SYSTEMIC ASSESSMENT – A NEW APPROACH IN THE EVALUATION OF THE INSTRUMENTAL OF MONITORING DAM

Fabián Restelli
Doctor en Ingeniería, Máster en Métodos Numéricos Aplicados, Ingeniero Civil
Jefe Departamento de Ingeniería - Hidroeléctrica El Chocón S.A.
frestelli@elchoconsa.com.ar

RESUMEN

La etapa de evaluación de mediciones de auscultación es realizada por el ingeniero valiéndose de la información registrada en las bases de datos y la representación de las mismas a través de diversos y variados tipos de gráficos. En función de la edad de las obras, la información que provee el sistema de medición es abundante y, por lo general, no ha sido explorada la utilización de técnicas que permitan trabajar de manera integral y desestructurada con el objetivo de extraer información valiosa y oculta de nuestras bases de datos.

A fin de dar continuidad a las tareas I+D+I (Investigación, Desarrollo e Innovación) comenzadas en el año 2010 en Hidroeléctrica El Chocón SA, se presentan nuevos avances a partir de la aplicación de métodos multivariantes, de modo de obtener un enfoque de características sistémicas en la evaluación de la auscultación.

Los resultados muestran que se dispone de una herramienta que supone un nuevo aporte a la etapa de análisis de mediciones de auscultación e introduce un cambio cualitativo en la valoración de los aspectos de la Seguridad de Presas.

ABSTRACT

The analysis of the monitoring measurements is performed by an engineer who will use the information stored in the databases and their representation through different and varied graphics. Depending on the age of the dam sites, the information provided by the measurement system is, in general, quite abundant and therefore it is necessary to explore the use of techniques so as to work in a more integrated and unstructured way in order to obtain valuable and hidden information from our databases.

In order to give continuity to the R+D+I (Research, Development and Innovation) tasks started in 2010 by Hidroeléctrica El Chocón SA, we present new advances and developments from the application of multivariate methods.

The results show that is available a new tool that improvement the analysis stage of auscultation measurements and introduces a qualitative change in the assessment aspects of Dams Safety.

INTRODUCTION

Depending on the age of the dam sites, monitoring databases often have a large amount of information. These databases permanently grow in size, both in terms of quantity of records and dimensionality, and they eventually reach considerable proportions.

On the other hand, the complexity of certain phenomena characteristic of dam engineering forces researchers to face difficulties where multiple variables and huge amounts of information take place and which require advanced tools and concepts for their integrated treatment and interpretation. However, we could pose the following questions:

- Is the current control methodology suitable?
- Is it possible to improve the processes involved?
- Is there any valuable information hidden in our databases?
- Is it possible to reduce the study domain without losing valuable information?

This paper includes new advances that incorporate the Holistic Approach to the analysis of the piezometric instrumentation installed at the instrumented sections of the dam and its foundation.

BRIEF CONCEPTUAL

Factor Analysis is the generic name given to a set of multivariate statistical methods whose purpose is that of discovering the underlying structure in a data matrix. Basically, it consists in analyzing the structures of the interrelations of a large number of variables – control and causal – without the need to make a distinction whether they are dependent and independent and, based on this information, determining a set of latent dimensions known as factors which try to explain such interrelations [Restelli, F.; 2010].

First, a classification is obtained by applying the Principal Component Method (FA, Factorial Analysis) and then the studies are complemented with the Cluster Analysis (CA, Cluster Analysis) with which a hierarchical classification of measurements is obtained for the selection of instrument clusters according to response (or behavior) patterns.

In fact, the FA is a data reduction technique since, if its hypotheses are true, the initial information can be expressed without greater distortions in a smaller number of dimensions represented by these factors.

The strategy of combining both methods is supported by the fact that the use of factor coordinates allows for a common framework to be used in the process during which conglomerates or clusters are formed. Thus, FA is a previous process to CA since it turns original data into continuous, non-correlated variables. Figure 1 show an algorithm that conceptually summarizes the applied holistic approach:

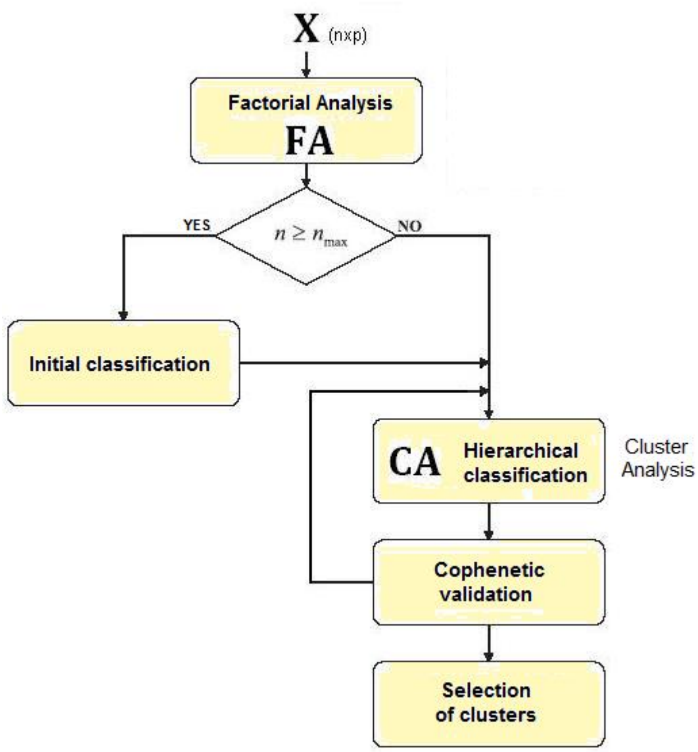


Figure 1. Algorithm used for the systemic analysis

The non-uniform spatial distribution of instrumentation (particularly on the abutments of the dam) and the different places in which piezometers are installed (core, contact and rock) make it necessary to give the same importance to all the variables, as a result of which it is convenient to work with standardized measurements.

Given a set of individuals (of *n* instruments) characterized by the information of *n* variables $X_{ij} = 1,2,..., p$ (where X_j represents the measurement for *p* dates), we are faced with the challenge of asking ourselves whether we will be capable of classifying measurements in a way that the individuals belonging to a cluster have characteristics that are as similar as possible to each other and, at the same time, as dissimilar as possible.

Pearson's Correlation Matrix *R* is used in graphic and numeric form in order to evaluate the distribution of measurements per instrument among each variable in relation to the rest. Measurement frequency histograms are displayed along the diagonal of the matrix with a domain divided into 10 units.

The principal components are determined according to *R* and presented in a factor matrix *C*. The eigenvalues of *C* (invariants) allow the methodology to be evaluated using the inertia graphs of eigenvalues (sorted in descending order) in which Jolliffe's Method is adopted with a 75% variance as the minimum cut.

Through the preparation of Biplot graphs (2D), an initial classification can be obtained by an arbitrary selection of instrument clusters (the Euclidean distance among instruments and the reservoir level represent the response delay). From these results, the exploratory analysis by CA starts in order to obtain an Agglomerative Hierarchical Classification, taking the standardized matrix *R* as the starting point. This exploratory method is also known with the generic name of unsupervised pattern recognition.

The clustering between two responses *i* and *j* is interpreted as a dissimilarity measure which will be represented by *d(i,j)*. The dissimilarity (or degree of divergence) measures the degree of similarity (or, more precisely, of dissimilarity between both responses) with respect to a certain number of inherent quantitative and/or qualitative characteristics.

The Euclidean distance is the easiest dissimilarity measure to be understood and determined. However and in spite of its simplicity, it presents certain difficulties:

- Sensitivity to the measurement units of the variables:

  The differences among high-value measurements have a greater contribution than the differences among low-value measurements. Consequently, the scale changes will determine changes in the distance among individuals.

- Nature of eigenvalues:

  If the variables used are correlated, a part of the analysis can be redundant. Some of the differences among the individual values of certain variables could be explained by the differences in other variables. Consequently, the Euclidean distance will increase the dissimilarity or divergence among individuals. Therefore, in order to perform the clustering the centroid method has been adopted in which the distance among clusters is determined as the Euclidean distance among their centroids.

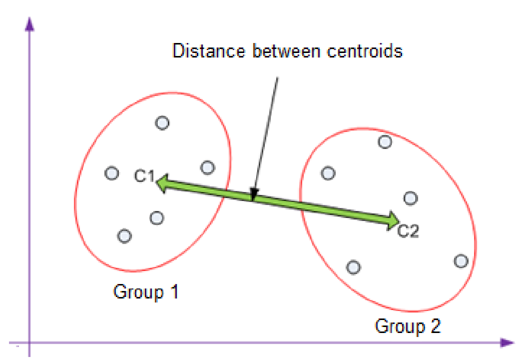
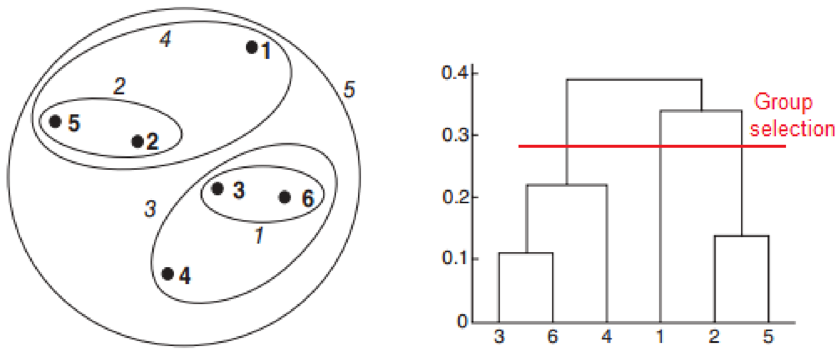


Figure 2. Centroid method criteria

The graph resulting from the application of CA for the hierarchical classification is called *Dendrogram,* and is a simple way to visualize results, compare them with those obtained by factor analysis and make the final classification of clusters. This structure has the shape of a nested tree which represents the similarity of each instrument's response before changes in the reservoir level.



CA - Clustering process   CA - Dendogram display

Figure 3. Cluster Analysis – Example of clustering visualization

The dendrogram shows the different stages of the analysis process (merging of clusters) and the existing distances among the merged elements.

The clustering level for each merge is provided by a numerical indicator called *Cophenetic Coefficient,* that varies between 0 to 1 and must be proportional to the dissimilarity (or divergence) considered in the merge. Its numerical expression is:

$$c = \frac{\sum_{i<j}(Y_{ij} - y)(Z_{ij} - z)}{\sqrt{\sum_{i<j}(Y_{ij} - y)^2 \sum_{i<j}(Z_{ij} - z)^2}} \quad (1)$$

Where:

Y_{ij} is the distance between targets i and j in Y.
Z_{ij} is the distance between the same targets of the distance matrix.
Z and z are the means of Y and Z, respectively.

The algorithm used in the application of the CA method is conceptually shown in the Figure 4.

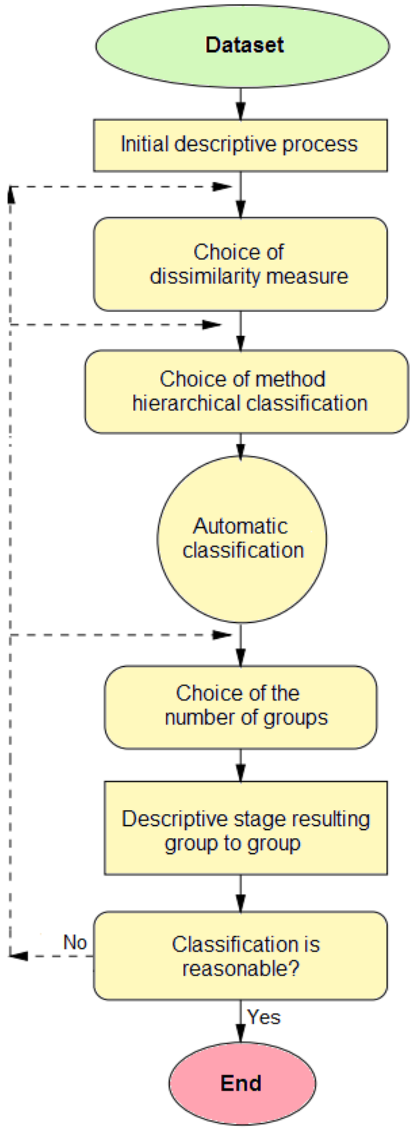


Figure 4. Algorithm used to perform the CA

There are two numerical criteria for the selection of the final clustering:

1) *A fixed rule*, whose objective is that of establishing an arbitrary and invariable level of affinity for the analysis of the entire cluster, assuming a certain "significance level".

2) *A variable rule*, which is simply the detailed analysis of the dendrogram.

However, the author considers that the final selection of clusters in the dendrogram should be the consequence of a holistic analysis of the response which includes the results obtained from temporal diagrams, cause-effect graphs, response rate, standardized data matrix analysis and comparison with the results of the initial classification by FA.

When the internodes of the dendrogram have clearly different lengths (thus delimitating discrete, well-differentiated clusters), the existence of natural clusters which can be isolated in a non-arbitrary way is reflected.

If this may not occur, it is necessary to perform a more careful analysis to form them considering that in such case we could be working with a non-natural classification, called *dissection*, which the engineer can perform when there is a homogeneous population with no natural clusters in it and which even so, and due to practical reasons, we prefer to divide into sub-clusters. In short, it is important to point out that the final classification will be an exclusive tool in which the engineer's experience and judgment become relevant.

Figure 5 summarizes the implemented conceptual model of the systemic instrumentation analysis:

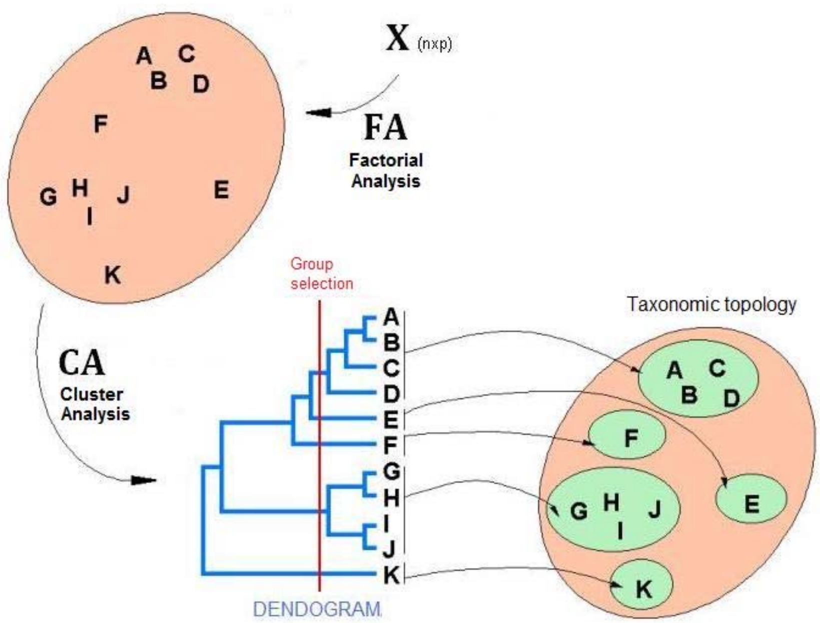


Figure 5. Conceptual Model of the Holistic Analysis

Application at El Chocón dam

El Chocón Dam is a 13 hm^3 embankment. The plan layout was chosen based on the topography conditions at dam-site, as show Figure 6.

The dam axis is straight across the main valley and curves slightly downstream over the left bank terraces to accommodate the intake structure, the penstocks and the Power Station.

Figure 6. Aerial view from right bank

Type of Dam: TE - Heterogeneous, with clay core (ICOLD classification)

The dam was substantially completed in December 1972 and first impounding took place from May-December on the same year.

Crest length: 2500 m.

Dam height: 86 m (above the lowest foundation level).

Foundation type: Rock (Sandstone) – Alluvium.

Spillway capacity: 8000 m^3/s (located on the right bank with centre line about 100 m from the right abutment).

Downstream population: 500.000 inhabitants.

The clay core, which is in the middle, it is founded in the rock and inclined toward downstream. This sloping core was selected to reduce the differential settlement between the clay core and the gravel shells (founded in alluvium) and to reduce downstream foundation shear stresses in the weak layers of argillite within the sandstone. The upstream slope is protected for waving erosion with a basaltic mantle of rip-rap.

The original impervious curtain in the foundation was made by cement injections through drills in the rock, crossing the entire dam under the core.

Several types of instruments were installed in the dam during construction. Later, during the treatments work new instruments were installed which were linked to the Automatic Monitoring System.

Today, El Chocón Dam has an extensive and complex instrumentation system for monitoring deformations, movements, pore pressure, seepage flow, moisture and water temperature.

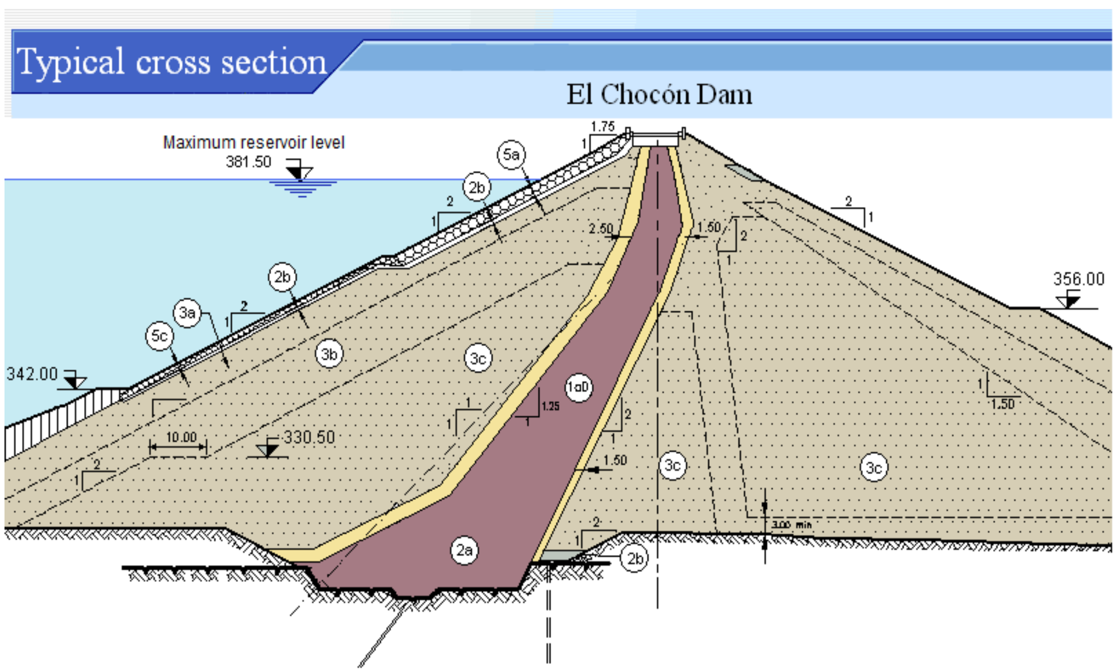


Figure 7. El Chocón Dam, Typical cross Section

El Chocón Dam has an automatic monitoring system (called SADA) which it is possible to evaluate the structural performance in real time (one measurement every 10 minutes by instrument). This way, SADA is a very important tool in early detection of any anomaly that may occur in the dam or their foundation.

The management of databases dam safety (data verification, reduction, storage, structural behavior analysis, graphical representation, alarm reports, etc.) is performed using an internally developed software called Fiona®.

The Figure 8 shows the instrumented cross section 1100.

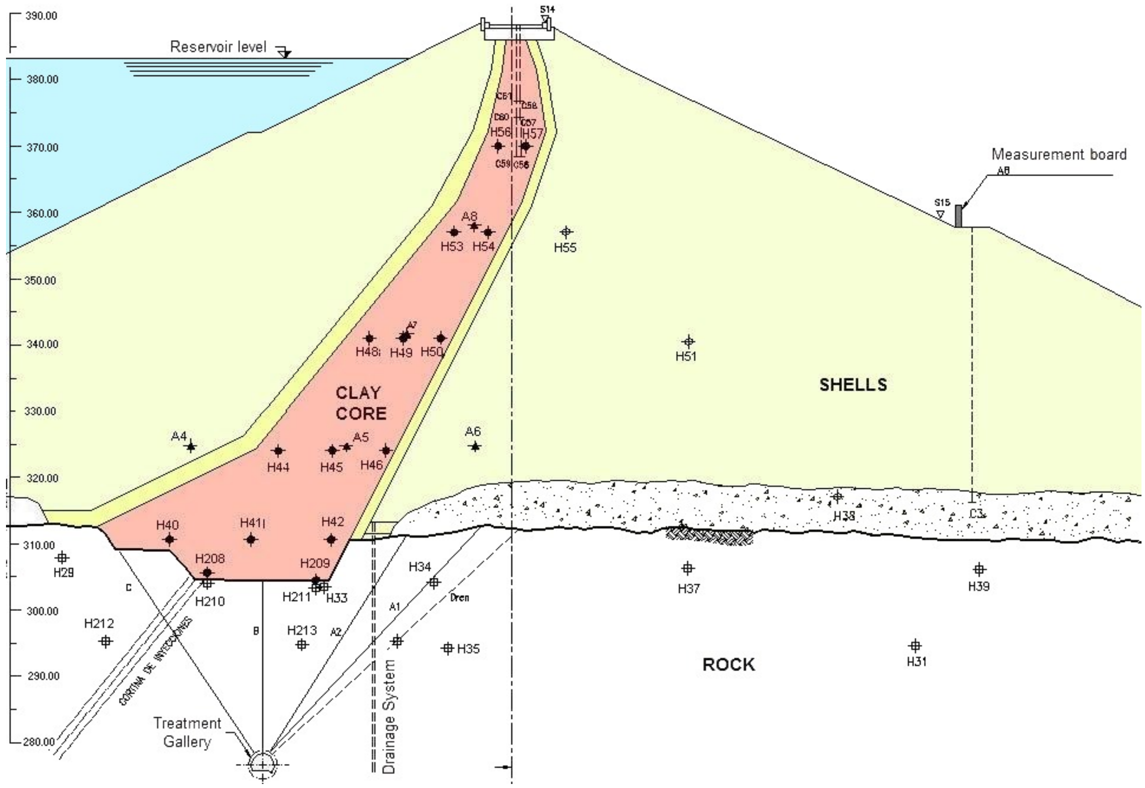


Figure 8. Instrumented cross section 1100

For analysis were taken only the instruments installed in the core and were compared with the evolution in the level of reservoir (NE) as the main causal variable.

A Matlab® Script was programmed with which all calculations and graphs for analysis were performed. The output graphics for the section analyzed are shown below.

Figure 9 shows the Pearson’s correlation matrix. In the diagonal are plotted the frequency histograms of measurements discretized into 10 units

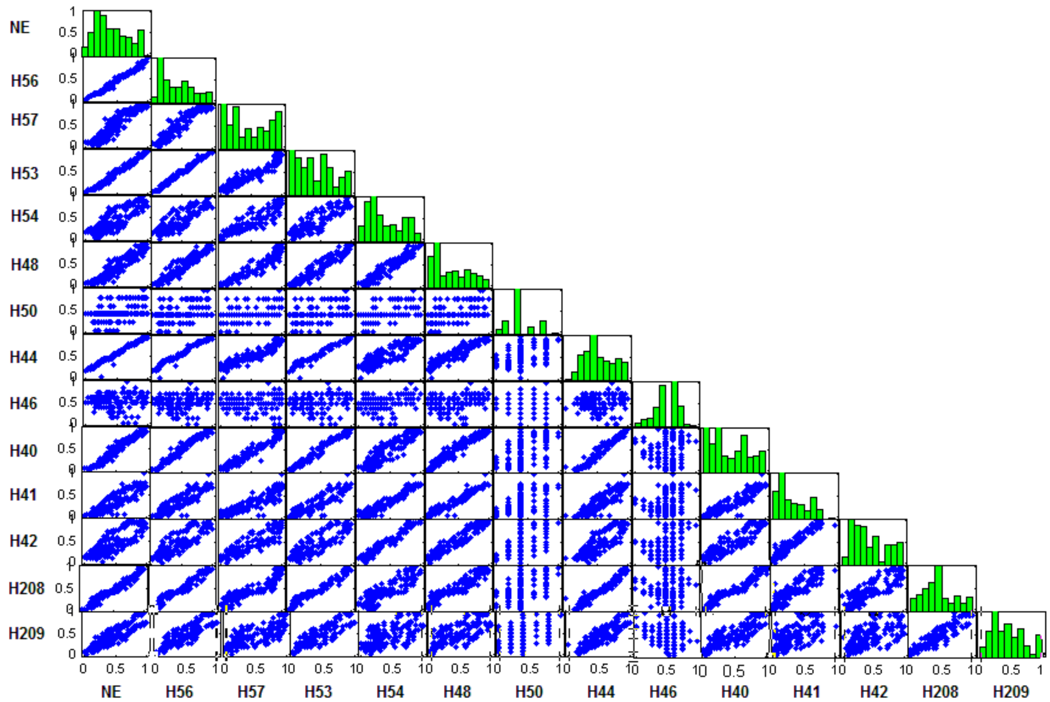


Figure 9. Pearson’s correlation matrix

Figure 10 shows the graphical output by Factorial Analysis.

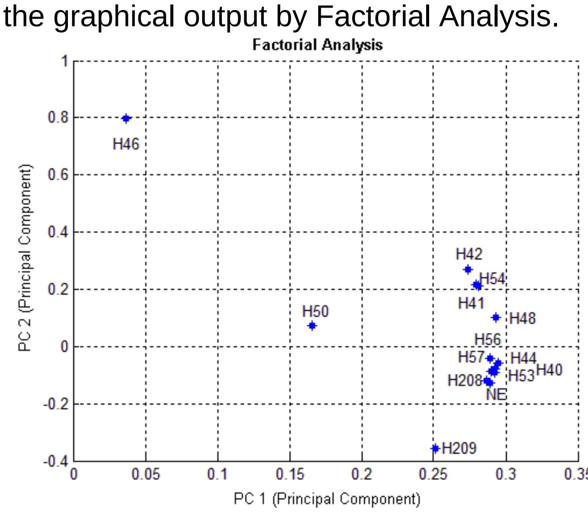


Figure 10. FA - Graphical and numerical output

Reducing the problem to 2D (principal component 1 and 2) is validated by Jolliffe’s Method, where an analysis of inertia of eigenvalues was made, as shown in Figure 11.

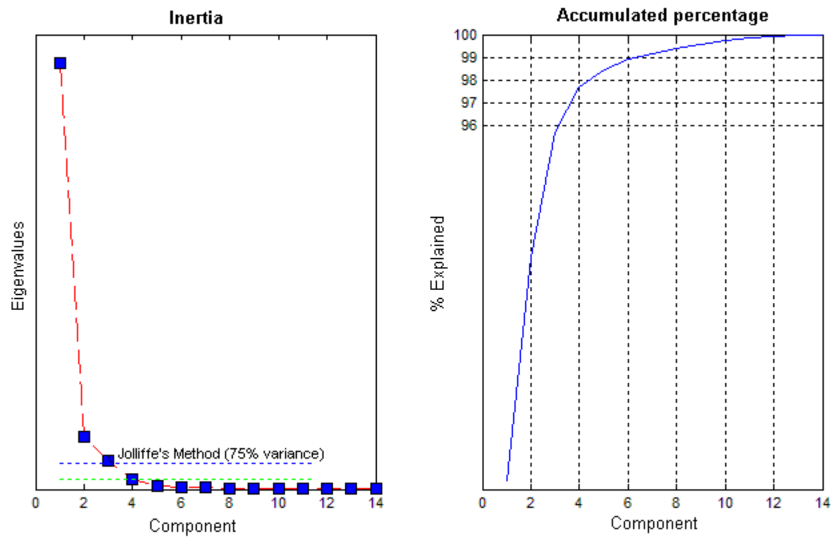


Figure 11. Inertia of eigenvalues

In general, systemic analysis of the entire core shows good behavior of the instrumented section. The Pearson's matrix shows that a single negative is observed in tandem H46-H209, the first with higher concentration of records than in the downstream heel.

Factorial analysis shows that the first two principal components explained 80% achieved the level set, while the Cluster analysis indicate a high cophenetic correlation, which allows grouping (given a good and varied spatial distribution) instruments in 4 patterns of piezometric response (see Figure 12 – Resulting Dendogram and cluster distribution):

1. A superior group and closer to the wet side (composed by H56-H57-H53-H44-H40-H208) except in the middle sector of the nucleus (level 340) responds very quickly to changes in the reservoir level with very small delays.
2. An additional group conformed by the upper-half of the core (H54-H48-H50) and the middle and downstream heel (H41-H42),
3. H209 downstream in the heel of the section, and
4. Finally, H46 (level 325, located in downstream sector).

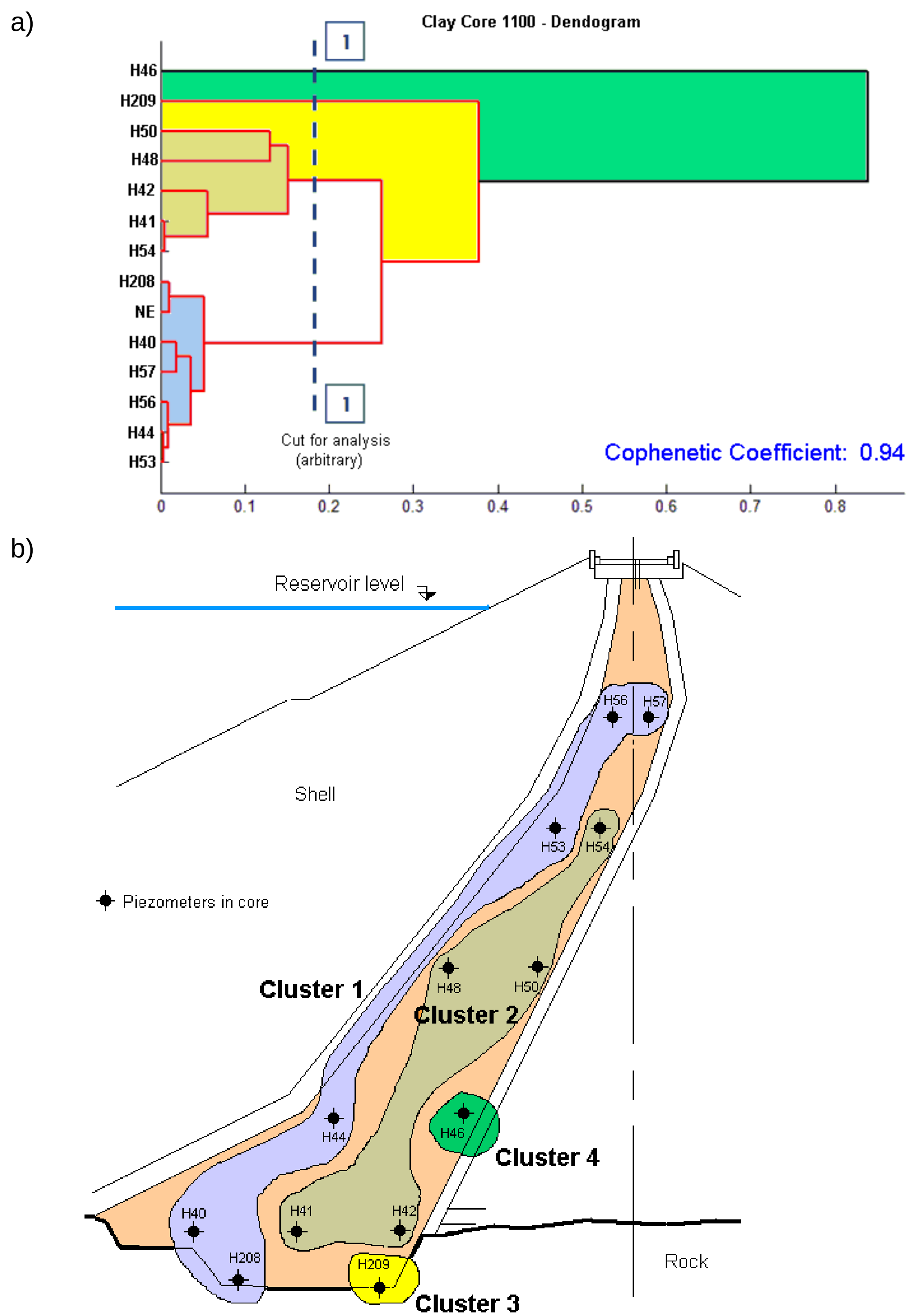


Figure 12. a) Resulting Dendogram b) Cluster distribution

Piezometer	Place	Cluster	Delay
H56	Clay core	1	0.05
H57		1	0.09

H53	1	0.04
H54	2	0.34
H48	2	0.23
H50	2	0.24
H44	1	0.04
H46	4	0.96
H40	1	0.07
H41	2	0.34
H42	2	0.40
H208	1	0.02
H209	3	0.39

Table1. Summary Table – Clusterization and numeric results

CONCLUSIONS

Research was carried out for a temporal period of 4 years during which the reservoir level remained within its normal operating range.

The methodology was used for the instrumented sections of the dam and its foundation and both abutments. In all cases, the high correlation rates among the instrumentation responses not only confirm the methodological validity of the Factorial Analysis, but also allow for interrelated clusters of instruments to be explored, selected and classified.

The investigation has been established using the method of distance between centroids to obtain the first classification by cluster analysis. However, the results may be somewhat different compared with other numerical grouping criteria (eg. nearest neighbors, more distant neighbors, averages across all points, etc.).

In order to equalize the changes and compute them systemically it is necessary to use standardized measurements.

The exploratory nature of the research allows us to meet the following goals:

- It is possible to eliminate redundant information contained in the original variables.
- It is possible to perform a transformation of the initial set of variables into a smaller one in order to enable the graphic interpretation and the quantification of delays.
- The proposed algorithm allows (depending on the needs assessment) to establish a cap on the "cut" the dendrogram to stop the classification process.
- This new tool enables the holistic analysis so as to have a more integrated vision when the response of dam monitoring instruments has to be analyzed.

Therefore, it seems reasonable to recommend making new efforts in order to broaden researches since this type of tools will be increasingly used in the next few years, giving a significant progress in the possibilities to interpret the results of the behavior of large dams.

References:

1. BALLABIO, O. – (2006). Chemometric characterisation of physical-chemical fingerprints of food products. PhD Thesis. Department of food science and microbiology. University of Milan. ITALY.
2. CHAGOYEN, M. et al – (2006). *"Discovering semantic features in the literature: a foundation for building functional associations"*. BMC Bioinformatics, 7:41.
3. HONG YU et al – (2010). *"Multivariate Analysis in Dam Monitoring with PCA"*, in Science China Technological Sciences, Volume 53, Number 4, Pp 1088-1097.
4. RESTELLI, F. – (2010). *"Técnicas del Análisis Factorial para la evaluación grupal de niveles piezométricos"*, en CACIER 2010, Posadas – ARGENTINA.
5. RESTELLI, F. – (2010). *"Evaluación sistémica de la instrumentación de Auscultación utilizando ACP"*, en Congreso Argentino de Presas, Neuquén – ARGENTINA.
6. RESTELLI, F. – (2013). *"Systemic evaluation of the response of large dams instrumentation"*, in ICOLD Proceeding 2013 – Seattle (USA).
7. WOLD, S. et al. - (1987). Principal component analysis. Chemometr. Intel. Lab. Syst., 2: 37-52.