J. HAKÁČ, E. BEDNÁROVÁ

APPLICATIONS OF MATHEMATICAL METHODS IN THE SUPERVISION OF THE TECHNICAL SAFETY OF WATER CONSTRUCTIONS

Ján Hakáč

email: jan.hakac@vzb.sk

Research fields: dams, numerical mathematics, time series.

Watermanagement Comp., s.e.

Karloveská 2, 842 04, Bratislava, Slovakia

Emília Bednárová

email: emilia.bednarova@stuba.sk

Research fields: dams, behaviour of dams, finite element analysis.

Department of Geotechnics,

Faculty of Civil Engineering,

Slovak University of Technology, Slovakia

ABSTRACT

In this paper, we present the use of time series (stochastic models) in order to predict water seepage in the wells of water constructions. The aim of this paper is to assess the agreement between the predicted and measured levels of the water seepage in the wells of the Liptovska Mara water construction. We assume that the predicted values will match the values that were measured. The ability to predict the level of the water in observation wells and the total filtration stability of water constructions by using real measures in a time series is a desirable step in ensuring adequate safety with the possibility of remedying potential defects at an early stage.

KEY WORDS

- *Smooth Transition Autoregressive Model,*
- *nonlinearity,*
- *linearity testing,*
- *water seepage*

1. INTRODUCTION

An evaluation and prediction of the filtration stability of water constructions is important for ensuring safety. However, this essential task is methodologically challenging. Applying mathematical models to predict events is not a particularly simple problem, when predictions for more than one period are required. A few attempts have been made to apply non-linear stochastic models to predict the stability of water constructions. The prerequisite for a mathematical model to be used for predictive purposes is the ability to exhibit limit cycle behavior. One of the most important nonlinear time-series models, which is capable of exhibiting limit cycle behavior, is the Self-Existing Threshold Autoregressive (SETAR) model. However, one limitation of this model is that transitions between various regimes take place in a discontinuous and sudden manner. For more realistic modeling, these transitions should be smooth. Here the nonlinear time-series model STAR (Smooth Transition Autoregressive) was suggested by Terasvirta (1994).

In this paper, we present the use of the STAR stochastic model in order to predict the water seepage in the wells of the Liptovska Mara water construction. The reliability of the results from the STAR model is compared with the results obtained by linear regression.

For the purposes of the study, the Liptovska Mara water construction was used to obtain the required measurements. Liptovska Mara was constructed from 1965-1975. The area of the reservoir is 22 km²; its max. depth is 45 m; and its capacity is 360 mil. m³. The Liptovska Mara earth fill water construction has as its main functions the balancing of the flows of the river Vah, protecting against floods, supplying drinking water, producing electric power from the flow and from the pumping over of the river Vah, producing electricity in other hydropower plants, and creating conditions for recreation, sports, and fish breeding in the Tatra region. Liptovska Mara supplies other hydropower plants of the Vah cascade (such as the Krpeľany and Žilina water constructions).

Fig. 1 Water levels in observation wells and corresponding level of the water in water construction Liptovska Mara.

The aim of this paper is to assess the agreement between the predicted levels of the water seepage by using the STAR model and the measured levels of the water seepage in the wells of the Liptovska Mara water construction. We hypothesize that the predicted values will agree with the values that were measured. The reliability of the results from the STAR model are compared with the results obtained by linear regression.

2. METHODS

Measures of the water level in the PS-21, PS-22 and PS-23 observation wells on the downstream slope of the Liptovska Mara water construction were performed over 204 months (1991 - 2007). The measures were done in one-week periods. A total of 887 measures were obtained.

The values from the first sixteen years (1991-2006) (Fig. 1) were used for setting the parameters of the STAR model for predicting the seepage in the observation wells in the cross-section II. of the Liptovska Mara water construction in the year 2007.

2.1 Linear Regression Model

Linear regression refers to any approach in modeling the relationship between one or more variables denoted y and one or more variables denoted x , such that the model depends linearly on the unknown parameters to be estimated from the data.

Given a data set $\{y_i, x_{i1}, \dots, x_{ip}\}_{i=1}^n$ of n statistical units, a linear regression model assumes that the relationship between the dependent variable y_i and the p -vector of the regressors x_i is approximately

linear. This approximate relationship is modeled through a so-called “disturbance term” ε_i -- an unobserved random variable that adds noise to the linear relationship between the dependent variable and the regressors. Thus, the model takes the form:

$$y_i = \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, \dots, n$$

where β is a p -dimensional parameter vector, and ε is called the error term.

2.2 Smooth Transition Autoregressive (STAR) Model

Nonlinear stochastic models were used for predicting the level of the water seepage in the water construction during one-week periods. The Smooth Transition AutoRegressive (STAR) model is defined as follows (Franses, et al., 2000):

$$y_t = G(z_t, y_t; \Psi) = \phi' w_t + (\Theta' w_t) G_k^L(\gamma, c; y_{t-d}) + \varepsilon_t$$

where $\{\varepsilon_t\}$ is a sequence of normal $(0, \sigma^2)$ independent errors, $\phi = (\phi_0, \phi_1, \dots, \phi_p)'$ and $\Theta = (\Theta_0, \Theta_1, \dots, \Theta_p)'$ are $(p+1) \times 1$ parameter vectors, $w_t = (1, y_{t-1}, \dots, y_{t-p})'$ is the vector consisting of an intercept, and the first p lags of y_t and $G_k^L(\gamma, c; y_{t-d})$ are known as transition functions. Depending upon the forms of the transition, different forms of STAR models are defined. The model used is defined by the transition function, where the transition function is the logistic function (Maringer, Meyer, 2006):

$$G_k^L(\gamma, c; y_{t-d}) = \left[1 + \exp \left\{ -\gamma \prod_{i=1}^k y_{t-d} - c_i \right\} \right]^{-1}, \quad \gamma > 0$$

In this equation parameter γ is the slope parameter, and $c = (c_1, c_2, \dots, c_k)'$ is the vector of the location parameters $c_1 < c_2 < \dots < c_k$. These restrictions, as well as restricting γ to be positive, are needed to identify the model. The transition function is a bounded function of y_{t-d} , which is continuous everywhere in the parameter space for any value of y_{t-d} (Terasvirta, et al., 2005).

The strategy for building a STAR model involves three steps:

- First, carry out the complete specification of a linear $AR(p)$ model. The maximum value of the latency p has to be determined from the data.
- Second, test the linearity for the different values of the delay parameter d . If linearity is rejected for more than one value of d , choose the one for which the p -value of the test is the lowest (Arango, Gonzalez, 2001).
- Third, consider the value of d as given and use a sequence of tests nested in the second step to choose between the ESTAR

Fig. 2 Scheme of cross section II. of the Liptovska Mara water construction

(Exponential Smooth Transition Autoregressive) and LSTAR (Logistic Smooth Transition Autoregressive) models.

2.3 Data analysis

1. The equations of the regression coefficient for the individual wells were developed using the Mathematica program. They are defined based on the STAR nonlinear stochastic model but also by linear regression. The regression problem was solved by the least squares method.
2. The one-week measures of the water level and the corresponding levels of the water in the water construction (used as an exogenous part in the time series) were used as input data for the analysis of the time series. The values from the first sixteen years (1991-2006) were used for setting the parameters of the STAR model for predicting the seepage in cross section II. (Fig. 2) of the wells at Liptovska Mara in 2007. The data were assumed to permit an optimal forecasting of the regression coefficient after fitting the least squares for creating the linear regression mathematical model. Utilization of the STAR model was assessed as follows: First, by decomposing the time series, the trend, cyclical components and seasonal components were eliminated from the input data. The time series was decomposed, because we can easily identify the regular behavior of a time series in particular elements. Decomposition is the first step in using the STAR mathematical model. This model uses residuals in which the systematic components are not recognizable.

The strength of the predicted values was obtained by comparing the STAR model with the measured values and also with the model based on a linear regression.

The actual measures from the year 2007 were then used for analyzing the agreement and calculating the errors of the predicted

measures assessed by the STAR model compared to the actual measured values. The accuracy of the predicted values obtained by the linear regression model and the STAR model was compared.

The Spearman correlation test was used to assess the correlation between the predicted and measured values. The agreement of the values was tested using the Bland-Altman statistics. A comparison of the predicted and actual values was attested by three self-existing criteria: 1. RMSE (Root Mean Square Error), 2. MSE (Mean Square Error) and 3. MAE (Mean Absolute Error). The prediction was made for the water level in the wells for a randomly generated water level in the reservoir by the Monte Carlo method in the first four months of 2008.

3. RESULTS

The equations for the regression coefficient detected by linear regression for the individual wells are listed in Tab.1.

In Fig. 3, the predicted values (red line – by the STAR model, green line – by linear regression) and the actual measured values (blue line) assessed in the time period in 2007 in the cross profile II. of the Liptovska Mara wells are depicted. The coefficient of the correlation between the values predicted by the STAR model and the actual measured values in the PS-21 well was $R = 0.77$; in well PS-22 and well PS-42 it was $R = 0.63$. These results show that in

Tab. 1 Equations of the regression coefficient detected by linear regression

Time series	Regression function
PS-21	$y_t = 505.78 + 0.03037x_t - 0.00104x_{t-1}$
PS-22	$y_t = 500.27 + 0.0314x_t - 0.0076x_{t-1}$
PS-23	$y_t = 504.41 + 0.0560x_t - 0.0246x_{t-1}$

Fig. 3 Predicted values (red line – by STAR model, green line – by linear regression) and measured values (blue line) in the time period of 2007 in the cross profile II. of the wells in Liptovska Mara.

cross section II., excellent forecasting results were obtained by choosing the STAR model. For the individual wells the correlation coefficients from the linear analyses are the following: PS-21: $R = 0.55$, PS-22: $R = 0.41$ and PS-23: $R = 0.46$.

Tab. 2 Calculated criteria in observation wells

Model	Well	RMSE [m]	MSE [m ²]	MAE [m]
STAR	PS-21	0.062	0.003	0.055
	PS-22	0.067	0.004	0.052
	PS-23	0.073	0.005	0.062
Linear regression	PS-21	0.291	0.085	0.284
	PS-22	0.320	0.102	0.309
	PS-23	0.248	0.061	0.238

Referring to Figure 3, it can be seen that a significant gap exists in the ability to predict real progress in the observation wells of water constructions by using a linear regression model. The coefficient of the determination reached the following values: $R^2 = 0.40$ (for wells PS-22 and PS-23) to $R^2 = 0.59$ (for well PS-21). From the listed results we can see that till 59%, the variability of the predicted values and the actual measured values are determined together (for well PS-21). The linear regression analyses for single observation wells in the cross profile II. allocation showed a lower coefficient of determination between the predicted and actual values compared to the STAR model. For the single observation wells the coefficient of determination was elicited from the linear regression analyses as follows: PS-21: $R^2 = 0.30$, PS-22: $R^2 = 0.17$ a PS-23: $R^2 = 0.21$.

In Tab. 2 the criteria RMSE (Root Mean Square Error), MSE (Mean Square Error) and MAE (Mean Absolute Error) are listed; they rate the effect of the models assimilation.

As shown in the Bland-Altman plot (Figs. 4, 5 and 6), 100% of the predicted values in the well PS-21 were within the agreed specified limits of ± 0.15 m from the absolute agreement between the predicted and measured values. For well PS-22 that was 96%, and for well PS-23 that was 98% of the predicted values, which were within the agreement limits.

In the final step the values of the water level by the STAR model were predicted in wells PS-21, PS-22 and PS-23 (Fig. 8) in the four-month period in 2008 (16 values), for randomly generated water levels in the reservoir at the Liptovska Mara dam (Fig. 7).

4. CONCLUSION

Using the STAR statistical model, the water level in the PS-21, PS-22 and PS-23 in cross section II observation wells at the Liptovska Mara dam was predicted through the verification period in 2007. The results obtained agree with the measured data. The STAR model developed was shown to be a useful method for predicting a short-term water level in observation wells.

Fig. 4 Correlation graph of the predicted values (P) and measured values (M) in the time period of 2007 and the Bland-Altman plot of the differences versus the average with the agreed limits (red line) in the cross profile II. of Liptovska Mara of the well PS-21.

Fig. 5 Correlation graph of the predicted values (P) and measured values (M) in the time period of 2007 and the Bland-Altman plot of the differences versus the average with the agreed limits (red line) in the cross profile II. of Liptovska Mara of the well PS-22.

Fig. 6 Correlation graph of the predicted values (P) and measured values (M) in the time period of 2007 and the Bland-Altman plot of the differences versus the average with the agreed limits (red line) in the cross profile II. of Liptovska Mara of the well PS-23.

Fig. 7. Randomly generated values of the water level at the reservoir of the Liptovska Mara dam for the first four months in 2008.

The presented paper shows the ability to use stochastic methods for predicting the actual water levels in the observation wells in profile II. at Liptovska Mara. According to the results, there was excellent agreement between the predicted and measured values in each well in cross profile II. of the dam construction.

From a large-scale set of the analysis of the data, it is evident that for a complex appraisal of the future progress of the seepage of water through the body and bedrock, the STAR model is preferable compared to the linear model. The results of the comparison from the STAR model with the actual dates of the water levels at the Liptovska Mara water construction suggest that the STAR model may be useful in predicting water levels in the near future. On the other hand, the prediction of the water levels using the linear regression model did not show such an agreement between the predicted and actual values. The reason for this observation may be based on the fact that in a linear regression model, both the stochastic and deterministic parts of the time series are predicted. In contrast to that STAR model, just the stochastic part of the time series is predicted. This may have an advantage in the prediction process.

We conclude that the STAR model can be used for prediction of the future levels of the seepage in wells of water constructions with a high degree of agreement. This study shows a method for predicting the filtration stability of a water construction. This can help in early identification of the anomalies of water constructions and early management of their potential negative impacts.

This paper was supported by grant project VEGA No. 1/0704/09.

Fig. 8. Predicted values of the water level in the Liptovska Mara wells for the first four months in 2008.

REFERENCES

- [1] ARANGO, L. E., GONZALEZ, A.: Some evidence of smooth transition nonlinearity in Colombian inflation. *Applied Economics*, 33. 155-162. 2001.
- [2] FRANCES, P.H., VAN DIJK, D., TERÄSVIRTA, T: Smooth Transition Autoregressive Models – A Survey of Recent Developments, *Econometric Institute Research Report EI 2000-23/A*, 2000.
- [3] MARINGER, D., MEYER, M.: Smooth Transition Autoregressive Models – New Approach to the Model Selection Problem, University of Essex, 2006.
- [4] TERÄSVIRTA, T.: Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89. 208-218. 1994.
- [5] TERÄSVIRTA, T., van DIJK, D., Medeiros, M. C.: Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination. *International Journal of Forecasting*, 21. 755-774. 2005.