

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 6, 1322-1327 (2020)

<u>www.astesj.com</u>

ASTESJ ISSN: 2415-6698

Special Issue on Multidisciplinary Innovation in Engineering Science & Technology

Evaluation of Simple Space Interpolation Methods for the Depth of Precipitation: Application for Boyacá, Colombia

Pedro Mauricio Acosta Castellanos^{*,1}, Alejandra Castro Ortegon², Hugo Fernando Guerrero Sierra³

¹Faculty of Civil Engineering, Santo Tomas University, Tunja, 150003, Colombia

²Faculty of Environmental Engineering, Santo Tomas University, Tunja, 150003, Colombia

³Faculty of Economics and Social Sciences, La Salle University, Bogotá, 110231, Colombia

ARTICLE INFO
Article history:
Received: 04 September, 2020
Accepted: 01 December, 2020
Online: 16 December, 2020

Keywords: Regionalization Interpolation Precipitation Hydrology

ABSTRACT

Interpolation tools are used daily in hydrology and climatology. With the purpose to regionalize spot's registration parameters, such as depth of precipitation, temperature, humidity, among others. The accuracy of these methods is not fully validated. This research presents a comparative study between the most used interpolation methods for the regionalization of hydrological and climatological parameters. Comparative analysis of spatial interpolation; it was carried out using the IDW, Kriging and Spline methods, for this the ARGIS software was used, because its widespread and widespread use. The depth of precipitation was considered as a parameter for the comparison. Accuracy was determined by cross validation. Also, for the coefficient of determination and the comparison for visual errors such as "bull's eyes". The research was spatially restricted to a politically delimited region; Boyacá, Colombia, South America. The best method for spatial interpolation in this case was shown to be Spline.

1. Introduction

Rainfall records are of great importance, because from this information can be obtained among other analyses of flows, design storms and changes in rainfall behavior. The records and their analysis usually come from a network of meteorological stations, each corresponding to a specific geographical point [1]. With this information it is possible to estimate and design structures for water control, economic evaluation of flood protection projects, planning and management of land use, water quality control, among others [2].

Because rainfall information represents the record of a single geographical point, it is common for information to be researched and developed for interpolation in the form of regionalization maps [3]. This information is more common than rainfall, due to its simplicity in the form of record, in both cases can be obtained from calculation treatments and analysis rainfall runoff models and obtain important information such as the curves intensity, duration and frequency [4]. In this sense, regionalization can be defined as the transfer of data or spatial information from one station to another [5]. The transfer can range from data characteristics to

*Corresponding Author: P. M. Acosta Castellanos, pedro.acosta@usantoto.edu.co

hydrological models, [6]. A factor which determines a better condition in regionalisation is homogeneity which essentially refers to the geographical proximity or contiguity and hydrological similarity of meteorological stations [7].

Geographical dispersion in measurement networks means that there are gaps in the spatiality of information, specifically in areas without station coverage [8]. It is evident that the same applies to rainy seasons, for this reason, it is usual that different investigations are presented that seek to find the best method of regionalization and interpolation of hydrological data in order to present a viable alternative for obtaining data in areas where there is no availability. This may influence better and more accurate hydrological models, especially in small and medium catchment areas [9].

This paper presents the evaluation of three interpolation methods; Interpolation with Weighted Inverse Distance (IDW), Interpolation by Gaussian Regression Processes (Kriging) and Segmental Interpolation (Spline), applied to the regionalization of rainfall data, looking for the best method for current and future research, determining the degree of reliability of each method. The interpolation and evaluation were carried out for the rainfall data of monthly and annual records, the application of the study was limited to the geographical area comprised by the entire department of Boyacá, Colombia. These three methods or interpolation techniques were evaluated because they are the most commonly used by consultancies and research [10].

2. Methodology

2.1. Interpolation

This research seeks to validate or give support to what is used daily, for this reason the ordinary kriging and the circular model were used because within the wide spectrum of the kriging method, these are the two most used, it is important to note that they can make and create the empirical models for semivariograms but the vast majority of references and even companies like ARGIS indicate that the ordinary kriging and the circular model are the most used [11–19].

Due to the fact that rainfall information is of a point type due to the need for a spatial range and not a point range, interpolation tools are often used, which can define regions with similar behavior [20]. The commonly used methods are Inverse Distance Weighting or IDW, Kriging and Spline, in this sense there are few comparative studies that seek to define the most suitable spatial interpolation method for climatic or environmental variables such as precipitation [21]. Spatial interpolation methods can be defined in four categories: regression method or model and trend surface, local methods, geostatistical methods and mixed methods [9], Only local methods (IDW, Spline) and a known geostadistic method Kriging were evaluated for this research. Likewise, only monthly and annual rainfall depth record values were interpolated for the region of Boyacá. Colombia, because it has a spatially acceptable coverage of rainfall stations, which has a direct bearing on the use of interpolation methods since it is necessary to have a regional coverage to carry out interpolation, since the techniques preside over the variable of interest in a specific place taking values from the surrounding region, resulting in homogeneity [22]. The project will be carried out in the department of Boyacá, Colombia, which has 123 municipalities and 116 rainfall stations of which 60 were useful, since they fulfilled a range of 20 years of historical records counting from 1998 to 2018.

2.2. IDW method

The IDW interpolation method is widely used because of its simplicity [23]. Which is based on the fact that the unknown or sought value of a point has the greatest influence of the nearby control points and the degree of influence or weight of them [24], this implies that it is directly proportional to the inverse of the distance between the points posed [25]. When defining a high power, greater emphasis will be on the nearest points and the result of the surface will be more detailed, but less smoothed. By specifying a lower power, greater importance will be given to the more distant points and therefore the resulting surface will be smoother. equation (1) represents the method

$$Z = \frac{\sum_{i=1}^{N} W_{i} Z_{i}}{\sum_{i=1}^{N} W_{i}} = \frac{\sum_{i=1}^{N} \left[\frac{Z_{i}}{D_{i}^{p}}\right]}{\sum_{i=1}^{N} \left[\frac{1}{D_{i}^{p}}\right]}$$
(1)

where Z is the interpolated value for a point with the unknown observed value, Wi is the weighting function that determines the

importance and relationship of each control point Zi, while Zi is the observed value of the control point i (i=1,...,n), which is the closest zone to the interpolated point and n is the total number of points used in interpolation [26]. The surfaces obtained by this method tend to be symmetrical over the points and reflect the mean value of the dispersion points in the regions between the points. In this method it is possible to control the importance of known points over interpolated values based on distance.

2.3. Kriging method

The method of interpolation Kriging, has a large number of variants that have been developed in recent times, the common is the Kriging Ordinary (KO). We have in this method that the observed or not observed search points are calculated by the linear weighted average of the observed or existing points [23], as shown in equation (2).

$$\widehat{Z_0} = \sum_{i=1}^n \lambda_i \, Z_i \tag{2}$$

where $\widehat{Z_0}$ represents the point searched or not observed, Zi means the attribute of the value of an observed point i, λi is the weight or significance of Zi.

The KO method presents the best linear prediction for points not observed, this means that in theory the value sought or point not observed is equal or very similar to the true value and the variance in the error is minimal [27].

2.4. Spline method

This interpolation method determines the unobserved points using a mathematical function that minimizes the general curvature of the surface, which has the particularity of being a smoothed surface. Many tests performed with Spline show their accuracy which means that the interpolated regions pass through the observed points [28]. There are two types of Splines mostly diffused, the Regularized and Stress, the first refers to softened regions or surfaces and the second creates less smoothed surfaces that are closer to the observed values, both cases result from the general equation of the equation method (3).

$$Z_{(x,y)} = T_{(x,y)} + \sum_{i=1}^{n} \lambda_i Z_{ri}$$
(3)

where Z is the value sought or point not observed, n, the number of points, λi coefficient of the system of linear equations, Zi is the distance of the point (x,y) to the point I [29].. T and R, are defined according to whether it is regularized or tension.

2.5. Hydrological Information

The information used for the assessment of interpolation methods is of the monthly and annual rainfall type. We reviewed the records of 116 stations located at the site of the study (Boyacá, Colombia), of which 60 had continuous records between 15 and 20 years. It should be recalled that rainy seasons record only the precipitation depth measured in (mm). The place where the research was applied is located in Colombia, South America, in the department of Boyacá, with an area of 23189 km², which represents 2.03% of the Colombian territory [30]. It is a region with three main hydrographic basins on the rivers Magdalena, Arauca and Meta, in the central area of Colombia, the figure 1, shows the area of study.



Figure 1: Location of rainfall seasons

Results and discussion 3.

3.1. Validation of the methods evaluated.

The validation of the interpolated points using the three methods was performed by cross-validation, technique used to evaluate the results of a statistical analysis and to ensure that they are independent of the partitioning of the data obtained in this case from rainfall stations. The estimated value must be similar to the actual value, that is the measured value, so that the error committed by the model, that is, the difference between the two, is as low as possible [31].

The difference indicates how out-of-date the precipitation depth data calculated by the software are, when compared with the actual or measured information presented in the study area, that can be analyzed by the coefficient of determination or R2.

3.2. Cross-validation

Cross-validation is a way of predicting the fit of a model to a hypothetical test data set when the explicit test data set is not available. The cross-validation process gives a precise approximation of performance and error prediction [32]. Crossvalidation in this study was performed by interpolating the totality of the rainfall stations and subtracting one weather station at a time, interpolating again, to verify by crossing the interpolated information against the actual or measured of each meteorological station. In other words, the interpolated data at the points where the measured data is held are verified and the actual or measured values and the interpolated values are compared.

3.3. Comparison of interpolation methods

Raster maps generated by a GIS were used in order to identify which interpolation method presents better results from different points of analysis; cross validation, coefficient of determination (R2) and the visual aspect, in the latter was sought to identify the error called "bull's eye". Based on the measured and calculated differences, statistical analyses were performed for the maximum annual and maximum monthly precipitation depth. The mean, median, minimum and standard deviation were compared, although it should be noted that greater importance was given to the statistical parameter of the mean, since it was considered to be the most relevant parameter in this case.

3.4. Interpolation of Annual precipitation depth.

It was observed that the Kriging technique has the least variation in the statistical parameter of the mean (54.70), of data www.astesj.com

standardization, but has the greatest dispersion of data with respect to the measured data (standard deviation), in addition it does not present an anisotropic behavior with spherical distribution behavior normalizing thus adjusting the plateau, data tolerance and bandwidth as shown in Figure 2.



Figure 2: Data adjustment

In the standardization of data, a circular or spherical model was used in the semivariogram, adjusting the data to an exponential trend Figure 3, thus obtaining the highest reliability to perform the interpolation of IDW, ordinary kriging and spline as shown in Figure 4, 5, and 6 for 2008 year.







Figure 4: Interpolation by IDW, january 2008



Figure 5: Interpolation by Kriging, january 2008



Figure 6: Interpolation by Spline, january 2008

On the other hand, the Spline technique also shows good performance in terms of average (54.66), but less than the Kriging value. The Spline method presented less dispersion than Kriging. Similarly, it was found that the IDW technique shows the greatest difference between the measured data and the interpolated data, in Table 1; the results of the analysis are evident.

Statistics	Data	IDW	Spline	Kriging
	measured			
Mean	56.81	53.02	54.66	54.70
Median	47.1	49.04	48.40	51.04
Estándar	30.51	20.12	20.99	16.26
deviation				
Minimun	10.50	25.10	27.40	34.12

Table 1: Statistical	analysis of annu	al precipitation	depth for 2008
	winer, bib of winner	ar preespication	aspen for move

For the analysis of the interpolation of annual precipitation, it was determined that the kriging technique is the best fit in the statistical analysis of data, since the mean with the Kriging technique has a more appropriate behavior than the other two evaluated techniques.

3.5. Interpolation of monthly precipitation depth.

In the case of the interpolation of the monthly precipitation depth data, the Spline technique presented better results in terms <u>www.astesj.com</u>

of the statistical parameters of the mean and median, compared with the other techniques, although with great closeness between each method. In the second place, the Kriging technique is found in Table 2; the results of the analysis are evident.

	,			
Statistics	Data measured	IDW	Spline	Kriging
Mean	66.31	62.22	63.95	63.86
Median	55.15	54.60	55.14	53.17
Estándar	32.04	22.58	24.75	25.20
deviation				
Minimun	28.80	35.13	31.70	34.44

Table 2: Statistical results of methods evaluated against mesuared monthy values, 2008

3.6. Análisis de contornos.

Raster maps because they represent the space in a regular array of pixels, where the representation of the elements is made by pointing out the existence or not of data within each pixel [33], in the case of the regionalization of precipitation depth facilitates the visual behavior of the three interpolation techniques used in this investigation. Where we visually compared which of the three techniques has better contours and a more detailed surface in areas where there are no weather stations. One of the characteristics that is intended to show both in presence and absence is the typical error of "ox eyes" where interpolation techniques usually create concentration points or form islands around interpolated or measuring points [34], in this case the rainfall seasons. This anomaly is evident that with the IDW method where concentric circles are presented, as can be seen in Figure 1, this anomaly was presented as a visual trend a large percentage of the generated maps.

Ordinary kriging was the technique that presented the best results, since visually it presents different contours of precipitation depth, which is more accommodated to the recorded precipitation, consistent with the cross validation performed. In addition to performing the ordinary kriging, an additional raster is also interpolated to validate the interpolation of standard prediction errors, where the less data is those that may reliability and the larger data are those that should be resampled, as shown in figure 7.



3.7. Coefficient of determination R2

Raster maps because they represent the space in a regular array of pixels, where the representation of the elements is made

by pointing out the existence or not of data within each pixel [33], in the case of the regionalization of precipitation depth facilitates the visual behaviour of the three interpolation techniques used in this investigation.

3.8. Measurement of R2 in precipitation interpolation Annual

In this sense the method of interpolation of Kriging presents better correlation in the interpolation with depth of precipitation, the coefficient of determination is close to 1, with respect to the other models (Idw and Spline), thus yielding a high degree of reliability for choosing this model and carrying out the interpolation with which the regionalization of precipitation depth can be generated, due to the nature of this method it can be observed that it has a better behaviour with a higher data density, since analyzing interpolation with depth of monthly precipitation shows that the so-called spline model presents a greater correlation between its data, this allows choosing it to generate monthly regionalization.

Interpolation Precipitation Ma	Depth ax. Annual	of	Interpolation Depth Max. Mo	Precipitation nthly
Method	R2		Method	R2
Kriging	0,5762		Kriging	0,6717
Spline	0,5484		Spline	0,6861
IDW	0,5308		IDW	0,655

Table 3: R2, para los métodos evaluados.

4. Conclusions

It was identified that visually the methods do not present an obvious difference, but in the case of the IDW technique generates the so-called portholes. For the regionalization of the maximum monthly precipitation depth it was possible to identify that the Spline method has a better behavior compared to the real data, since the R2, had a higher value (0.6861) compared to the other two methods evaluated.

In the case of the maximum monthly depth, the method that had a greater difference compared to the real data was the IDW, added to the visual errors (portholes) it is not advisable to use this method to regionalize hydrological data. Based on the analysis of both visual and cross-validation results, it can be stated that the kriging technique has a high degree of reliability when performing regionalizations with annual precipitation information.

In addition to the comparative results of the methods evaluated, the resulting information can be used as input for the determination of hydrological models in areas or locations that do not have their own information or measurement. In other words, they do not have a hydrological station. It could be shown that it is possible to use regionalized information with a high degree of confidence provided that it comes from data interpolated by the Spline method or the Kriging method

Twenty-four raster maps were generated, with maximum monthly and maximum annual rainfall depths, where twelve of these maps belong to the Spline technique which generates greater reliability with the maximum monthly values, and the remaining twelve were generated by the Kriging technique for maximum annual precipitation depth values. Maps generated by the IDW technique were discarded. It is evident that having a greater number of stations in optimum condition could make a more precise regionalization since interpolation would be more accurate, as well as having support stations would minimize the effect of "bull's eye".

References

- T.S. Kokkonen, A.J. Jakeman, P.C. Young, H.J. Koivusalo, «Predicting daily flows in ungauged catchments: Model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina», Hydrological Processes, 17(11), 2219-2238, 2003, doi:10.1002/hyp.1329.
- [2] M. Jessen, T. Einfalt, A. Stoffer, B. Mehlig, «Analysis of heavy rainfall events in North Rhine-Westphalia with radar and raingauge data», Atmospheric Research, 77(1-4 SPEC. ISS.), 337-346, 2005, doi:10.1016/j.atmosres.2004.11.031.
- [3] J.M. Mirás-Avalos, R.M. Mestas-Valero, P. Sande-Fouz, A. Paz-González, «Consistency analysis of pluviometric information in Galicia (NW Spain)», Atmospheric Research, 2009, doi:10.1016/j.atmosres.2009.05.011.
- [4] G. Blöschl, M. Sivapalan, «Scale issues in hydrological modelling: A review», Hydrological Processes, 9(3-4), 251-290, 1995, doi:10.1002/hyp.3360090305.
- [5] Problems of rainfall-runoff modelling in arid and semiarid regions, oct. 2020, doi:10.1080/02626668809491261.
- [6] R.J. Nathan, T.A. McMahon, «Evaluation of automated techniques for base flow and recession analyses», Water Resources Research, 26(7), 1465-1473, 1990, doi:10.1029/WR026i007p01465.
- [7] D.A. Post, A.J. Jakeman, "Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model", Ecological Modelling, 123(2-3), 91-104, 1999, doi:10.1016/S0304-3800(99)00125-8.
- [8] A. Agarwal, R. Maheswaran, V. Sehgal, R. Khosa, B. Sivakumar, C. Bernhofer, «Hydrologic regionalization using wavelet-based multiscale entropy method», Journal of Hydrology, 538(538), 22-32, 2016, doi:10.1016/j.jhydrol.2016.03.023.
- [9] C.C.F. Plouffe, C. Robertson, L. Chandrapala, «Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: A case study of Sri Lanka», Environmental Modelling and Software, 67, 57-71, 2015, doi:10.1016/j.envsoft.2015.01.011.
- [10] Y. Gyasi-Agyei, G. Pegram, «Interpolation of daily rainfall networks using simulated radar fields for realistic hydrological modelling of spatial rain field ensembles», Journal of Hydrology, **519**(PA), 777-791, 2014, doi:10.1016/j.jhydrol.2014.08.006.
- [11] M.A. Oliver, R. Webster, «Kriging: A method of interpolation for geographical information systems», International Journal of Geographical Information Systems, 4(3), 313-332, 1990, doi:10.1080/02693799008941549.
- [12] Cómo funciona Kriging—Nuevas características y mejoras que se han agregado en esta versión. | Documentación, oct. 2020.
- [13] A. Mazzella, A. Mazzella, «The importance of the model choice for experimental Semivariogram modeling and its consequence in evaluation process», Journal of Engineering (United Kingdom), 2013, 2013, doi:10.1155/2013/960105.
- [14] I.K. Tsanis, M.A. Gad, «A GIS precipitation method for analysis of storm kinematics», Environmental Modelling and Software, 16(3), 273-281, 2001, doi:10.1016/S1364-8152(00)00068-2.
- [15] D. Koutsoyiannis, «Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation», Hydrological Sciences Journal, 49(4), 575-590, 2004, doi:10.1623/hysj.49.4.575.54430.
- [16] E. Rojas, B. Arce, A. Peña, F. Boshell, M. Ayarza, «Cuantificación e interpolación de tendencias locales de temperatura y precipitación en zonas alto andinas de Cundinamarca y Boyacá (Colombia)», Corpoica Ciencia y Tecnología Agropecuaria, 11(2), 173, 2010, doi:10.21930/rcta.vol11 num2 art:209.
- [17] M. Sivapalan, G. Blöschl, «Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves», Journal of Hydrology, 204(1-4), 150-167, 1998, doi:10.1016/S0022-1694(97)00117-0.
- [18] R.J. Moreano Viteri, SISTEMA DE INFORMACION PARA LA INTERPOLACION ESPACIAL Y TEMPORAL DE DATOS SOBRE EL TIEMPO ATMOSFERICO Y EL CLIMA DEL ECUADOR, QUITO/ EPN/ 2008, 2008.
- [19] Ana Carolina SantosAndrés Vargas-LunaAndrés Vargas-LunaNelson

Obregón-NeiraNelson Obregón-NeiraEDER CÁRDENAS, «Análisis de la distribución e interpolación espacial de las lluvias en Bogotá, Colombia», Dyna, 151-159, 2010.

- [20] G. Gong, S. Mattevada, S.E. O'Bryant, «Comparison of the accuracy of kriging and IDW interpolations in estimating groundwater arsenic concentrations in Texas», Environmental Research, 130, 59-69, 2014, doi:10.1016/j.envres.2013.12.005.
- [21] D. Zimmerman, C. Pavlik, A. Ruggles, M.P. Armstrong, «An experimental comparison of ordinary and universal kriging and inverse distance weighting», Mathematical Geology, **31**(4), 375-390, 1999, doi:10.1023/A:1007586507433.
- [22] S.M. Vicente-Serrano, M.A. Saz-Sánchez, J.M. Cuadrat, «Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): Application to annual precipitation and temperature», Climate Research, 24(2), 161-180, 2003, doi:10.3354/cr024161.
- [23] F. Huang, D. Liu, X. Tan, J. Wang, Y. Chen, B. He, «Explorations of the implementation of a parallel IDW interpolation algorithm in a Linux clusterbased parallel GIS», Computers and Geosciences, 37(4), 426-434, 2011, doi:10.1016/j.cageo.2010.05.024.
- [24] P.M. Bartier, C.P. Keller, «Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW)», Computers and Geosciences, 22(7), 795-799, 1996, doi:10.1016/0098-3004(96)00021-0.
- [25] M.P. Armstrong, R.J. Marciano, «Local interpolation using a distributed parallel supercomputer», International Journal of Geographical Information Systems, 10(6), 713-729, 1996, doi:10.1080/02693799608902106.
- [26] Q. Huang, C. Yang, «Optimizing grid computing configuration and scheduling for geospatial analysis: An example with interpolating DEM», Computers and Geosciences, 37(2), 165-176, 2011, doi:10.1016/j.cageo.2010.05.015.
- [27] H. Hu, H. Shu, «An improved coarse-grained parallel algorithm for computational acceleration of ordinary Kriging interpolation», Computers and Geosciences, 78, 44-52, 2015, doi:10.1016/j.cageo.2015.02.011.
- [28] N.C.-N.Y.A.W.-I. Publicatiorq, undefined 1993, «Statistics for spatial data. revised edition».
- [29] M.F. Hutchinson, «A new procedure for gridding elevation and stream line data with automatic removal of spurious pits», Journal of Hydrology, **106**(3-4), 211-232, 1989, doi:10.1016/0022-1694(89)90073-5.
- [30] R. Franke, «Smooth interpolation of scattered data by local thin plate splines», Computers and Mathematics with Applications, 8(4), 273-281, 1982, doi:10.1016/0898-1221(82)90009-8.
- [31] J.A. Becerra Oviedo, L.F. Sánchez Mazorca, P.M. Acosta Castellano, J.L. Díaz Arévalo, «Regionalización de curvas IDF para el uso de modelos hidrometeorológicos en la Sabana Occidental del departamento de Cundinamarca», Ingeniería y Región, 14(2), 143, 2016, doi:10.25054/22161325.701.
- [32] P.M. Acosta Castellanos, L.X. Sierra APonte, «Evaluación de métodos de construcción de curvas IDF a partir de distribuciones de probabilidad y parámetros de ajuste», REVISTA FACULTAD DE INGENIERÍA, 22(35), 25, 2013, doi:10.19053/01211129.2512.
- [33] Colombia: Diagnóstico departamental de situación de DDHH y DIH -Boyacá - Colombia | ReliefWeb, oct. 2020.
- [34] T Izquierdo., Álvaro Márquez, «Estudio comparativo de diferentes métodos de interpolación para la realización de mapas de precipitación en la isla de La Gomera (Islas Canarias): Evaluación de la incorporación de la altura -Dialnet», Geogaceta, 131-134, 2007.