RESEARCH ARTICLE OPEN ACCESS

Mapping Gradex values on the Tensift basin (Morocco)

Jihane Ahattab*, Najat Serhir**, El Khadir Lakhal*
*(Department of Physics, Semlalia sciences Faculty, Cadi Ayyad University, Marrakech, Morocco)
** (Department HEC, Ecole Hassania des Travaux Publics, Casablanca, Morocco)

ABSTRACT
The aim of this study is to elaborate the cartography of Gradex parameter used in the Gradex method for estimating flood peaks in order to size hydraulic structures.
Map of spatial variation is elaborated using the geostatistical method of kriging. Several reference functions (exponential model, spherical, linear, Gaussian and cubic) were used for modeling the kriging variogram. Cross-validation enabled a comparison between the results of these models and choice of spherical model with anisotropy and trend fit by a second-order polynomial as the most suitable.
The use of available series of annual maximum daily rainfall recorded at 23 rainfall stations, distributed over the Tensift basin, led also to develop the cartography of standard prediction errors' values associated to the predicted parameter for each point of Tensift basin. These errors vary from acceptable values (16.8%) to very high ones depending on the density of the rainfall stations at the desired site.
***Keywords* -** Cartography, flood peaks, Gradex, IDF, kriging.

I. INTRODUCTION

Gradex method is one of the most used methods for estimating flood peaks in order to size hydraulic structures in Morocco. It is based on a regional parameter calculated from the information provided by the rain. But precipitation data usually comes from observations registered at rainfall stations and which do not allow calculating this parameter at any given point (the site of a project).

Thus, the engineer resorts to the usual interpolation methods for estimating this parameter for the desired site: Thiessen polygons, linear interpolation IDWA (weighting with the inverse square distance). These interpolation methods ignore the spatial structure of the variable and may omit very specific local situations (areas of high or very low values). Also, no statistical criterion to judge the accuracy of the results obtained is formulated [1].

So the aim of this study is developing an accurate mapping of the Gradex parameter using the method of kriging. The overall purpose is to make the usual methods used to estimate flood peaks more reliable in order to have a more secure design of hydraulic structures (dams, bridges, Road crossing…) [2].

The choice of this geostatistical tool is based on the fact that the Kriging is a spatial interpolation method that takes into account both the geometrical configuration of the observed points and the spatial dependence of the data structure, unlike interpolations regression. This method has also the advantage of quantifying the errors related to the predicted values [3].

II. STUDY AREA AND DATA

The Tensift basin is located between latitudes 32°10' and 30°50' North and longitudes 9°25' and 7°12' West, around the city of Marrakech in the western center of Morocco. It is drained by the Tensift River which flows from east to west for over 260 km. The basin extends over 19400 km^2. Its vegetation is generally poor and depends on the topography and the nature of land. The climate is semi-arid influenced by the presence of high altitudes (the High Atlas). The altitude ranges from 0 to 4167m NGM with an average altitude of 2014m. "Fig.1"

Rainfall is generally low and characterized by high spatial and temporal variability. The annual average rainfall is about 200 mm in the plains and more than 800 mm on the peaks of the Atlas [4].

Available data used are the series of annual maximum rainfall available for durations of 1day, 2 days, 3 days, 4 days and 5 days. These series were recorded using rain gauges at the 23 rainfall stations that are located all aver the Tensift basin at altitudes ranging from 53m to 2230m NGM. The series' lengths vary from 14 years to 44 years (since 1967 until 2011). The location of these stations is shown in "Fig.1".

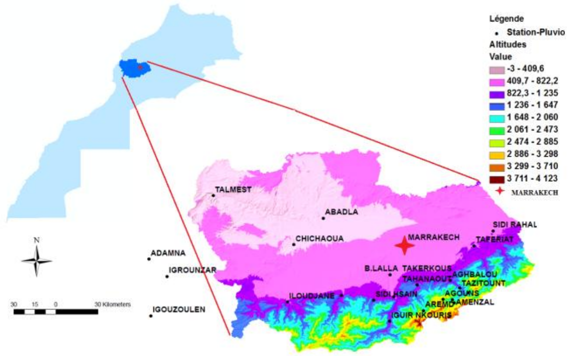


Figure 1: Location of the Tensift basin and the rainfall stations

III. STATISTICAL DISTRIBUTION OF EXTREME RAINFALL SERIES AND COMPUTATION OF GRADEX

Gradex method was developed by the EDF for estimating flood peaks with a frequency of 1% to 0.01% [5]. This method measures the increase in the risk of heavy precipitation according to the return period. It allows deducing the probabilistic behavior of the series of flows from the information provided by the probability distribution of the time series of rainfall for the high values of return period.

This method is essentially based on seasonal and local climate invariant defined by the gradient of extreme rainfall values and called Gradex [6].

We used the available series of annual maximum daily rainfall recorded at the 23 rainfall stations. All these series were adjusted to the Gumbel law with the distribution function F (x) expressed by [5]:

$$F(x) = \exp\left(-\exp\left(-\frac{x-\alpha}{\beta}\right)\right) \qquad (1)$$

α and β are the parameters of the probabilistic model of Gumbel.

Thus Gradex of 24h rain, at each station, is the slope of the frequency distribution on a Gumbel graphic scale of probability, with a linear relationship between the quantile x(F) and the reduced variable of Gumbel u(F) defined by [7]:

$$x(F) = \alpha + \beta * u(F) \qquad (2)$$

$$u(F) = -\ln(-\ln(F(x))) \qquad (3)$$

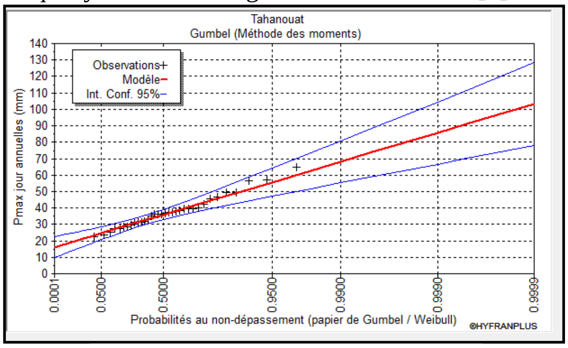
The parameters α and β can be estimated by the method of moments using the following equations:

$$\hat{\beta} = (\sqrt{6}/\pi) * S \ \ [mm] \qquad (4)$$

$$\hat{\alpha} = m - \hat{\beta} * \gamma \ \ [mm] \qquad (5)$$

- S and m are respectively the standard deviation and the mean of the series of annual maximum rainfall of 24h.
- γ: Euler's constant (0.5772156 ...).

The following figure shows the adjustment, of the maximal daily rainfall series (registered at the Tahanaout rainfall stations shown in “Fig.1”) to Gumbel law. All adjustments were made by the HYFRAN software [8], which allows, confirming the adequacy of the fit using tests such as XIII 2 [9].



LEGENDE

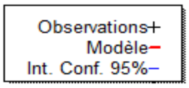


Figure 1 : Adjustment of the maximum daily rainfall by Gumbel law for Tahanaout rainfall station

The following table shows the values of Gradex of 24h rain calculated at all stations:

Table 1 : Calculated values of 24h Gradex at the 23 rainfall stations

Rainfall stations	***Gradex mm/24h***	***Rainfall stations***	***Gradex mm/24h***
Abadla	8.77	Taferiat	23.7
Adamna	18.74	Tahanaout	8.5
Aghbalou	10.6	Talmest	14.6
Chichaoua	9.1	Tazitount	20.3
Igrounzar	16.3	Amenzal	28.5
Iguir N’kouris	10	Marrakech	25
Iloudjane	10.28	Lalla Takerkoust	11.9
Imine El Hammam	12.4	Armed	22.3
S.B.Othmane	10.6	Igouzoulen	10.3
Sidi Hssain	8.9	Agouns	12.7
Sidi Rahal	13.5	Tourcht	21.7
Touirdiou	15.8		

IV. PRINCIPLE OF KRIGING

The principle of kriging is to assess at each point of the basin s_0 the value of the variable (Gradex values) $\hat{z}(s_0)$ from the weighting λi of the measured values around this point in terms of rainfall stations (neighborhood of s_0 noted $V(s_0)$), knowing the spatial structure of this variable [1]:

$$\hat{z}(s_0) = \sum_{i \in V(s_0)} \lambda_i \, z(s_i) \qquad (6)$$

This spatial structure is characterized by a feature called variogram (or semivariogram) estimated from pairs of experimental values at the rainfall stations by the equation:

$$\hat{\gamma} = \frac{1}{|2N(h)|} \sum_{N(h)} \left[z(s_i) - z(s_j)\right]^2 \qquad (7)$$

- h is a distance between the rainfall stations,
- z (si) is the value of the regionalized variable (Gradex values) at the rainfall stations.
- N (H) = {(i, j) such that (si - sj) = h} and | N (h) | is the number of distinct pairs of the set N(h). [10]

V. KRIGING STEPS

1. Exploratory analysis

The data used are the Gradex values of 24h rain calculated at the rainfall stations. The steps described in the following were performed using the tool "Geostatistical Analyst" [11] of Arcgis software [12].

The first step is to make an exploratory analysis of the data used, ie values of Gradex calculated at the 23 rainfall stations in order to gain a better understanding of it and to be able to estimate the quality of the results. To do this we began by representing the histogram and the QQplot of the database ie Gradex calculated at the rainfall stations:

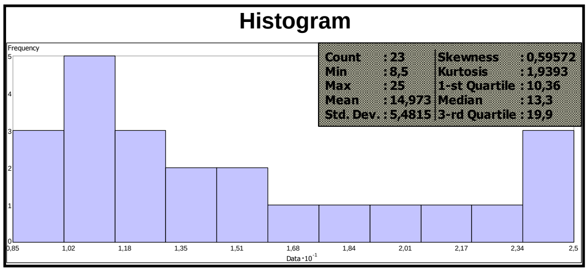


Figure 2 : Histogram of the values of Gradex of 24h rain

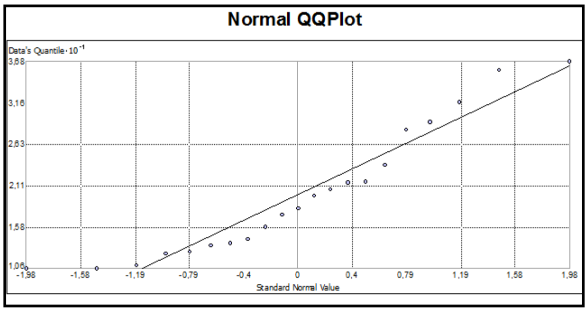


Figure 3 : QQplot of the values of Gradex of 24h rain

The distribution of the Gradex values is depicted in the histogram with the range of values split into 10 classes. The frequency of data within each class is represented by the height of each bar. The values of the mean and the median are different which indicates that is not normally distributed. This can be confirmed by the QQ plot. It is a graph on which the quantiles from two distributions are plotted versus each other. For two identical distributions, the QQ plot will be a straight line. Figure 4 shows that the plot is not very close to being a straight line so data does not exhibit a normal distribution.

The next step is trend analysis with the objective of extracting an underlying pattern of behavior in the data. If a trend exists, it is a nonrandom (deterministic) component of a surface that can be represented by a mathematical formula. This formula may produce the representation of the surface and identify which order of polynomial fits the trend best. [11]

The following figure shows the result of trend analysis on the data:

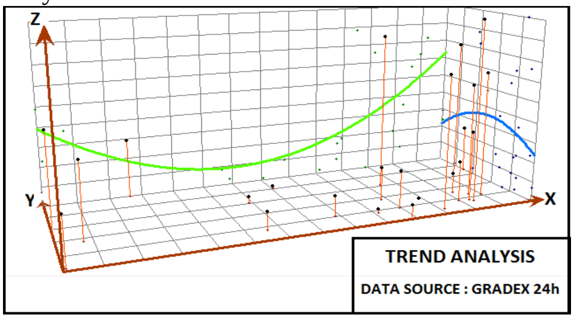


Figure 4 : Trend analysis of the values of Gradex of 24h rain

Each vertical stick in the trend analysis plot represents the location and value (height) of each Gradex value. The data points are projected onto the perpendicular planes, an east–west and a north–south plane. A best-fit line (a polynomial) is drawn through the projected points, showing trends in specific directions. The light green line starts out with high values, decreases as it moves toward the center of the x-axis, then increases. The blue line is on the contrary increasing as it moves north and decreases starting from the center of the basin. Because the trend is U shaped, a second-order polynomial is a good choice to use as a global trend model [11].

The third step is to determine the spatial autocorrelation between dataset and directional influences by representing the semivariogram/covariance. A semivariogram value, which is the difference squared between the values of each pair of locations, is plotted on the y-axis relative to the distance separating each pair of measurements, which is plotted on the x-axis:

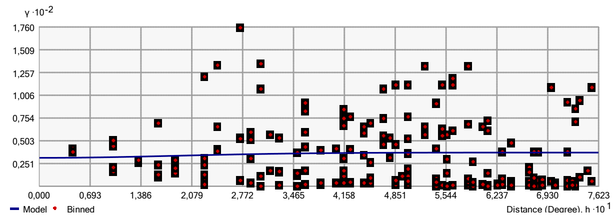


Figure 5 : Semivariogram of the values of Gradex of 24h rain

Each red dot in the semivariogram/covariance cloud represents a pair of locations. Since locations that are close to each other should be more alike. In the semivariogram plot the locations that are closest

(on the far left on the x-axis) should have small semivariogram values (low values on the y-axis). As the distance between the pairs of locations increases (moving right on the x-axis), the semivariogram values should also increase (move up on the y-axis). However, a certain distance is reached where the cloud flattens out, indicating that the values of the pairs of points separated by more than this distance are no longer correlated.
The exploratory analysis reveals that the data does not exhibit a normal distribution. A data transformation may be necessary. It also shows that the data exhibited a trend and, once refined, identified that the trend would be best fit by a second-order polynomial. The semivariogram/covariance indicates that the interpolation model should account for anisotropy.

2. Applying the method

The semivariogram/covariance cloud showed one red point for every pair of points in the dataset “Fig.7”. The next step is to fit a curve through those points. To have a clearer picture of the semivariogram values, the empirical semivariogram values (red points) are grouped according to the separation distance they are associated with. The points are split into bins (or lags), and the lag size determines how wide each interval (bin) will be. By default, optimal parameter values are calculated for an omnidirectional (all directions) stable semivariogram model. Parameter values for the omnidirectional stable semivariogram model are the nugget, range, partial sill, and shape. The nugget effect is the limit of the variogram at zero and represents the variation between two closely related measures. The range is the distance beyond which no spatial dependency between the data exists, [13]. The following figure is an illustration of these three parameters:

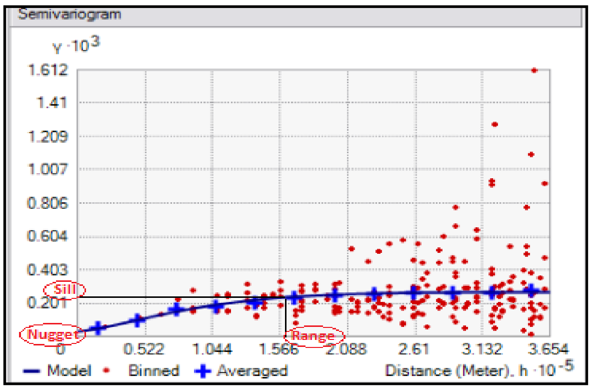


Figure 6 : Variogram showing the sill, the range and the nugget

After reporting the variogram calculated only for h distances less than or equal to the range, a reference function is used to model the diagram: exponential model, spherical, linear, nugget...

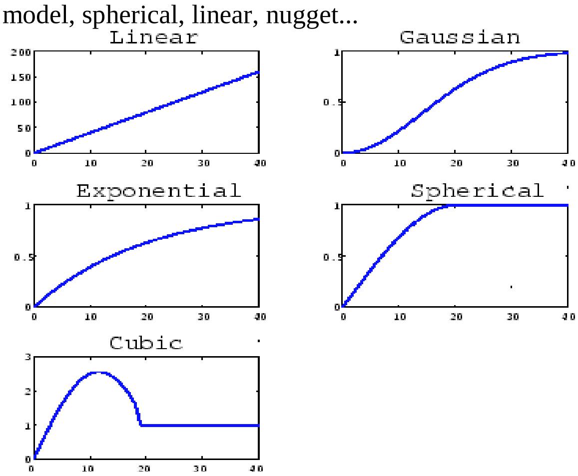


Figure 7 : Functions used to model the variogram [14]

To estimate Gradex values at any point of the basin, the weight associated with each measurement is calculated using different functions.

Finally, to choose the most suitable model cross-validation or "leave-one-out" [15] is used:

- Observation **i** is removed from the sample and the model is started with the new sample.
- Then using this new model, we calculate the predicted value of observation **i** removed.
- The process is repeated for all the observations available. One then obtains a set of predictors that can be compared to values observed by calculating the residuals of the model.

With this procedure, each sampled point is excluded and considered. A graph of the estimated values from the actual values is then constructed as shown in the following figure:

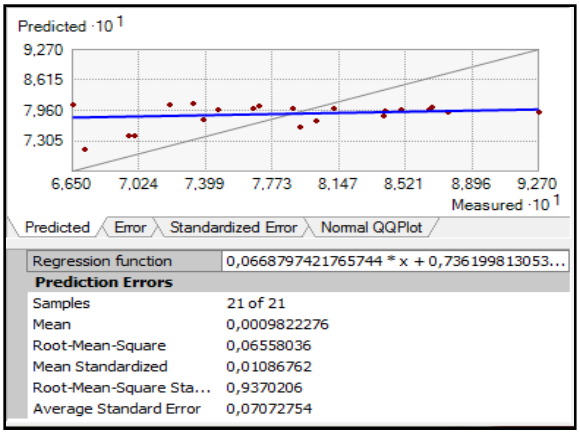


Figure 8 : Illustration of results of cross validation

The slope of the regression line is a measure of model fit to the data. A value of 1 indicates that the regression model passes through the first bisector 45º and the estimated values are the actual values.

The best model is the one that satisfy these conditions:

- An average error close to zero;
- A small root mean square prediction error;
- An average standard error similar to the root mean square prediction error;
- Root mean square standardized close to 1;
- A standardized mean prediction error near 0.

The following table summarizes the results of cross-validation of ordinary kriging models with anisotropy and trend fit by a second-order polynomial. These results were obtained for the 23 stations for different models of variograms (Linear, spherical, exponential, Gaussian and cubic)

Table 2 : Results of cross validation for different variogram models

Linear

Regression function	0,0268762255852461 * x + 13,9403411511288
Prediction Errors	
Samples	23 of 23
Mean	-0,3546049
Root-Mean-Square	1,45328
Mean Standardized	-0,06030997
Root-Mean-Square Standardized	0,9352536
Average Standard Error	1,859568

Exponential

Regression function	0,285652264692182 * x + 9,95562661...
Prediction Errors	
Samples	23 of 23
Mean	-0,2240597
Root-Mean-Square	1,456347
Mean Standardized	-0,05001975
Root-Mean-Square Standardized	1,104371
Average Standard Error	1,342101

Spherical

Regression function	0,351039000770897 * x + 9,69988598...
Prediction Errors	
Samples	23 of 23
Mean	0,1709148
Root-Mean-Square	0,62742
Mean Standardized	0,03645469
Root-Mean-Square Standardized	1,190618
Average Standard Error	0,759034

Gaussian

Regression function	0,00966087525588809 * x + 14,42985...
Prediction Errors	
Samples	23 of 23
Mean	-0,3985321
Root-Mean-Square	1,45913
Mean Standardized	-0,0647386
Root-Mean-Square Standardized	0,9226276
Average Standard Error	1,930889

Cubic

Regression function	0,0407340926648163 * x + 13,992820...
Prediction Errors	
Samples	23 of 23
Mean	-0,37031
Root-Mean-Square	1,367566
Mean Standardized	-0,06236219
Root-Mean-Square Standardized	0,9279287
Average Standard Error	1,805238

Comparison between these models shows that spherical model is the best one because it has the smallest mean and standardized mean and shows the least difference between average standard error root mean square prediction errors.

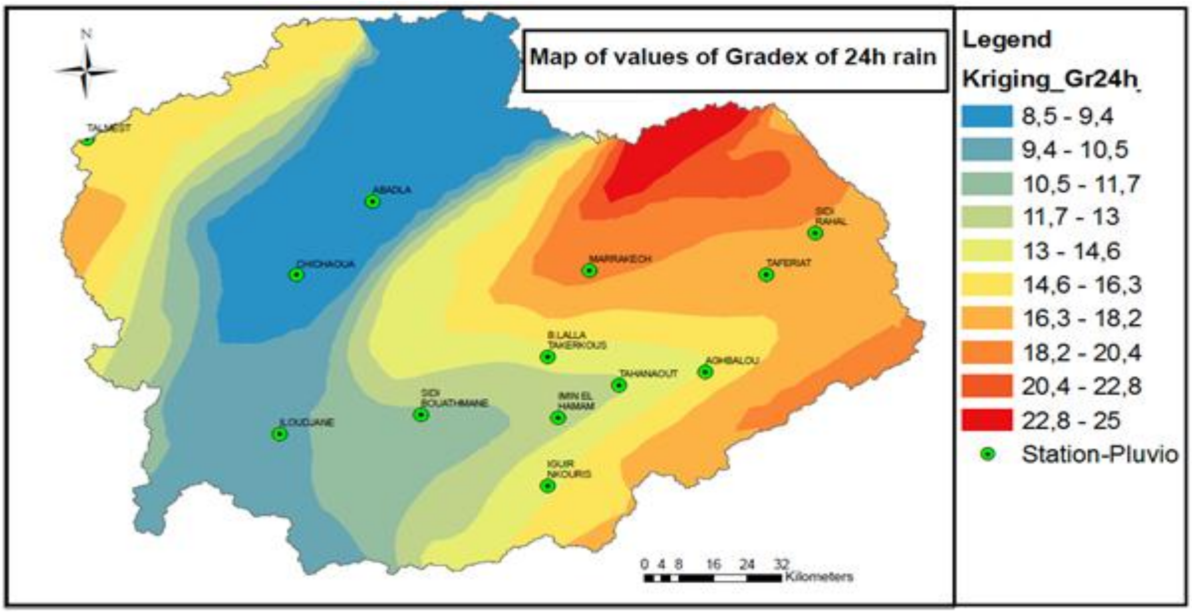


Figure 9: Map of Gradex of 24h rain on Tensift basin

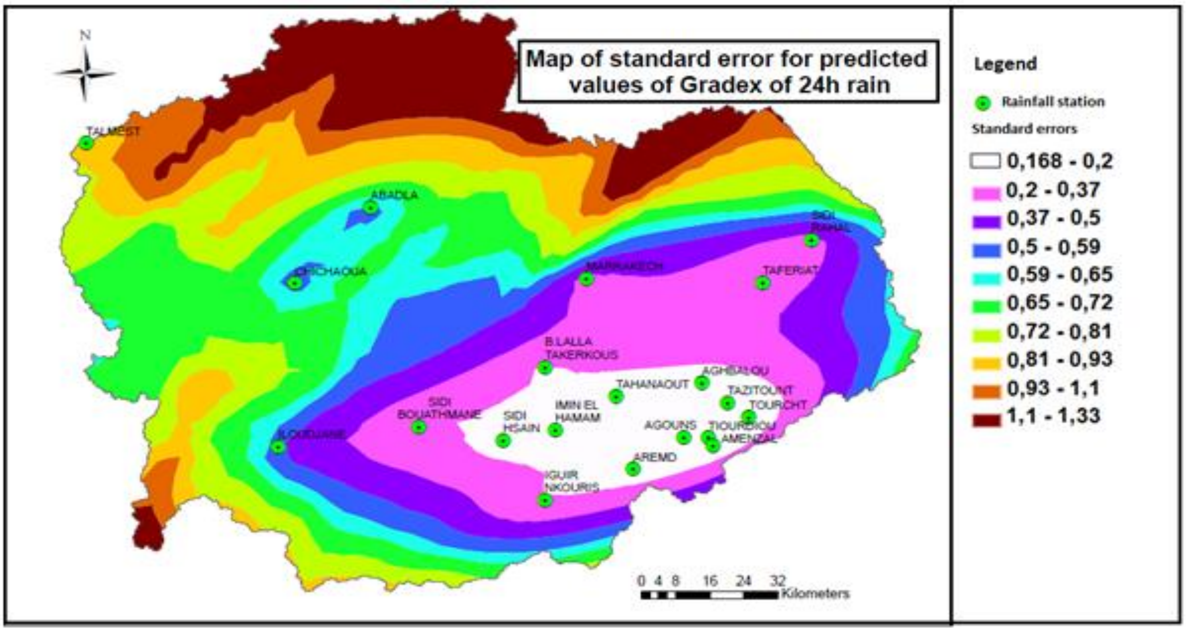


Figure 10: Map of standard errors for predicted values of Gradex of 24h rain on Tensift basin

VI. RESULTS

1- Map of geostatistical modeling of Gradex of 24h rain

The map of Gradex gives the values of Gradex of 24h rain at each point of the Tensift basin “Fig.9”.

We Note that values of Gradex of 24h rain are the highest west the basin, decrease in the middle and increase again as we head east. We deduce accordingly, that the location of slow floods (under the constraint of using 24h rain) is upstream and downstream the basin.

2- Map of standard errors of prediction for the values of Gradex of 24h rain

Although geostatistical modeling using kriging is able to give fairly accurate results, no elaborate map unfortunately is error-free and this work is no exception.

In this sense we have developed a map that assesses the standard errors of prediction associated with the modeled Gradex at each point of Tensift basin ‘”Fig.10”.

We note that this error vary from 16.8% to 133%, depends on the network density and increases as we recede from the zones with high density of rainfall stations.

VII. CONCLUSION

The Gradex map shows that risks of floods (highest values) increase when moving from the center of the basin to the east and the west. So these results represent the location of the rather slow flooding (with propagation time that can range up to several days) due to the use of the 24h rain.

However, it is appropriate to develop maps based on the duration of rainfall data less than 24 hours (1 hour to 12hours) using the methodology elaborated in this paper in order to visualize the flood risk in mountain areas. Indeed, the high slopes in these areas make concentration times for flood lower.

The map of standard errors of prediction allows assessing the scale of the uncertainties that are involved in this model. The values range up from acceptable values to very high ones depending on the density of the hydrological network. We point out, that the results can be improved by setting up a more dense measurement network especially in areas where standard errors are quite high.

Finally, the kriging remains one of the best geostatistical modeling methods since it has the advantage of quantifying the uncertainties associated with the estimated parameters when conventional interpolation methods can’t.

REFERENCES

[1] Baillargeon, S. (2005). *Le krigeage : revue de la théorie et application de l'interpolation spatiale de données de précipitations*, doctoral diss., - QUEBEC : Faculté des sciences et de génie UNIVERSITEE LAVAL, 2325 Rue de l'Université, Canada .

[2] Ahattab. J., Serhir. N., & Lakhal, E.K.. 2015. Vers l’élaboration d’un système d’aide à la décision pour le choix des méthodes d’estimation des débits max des crues : Réadaptation aux données hydrogéologiques récentes. [Revue]. - *Revue Internationale de l’Eau, La houille Blanche*, section SimHe, n° 1, 2015, p. 63-70.

[3] Renard, F., & Sarr, M. A. (2009). Quantification spatiale de la pluie en milieu rural sahélien (Ferlo, Sénégal) et en milieu

urbain tempéré (Grand Lyon, France). *Séchresse*, 20 (3).

[4] ABHT. (Mars 2007). *Etude du plan de gestion intégrée des ressources en eau dans la plaine du Haouz*.

[5] Guillot, P., & Duband, D. (1967). *La méthode du gradex*.

[6] Comité Français des grands Barrages. (1994, Novembre). *Les crues de pojet des barrages : méthode de Gradex. Barrages et réservoirs* (2).

[7] Musy, A., & Higy, C. (1998). *Hydrologie Appliquée*. Bucarest: *H*G*A*.

[8] INRS-EAU HYFRAN [Patent]. - Canada, 1998.

[9] Droesbeke Jean Jacques, Lejeune Michel et Saporta Gilbert. *Modèles statistiques pour données qualitatives* [Book]. - [s.l.] (OPHRYS, 2005).

[10] Obled CH *Introduction au krigeage à l'usage des hydrologues* [Section du livre]. - Gronoble : [s.n.], 1987.

[11] ESRI *Geostatistical Analyst Tutorial* [Report]. - [s.l.] : ESRI INC, 2010.

[12] ESRI *Aide Arcgis* [Report]. - [s.l.] : ArcGIS Ressources, 2014.

[13] Bailey T et Gatrell A Interactive spatial data analysis [Revue]. - Harlow : *Taylor and Francis, 1996. - 4 : Vol. 10.*

[14] Sierra Raimundo Nonrigid *Registration of diffusion tensor images* [Rapport]. - Switzerland : Swiss Federal Institute of Technology, 2001.

[15] Isaaks E H et Srivastava R M *Applied Geostatistics* [Livre]. - [s.l.] : Oxford University Press, 1989.