Utility of Graphic and Analytical Representation of Simple Correlation, Linear and Nonlinear in Hydro-climatic Data Analysis

Elena GRIGORE Ph.D. Assistant
University of Bucharest, Meteorology-Hydrology Dept.
Teodor Ioniţ POIANĂ Student
University of Bucharest

Abstract
Through the appearance of analysis models was possible to simplify the graphical representation of the various processes of the hydrology, and also from the climatologic domain. Current scientific research imposed the deepening of operating mechanism and the way of manifestation of these hydrological and climate systems, because any change in their evolution affects the components and relations between them. Correlations generally, express the relations between two or more variables, being such as explanatory and explained, more precisely some are independent, others dependent. Correlations are particularly useful in hydro-climatic research, because they can check and complete statistical data series and determine new hydrological parameters.

Key words: correlation, relations, analysis, statistics, graphics, simulation.

In the specialized literature has been specified, analyse and showed that any change or modification manifested in the natural environment, as well as the anthropic, is necessary to be seen as a vast and complex process, with an impact visible or less visible in the manifestations mechanism of the components of that system. For this purpose, along time, have been developed different methods and models useful in quantitative and qualitative evaluation of the impact of such changes or modifications. Thereby, currently, evaluations are performed using mathematical models elaborated on the idea that their operating principle is based on the quality of the input data in the model, thereby ensuring the quality of the type of response provided by it, because they offer the possibility of the original
data to be modified. Are preferable these types of models because they provide more answer variants, which makes achievable generation of scenarios which fulfils demands, but also obtaining useful information in restoring of a disturbed balance, either natural or anthropic. The emergence and refinement of mathematical models for analysis, presented both as physical or mathematical form, offered the possibility to simplify the graphical representation methods useful in the analysis, these bringing more expressiveness to visual playback, while imposing deepening of operating mechanism and the manifestation of hydrological and climatic systems, because any modification or change in their evolution affects the components and relations of the entire analyse system.

Through means of this analysis is attempted to raise awareness of student community, but not only, by deepening the mechanisms of modern research and the use of these methods in different studies. For this reason, this study attempts to synthesize methods of analysis and graphical representation of hydrological and climate data, as hydro-climatic changes and risks are common, with an impact visible on nature and anthropic environment, being, also, widely publicized and analyse.

1. Input data and methods

Useful in hydrological and climatic analysis are, in the first place, Initial statistical databases provided by the points of observation and instrumental measurements, also being included the way for their management, but also the model type of graphical representation used in hydro-climatic modelling depending on the nature of the variables and physical, hydrological and climatologically parameters used. The models used have been developed at both a regional and local scale, these being a physical nature models, analogical models, mathematical or numerical models, but also the equilibrium models, transition models etc. In a synthetic expression, correlations, according to various scientific sources, can be seen as a final product essentially expressing the relationship between two or more variables or the result that shows and points the degree to which a variable is dependent on another variable or several, conclusion in both cases being likely explanatory and explained.

In this analysis the correlation method was chosen for presentation in terms of the quantitative relationship of a repository, as a bridge between the real world and a system which confirms the statistical dependence between quantitative or ordinal variables of the analyse case. Regarding the geographical field, correlation is built on the grounds of statistical data of a geographical nature, which represents the connection inside but also the
outside of this phenomenon, connection supported and confirmed by the specialized arguments, but who presents a false correlation due to data independence. For this reason the analysis of hydro-climatic correlations role of simulation of the evolution of natural/anthropic processes regardless of complexity (even if the models are not perfect) are useful in practical analysis of hydro-climatic data. The elements are analyse by analytical expression of the links, which may be linear or non-linear (curvilinear) links, by the indicators that determine the intensity of the link as: covariation, correlation coefficient and the correlation ratio etc., that determine the final form of the correlation to be expressed practically by the method of correlation table and graphical method.

2. Results and discussions

In geographical practice it is frequently used simple correlation, Linear and nonlinear (Table 1), the basis of analysis being given by the correlation that express the connection that exist between the elements analyse, one of which may be appreciated or quantified. For example, in the analysis of a watercourse is noted that no direct measurements were made, correlations are useful in highlighting some relations such as: relationship between river flow and water level; relationship between rainfall - liquid flow of the river etc., because these relationships are influenced in these cases by the external factors.

<table>
<thead>
<tr>
<th>Type of correlation</th>
<th>Linear</th>
<th>Nonlinear monotone</th>
<th>Nonlinear nonmonotonic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
</tr>
<tr>
<td>Inverse</td>
<td><img src="image4" alt="Diagram" /></td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
<tr>
<td>Strenght measuring instruments:</td>
<td>$R_{Bravais-Pearson}$; $R^2$</td>
<td>$R_S$ Spearman</td>
<td>Special tools, little known, often missapplied</td>
</tr>
</tbody>
</table>

| Table 1 | Types of correlation and their intensity |
From mathematical point of view, graphic expression is given by a regression line, resulted from the mathematical relations: \( X = f(Y) \) or \( Y = f(X) \) or by the regression curve, representing virtually a function expressed by a parabola, hyperbola, exponential etc. For example, correlations are particularly useful in hydrological research, because they can check and complete statistical data series and determine new hydrological parameters. Graphical representation expressing the correlation is built based on the pairs of observed values \( x \) and \( y \), represented in a system of rectangular axes, where OX axis presents the independent variable \( x \), and the OY axis the dependent variable \( y \). In the following collage of images, were detailed using images saved in .jpg the steps useful in achieving the coefficient of correlation and exemplified various types of graphics resulted from the application of the correlation function. In practical analysis Excel is most commonly used, by accessing the Insert Function menu respectively CORREL function, which according to developer is useful in the calculation of correlation coefficient (Function Arguments menu), where \( array1 \) is the first value series, and \( array2 \) the second set of values, being necessarily the same number of cells in both ranges of values (Fig.nr.1).

![Fig.1 The calculation method of the variables by using CORREL function and obtaining the coefficient of correlation](image1.jpg)
By calculating in this way we obtain linear correlation between two variables, and after calculating the correlation coefficient, using Chart Tools menu we achieve different graphical forms, being useful the list provided by the program (Format Trendline menu), which offers the possibility to draw linear, exponential, logarithmic graphs etc. To facilitate the interpretation of these representations, we will highlight the various calculation formulas for the presented graphics: linear chart represents, in fact, a line model: $y = a + bx$; polynomial chart, of different order: $y = a_0 + a_1x + a_2x^2 + \ldots + a_kx^k$; logarithmic chart: $y = a + b \ln x$; exponential chart: $y = a e^{bx}$; power chart: $y = a x^b$; moving average chart, represents the model type of sliding average, in which can be calculated a new series with values obtained as the arithmetic mean of the values from the original series, based on the formula $y_n = \frac{x_n + x_{n-1} + \ldots + x_{n-k+1}}{k}$, where $k$ is the model order and represents, generally, the model which removes short-term influences.

Fig. 2 Graphical representation of correlations with Chart Tools menu and Format Trendline (linear, logarithmic, polynomial, power, moving average)
Regarding the **Order** command it is active for the polynomial model, because it establishes the model order, which can be up to 6, and **Period** command is active in the case of **Moving Average**, thus being the possibility of establishing the model order, more correctly expressed how many elements are needed to calculate the arithmetic mean. Also, **Based on series** command allows designation of represented series, estimated by the method of least squares of the selected model. The effect of the procedure for adding the representation line (trend line) on the surface of the graph will appear as a separate line, a hypothetical series obtained by calculating the trend.

For a better choice, scientific and practical correct, we can use prior information or those visually provided by the aspect of graphical representation and the arrangement of points. The most popular and commonly used correlations can be listed correlations between/among: maximum specific flow $q_{\text{max}}$ measured in l/sec/km² and the average altitude of hydrographical basin: $H_{\text{med}}$, measured in m; maximum specific flow $q_{\text{max}}$ measured in l/sec/km² and surface of hydrographical basin $F$ measured in km² (logarithmic correlation: $\lg$); $T_{\text{cr}}$ or $T_{\text{tot}}$ and length of the river until the section of calculation; $T_{\text{cr}}$ or $T_{\text{tot}}$ and ratio of the length of the river and its slope $L/l_r$; rainfall in mm and altitude from analyse region in m; air temperature in °C and the altitude in m; evaporation and altitude; different pollutants etc.

Analysis based on graphical visual expression indicates that all correlations have territorial validity, observing that it is not just a simple curve that ensures the connection between two variables, but several more parallel curves, which may be linked to each other. We recall that, usually in the hydro-climatic data analysis, points may show greater scatter in the graph field of correlation, occurred due to numerous local factors influencing elements at the correlation. Also, drawing curves or lines of correlation require knowledge of reality, control factors, but also experience. In its turn a lower scattering of points on correlations can be seen in the case of insured values or with different probabilities, possible situation in real cases, where local factors are well defined and expressed, so, in the insured cases, local influences attenuate the phenomenon. Useful models in the hydro-climatic analysis presents advantages and disadvantages, their explanations being summarized in other specialized papers. It is recommended that in the analysis should be used various methods, because it is vital to be executed a better estimation of uncertainties.
Bibliography