognostic problems in various fields [10].

3 Case study

3.1 Earthen dam description

The earthen dam under examination is the main dam of the Kyiv hydropower plant. The HPP is located upstream of the city of Kyiv, Ukraine, on the Dnipro river. The dam is distinctive due to its considerable length, which is 17.2 km. The dam abuts upon 52 km-long left-bank earth dike. In its entirety, it is the longest water-retaining structure in Ukraine and one of the longest dams in the world.

The dam is of a homogeneous hydraulic fill-type. The maximum height of the dam is 13.5 m; the dam crest is 7 m wide. The upstream slope ratio varies from 1:3.75 to 1:4.5; the downstream slope ratio is 1:3.5 and 1:15 in different sections. The upstream slope is faced with concrete slabs. A toe pipe drain is employed to catch seepage water. The open drainage canal is 20-60 m wide with an average depth of 6 m and collects seepage water beneath the dam.

From the beginning of the dam operation, seepage monitoring is performed including piezometric levels and seepage quantities measurements. In 2002 the dam was equipped with AMS [2, 14]. In total, the AMS monitors 107 piezometers in 28 piezometric cross-sections including 13 cross-sections in the left-bank earth dike. Measurements are performed hourly.

To represent the new approach one piezometer was chosen arbitrarily, without any special considerations. However, additional research concerning other piezometers installed in the dam was also performed in order to verify the validity of the approach. The input data for modelling are time series, specifically, the time series of observed piezometric levels in the control piezometer and reservoir water levels. The time series cover daily automated observations for the period 2002-2009. Based on this data, preliminary investigations and the calibration of interpretation and prognostic models were performed.

3.2 Background of studies

Preliminary research (or pre-modelling) is an important part of the modelling process based on monitoring data. Pre-modelling involves choice of model variables as a crucial part of ensuring the adequacy of the models. It can be applied to both dependent and independent variables (predictors). Pre-modelling may result in the removal of some predictors in order to facilitate the simplification of the model structure. First of all, it is about the removal of intercorrelated predictors. Often, if predictors are highly intercorrelated, it means that they reproduce the same underlying information. The main advantage of removing predictors in regression modelling [10] is that it can facilitate model performance as well as solution stability. However, all considered simplifications should be justified in terms of predictive validity.

Pre-modelling shows [2, 9, 11-13, 15-17] that the most appropriate dependent variable for modelling water level in a piezometer is the hydraulic head difference (or hydraulic head drop) h along the seepage path from the entry point to the piezometer:

$$h = USWL - PWL, \tag{1}$$

where $USWL$ is the upstream water level; PWL is the water level in the piezometer.

Therefore, the predicted water level PWL_p in the piezometer is given by:

$$PWL_p = USWL - h_p, \tag{2}$$

where h_p is the predicted head drop.

The choice of h (1) as a dependent variable is explained by its high correlation (Fig. 1 b) with $USWL$ that is stronger than the correlation of PWL with $USWL$ (Fig. 1 a).

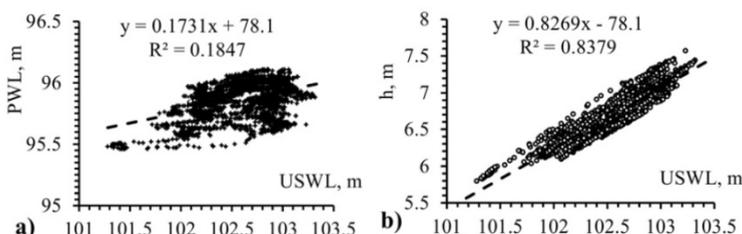


Fig. 1. Scatter plot and regression between: (a) PWL and $USWL$; (b) h and $USWL$ (data 2002-2008).

Large time series can be divided into separate time segments to meet the appropriate constraints and assumptions of regression modelling. The shorter the obtained samples, the more monotonous and homogeneous they may be. This can mainly be explained by the fact

that fewer factors remain influential within shorter time intervals [2]. Segmentation was made according to various behaviours of regression models at different time intervals. Time series of modified dependent variable h before and after segmentation are presented in Fig. 2.

The implementation of these techniques provides the simplest single-factor linear regression models, representing strong dependence between variables (with determination coefficient $R^2 = 0.9906$).

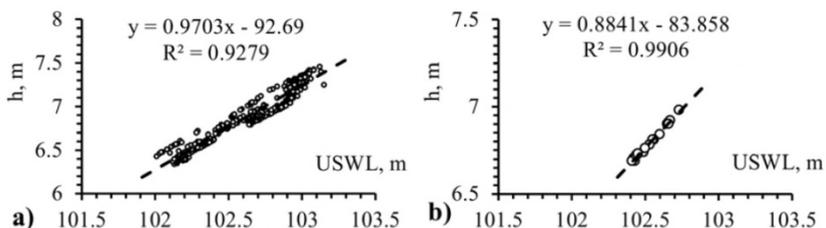


Fig. 2. Scatter plot and regression of h depending on $USWL$ (a) for full time series for 2004; (b) for separate segment of time series for the period 20.11.2004-08.12.2004.

4 Methods

The combination of situational and inductive modelling is implemented in the approach.

In situational modelling it's considered that sequential situations describe system evolution [18]. Complete description by an infinite set of all possible situations is replaced by a certain finite set of generalised model situations that reproduce possible states of the system [18, 19]. To describe the evolution of the system it is assumed [18] that on the basis of past facts, through which past model situations are described, it is possible to predict some similar situations that will appear in the future.

In our study situational models are presented by the simplest single-factor regression models of piezometric head drop (1) in which every regression is adapted to a specific model situation corresponding to a limited time span. Other uncontrollable factors taken all together are considered as a specific predictive background [2, 11-13]. The predictive background is determined as a set of external and (or) internal conditions stipulating the resultant regression model. Therefore, various regression models corresponding to various predictive backgrounds (or situations) are adopted as situational models.

In our study, the definition "inductive model" is related to a generalised model obtained on the basis of a set of situational models [20]. Actually, inductive models determine the behaviour of the dependent variable h through outcomes of situational modelling for some fixed "levels" of predictors (given discrete values of $USWL$) [11-13, 20]. Inductive models can be presented as a set of trends if the time factor is considered, or as a set of regression models [13], if the time factor is not essential. More generally, inductive models may be a composition of trends and random "balances" (residuals obtained after trend extraction), or trends and regression models for random "balances" [12, 20].

5 Results and discussion

5.1 Retrospective modelling

To calibrate the models retrospective modelling was performed. Daily monitoring data of the piezometric levels (PWL) and head water levels ($USWL$) from 2002 to 2009 were used for the modelling. In total, 83 specific model situations were considered. For all situations,

the best corresponding situational regression model is the following linear regressions:

$$h_i = a_i + b_i \cdot USWL, i = 1, \dots, 83, \tag{3}$$

where i is the number of the model situation; h_i is the head drop (1); a_i, b_i are regression coefficients.

Linear regressions (3) for all model situations show high values of determination coefficients R^2 (from 0.85 to 0.99). Figure 3 shows a set of “seasonal” situational models corresponding to the same time interval (from the end of November to the 1st third of December) for each year of the period of monitoring. Below, Figs. 4 and 5 present sets of time series and trends of outcomes of retrospective situational modelling, where each time series reflects the behaviour of simulated value h_m according to the chosen situational regression h_i (3) for three given values of the independent variable $USWL$.

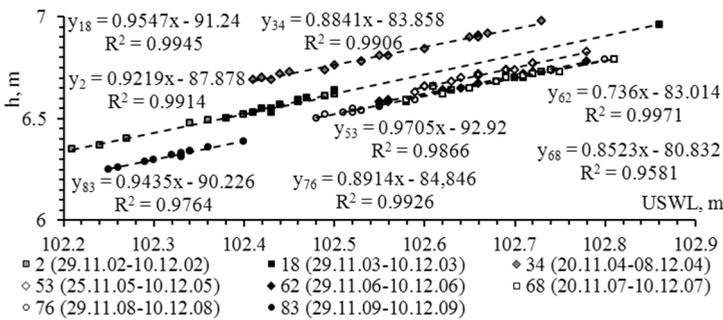


Fig. 3. A set of situational models for seasonally appearing situations.

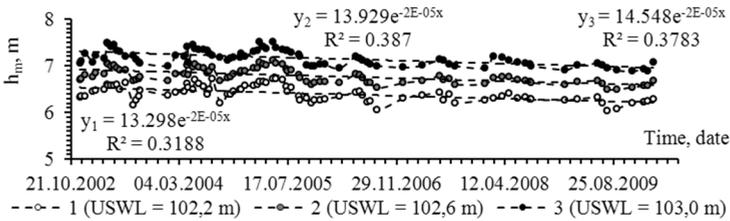


Fig. 4. Full time series of simulated head drop h_m and trends based on 83 situational models.

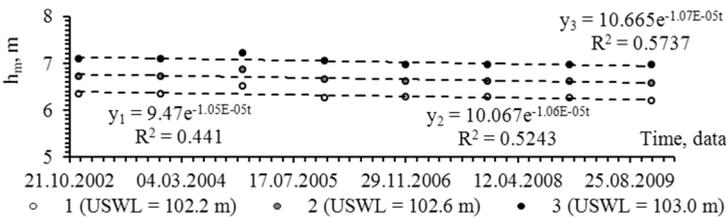


Fig. 5. Time series and trends of the results of “seasonal” modeling based on models on Fig. 3.

According to the results of retrospective situational modelling, the following general structure for an inductive model of piezometric head drop h is proposed. This is a set of equations representing time trends of a modelled variable $h_{m,k}$ at given k -th value of $USWL$:

$$h_{m,k}(t) = T(h_{m,k}(t)) + R(\Delta h_{m,k}(Z_m)), k = 1, \dots, K, \tag{4}$$

where $h_{m,k}(t)$ is presented with a k -th time series (Figs. 4 and 5) of simulated data inferred from situational models of the head drop h ; $T(h_{m,k}(t))$ is the trend of the variable $h_{m,k}(t)$; $R(\Delta h_{m,k}(Z_m))$ is a regression for “balances” (residuals) $\Delta h_{m,j,k} = h_{m,j,k} - T(h_{m,k}(t_j))$, where $h_{m,j,k}$

is the value of $h_{m,k}(t)$ at time t_j ; $T(h_{m,k}(t_j))$ is the value of the function $T(h_{m,k}(t))$ at time t_j ; Z_m is some explanatory variable, used for modelling “balances”. In this case study this explanatory variable is the variable considering lag effects between previous ($i-1$) and current (i) model situations.

5.2 Prediction results and verification

Below, Figs. 6 and 7 provide results of the prognostic modelling of the piezometric head drop (h) and, thus, the piezometer water level (PWL) depending on the head water level using composed inductive models of type (4). Prognostic situational models cover different parts of December 2010, 2011, and 2012.

To calibrate regressions $R(\Delta h_{m,k}(Z_m))$, monitoring data of adjacent previous periods October-November of 2010, 2011, and 2012 were used. Relevant pre-forecast situations were revealed to determine the model variable Z_m .

To measure the quality of the prediction performed with the proposed approach, and to verify whether forecast results could adequately represent the observed data, the average absolute standard and relative error estimators and the Nash-Sutcliff model efficiency criterion [21] were used. The average absolute standard and relative errors of the PWL prediction turned out to be quite small. The average absolute standard errors are around 0.006 m (0.0059 m, 0.0054 m, and 0.006 m for 2010, 2011 and 2012, respectively) and the relative errors do not exceed 2.3% (2.3%, 1.4%, and 1.7% for 2010, 2011 and 2012, respectively).

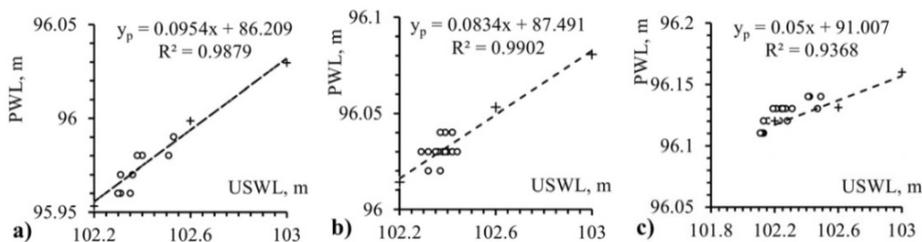


Fig. 6. Scattering plots and situational models of head drop h . Forecasts (a) for period 01.12.2010-10.12.2010; (b) for period 11.12.2011-29.12.2011; (c) for period 01.12.2011-21.12.2012. Index “o” relates to models based on observed data, index “p” – relates to prognostic situational models.

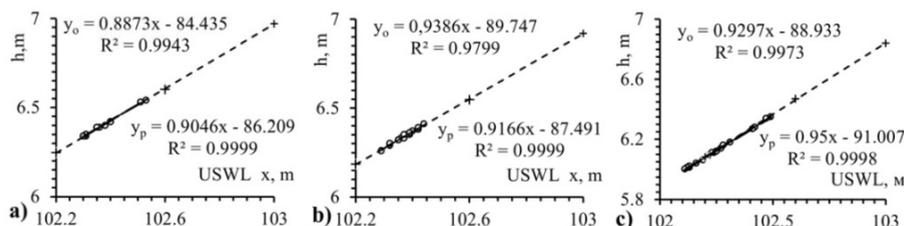


Fig. 7. Scattering plots and situational models of PWL (a, b, c) same as in Fig. 6.

The Nash-Sutcliff model efficiency criterion (NSE) [21] that is widely used for assessment of the predictive power of hydrological models was also employed. Predictions with NSE above 0.8 can be considered as being very good; with NSE above 0.65 as being average and slightly good. Below, Fig. 8 shows graphical illustrations of the NSE assessment of the prediction outcomes for several years. According to all applied criteria the predictive power of the model obtained with the proposed approach can be considered as being acceptable.

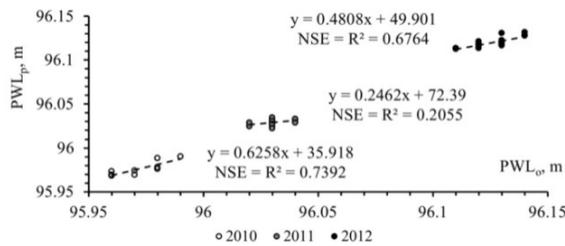


Fig. 8. Results of the assessment of the Nash-Sutcliffe model efficiency criteria according to prognosis results of different years: Index “o” relates to observed data, index “p” – to predictions.

5.3 Discussion

The task of situational prediction based on observational data is an extrapolation task, which can be defined as the identification of the most probable situational model that will meet some expected situation in the future on the basis of situations in the past. Inductive models create an effective basis for predicting perspective situational models. Either unambiguous situational models or a set of perspective situational models corresponding to various expected situations in the future can be the outcome of prediction. When applying composed models of type (4), the previous situation can definitively determine the following situational model if there is a transport lag between adjacent situations.

It is assumed that the shorter the time intervals corresponding to the model situation are the more likely it is that the relevant predictive background would be constant within it. Therefore, more accurate inputs for predicting can be provided. Our study shows that by having complete monitoring data sets it is easier to provide monotony and homogeneity of data for short time intervals, and to optimise models through the elimination of unimportant factors.

It should also be noted that the accuracy of the situational prediction made by inductive models can be improved when inductive models are built on the basis of situational models corresponding to similar segments of actual data (for example, related to the same period (month, season) of each year of observations).

Single-factor regression model is determined as the most efficient type for situational modelling in our research. The most appropriate independent variable for situational modelling is the reservoir water level. Besides, it is the only independent variable (from others considered as inducing *PWL* changes), which is adjustable during dam operation if required.

6 Conclusions

Having complete monitoring data provided by AMS on earthen dams creates both new opportunities and new challenges for modelling and especially for the prediction of dam behaviour as data monotony reduces. The selection of the appropriate dependent variable along with time series segmentation may be prerequisites for the essential simplification of the regression models traditionally used for modelling based on monitoring data.

A combination of situational and inductive modelling is implemented in the approach. Situational models representing dam behaviour within short time intervals are the simplest single-factor regressions. Inductive models developed on the basis of situational models are efficient as predictive models. Future situational models are the outcome of prediction according to this approach.

The predictive validity of the proposed approach is justified by the goodness of the fit criteria.

References

1. L. W. Mays, *Water Resources Engineering*, (NY, John Wiley & Sons, 2001)
2. D. V. Stefanyshyn, *Hydrop. of Ukr.* **3**, 27-32 (2012)
3. K. Hood, R. A. Perez, H. E. Cieplinski, T. V. Hromadka II, G. E. Moglen, and H. D. McInvale, *JAWRA* **55**(1), 89-101 (2019)
4. A. V. Demianiuk, D. V. Stefanyshyn, *Env. Saf. Nat. Resour.* **30**, 30-43 (2019)
5. *Dam safety – Guidelines*, (ICOLD, Bull. 59, 1989)
6. *Automated dam monitoring systems – guidelines and case histories*, (ICOLD, Bull. 118, 2000)
7. P. Choquet, R. Taylor, *Geo-Congress ASCE*, (180-191, 2014)
8. R. A. Collacott, *Structural Integrity Monitoring*, (Chapman and Hall, 1985)
9. S. Bonelli, and P. Royet, *Proc. Int. Symp. on Dam Safety*, (Geiranger, Balkema Rotterdam, 91-100, 2001)
10. M. Kuhn, K. Johnson, *Applied Predictive Modeling*, (Spr. Sc.+Bus. Media, NY, 2013)
11. D. V. Stefanyshyn, *Proc. of 4th ICIM*, (221-224, 2013)
12. D. V. Stefanyshyn, A. V. Demianyuk, *Env. Saf. Nat. Resour.* **24**, 138-147 (2017) (in Ukrainian)
13. D. V. Stefanyshyn, *MME*, **9**, 50-61 (2017)
14. https://www.bkc.com.ua/en/direction_monitoring/
15. G. Tayfur, D. Swiatek, A. Wita, and V. P. Singh, *J. Hydraul. Eng.* **131**(6), 431-440 (2005)
16. I. S. Jung, M. Berges, J. H. Garrett Jr, Ch. J. Kelly, *Comp. in Civil Eng.* 25-32 (2013)
17. S. Wang, Y. Xu, C. Gu, T. Bao, *WSE*, **11**(4), 344-354 (2018)
18. J. McCarthy, *Situations, actions, and causal laws, Memo 2*, (Stanford University Artificial Intelligence Project, 1963)
19. R. Reiter, *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems* (MIT Press, 2001)
20. D. V. Stefanyshyn, *MME*, **17**, 69-81 (2019)
21. J. E. Nash, J. V. Sutcliffe, *J. Hydrol.*, **10**(3), 282-290 (1970)