Determination of an empirical model for the deformation of a dam on the basis of the hydrostatic thrust and temperature

[Julio Manuel de Luis Ruiz, Raúl Pereda García, Rubén Pérez Álvarez, Felipe Piña García, Elena Castillo López]

Abstract— **The deformational control of structures has its main exponent in the geometric control of large dams. Considering the constructive typology, they can suffer deformation, which is function of the hydrostatic thrust and the temperature that supports the dam under usual circumstances.**

The project designer usually calculates the deformation that the dam will suffer on the basis of thrust and temperature, but this calculation is theoretical, and it must be contrasted with the real data about deformation that are periodically observed in the dam, in order to certificate the adequate evolution of the structure through time.

This research work is focused on the determination of an empirical model that allows the calculation of the deformation on the basis of the simultaneous observations of a direct pendulum that have been recorded with the correspondent temperatures during 14 years. This empirical model is later contrasted and validated with the records taken during the subsequent year. Although it is only valid for the dam where the tests were held (La Cohilla Dam, Cantabria, Spain), the methodology is suitable for any other structure.

The main contribution of this research is the methodology itself, which allows obtaining an empirical model that determines the expected deformation, whatever the conditions of hydrostatic thrust and temperature are. It complements the theoretical model that the project designer has established, which is done just for a certain set of conditions of these variables.

Keywords— **Auscultation, geometric control, pendulum hydrostatic thrust, temperature.**

Julio Manuel de Luis Ruiz. Researcher ID: B-4956-2015

School of Mining and Energy Engineering / University of Cantabria / Spain

Raúl Pereda García. Researcher ID: B-5240-2015

School of Civil Engineering / University of Cantabria / Spain

Rubén Pérez Álvarez. Researcher ID: B-8959-2015

School of Mining and Energy Engineering / University of Cantabria / Spain

Felipe Piña García. Researcher ID: B-5710-2015

School of Mining and Energy Engineering / University of Cantabria / Spain

Elena Castillo López. Researcher ID: M-1310-2015

School of Civil Engineering / University of Cantabria / Spain

I. Introduction.

Current laws in Spain require companies that operate large dams to apply periodical controls to assess a proper structural behavior. For this purpose, suitable devices are installed on the dams to assess their geometric state, which is intimately bound up with the stress state of the structure, and to measure the major variables that determine that state.

The most common device for the definition of the geometric state of the structure is the pendulum or plummet. Independently of being direct or indirect, it continuously registers the movement of the point of the structure where it is placed, usually at the center, on top of the dam. Surveying devices are usually applied to measure the variables on which the displacement depends. They allow determining the height of the stored water, and therefore, the hydrostatic thrust that is applied to the structure. Meteorological stations permit assessing the thermic variations that the structure suffers, and their consequences.

In addition to the main variables on which the stress state of the structure depends, hydrostatic thrust and temperature [1], the availability of other devices aimed to develop geotechnical monitoring in the surroundings of the dam, such as piezometers, strain gauges, inclinometers, etc., is common. Except for specific cases, they provide very little information during the stage of exploitation of a certain structure.

During the development of the project of a dam, the engineer usually calculates the theoretical movement that the dam should have according to a hypothesis of hydrostatic thrust and temperature in case that it should exist, given that not every constructive typology requires the movement of the structure under external demands [2]. One first drawback arises here, as the availability of an estimated theoretical displacement that is only valid under certain conditions of the variables on which the movement depends is not very useful for the operating company, since the geometric control must be developed in a continuous way, with independence of the state of the mentioned variables. In addition to this, the theoretical displacement is generally determined according to the theoretically projected structure, which is not always identical to the one that has been really executed, due to imponderables that arise during the stage of construction, and are a source of uncertainty, and therefore of inconvenience for the company that must control the structure [3].

Given the above, it is necessary to establish the actual framework of the movement obtained with the really executed structure, whatever the conditions of hydrostatic thrust and

temperature are. This research aims to analyze the relation between the real planimetric movements of a structure under these two variables, so as to define a hypothetical mathematic correlation that allows foreseeing the movements under any initial hypothesis.

It is worth noting that this research is the continuation of the initial one, which was initially focused on the linear regression between the displacement and the hydrostatic thrust [4], and subsequently advanced towards the linear regression between displacement and temperature [5]. This work proposes a multiple linear regression between the displacement and the couple height of stored water-temperature, so as to find the regression that provides the best fit to reality and solves the problem posed.

II. **Instrumental and Materials.**

All the data that have been applied to develop this research belong to the observations that have been held in a large dam which is known as "La Cohilla Dam" (Cantabria, Spain). It sets up a water storage of 10 cubic hectometers that constitutes the header dam of a hydroelectric facility that is called "Saltos del Nansa", which comprises other three dams located downstream along River Nansa, that constitute the supply itself.

Due to its constructive typology and to the changes in the conditions of the variables that determine its stress state, movements of up to 12 mm are registered in the center of the dam crown. As it counts with the instrumental (pendulum and meteorological station) and the required records, the dam is ideal for the development of this research. Figure 1 gives an idea of the dimensions of this structure, which is an arch dam with an approximate size of 240 m in width and 110 m in height.

Figure 1.   La Cohilla Dam (Cantabria,Spain).

The pendulum basically comprises a very long wire with a mass at its lower end. When the gravitational field operates on the mass, the gravity force that is generated tightens the wire and makes it follow the directions of the field lines (vertical). Considering the case of La Cohilla Dam, the pendulum is direct, and it is placed at the center of the dam, on its top. Since the pendulum hangs from its upper side, any movement that could happen is transmitted to the lower end, where the reading micrometer is placed.

Although the pendulum determines the position of a point in a continuous way, the records are taken every 3 or 4 days, since reading is not automated and the person in charge of the dam must get to the location of the micrometer. This is not always possible, due to the adverse weather conditions that are frequent in this area, specially in winter.

The pendulum determines any kind of movement that happens at the center of the dam crown, although due to the geometry of the structure itself and its circular base, this displacement is mainly radial. Figure 2 shows the pendulum reference system and the characterization of the tangential and radial movements.
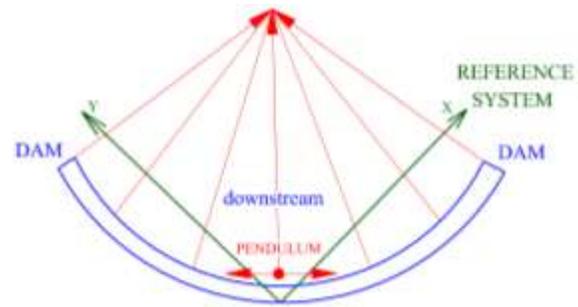

Figure 2.   Pendulum reference system.

The determination of the stored water height is held by means of a continuous register based on a level staff, and the temperature is measured with a meteorological station located on the dam crown. This measurement is taken right after reading the pendulum.

III. **Fundamentals.**

A. *Relation techniques.*

Relation techniques aim to determine a possible relationship among a dependent variable and one or some other independent variables. It is also known as correlation. On a first stage, the existence or absence of relationship between two variables is evaluated (if this relation is statistically significant), and after that, the possibility to classify that relation as belonging to a certain type is assessed (direct or indirect relationship, etc.). Finally, a standardized process permits quantifying it by delimiting the values between which it is operational.

It is worth noting that this research is focused on quantitative variables. Had the relationship among variables been determined, it can be mathematically modelled by applying regression procedures. [6].

B. *Correlation.*

The mostly applied correlation for quantitative variables with a normal distribution is Pearson correlation. If the variables do not show a normal distribution, Spearman or Kendall correlations are frequently used [6]. Both correlations have become common in most statistical software, but checking on a first stage whether the variables are normally

distributed in order to define the type of correlation to be applied is important.

There are three main elements to consider when analyzing a correlation:

- Sign: The correlation, expressed as the grade of relationship between two variables, is quantified by a value which is known as "Correlation Coefficient". It is comprised between -1 and +1, which implies positive and negative correlations. A positive correlation implies that the variables have a direct relationship, while a negative correlation means that it is inverse.

- Magnitude: The second aspect to consider in a correlation is the magnitude, which is defined by the absolute value of the Correlation Coefficient. The closer it is to the interval limits (-1, +1), the bigger the correlation between variables is.

- Significance: The definition of a correlation as statistically "significant" implies that the result is not very likely to have occurred as a consequence of randomness in a certain sampling, while a "non-significant" correlation doubts of the representativeness of the sampling. P-value measures the statistical significance. It ranges between 0 and 1, with two well differentiated zones in this interval. A correlation is significant if its p-value is less than 0.05. If it is greater, it is considered as non-significant and the considered correlation coefficient must be zero.

C. *Regression.*

When a significant relationship among two or more variables is detected, modeling it by means of a mathematic equation that allows predicting the dependent variable with respect to values of the independent variable is usual. Regression consists in the development of an equation that models the relation among two or more variables, which is not exact.

In order to characterize the different types of regression, a first classification is usually set by considering the number of dependent variables. Single-variant and Multi-variant Regressions are defined by the existence of one or more dependent variables. The second stage of characterization distinguished between Linear and Non-linear Regression. The first is modeled by the equation of a linear function, such as a straight line, plane or hyperplane, depending on the number of variables, while the suitable function for the latter could be non-linear, such as exponential or logarithmical expressions. On a third step, they can be classified according to the number of independent variables. If there is just an only independent variable, they are called Simple Regressions, while if there are two or more, they are known as Multiple Regressions [6].

IV. **Methods.**

So as to assure that the data used for this research is significant enough, those obtained during 14 years (from 1st

January, 2000 to 31st December, 2013) have been processed. This implies a sample of 1242 records which can be considered as statistically consistent.

On a first step, coordinates have been turned into differences of coordinates with respect to a certain position in order to work with displacement instead of coordinates, and to be able to develop a mathematical fit. The rearmost position of the dam, which is fixed by the position with the lowest hydrostatic thrust, is adopted as Po location to ease the interpretation of this displacement. Had the field observables been checked, it was determined that log 4903, from 14th September, 2014, provided the rearmost position of the dam among the whole dataset [1.27/4.50]. It was due to a malfunctioning at the bottom drainage that left the water level at 9.02 m.

Applying the following expression:

$$D = \sqrt{(X - 1.27)^2 + (Y - 4.50)^2} \qquad (1)$$

The displacement with respect to the initial position is obtained, which, along with the height of the stored water and the average temperature measured during the previous 30 days, comprise the data that will be aimed to be correlated later. It is worth noting that the use of the average temperature for the previous 30 days could be discussed, but it tries to palliate the great variability of temperatures that can be registered in so little intervals of time. The structure is not able to assume them, and therefore the effects of the thermal gradient can are softened. Table I shows the variables that are analyzed in this research.

TABLE I. PENDULUM OBSERVABLES, TEMPERATURE AND WATER HEIGHT.

Record	Date	X Coord. (mm)	Y Coord. (mm)	Displac. (mm)	Temp.	Water height (m)
3777	04.01.00	8.16	11.84	10.1	0.0	58.9
3778	11.01.00	7.57	11.10	9.1	-0.4	53.1
3779	14.01.00	7.13	10.69	8.5	-0.3	49.9
3780	18.01.00	6.09	9.66	7.1	-0.7	44.8
3781	21.01.00	5.63	9.19	6.4	-1.1	40.2
3782	28.01.00	4.19	7.54	4.2	-1.2	31.6
….	….	….	….	….	….	….

The first column shows a number that indicates the order of observation. The second one specifies the date in 'day.month.year' format. X and Y coordinates are expressed in mm and defined according to the local reference system, which was set when the pendulum was instaled in the dam with the geometry shown in Figure 4. The displacement is determined on the basis of the increment of coordinates established by the pendulum (mm). The following column shows the average temperature for the previous 30 days in ºC,

while the last one comprises the stored water height, and therefore, it represents the hydrostatic thrust.

Had the dataset been processed, the Multiple Regression is developed by considering the displacement as the dependent variable, and the stored water height and the temperature as independent variables. After that, they are contrasted and analyzed with respect to the Simple Regressions developed with the stored water height and the temperature.

V. **Results.**

In [4] it is justified that:

$$D = 0.2174 \cdot Z - 2.3652 \qquad (2)$$

In [5], it is justified that:

$$D = 9.0737 - 0.3935 \cdot T \qquad (3)$$

As it is described in [5], the application of Anderson-Darling test probes that the sample of data that represents the independent variable temperature does not have a normal distribution. Hence, Spearman correlation is required. By applying this correlation to both the dependent and independent variables, the following expression is obtained:

$$D = 0.218 \cdot Z + 0.003 \cdot T - 2.411 \qquad (4)$$

Where:

D = Displacement (mm.)

Z = Stored water height (m.)

T = Temperature (ºC)

Table II shows the results of the three linear regressions. The linear regression that has been applied in this research can be noticed on top of the table. Below it, the table lists the results of the linear regressions that were obtained individually with the stored water height and the temperature in previous works.

TABLE II.　　RESULTS OF THE LINEAR REGRESSIONS

MULTIPLE LINEAR REGRESSION				
	Coeficients	Standard Error	Statistical parameter t	Probability
Interception	-2.411434	0.049428	-48.786946	1.111E-290
T Variable	0.003187	0.002756	1.156264	2.478E-01
Z Variable	0.217959	0.000868	251.126391	0
SIMPLE LINEAR REGRESSION (T)				
Interception	9.073726	0.135003	67.211052	0
T Variable	-0.393529	0.016262	-24.199761	2.799E-106
SIMPLE LINEAR REGRESSION (Z)				
Intercepción	-2.365215	0.029078	-81.341448	0
Variable Z	0.217383	0.000711	305.625668	0

The review of the standard error, the statistical parameter t and the probability, indicate that the relation has a very little

variability around the line of regression, in addition to its significance. An analysis of the previous table leads to afirm that the new Multiple Linear Regression is significant as a whole. Table III shows the most representative statistic values provided by the development of the same regressions.

TABLE III.　　RESULTS OF THE STATISTIC VALUES OF THE REGRESSIONS.

MULTIPLE LINEAR REGRESSION			
Linear correlation coefficient	Coefficient of determination R^2	Fitted R^2	Standard Error
0.99343486	0.98691282	0.98689169	0.29399807
SIMPLE LINEAR REGRESSION (T)			
0.56637605	0.32078183	0.32023407	2.11714604
SIMPLE LINEAR REGRESSION (Z)			
0.99342775	0.98689869	0.98688813	0.29403801

A review of the values of the correlation coefficient, the coefficient of determination, the fitted coefficient of determination and the standard error, allows stating that the new correlation can be considered as positive and very high. In addition to this, it can be noticed that the Multiple Linear Regression improves. The improvement is minimal with respect to the Simple Linear Regression based on the stored water height, but it is remarkable when compared with that of the temperature, whose relation could be considered as simply moderated [7].

The analysis of Table III indicates that the new Multiple Linear Regression is strongly correlated as a whole. Table IV shows the results of the analysis of variance (Anova) developed for the three regressions [8].

TABLE IV.　　RESULTS OF THE ANALYSIS OF VARIANCE OF THE REGRESSIONS.

ANALYSIS OF VARIANCE, MULTIPLE LINEAR REGRESSION				
	Degrees of freedom	Sums of squares	Average of squares	F
Regression	2	8075.9349	4037.9674	46716.8839
Residuals	1239	107.0928	0.0864	
Total	1241	8183.0277		
ANALYSIS OF VARIANCE, SIMPLE LINEAR REGRESSION (T)				
Regression	1	2624.9666	2624.9666	585.6284
Residuals	1240	5558.0611	4.4823	
Total	1241	8183.0277		
ANALYSIS OF VARIANCE, SIMPLE LINEAR REGRESSION (Z)				
Regression	1	8075.8193	8075.8193	93407.0491
Residuals	1240	107.2084	0.0865	
Total	1241	8183.0277		

Considering the Anova analysis developed for the three regressions, it can be stated that the statistical parameter F allows contrasting the null hypotesis, which implies that the population value of R is zero in the Multiple Linear Regression, and in the case of the Simple Linear Regression it is equivalent to contrasting the hypothesis that the slope of the line of regression is equal to zero [9].

VI. **Discussion.**

The first line of discussion to consider about this research is focused on the type of regression adjustment that has been applied. A multiple correlation with two independent variables (Stored Water Height and Temperature) has been developed. Both variables have been considered as the only causes of deformation in the dam, at the expense of other possible multi-variant correlations that could be related to the deformation of the dam.

The second line of discussion is centered on the possibility to use other non-linear correlations. In this regard, it is worth noting that, among all the developed adjustments, the linear correlation has provided the best correlation coefficient. Although other alternatives, such as polynomials of different grades, exponential expressions, potential adjustment, etc., could be applied, the distribution of the observables clearly suggest a linear adjustment. In addition to this, this research is derived from other works developed by the authors that properly justified the linear regression as the most suitable one.

The third line of discussion that arises from this research is related to the typology of the deformation with which the research itself is developed. The decomposition of the planimetric displacement in radial and tangential implies that most of the deformation due to temperature is exclusively tangential deformation. This justifies that the relation obtained with the Linear Regression is simply moderated. However, the displacement due to the hydrostatic thrust (stored water height) has a mainly radial component, which causes that the Simple Linear Regression with the stored water height has a very high relation. This leads to think that working with a Multiple Linear Regression is more complicated than with a Simple Linear Regression, being the error committed with the latter perfectly assumable.

The last line of discussion arises from the observation of temperature, which is measured outside the dam, and therefore, it does not match with the temperature within the structure. In this regard, the average value of the environmental temperature during the previous 30 days was adopted. This thermal gradient fits better with the time that takes the structure to modify its temperature. It is worth noting that other averages could have been adopted, but the references consulted recommend 30 days.

VII. **Conclusions.**

In view of the developed methodology, it is worth presenting as the main conclusion the easiness provided nowadays by software for the management of massive amounts of data and the development predictive mathematic tools that allow the person in charge of the dam to know the a priori behavior that the structure must have, and even to obtain behavioral models for critic states.

The results obtained allow setting a conclusion that is focused in the case study. An expression that allows defining the "a priori" displacement suffered by a pendulum that is located at the center of the dam crown has been obtained. The contrast of the "a priori" value thus obtained with the one observed in field allows defining another indication about the structural safety of the dam, which is always a reason to worry about for people in charge of the dam.

This research has developed a methodology to relate the variables that affect the possible deformation of a structure. In this case, it has been applied to a certain structure, but it could be applied to other types or variables, in order to establish predictive models that delimit their deformation under extreme conditions.

References.

[1] Sundakov, Y.A., "Trabajos geodésicos en la construcción de grandes obras industriales y altos edificios", Editorial Mir, 1980.

[2] Ferrer Torio, R., et al., "Geodetic Deformation Monitoring: From Geophysical to Engineering Roles", International Association of Geodesy Symposia, Vol. 131, 2005, pp. 270-276

[3] De Luis Ruiz, J.M. "Auscultaciones geodésicas por métodos clásicos y con Láser Escáner". Editorial Académica Española, 2012, pp. 179–184.

[4] De Luis Ruiz, J.M., Pereda García, R., Piña García, F., Pérez Álvarez, R., "Determination of an empirical model for calculating the strain of a dam in terms of hydrostatic thrust". ISSN 2372-3971

[5] De Luis Ruiz, J.M., Pereda García, R., Piña García, F., Pérez Álvarez, R., Analysis of the correlation between deformation and temperature in a concrete dam". ISSN 2372-4102

[6] Llopis Pérez, J., "La estadística: una orquesta hecha instrumento", Editorial Ariel, 1996.

[7] Fan, H., "Theory of Errors and Least Squares Adjustment". Royal Institute of Technology. Stockholm. 2005.

[8] Luceño Vázquez, A., et al., "Métodos estadísticos para medir, describir y controlar la variabilidad", Editorial Universidad de Cantabria, 2005, pp.68–73.

[9] Sevilla de Lerma, M.J., "Colocación mínimos cuadrados". IV Curso de Geodesia Superior, Instituto de Astronomía y Geofísica, Madrid, 1987.