ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 11, Issue, 12, pp. 52792-52797, December, 2021
https://doi.org/10.37118/ijdr.23496.12.2021

RESEARCH ARTICLE OPEN ACCESS

STATISTICAL MODELING AND PREDICTION OF THE BEHAVIOR OF A BUTTRESS BLOCK AT THE ITAIPU DAM

Sheila R. Oro*, Suellen R. P. Garcia, Dimilson P. Coelho, Francisco A. F. Reinaldo, Guilherme Lopes de Campos, Maiquiel Schmidt de Oliveira, Liliane Hellmann and Fabiana Fratta Furlan Peres

[1]Department of Physics, Statistics and Mathematics, University of Technology - Paraná, Francisco Beltrão, Paraná, Brazil; [2]Department of Mathematics, University of Technology - Paraná, Toledo/PR, Brazil; [3]Department of Civil Engineering and Architecture, Itaipu Binacional, Foz Iguaçu/PR, Brazil; [4]Department of Informatics, University of Technology - Paraná, Francisco Beltrão/PR, Brazil

ARTICLE INFO

Article History:

Received 10th September, 2021
Received in revised form
18th October, 2021
Accepted 24th November, 2021
Published online 28th December, 2021

Key Words:

Dam Monitoring, Factor Analysis, Time-Series, Auto-Regressive Distributed Lag.

***Corresponding author: Sheila R. Oro**

ABSTRACT

In this paper, the data from piezometers, extensometer bases, plumblines and multiple-rod extensometers, used in the monitoring of a buttresses block of the Itaipu Dam, were treated and analyzed using a set of models and statistical techniques, with the objective of obtaining a statistical index that represents the global behavior of the block's responses, in relation to the oscillations in the environmental conditions of its surroundings. The joint monitoring index of the dam block responses was constructed based on the five factors identified through factor analysis, which together explained about 85% of the variability. The forecasts of the values of this index for the 48-month horizon were obtained through dynamic linear models, having as regressors the water level in the reservoir and the air temperature around the dam, considering the lags in relation to the dependent variable. The adjusted model passed all specification and adequacy tests, demonstrating a high quality of estimates (adjusted $R^2 = 0.87$) and good accuracy (RMSE = 0.22). The index proposed in this study can be used in real-time monitoring of the dam and assist in the decision-making process.

Citation: ***Sheila R. Oro, Suellen R. P. Garcia, Dimilson P. Coelho, Francisco A. F. Reinaldo, Guilherme Lopes de Campos, Maiquiel Schmidt de Oliveira, Liliane Hellmann and Fabiana Fratta Furlan Peres.*** "Statistical modeling and prediction of the behavior of a buttress block at the Itaipu Dam*", International Journal of Development Research*, 11, (12), 52792-52797.

INTRODUCTION

Assessing the safety of a dam is a complex task, as it involves monitoring the responses resulting from the interaction of its structures and foundations with the environment and other sources of disturbance. According to Cheng and Zheng (2013), the relationship between a concrete dam and the environment in which it is located can be represented by an input-output system, in which the inputs are the environmental conditions: ambient temperature around the dam, concrete temperature, water temperature, reservoir water level, downstream water level, rainfall, seismic activity, time (age of works), among others; and the outputs are the responses of the dam, such as: stresses, strains, displacements, uplift pressure, increased flow, and so on. Modeling the behavior of a dam structure is a fundamental action of a safety system, especially as regards assessing the behavior over time. This requires a combination of knowledge - especially in engineering, computing, mathematics and statistics - and prior experience of the technical personnel responsible for analyzing these data (Villwock, Steiner, Dyminski and Chaves Neto, 2013).

In this work, the data from different instruments used in the structural monitoring of a concrete dam were treated and analyzed through a set of models and statistical techniques, in order to obtain an index that represents the global behavior of the responses of the concrete dam in relation to fluctuations in the environmental conditions of its surroundings.

Factor Analysis: The set of measurements carried out in a given period by the various instruments used in monitoring dams makes up a multivariate sample. When the random measurement of these instruments are correlated with each other and subjected to factor analysis, it is possible to form a new set of unobservable independent variables, called factors. Factor analysis allows the exploratory description of matrices and the identification of the main relationships of an extensive set of instruments with each other or the relationships between samples and variables. We take X_{px1} as a vector the elements of which are the instruments installed in a dam, with normal p-varied distribution, with covariance matrix$Cov(X) = \Sigma_{pxp}$ and vector of averagesμ_{px1}. Generally, these instruments have sharp discrepancies between the variances, caused by different measurement units. In this case, performing a standardization takes all data to the same measurement scale, through the $Z = (V^{1/2})^{-1}(X - \mu)$ transformation, where$V_{pxp}^{1/2}$ is the diagonal matrix of the instruments' standard deviations. In this way,Z_{px1} will also have normal p-varied distribution, vector of averages $E(Z) = 0_{px1}$ and covariance matrix $Cov(Z) = (V^{-1/2})^{-1}\Sigma(V^{-1/2})^{-1} = \rho_{pxp}$. In this case,$\rho_{pxp}$ is the correlation matrix of the instruments.

A presupposition for the application of factor analysis is that the instruments' measurements are correlated with each other. Bartlett's test of sphericity can be used to verify the significance of correlations (Johnson and Wichern, 2019). The adequacy of the sample for the adjustment of the orthogonal factorial model can be assessed using the Kaiser-Meyer-Olkin criterion, the values of which vary between 0 and 1, where the closer to the unit, the greater the adequacy of the adjustment of the factorial model to represent the multivariate set (Johnson and Wichern, 2019).

In the factorial model,

$$Z_{px1} = L_{pxm}\ F_{mx1}\ +\ \epsilon_{px1}$$

the standardized p instruments (Z_{px1}) are linearly dependent on unobservable random $m$variables ($m < p$), called common factors (F_{mx1}) and other sources of variation, called specific factors (ϵ_{mx1}). Correlations between factors and instruments are measured by Lfactor loadings and can be estimated by the maximum likelihood method (Johnson and Wichern, 2019), which consists of maximizing the likelihood function defined by,

$$LV(\mathbf{0}, \rho) = \frac{e^{-\frac{1}{2}\sum_{j=1}^{n}(z_j)'\ (LL'+\Psi)^{-1}\ (z_j)}}{(2\pi)^{np/2}\ |LL'+\Psi|^{n/2}}$$

The maximization of the maximum likelihood function is done by numerical procedures, from a random sample of the $n$size of the Zvector, for a fixed value of$m$, which can be obtained using Horn's Parallel Analysis Method (Çokluk and Koçak, 2016). Performing the varimax rotation, which is an orthogonal transformation, allows obtaining a factorial structure in which each factor has a group of instruments highly correlated with it and with negligible correlation with the other factors (Johnson and Wichern, 2019). This transformation consists of finding a certain Tmatrix, which maximizes$V$, such that$LL' = LTT'L' = \hat{L}\hat{L}'$, where

$$V = \frac{1}{p}\sum_{j=1}^{m}\left[\sum_{i=1}^{p}\hat{l}_{ij}^{4} - \frac{\left(\sum_{i=1}^{p}\hat{l}_{ij}^{2}\right)^{2}}{p}\right]^{3}$$

For each of the $m$factors, the numerical values for each sample element are calculated, called factor scores (f_j, with$j = 1, \cdots, n$), which can be estimated by the weighted least-squares method, by this method (Johnson and Wichern, 2019):

$$f_j = (L'\ \Psi^{-1}\ L)^{-1}\ L'\ \Psi^{-1}\ z_j$$

The measurements of each instrument also form a time-series, as these are observations accumulated sequentially over time in relation to a random variable (instrument), with the characteristic of serial dependence, and may present systematic components (trend, cycle and seasonality) and non-systematic (random noises) (Box and Jenkins, 1976). The various types of components can act either independently or together, so that predictions based on time-series require that the properties of their components remain relatively stable during the time the prediction is made, that is, that the series be stationary (Barbão, 2007).

A process is said to be stationary when it oscillates around a constant average and with a constant variance, such that its development in time does not depend on the choice of an origin of time. Generally, the time-series of instruments used in monitoring dams are not stationary (Li, Wang and Liu, 2013). Another relevant aspect is that a concrete dam responds with some delay to external requests (Lombardi, Amberg, and Darbre, 2008; Ribeiro, et. al, 2019). For these cases, the AutoRegressive Distributed Lag (ARDL) and Error Correction Model (ECM) dynamic linear models can be used to represent and predict time-series values.

Dynamic models: The AutoRegressive Distributed Lag (ARDL) contains among the regressors the lagged values of the dependent variable and the current and lagged values of the independent variables. For example, an ARDL(r,s) model with only one independent variable,$\boldsymbol{x_t}$, can be represented by:

$$y_t = \mu + \sum_{i=1}^{r}\alpha_i y_{t-i} + \sum_{i=0}^{s}\beta_i x_{t-i} + \epsilon_t, \quad \text{iid}(0, \sigma^2)$$

where: μ represents the independent term, and the r and$s$ indices represent, respectively, the maximum number of lag for the dependent variable,y_t, and the independent variable, x_t. Like any linear model, ARDL assumes independent residuals, identically distributed and with constant variance. When specifying the model, using the Bounds Testing approach, two aspects are fundamental: the determination of lag orders (r and s) and the estimation of the coefficients. As for the determination of the lag order, some alternatives include the choice of the model that maximizes the determination coefficient or minimizes the error variance estimate, and the Akaike information criterion (AIC). As for the estimation, linear transformations are applied in the ARDL model in order to alleviate the problems of multicollinearity between the predictor variables, and the ordinary least-squares estimation method is used, which, in addition to being consistent, is invariant to these transformations (Arone, 2014). Using the lag operator and polynomial deductions, the unrestricted error correction model (ECM-I) is obtained as follows:

$$\Delta y_t = \mu + \phi\left[y_{t-1} - \frac{B(1)}{A(1)}x_{t-1}\right] + \sum_{i=1}^{r-1}\delta_i\Delta y_{t-i} + \sum_{i=0}^{s-1}\gamma_i\Delta x_{t-i} + \epsilon_t$$

This reparametrization consists of linear transformations applied to the ARDL model, without the imposition of any restrictions. This procedure offers advantages in estimation and the ordinary least-squares estimator is invariant to linear transformations (Lopes, 1999). According to Hassler and Wolters (2006), this differentiation process transforms a linear combination of non-stationary variables into a model with stationary series. The estimation of the parameters of the ECM-I model can be done by ordinary least-squares, as follows:

$$\Delta y_t = \mu + \theta_1 y_{t-1} - \theta_2 x_{t-1} + \sum_{i=1}^{r-1}\delta\Delta y_{t-i} + \sum_{i=0}^{s-1}\gamma\Delta x_{t-i} + \epsilon_t$$

Generally, the adjusted model is submitted to some tests to verify the specification, adequacy, and quality of the estimates, among others. Godfrey (1978) suggests the application of diagnostic tests for the model, as some tests are sensitive to the presence of integrated variables. Among the tests that can be used for this purpose are: Regression Specification Error Test (Ramsey, 1969; Wooldridge, 2015), that of the Breusch-Godfrey autocorrelation (Asteriou and Hall, 2015; Wooldridge, 2015), that of the Breusch-Pagan-Godfrey heteroscedasticity (Breusch and Pagan, 1979; Godfrey, 1978; Gujarati and Porter, 2011; Wooldridge, 2015) and the Variance Inflation Factor (Gujarati and Porter, 2011; Vu, Muttaqi and Agalgaonkar, 2015). Segundo (Salazar et al., 2017), most predictive models for the behavior of dams do not mention which forecasting approach is used.

Predictions can be performed in or out of the sample. Dynamic forecasts (multi-step ahead), carried out outside the sample, have potential for application in the context of dam monitoring. In this type of forecast, previously predicted values of the dependent variable are used and not those observed (Noble, 2011).

MATERIALS AND METHODS

The data set used in this study was made available by the Center for Advanced Studies in Dam Safety (Centro de Estudos Avançadosem Segurança de Barragens - CEASB), Itaipu Binacional, referring to the values measured by 61 instrumentation sensors installed in two buttress-type blocks of the right-bank lateral dam, in the period from January 1990 to December 2017. Table 1 shows the quantity of instruments taken into account in this study and the respective phenomena monitored by these. Data recorded in this period for the air temperature around the dam and the water level in the reservoir were also considered. Analyzes and implementations were carried out with the help of R and EViews software. Initially, a data matrix was constructed with the monthly averages of the observations of each instrument sensor, consisting of 336 lines (monthly averages) and 63 columns (sensors). We chose to standardize the data due to the use of different kinds of instruments, with different values in terms of magnitude and measurement scale.

Then, the standardized data from 61 sensors (except air temperature and reservoir water level) were subjected to factor analysis. The factor scores formed the input data-set used in modeling the Dam Blocks Joint Response Monitoring Index (IMCRB). The model for the IMCRB was created through the weighted average of the factors, with weights derived from the eigenvalues of the sensors' correlation matrix. The time-series originated by the IMCRB has dimensionless values. Several statistical models were investigated, considering the stationarity of the series and the time lag of the IMCRB and the hydrometeorological variables (air temperature and water level in the reservoir), and the ARDL model was chosen. As this model is sensitive to the presence of an outlier in the series, the period from January 2001 to December 2013 was selected for the adjustment. Data from January 2014 to December 2017 were left out of the sample, for the validation of the forecast results. Due to the presence of multicollinearity, reparametrization was performed and the unrestricted error correction model (ECM-I) was generated). Predictions were performed only with the selected model that passed all verification tests.Accuracy was evaluated using the root-mean-square error (RMSE) measure as a reference).

OUTCOMES

The sensors were considered significantly correlated with each other by the Bartlett sphericity test, as$T=53949.51$ was above the critical value, at 95% confidence ($Tc=1930.63$). The data sample was considered adequate ($KMO=0.92$) for the application of factor analysis. Therefore, Horn's Parallel Analysis was performed, which pointed to the extraction of 5 factors. The resulting factorial model was able to explain approximately 85.21% of the variability observed in the sensor data set of the two buttress blocks considered in this study. Thus, there was a reduction in the dimension of the problem, with minimal loss of information, as the 5 factors came to represent the set of 61 sensors. The importance of each factor was measured through the eigenvalues and the respective proportion of the total explained variability (Table 2). The factors were interpreted according to the sensors most correlated with them, according to their respective factor loadings estimated by the maximum likelihood method.

In the first factor, most of the multiple-rod extensometers were considered. In the second, the sensors of the extensometer bases that measure the gap between blocks were considered. Most piezometers had higher correlations with the third factor. The fourth factor included plumbline sensors that measure radial displacements. Only two piezometers were considered in the last factor. Thus, it can be stated that for the analyzed dataset, most of the variability came from foundation creep, measured by extensometers with high factor loadings in F_1 (39.56%) and from the openings/closings of the contraction joints between blocks, measured by the extensometer bases strongly correlated with F_2 (27.68%). The time-series of these two factors (Figure 1) showed seasonal movements resulting from thermal variation with maximum annual values in the winter period and minimum in the summer. In addition, F_1 and F_2 showed downward and upward trends, respectively, similarly to the sensors most correlated with each factor. The factor, being a latent variable, cannot be measured directly. Thus, the factor values, called factor scores, were estimated based on the factor loadings and sensor data. Considering the factors as quantitative variables, which jointly represent the variability of the dam's responses to the various phenomena acting on it, these were used as parameters for the generation of the Joint Response Monitoring Index of the Dam Blocks (IMCRB). The coefficients were estimated based on the proportion of variance explained by each factor, resulting in the following model

$$IMCRB = 0.3956\ F_1 + 0.2768\ F_2 + 0.0858\ F_3 + 0.0569\ F_4 + 0.036$$

The IMCRB time-series showed combined characteristics of the factors, with similar seasonality to factors F_1 and F_2 and a slight downward trend until the year 2012. Observing the instruments that had a greater representation according to the factorial model, it was possible to list the phenomena that most influenced the behavior of the dam blocks represented by this index, namely: the variation in the level of the reservoir, environmental thermal influences and eventual strains in the rock-mass. To expand the contribution to the monitoring of the dam, a model was fitted to the IMCRB that would allow the forecast of values. At this time, it was decided to include the variables reservoir water level (*Level*) and air temperature (*Temp*) around the dam, considering that these were the main hydrometeorological conditions that influenced the behavior of the dam blocks in the period under study. The unit root test was applied to the IMCRB, Level and Temp time-series and indicated stationarity only for the two hydrometeorological variables, while the IMCRB series was considered to be integrated of order one. This miscellany of stationary and integrated series led to the use of the AutoRegressive Distributed Lag (ARDL) model, through the Bounds Testing approach. The selected model was ARDL (1, 2, 6), as represented by

$$IMCRB_t = \alpha_1 IMCRB_{t-1} + \sum_{i=0}^{2} \beta_i Level_{t-i} + \sum_{i=0}^{6} \gamma_i Temp_{t-i} + \sigma t + \mu$$

that is, an autoregressive model with a lag for the dependent variable IMCRB, two for the predictor variable reservoir water level (*Level*), six for the air temperature (*Temp*), with trend (t) and a constant term. This model was chosen from among the 2028 simulated ones, as it presented the lowest AIC (-1.67) and the highest coefficient of determination (0.95), being considered a parsimonious model with few parameters.

The statistic $F=0.003$ (*p-value* = 0.97) was obtained in the first RESET test and $F=1.19$ (*p-value* = 0.31) in the second. As both tests resulted in a probability greater than 5%, the model was considered well specified. Two parameters presented VIF greater than 10 in the ARDL model (1, 2, 6), indicating the presence of multicollinearity. Therefore, we chose to use the ECM-I, with the estimation of the coefficients using the ordinary least-squares method.

The reparametrized model, presented below, did not change the result of the AIC and presented an excellent coefficient of determination (0.87), in addition to a low RMSE (0.10), indicating excellent quality for the representation of the IMCRB data.

Table 1. Instruments taken into account in this study and respective monitored phenomena

Phenomenon	Instrument	Sensors	Unit of measurement
Radial displacement	Plumblines	5	mm
Tangential offset	Plumblines	5	mm
Opening and closing of joints between blocks	Extensometer bases	6	10^{-3}mm
Horizontal sliding between blocks	Extensometer bases	4	10^{-3}mm
Differential settlement between blocks	Extensometer bases	2	10^{-3}mm
Rock-mass deformations	Multiple-rod extensometers	20	mm
Uplift	Piezometers	19	m

Table 2. Eigenvalues and percentage of variability explained by each factor extracted

Factor	Eigenvalue	Explained variation (%)
F1	24.13	39.56
F2	16.87	27.68
F3	5.24	8.58
F4	3.47	5.69
F5	2.25	3.69

Table 3. Test results applied to the ECM-I-Mod model

Verification	Test	Statistic	Prob.
Residuals normality	Jaque-Berra	5.86	0.05
Independence of residuals	Breusch-Godfrey Serial Correlation (LM)	0.67	0.81
Residuel homoscedasticity	Breusch-Pagan-Godfrey	1.43	0.18
Model specification	RESET - quadratic term	2.87	0.09
	RESET - quadratic and cubic terms	1.90	0.15
Long term relationship	Wald (c1 = c2 = c3 = 0)	60.81	0.00
	Wald (c1 = 0)	-12.37	0.00

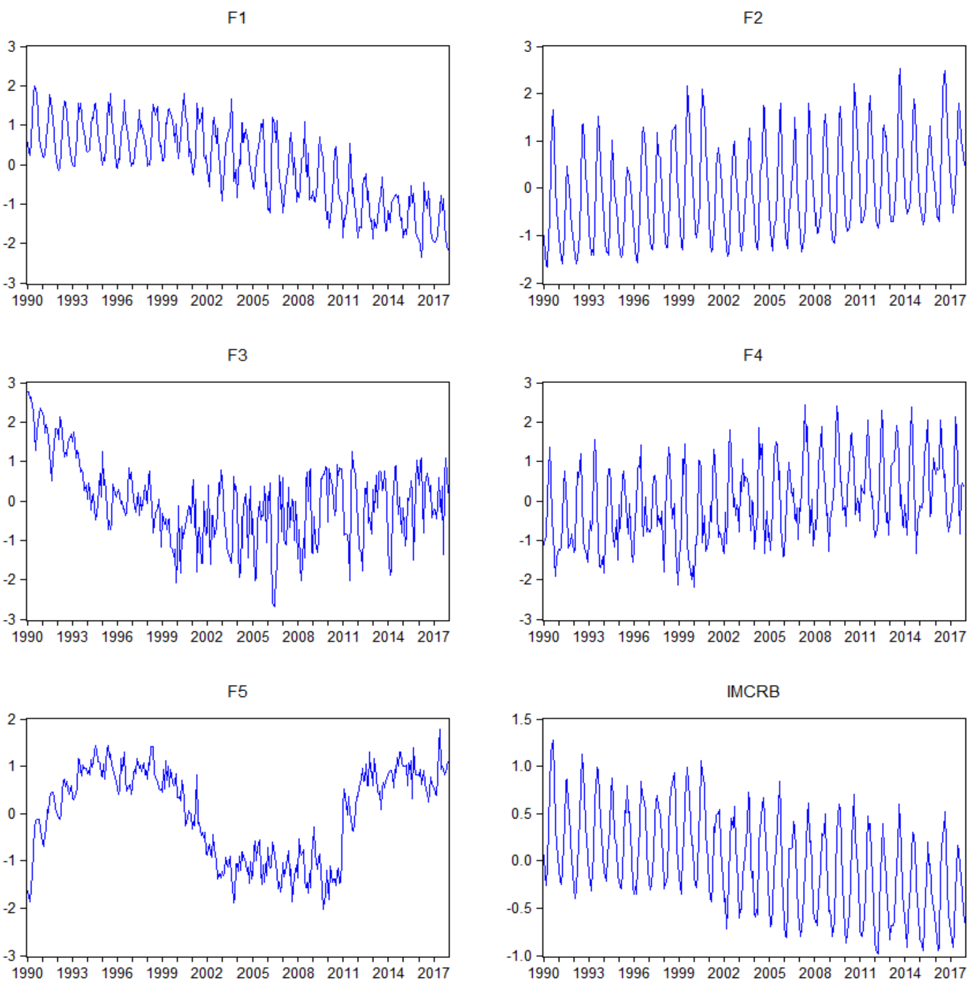

Figure 1. Time-series of factors and of the IMCRB

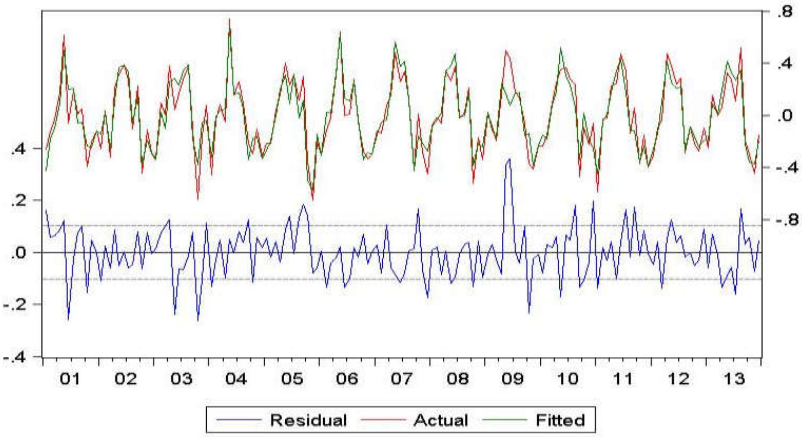

Figure 2. Graphic comparison of IMCRB with ECM-I-Mod model

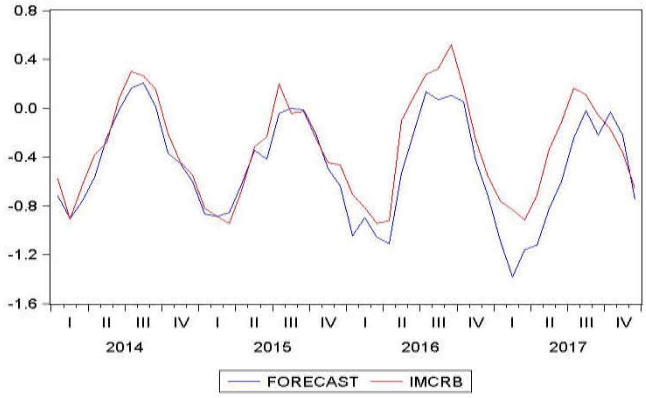

Figure 3. Monthly values predicted by the ECM-I-Mod model and actual values of the IMCRB, from 2014 to 2017

$$\Delta IMCRB_t = \theta_1 IMCRB_{t-1} + \theta_2 Level_{t-1} + \theta_3 Temp_{t-1} + \sum_{i=0}^{1} \gamma_i \Delta Level_{t-i} + \sum_{i=0}^{5} \delta_i \Delta Temp_{t-i} + \sigma t + \mu$$

Normality was verified (*p-value* 0.09) using the Jaque-Berra statistic (4.74). Independence was confirmed with the serial autocorrelation test (Breusch-Godfrey) up to lag 15 (*F=0.50*), which were considered non-significant (p-value = 0.94). Finally, homogeneity was confirmed (*p-value* = 0.19) through the Breusch-Pagan-Godfrey test (*F=1.37*). When applying the RESET to the ECM-I model, with a quadratic term (*p-value* = 0.21) and a cubic term (*p-value* = 0.40), the results did not reject the hypothesis of good specification. The application of the Wald Test pointed to the significance (*p-value=0*) of the first three coefficients (θ_1,θ_2,θ_3,) with*F=34.18*. The long-term relationship between the IMCRB and predictor variables (*Level* and *Temp*) was confirmed (*p-value=0*) by applying the Wald Test to the IMCRB lag coefficient ($\square_1$), with*t=-9.53*. The same two parameters of the ARDL model (1, 2, 6) presented VIF higher than 10 in the ECM-I model, indicating the continued existence of multicollinearity, which is detrimental to the realization of value predictions for the IMCRB. Therefore, new ECM-I models were performed, removing one term at a time and redoing the tests. The best result was obtained with the following specification for the model

$$\Delta IMCRB_t = -0.73 IMCRB_{t-1} + 0.03 Level_{t-1} - 0.10 Temp_{t-1} - 0.02 \Delta IMCRB_{t-1} + \\ - 0.03 \Delta Level_t - 0.08 \Delta Temp_t - 0.02 \Delta Temp_{t-3} - 0.01 \Delta Temp_{t-5} - 0.002t + 9.33$$

This model, called ECM-I-Mod, did not show multicollinearity between variables and passed all validation tests (Table 3), maintaining an excellent coefficient of determination (0.86) and low values for the AIC (-1.64) and RMSE (0.10), indicating excellent quality for the representation of IMCRB data. Figure 2 shows the graphics of the IMCRB (Actual) and the adjusted ECM-I model (Fitted), as well as the respective residuals. It is noted that, in general, the model is able to satisfactorily represent the behavior of the IMCRB series, with the largest errors occurring at the extreme points (local maximums and minimums) of the seasonality periods.

The forecast for the 48-month horizon was made and is shown in Figure 3. In general, the model was able to well predict the behavior of the dam blocks in the medium term (four years), as it satisfactorily represented the seasonality. Regarding accuracy, predictions were more accurate in the short term, with a RMSE of approximately 0.11 for the first 12 months, and greater errors the further away from the last reference value of the IMCRB (RMSE = 0.22). Knowing that the dataset was quite heterogeneous - consisting of 156 values of 61 observable variables, which were represented by 5 latent variables, which combined together gave rise to the IMCRB, taking into account the influence of the two environmental variables on this index, and the low magnitude of the errors - the model used for the forecast was considered valid. Thus, the ECM-I-Mod model was able to represent the medium-term variation in the responses of the dam blocks under the influence of the variation in the reservoir water level and the ambient temperature.

CONCLUSION

The joint monitoring index of the responses of the dam block, resulting from the factor analysis of data from piezometers, extensometer bases, plumblines and multiple-rod extensometers, was able to represent the global behavior of a buttress block, under the influence of variations in the reservoir water level and ambient temperature. This index can help in structural monitoring, as a single graph concentrates all instrumentation data and any change in behavior can be observed immediately, through the difference between what was planned and what happened (measured), speeding up the decision-making process on the part of the responsible personnel.

ACKNOWLEDGEMENTS

This research was carried out with the support of the Itaipu Technological Park (PTI) and the Center for Advanced Studies in Dam Safety (CEASB).

REFERENCES

Arone, S.G. 2014. Impacto de meios electrónicos de pagamento sobre as notas e moedas em circulação: o caso de Moçambique. Ph.D. Thesis. Instituto Superior de Economia e Gestão.

Asteriou, D. and Hall, S.G. 2015. Applied econometrics. Macmillan International Higher Education.

Barbão, J. 2007. Análise de componentes cíclicas em séries temporais uni e multivariadas via filtros HP modificados e outros métodos. Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica, Campinas, SP.

Box, G.P. and Jenkins, G.M. 1976. time-series Analysis: Forecasting and Control. Holden-Day.

Breusch, T.S. and Pagan, A.R. 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica: Journal of the Econometric Society, 1287–1294.

Cheng, L. and Zheng, D. 2013. Two online dam safety monitoring models based on the process of extracting environmental effect. Advances in Engineering Software 57, 48 – 56.

Çokluk, O.and Koçak, D. 2016. Using Horn's parallel analysis method in exploratory factor analysis for determining the number of factors. Educational Sciences: Theory & Practice 16, 537–552.

Godfrey, L.G. 1978. Testing for multiplicative heteroskedasticity. Journal of Econometrics 8, 227–236.

Gujarati, D.N. and Porter, D.C. 2011. Econometria Básica. Amgh Editora.

Hassler, U. andWolters, J. 2006. Autoregressive distributed lag models and cointegration, in: Modern Econometric Analysis. Springer, pp. 57–72.

Johnson, R.A. and Wichern, D.W. 2019. Applied Multivariate Statistical Analysis. Upper Saddle River, New Jersey: Pearson.

Li, F., Wang, Z. and Liu, G. 2013. Towards an error correction model for dam monitoring data analysis based on cointegration theory. Structural Safety 43, 12–20.

Lombardi, G., Amberg, F. and Darbre, G. 2008. Algorithm for the prediction of functional delays in the behaviour of concrete dams. The InternationalJournalonHydropower&Dams 15, 111.

Lopes, A.C. 1999. Modelos dl e adl, raízes unitárias e cointegração: uma introdução. CEMAPRE, ISEG.

Noble, N.R. 2011. EViews Tutorials: Forecasting. IHS Global Inc.

Ramsey, J.B. 1969. Tests for specification errors in classical linear least-squares regression analysis. Journal of the Royal Statistical Society: Series B (Methodological) 31, 350–371.

Ribeiro, L.S., Wilhelm, V.E., Étore F. Faria, Correa, J.M. and dos Santos, A.C.P. 2019. A comparative analysis of long-term concrete deformation models of a buttress dam. Engineering Structures 193, 301 – 307.

Salazar, F., Morán, R., Toledo, M.Á. and Oñate, E. 2017. Data-based models for the prediction of dam behaviour: a review and some methodological considerations. Archives of Computational Methods in Engineering 24, 1–21.

Villwock, R., Steiner, M.T.A., Dyminski, A.S. and Chaves Neto, A. 2013. Itaipu Hydroelectric Power Plant Structural Geotechnical Instrumentation Temporal Data Under the Application of Multivariate Analysis-Grouping and Ranking Techniques. InTech.

Vu, D., Muttaqi, K. and Agalgaonkar, A. 2015. A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. Applied Energy 140, 385 – 394.

Wooldridge, J.M., 2015. Introductory econometrics: A modern approach. 4 ed., Nelson Education.

****