STUDY OF THE INFLUENCE OF TEMPERATURE AND WATER LEVEL OF THE RESERVOIR ABOUT THE DISPLACEMENT OF A CONCRETE DAM

S.R. ORO[*] and T.R. MAFIOLETI
Department of Physics, Statistics and Mathematics
Federal Technological University of Paraná
Francisco Beltrão, Paraná, BRAZIL
E-mail: sheilaro@utfpr.edu.br

A. CHAVES NETO
Department of Statistics
Federal University of Paraná
Curitiba, Paraná, BRAZIL

S.R.P. GARCIA
Coordination of the Degree in Mathematics
Federal Technological University of Paraná
Toledo, Paraná, BRAZIL

C. NEUMANN JÚNIOR
Division of Civil Engineering and Architecture
Itaipu, Foz do Iguaçu, Paraná, BRAZIL

Multivariate techniques are used in this study to analyze the monitoring data displacements of a concrete dam, measured by means of pendulums, extensometer bases and multiple rod extensometers, taking into account the action of environmental conditions, bounded by the surface temperature of the concrete at ambient temperature and the tank water level. The canonical correlation analysis is used to evaluate the influence of environmental variables in the displacement of structures and dam foundations. The factor analysis is used to identify data sources of variability and order the sensors according to the action of factors. The dates of the measurements are grouped according to similarities in the present observations, by applying the cluster analysis. Then the discriminant analysis is used to assess the groups as to their homogeneity. The results demonstrate that the techniques used for distinguishing the dam responses and identify the effects of changes in environmental conditions on the displacements of the structures and dam foundations.

Key words: structural monitoring; concrete dam; multivariate analysis.

1. Introduction

The concrete dam structures are subject to changes caused by the incidence of phenomena, such as displacements, strains, stresses, pressures, etc. [1]. This occurs because these structures have strong interaction with environmental, hydraulic and geomechanical factors, as the temperature of the concrete, the hydrostatic pressure and the effect of time [2]. Therefore, these factors should be taken into account during structural evaluation.

---

[*] To whom correspondence should be addressed

Structural Monitoring is accomplished by visual inspection, geodetic measurement using vertical and/or horizontal displacements, bathymetric surveys and monitoring instrumentation [3]. The instruments used in this monitoring include pendulums, extensometer bases, triple orthogonal meters, flow meters, piezometers, multiple rod extensometers [4].

The instrumentation data set is useful for assessing the safety of dam's performance, especially if the current measures are compared with the entire series data through statistical and structural identification tools [5]. Reports on instrumentation and visual inspections are useful in this case because they cover all aspects of the dam since its construction up to the operational phase [6].

A detailed analysis of instrumentation data requires a combination of knowledge, especially of engineering, mathematics and statistics, and should be done by an experienced technical team, with the assistance of computational resources [7].

This work presents a helper method in the structural monitoring of concrete dams, combining multivariate statistical techniques for: (1) quantifying the influence of environmental conditions on displacements of structures and foundations; (2) identification of the most relevant sensors with respect to the variability of the data; (3) grouping of the dates of measurements, according to the similarities.

The text structure is composed of six sections, with Introduction being the first one. The second section presents the review of the literature on the use of statistical techniques in the structural monitoring of concrete dams. Section 3 discusses the theory related to multivariate techniques of Canonical Correlation Analysis, Cluster Analysis, Discriminant Analysis and Factorial Analysis. The proposed method is described in the fourth section. Then, the results obtained in applying the method and discussed. The last section grants the conclusion of the article and suggestions for future research.

2.   Structural security

The interest in perfecting techniques for structural monitoring of dams has grown in recent decades mainly due to the need for greater security in order to avoid the consequences of disasters caused by structural problems [8].

The methods used to monitor the structural safety of dams usually consist of comparing loads and safety factors used in the projects of dams with the behavior of all structures over the years [6].

The selection of data, during the process of monitoring the behavior of structures, involves the choice and type of method, number and location of the sensor, and hardware acquisition/storage/data transmission. This process is specific to each application. Economic issues play an important role in making these decisions. The time interval in which data should be collected is another point that should be considered [9].

In situations that have uncertainties inherent in the system adopted, the statistical analysis is indicated [9] to sort an amendment of the parameters as from the structural condition change (failure) or modification of environmental and/or operating conditions.

As the data may represent measures of different natures and scales, it is important to standardize the data to enable the damage identification process. Accordingly, as pointed out in [10] standardization of data is an inherent procedure for data acquisition, feature extraction and statistical modeling in structural monitoring process, such as to remove the effects of operational and environmental variables with the extracted resources.

In addition, in [9] the need to identify and minimize sources of variability in the data acquisition process and in the monitored system is stressed. However, not all sources of variability can be eliminated, for example, the variation caused by various environmental conditions such as temperature, humidity, loading and boundary conditions. Therefore, it is necessary to make appropriate measures so that these resources can be quantified statistically.

According to the literature, the statistical modeling of the structures monitoring data has been applied, mainly, to achieve classification, association, forecast values and outliers detection. Among multivariate statistical techniques, the discriminant analysis, canonical correlation, multiple linear regression and principal component analysis are most frequently used for this purpose. Reports and discussions of these applications can be found in [11; 5; 12; 13; 14; 15; 10; 16; 2; 17; 18; 19].

3. Theoretical

3.1. Canonical correlation analysis

The Canonical Correlation Analysis is an interdependence analysis technique that allows researchers to identify and quantify the associations between two groups of variables (X and Y). The basic idea is to find the linear combination of variables X and linear combination of variables Y that produce the highest correlation between the two groups [20].

The first group (X) is composed of p decision variables, also called explanatory variables (independents). While the second (Y) is formed by q response variables (dependents on explanatory).
The vectors X and Y have covariance matrices respectively, and the relationship is summarized in the cross-covariance matrix between these vectors.

Given that U and V are linear combinations of the vectors in X and Y Eq.(3.1), respectively, the canonical problem is to obtain the vectors of coefficients a and b that maximize the correlation between U and V Eq.(3.2). The vectors a and b, in this case, are solutions of a system of equations Eq.(3.3). The linear combinations U and V, in this case, are called canonical variables.

$$U = a'X, \qquad V = b'Y, \tag{3.1}$$

$$Corr(U,V) = \sqrt{\lambda} = \frac{a'\Sigma_{XY}b}{\sqrt{a'\Sigma_X a}\sqrt{b'\Sigma_Y b}}, \tag{3.2}$$

$$\left(\Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX} - \lambda\Sigma_X\right)a = 0,$$

$$\tag{3.3}$$

$$\left(\Sigma_{YX}\Sigma_X^{-1}\Sigma_{XY} - \lambda\Sigma_X\right)b = 0$$

where λ is the largest eigenvalue of the matrix or, equivalently, of the matrix $\left(\Sigma_{YY}^{-1}\Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}\right)$.

Thus, each pair of canonical variables has unit variance, maximum correlation and is not correlated with others pairs of canonical variables. The number of pairs of canonical variables that can be obtained is equal to the lowest value of p and q. In general, we try to get a few pairs of canonical variables that explain much of the interdependence between the two sets of observable variables.

3.2. Factorial Analysis

The application of multivariate technique of the Factorial Analysis allows the explanation of the correlations between many variables of a set of data through a limited number of unobservable random variables, called factors. [20].

Verification of the viability of the factor model used is made by applying the Bartlett Test Eq.(3.4), and the quality of fit of the model to the data set is estimated by the Kaiser-Meyer-Olkin Criterion (KMO). The KMO coefficient Eq.(3.5) varies between 0 and 1. The closer to 1, the better the adjustment factor model to the data.

$$H_0{:}\rho = I_{p\times p} \times H_A{:}\rho \neq I_{p\times p}, \tag{3.4}$$

$$T = -\left[n - \frac{1}{6}(2p+11)\right]\sum_{j=1}^{p} \ln\lambda_j, comT \quad \chi_v^2, v = \frac{p}{2}(p-1),$$

$$KMO = \frac{\sum\limits_{i \neq j} r_{ij}^2}{\sum\limits_{i \neq j} r_{ij}^2 + \sum\limits_{i \neq j} q_{ij}^2} \tag{3.5}$$

where q_{ij} is the element belonging to the i-th row and j-th column of the matrix, $Q = DR^{-1}D$ with $D = \left(\sqrt{diagonal\left(R^{-1}\right)} \right)^{-1}$.

The factor model Eq.(3.6) considers that each variable can be written as a linear combination of the common factors (F_k) and specific factors. During the process of obtaining the factors, that following quantities are estimated: the factor loadings (l_{ji}), the commonalities (h_i), the specific variances (ε_j) and the factorial scores (f_{jk}) that are measurements with explanatory properties of great interest to the researcher.

The load factor is a measure of the variable correlation with the factor. The commonality is the portion of the variance of each original variable from the extracted factors. The remainder of the variability, which can be owed to other factors, is measured by specific variance.

$$Z_1 = l_{11}F_1 + l_{12}F_2 + \ldots + l_{1m}F_m + \varepsilon_1,$$

$$Z_2 = l_{21}F_1 + l_{22}F_2 + \ldots + l_{2m}F_m + \varepsilon_2,$$
$$\vdots$$
$$Z_j = l_{j1}F_1 + l_{j2}F_2 + \ldots + l_{jm}F_m + \varepsilon_j, \tag{3.6}$$
$$\vdots$$
$$Z_p = l_{p1}F_1 + l_{p2}F_2 + \ldots + l_{pm}F_m + \varepsilon_p.$$

The factor scores Eq.(3.7) are estimates for the values of the factors for each sample element. They may be used to sort sample components or as input variables for further statistical analysis.

$$\underline{f}_i = \left(L'_z L_z \right)^{-1} L'_z \left(\underline{x}_i - \underline{x} \right), \qquad i = 1, 2, \ldots, n. \tag{3.7}$$

3.3. ClusterAnalysis

The use of the Cluster Analysis seeks to find, within a heterogeneous set of data, a small number of homogeneous groups, whose variation within the group is substantially smaller than the total variability of the data set.

Initially, in the hierarchical agglomerative method, each observation forms a separate group. At each step of the process, the groups join according to the similarities, forming new groups, until only one group with the total number of observations included remains.

Similarity is a measure of proximity between two groups. One way of calculating this measure is the Mahalanobis distance Eq.(3.8).

$$D_{ij}^2 = \left(x_i - x_j \right)' \Sigma^{-1} \left(x_i - x_j \right) \tag{3.8}$$

where Σ is the complete data set of covariance matrix X.

3.4. Discriminant Analysis

The discriminant Analysis is a technique that enables, starting from independent variables, to study the profile, performing classification and differentiation of two or more group elements. The number of groups should be known in advance. The discrimination is made based on a mathematical rule, which minimizes the likelihood of incorrect classification errors.

In the perspective of Mahalanobis Eq.(3.9) is calculated the distance $\left(D_g^2 \right)$ of each observation to the centroid of each group $\left(\overline{x}_g \right)$. Then, the observation is allocated to the nearest centroid group.

$$D_g^2 = \left(x - \overline{x}_g \right)' \Sigma_W^{-1} \left(x - \overline{x}_g \right) \tag{3.9}$$

where Σ_W is the covariance matrix within the group between the independent variables.

4. Methodology

The evaluation of the responses of a dam structure, considering the instrumentation data and taking into account interaction with the environment, is a problem composed of various dimensions. Therefore, it is necessary to use techniques that allow the joint analysis of monitoring data, reduce the magnitude of the problem and assist decision-making.

The multivariate statistical techniques, called the Canonical Correlation Analysis, Factorial Analysis, Cluster Analysis and Discriminant Analysis, meet these requirements and, therefore, constitute the method used in this work.

As illustrated in Fig.1, the method consists of five steps. The first is the selection of instruments, the definition of the time period and data collection. If there is a difference in frequency of the measurement instruments, it is necessary to equate the periods, using for example, the monthly average of observations. Furthermore, the gaps coming from the absence of readings in the period are identified and filled in this step. The filling is made with modeling and forecasting time series.

Then, the Box-Plot and Scatterplot graphics are used to identify the occurrence of outliers. For each detected outlier, it is needed to evaluate its maintenance or exclusion from the data set. If you choose to exclude it, you must make a new value forecast, using the same procedure of filling the gaps.

Composition of the data matrix (*Z*)

↓

Canonical Correlation Analysis of displacement and environmental conditions

↓

Factorial Analysis of the data set

↓

Cluster Analysis of the dates of the readings
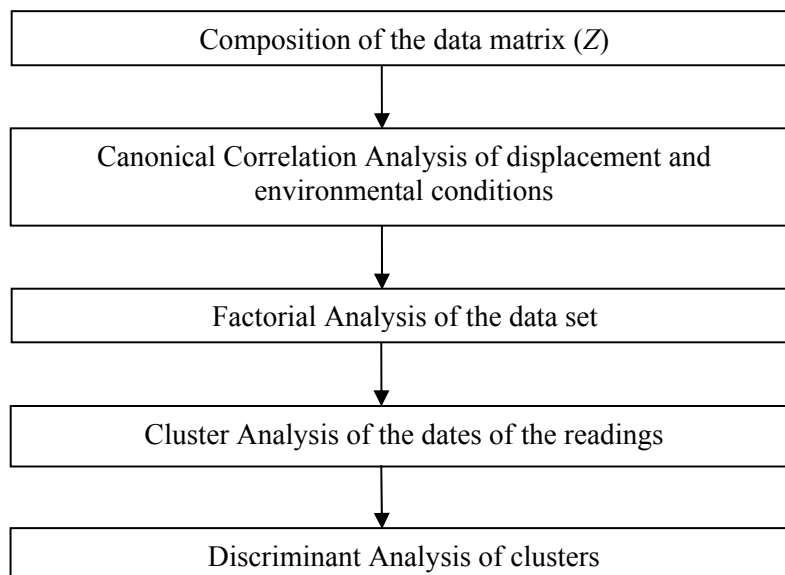
↓

Discriminant Analysis of clusters

Fig.1. Flow chart of model.

Thus, a sample data matrix (X) is composed, with n lines (monthly average) and p columns (sensors of the instrument). If there are differences of magnitude and in the scale of the observations, owed to the use of different kinds of instruments, the data must be standardized. The standardized data matrix (Z) is, then, used as input to the procedures listed in the following steps.

The second step is the application of the Canonical Correlation Analysis to study the relationship between the group of sensors, which measure the displacements of the dam structures and their foundations, and the group of indicators of environmental conditions.

Then, in step 3, the Factorial Analysis is used to estimate the influence of environmental conditions in shifts, perform the ranking of the instruments according to their importance in the factor model and identify factors that can be used as criteria for the overall assessment of displacements.

In the next stage, the scores of factors are subjected to the Cluster Analysis for the formation of homogeneous groups of measurements dates.

In the last step, the Discriminant Analysis tests the formed groups, with reference to the sensors with higher ranking in the Factorial Analysis.

5. Application in the Context of Structural Monitoring of Dams

The proposed method was applied to a structural monitoring process of a concrete dam. The data set used consists of the observations recorded in the period between January 1990 and December 2013, which were obtained through manual measurements of the installed instrumentation in blocks D7 and D8 keys, the D portion (Dam Right Side, built in blocks buttresses) of the Itaipu dam, and the hydrometeorological data from the same period.

The instrumentation consists of 40 sensors (Tab.1), including direct pendulums, inverted pendulums, extensometer bases, surface thermometers and multiple rods extensometers. Hydrometeorological data were also considered: reservoir water level and temperature at the dam surroundings. Thus, 42 variables were computed, related to sensors considered in this study.

Table 1. Phenomena monitored by instruments and respective sensors.

| Phenomena	Instrument	Sensors
Radial displacement	Direct pendulum	Z1, Z5
Radial displacement	Inverted pendulum	Z3
Tangential displacement	Direct pendulum	Z2, Z6
Tangential displacement	Inverted pendulum	Z4
Opening and closing of joints between blocks	Extensometer base	Z7, Z9, Z11, Z13, Z15, Z17
Horizontal sliding between blocks	Extensometer base	Z8, Z12, Z14, Z18
Differential settlement between blocks	Extensometer base	Z10, Z16
Surface temperature of block	Concrete thermometer – downstream	Z19
Surface temperature of block and water reservoir	Concrete thermometer – upstream	Z20
Deformations of rocky massive	Multiple rod extensometer	Z21, ..., Z40
Ambient temperature	Thermometer	Z41
Water level of the reservoir	Limnimetric Ruler	Z42

5.1. Composition of the Sample Data Matrix

Given that the frequency of measurements gauged with the different instruments (sensors) was not the same, it was decided to use the monthly average of the observations. The implementation of a

computational procedure, performed with the help of Matlab software (Matlab R2013, 2013), allowed to create a monthly average observation of each of the 42 sensors and to identify the existence and location of gaps, corresponding to periods (months) that measurements were not performed.

It was found that there are eight incomplete series, totaling 15 deficiencies resulting from the lack of measurements at some point. There were created two (sub) series for each sensor that detects the occurrence of gaps: one with previous observations to the missing data and the other with subsequent information to the same issue. The forecast of this data was performed using the forecasting/backforecasting procedure, which consists in modeling each (sub) series, making value forecast and filling the gap with the average value of the two forecasts. The ARIMA models were used for the realization of the forecasts, with the help of Statgraphics software (Statgraphics Centurion XVI, 2010).

After completing the sensor data series, we proceeded with the analysis of the Box-plot and Scatterplot graphics of each series, in search of the occurrence of outliers. Outliers were identified in the data reservoir water level (Fig.2). The cause of this occurrence was the incidence of drought (low rainfall) in the months of the summer season of the years 1999/2000 and 2012/2013. It was decided to keep the values as observed, to check the influence of this occurrence in the dam's answers.
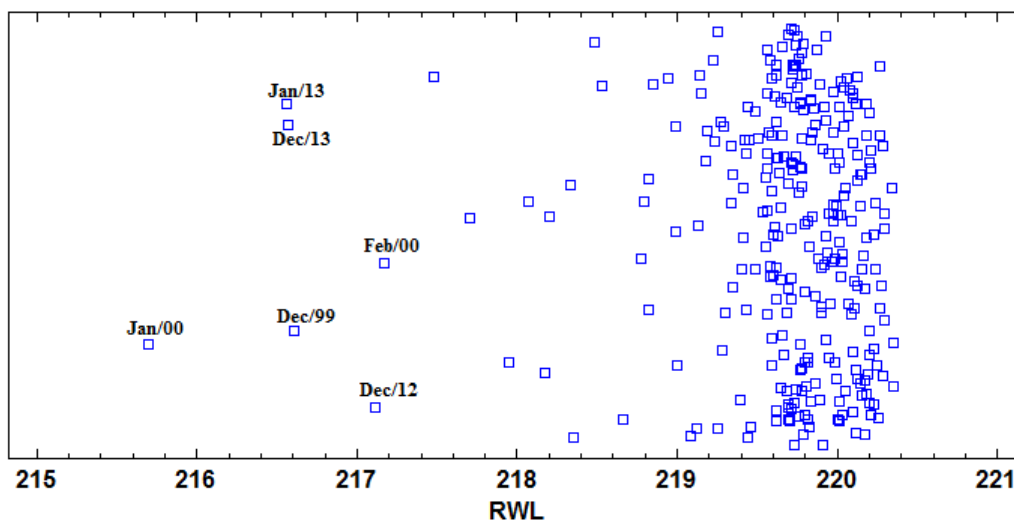
(a)
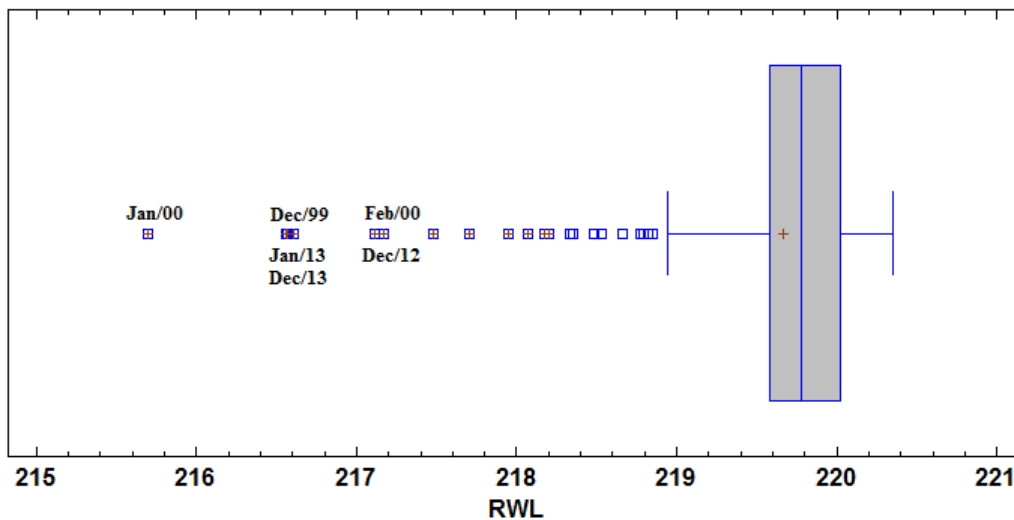
(b)

Fig.2. (a) Box-Plot and (b) Scatterplot graphic of the reservoir water level (RWL) from Itaipu dam, from Jan /90 to Dec /13.

Owing to differences in quantities and measuring scale on the variables, because of the nature of the sensors, it was necessary to carry out the standardization of data.

Thus, sampling of the data matrix (Z), order 288×42 was mode, whose lines correspond to the dates (month/year), the columns to the variables (sensors) and the elements to standardized data.

5.2. Canonical Correlation Analysis of displacements and Environmental Conditions

The Canonical Correlation Analysis was used to study the relationship between the sensors clusters that measure displacement (Z1 to Z18 and Z21 to Z40) and indicators of environmental conditions (Z18, Z19, Z41 and Z42). As shown in Tab.2, all the eigenvalues were considered significant at a confidence level of *95%* (*p*-value *<0.05*). It was decided to discuss the results of the canonical correlation of higher value.

The canonical correlation between the two groups was *0.986*. This measure indicates the strong influence that the environmental conditions (V1) have on the set of sensors that measure the displacements of the dam's structures (U1).

Table 2. Canonical correlation of displacement sensors vs environmental conditions.

Eigenvalue	Canonical Correlation	Wilks' lambda	χ^2	Degree of Freedom	*p*-value
$\lambda_1 = 0.973$	*0.986*	*0.003*	*1537.080*	*152*	*0*
$\lambda_2 = 0.717$	*0.847*	*0.114*	*577.122*	*111*	*0*
$\lambda_3 = 0.437$	*0.661*	*0.402*	*242.062*	*72*	*0*
$\lambda_4 = 0.286$	*0.535*	*0.714*	*89.492*	*35*	*0*

The correlations between the 38 variables of the first group and four of the second were estimated. Table 3 lists the variables with the highest correlations ($| \rho | > | 0.8 |$ and *p*-value *< 0.05*), with predominance of multiple rod extensometers (Z23, Z24, Z35, Z38, Z40) related to temperatures (Z19, Z20 and Z41). Three sensors pendulums (Z2, Z3 and Z5) associated to the block surface temperatures in the upstream (Z20) and environment (Z41) appear.

Table 3. Strongly correlated variables.

Sensors	Correlations
Z35 – Z20	*-0.913*
Z24 – Z20	*-0.910*
Z24 – Z41	*-0.889*
Z23 – Z20	*-0.888*
Z23 – Z41	*-0.872*
Z5 – Z41	*-0.855*
Z40 – Z19	*-0.855*
Z40 – Z41	*-0.843*
Z3 – Z41	*-0.829*
Z2 – Z20	*-0.820*
Z35 – Z41	*-0.815*
Z23 – Z19	*-0.814*
Z24 – Z19	*-0.812*
Z38 – Z19	*-0.811*

In the region where the Itaipu dam is located, the range of monthly average ambient temperatures, observed in the same year, may reach *20° C.* The strong correlation of the sensors with temperature variables

confirmed this. The negative correlation, with dominant presence in the first rows of the table, indicated an inverse relationship between the variables. That is, in periods of low temperature, the displacements were larger than those registered in periods of high temperature.

Moreover, the reservoir water level alone showed a small positive correlation with just a few sensors. Possibly, because the low variability in this variable introduced the forces acting on the dam almost constant. Another reason may be the need for the reservoir water level interaction with the temperature to influence the displacements.

Confronting the canonical variables U1 and V1, through the Scatterplot (Fig.3), the existing linear relationship between these variables was confirmed, showing the possibility to predict the dam's structural performance in a given time, depending on the measuring sensors offsets. The distance of a point compared to the others was also recorded. This point was referring to measurements made in July 2000, when it was recorded one of the monthly average lower to room temperature.
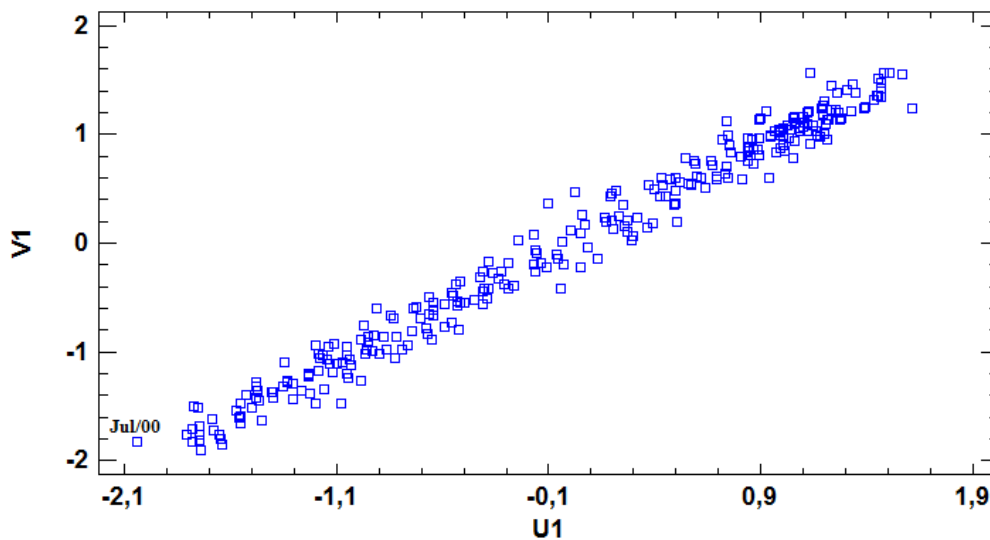

Fig.3. Scatterplot of the canonical variables U1 and V1.

Table 4 presents the most correlated sensors ($| \rho |> | 0.9 |$ e p-value < 0.05) with the first pair of canonical variables. It is observed that the rod of multiple rod extensometers Z23, Z24, and Z35 were the sensors more influenced by temperatures Z20, Z41.

Table 4. Key correlations between canonical variables and sensors of each group.

Sensor	U1	Sensor	V1
Z23	-0.924	Z20	0.989
Z24	-0.943	Z41	0.936
Z35	-0.926		

Quality assessment of the potential of canonical variables was based on the proportion of variance explained by the canonical variables for each group. The canonical variable U1 explained *38.7%* of the variance observed in shifts, while the proportion of the variance explained by V1 to the "Environmental Conditions" group was *65.6%*. Thus, the groups "Displacement" and "Environmental Conditions" were well represented by the first pair of canonical variables, since the canonical correlation between these groups was *0.986*, while the other pairs have lower values.

Therefore, if the "Environmental Conditions" group was the cause of the variability observed in the group "Displacement", then U1 can be used as the best predictor and V1 is the most likely criterion for the realization of the dam's structural performance prediction.

5.3. Factorial Analysis of displacements and Environmental Conditions

The application of the Factorial Analysis by principal components resulted in a model composed of five factors, identified based on the greatest factor loadings of sensors, able to explain *91.12*% of the variance of the set of comments. Figure 4 shows the series of identified factors.

The factors were named according to the sensors more correlated with them, that is, as the greatest factor loadings. The first factor, due to its positive correlation with most of the stems of multiple rod extensometers, especially with Z21, Z22, Z25, Z26, Z29, Z30, corresponds to the Foundation's Movement (MF). This was the most important among factors identified because accounts for *45.88*% of the observed variability in the data set.

The second factor, named the Horizontal Movement Structure in Normal Direction (MHESN), explained *30.83*% of the variance and is associated with the openings and joints between the dam block and the horizontal displacement in the normal direction (perpendicular to the direction of water flow). The Horizontal Movement Structure in the Flow Direction (MHESF), the third factor identified, accounted for *9.66*% of the variability. The dominant variables in this factor were the temperatures (Z19, Z20, Z41), with the multiple rod extensometers (Z23, Z24, Z38, Z40) and the horizontal displacements in the direction of water flow of D8 block (Z3 and Z5).

The fourth and fifth factors, called the Block D7 Structure Geometry (GEBD7) and Hydrostatic Pressure (PHI) were related to horizontal displacements in the direction of water flow of D7 block (Z1) and the reservoir water level (Z42), respectively. Furthermore, the "Block D7 Structure Geometry" accounted for *2.41*% of the total variability, while the influence of the "Hydrostatic Pressure" was *2.34*%.
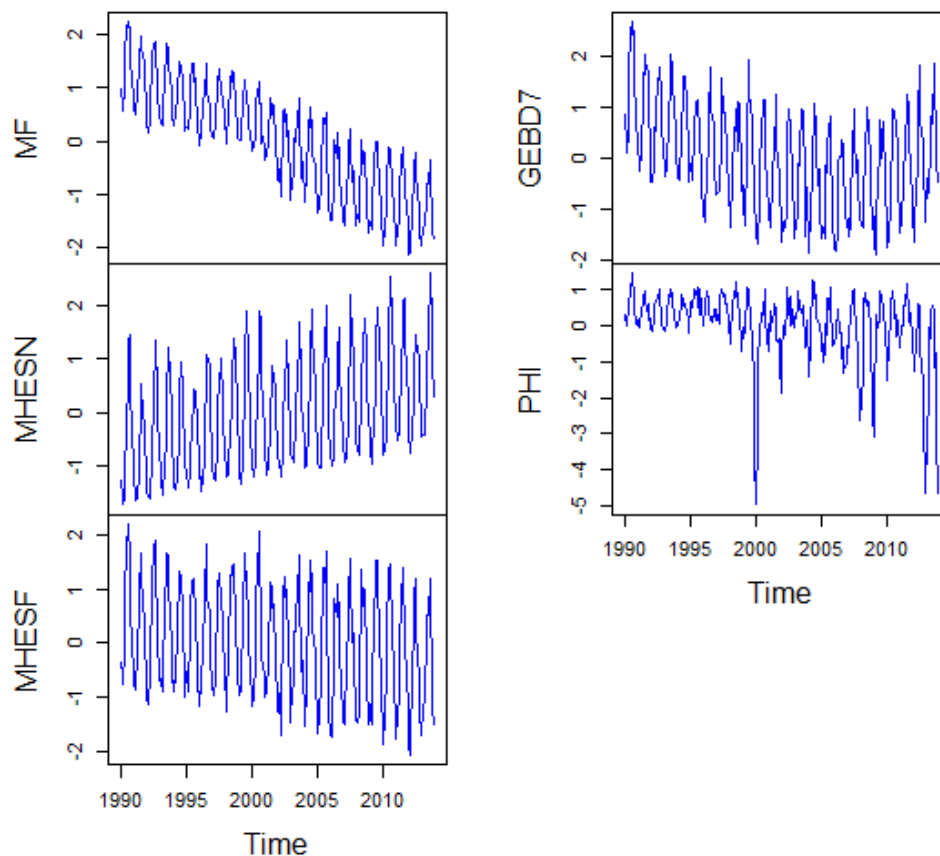

Fig.4. Time series of the identified factors.

The variability in the readings of each sensor, arising from identified factors, were estimated by commonality. Thus, a low commonality (less than *0.60*) would indicate that the variable would not be sufficiently explained by the model and could be discarded. The results pointed to the preservation of all sensors considered in this study.

Using as a measure of commonality as a measure of importance of each variable to the factorial model, the ranking of the sensors was obtained. Therefore, the most important instruments for the D7 and D8 blocks were respectively rods extensometer multiple Z33, Z34, Z35 (even borehole) and extensometer bases (opening) Z7, Z9. Furthermore, the reservoir water level (Z42) was the highest classified variable related to the environmental conditions.

The factor, being a latent variable, cannot be measured directly. However, the values of factors, called factor scores, are estimated based on factor loadings and the sensor values that dominate this factor.

5.4. Grouping the dates of the measurements

Using the method of the group average and the Mahalanobis distance, we identified three homogeneous groups of dates (Fig.5), adding 190 elements in the first (G1), 83 in the second (G2) and 15 in the third (G3). The first group comprised essentially the months from November to May, period in which the highest temperatures were recorded, while most months, with lower temperatures, were gathered in G2. The third group brought together the dates in which the lowest water levels of the reservoir were recorded.
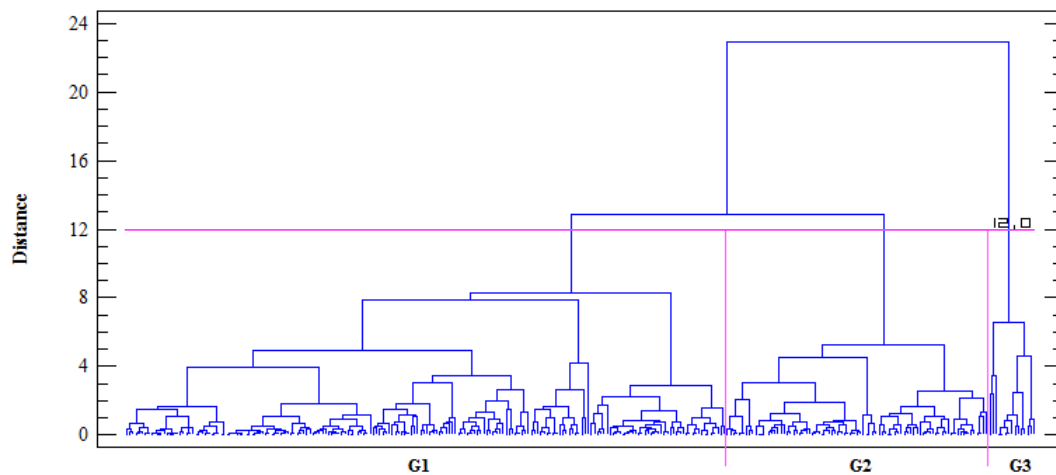

Fig.5. Groupings by the middle connection method, using the Mahalanobis distance.

Discriminant Analysis, considering the sensors Z7, Z9, Z33, Z34, Z35 and Z42, chosen because they have the greatest commonalities, tested the classification of 288 observations. Due to the large difference in size of the groups, we considered the proportionality of the number of observations per group. The results are in Tab.5. The high percentage of correct classification, *94.1%*, has confirmed the discriminating power of the sensors regarded in the analysis.

Table 5. Classification of the dates of the measurements into three groups.

Current Group	Size of the Group	Proportion	Provided Group			Cumulative percentage
			1	2	3	
1	190	*65.97%*	*178*	*11*	*1*	*93.68%*
2	83	*28.82%*	*5*	*78*	*0*	*93.98%*
3	15	*5.21%*	*0*	*0*	*15*	*100%*
TOTAL	288	*100%*	*183*	*89*	*16*	*94.10%*

Two functions were considered statistically significant, at the confidence level of *99%*, to distinguish the observations belonging to each group (Fig.6). The first function discriminates High Temperature and Low Temperature groups, while the second group discriminates the Low Water Level of Reservoir (LWLR) group from the others.
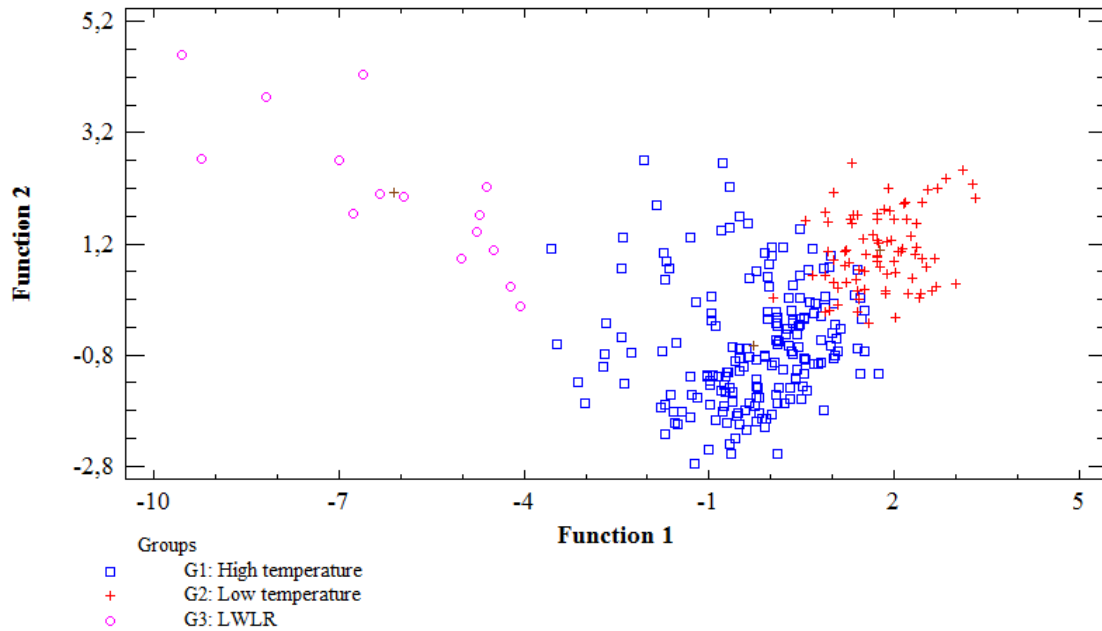

Fig.6. Dispersion of the elements according to discriminant functions between groups.

Ranking functions of observations in groups Eq.(5.1) are linear combinations of the sensors. These functions can be used to classify new dates readings. To do so, simply calculate the scores of each new element in each group and, then, allocate it in that with highest score.

$$G1 = -0.66 + 0.66 \times Z34 - 1.33 \times Z33 - 0.08 \times Z35 + 0.32 \times Z9 + 0.13 \times Z42 - 0.88 \times Z7,$$

$$G2 = -3.40 - 0.45 \times Z34 + 2.87 \times Z33 + 0.07 \times Z35 - 1.28 \times Z9 + 1.66 \times Z42 + 2.67 \times Z7, \quad (5.1)$$

$$G3 = -23.81 + 5.94 \times Z34 + 1.01 \times Z33 + 0.54 \times Z35 + 3.01 \times Z9 - 10.87 \times Z42 - 3.64 \times Z7.$$

6. Conclusion

The method proposed in this paper consists in the multivariate analysis of the displacements of the structures and foundations of a concrete dam, taking into account the interaction with the environment.

The results of the Canonical Correlation Analysis allow us to infer that these shifts are strongly influenced by environmental conditions. In general, the instrumentation registers larger dislocations during periods of low temperatures. The set of instruments that comprises pendulums, extensometer bases and multiple rods extensometer can be used to predict the structural performance of a dam, with respect to displacement, according to the variability criteria of environmental conditions.

The dates of the observations recorded by the instrumentation, when subjected to Cluster and Discriminant analyses, can be grouped into "High Temperature", "Low Temperature" and "Lower Reservoir Water Level".

Most of the measurement data variability is due to the following factors: Foundation's Movement; Horizontal Movement Structure in Normal Direction; Horizontal Movement Structure in the Flow Direction; Block D7 Structure Geometry and Hydrostatic Pressure.

Acknowledgments

Nomenclature

(U, V) – canonical variables
X – data matrix of decision variables
Y – data matrix of response variables
p – number of the decision variables
q – number of the response variables
λ – eingenvalue
Σ – matrix of variance / covariance
r_{ij} – correlation between the variables x_i and x_j
F – common factor
l_{ji} – factor loading
h_i – commonalities
ε_j – specific variances
f_{jk} – factor scores

References

[1] Carvalho J.V. and Romanel C (2007): *Temporal neural networks applied to the monitoring of dams*. – Rev. El. Sist. Inf., vol.6, No.1, pp.1-9. (In Portuguese)

[2] Li F., Wang Z.Z. and Liu G. (2013): *Towards an error correction model for dam monitoring data analysis based on cointegration theory*. – Struct. Saf., vol.43, pp.12–20.

[3] Cruz P.T. (2006): *100 Brazilian dams: case histories, building materials, project*. – Oficina dos Textos, São Paulo. (In Portuguese)

[4] Matos S.F. (2002): *Assessment tools for auscultation concrete dam. Case Study: deformeter and stress meter for concrete in the Itaipu Dam*. – Masters dissertation, UFPR. (In Portuguese)

[5] De Sortis A. and Paoliani P. (2007): *Statistical analysis and structural identification in concrete dam monitoring*. – Eng. Struct., vol.29, pp.110–120.

[6] Kuperman S.C., Moretti M.R., Cifu S., Celestino T.B., Re G. and Zoellner K. (2005): *Criteria to establish limit values of instrumentation readings for old embankment and concrete dams*.

[7] 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*. – In: Multivariate Analysis in Management, Engineering and the Sciences. InTech, pp.81–102.

[8] Medeiros C.H. and Lopes M.G.M. (2011): *The risk of dams rating by risk category, based factors weighting method*. – In. XXVIII Seminário Nacional de Grandes Barragens. Rio de Janeiro. (In Portuguese)

[9] Farrar C.R. and Worden K. (2007): *An introduction to structural health monitoring*. – Philos. Trans. R. Soc. A, vol.365, pp.303–315.

[10] Figueiredo E., Park G., Farrar C.R., Worden K. and Figueiras J. (2011): *Machine learning algorithms for damage detection under operational and environmental variability*. – Struct. Health Monit, vol.10, pp.559–572.

[11] Buzzi M.F. (2007): *Evaluation of time-series correlations readings of geotechnical and structural monitoring instruments and environmental variables in dams - Itaipu case study*. - Masters dissertation, UFPR. (In Portuguese)

[12] Guedes Q.M. and De Faria É.F. (2007): *Statistical model control dislocation monitored the bark dam UHE Funil*. - In. XXVII Seminário Nacional de Grandes Barragens, Belém. (In Portuguese)

[13] Deng N., Wang J.-G. and Szostak-Chrzanowski A. (2008): *Dam Deformation Analysis Using the Partial Least Squares Method*. - In: 13th FIG Int. Symp. on Deformation Measurements and Analysis and 4th IAG Symp. on Geodesy for Geotechnical and Structural Engineering, Lisbon.

[14] Jin-Ping H., Yu-Qun S. (2011): *Study on TMTD Statistical Model of Arch Dam Deformation Monitoring*. - Procedia Eng., vol.15, pp.2139–2144.

[15] Mata J. (2011) *Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models*. - Elsevier Eng. Struct., vol.33, pp.903–910.

[16] Xu C., Yue D. and Deng C. (2012): *Hybrid GA/SIMPLS as alternative regression model in dam deformation analysis*. - Eng. Appl. Artif. Intell., vol.25, pp.468–475.

[17] Mata J., Tavares de Castro A. and Sá da Costa J. (2013): *Time–frequency analysis for concrete dam safety control: Correlation between the daily variation of structural response and air temperature*. - Eng. Struct., vol.48, pp.658–665.

[18] Cheng L. and Zheng D. (2013): *Two online dam safety monitoring models based on the process of extracting environmental effect*. - Adv. Eng. Soft., vol.57, pp.48–56.

[19] Mujica L.E., Ruiz M., Pozo F., et al (2014): *A structural damage detection indicator based on principal component analysis and statistical hypothesis testing*. - Smart. Mater. Struct., vol.23, pp.25014–25025.

[20] Johnson R.A. and Wichern D.W. (2007): *Applied Multivariate Statistical Analysis*. - 6th edn. Pearson.