International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 21 (2018) pp. 15118-15123

ARIMA Analysis of the Runoff Curve Number Method in a Sub-urban River Basin

Daniel Aguilar-Gómez[1,*], Lizeth Páez-Rivera[1,], Carlos Zafra-Mejía[1,***]**

[1]Environmental Engineering Research Group (GIIAUD), Facultad de Medio Ambiente y Recursos Naturales, Universidad Distrital Francisco José de Caldas, Avda. Circunvalar Venado de Oro, E-111711, Bogotá D.C., Colombia.

Abstract

Global changes in the study approaches of environmental variables have allowed the time series analysis to be one of the main science tools to propose new environmental simulation models and to evaluate the models previously developed. This paper shows a temporal analysis using ARIMA models of the river flows generated with Runoff Curve Number Method (CN method, U.S. Soil Conservation Service) in a sub-urban river basin of Bogota city (Colombia). We analyzed and compared the river flow series foretold with CN method and those observed in a hydrometric station located at the basin exit. On average, the results showed that CN method tended to overestimate the daily flows observed in the hydrometric station (2.91 times). In torrential basins, ARIMA analysis showed a low order in polynomials of the developed models: [0,1,3], [1,0,3], and [2,1,2]. These models also suggested a greater variability of the daily flows foretold with CN method in relation to those observed in the hydrometric station. The results of this study will be useful for institutions responsible for managing water resources systems, specifically, to propose new methodologies for river flow analysis in projects of exploitation and risk management.

Keywords: ARIMA model, Runoff curve number, Time series, Torrential basin.

INTRODUCTION

In developing countries due to limited instrumentation the climatological and hydrologic information available in river basins is scarce, making useful the application of the Runoff Curve Number (CN) method as a tool for forecasting flows. Moreover, in recent decades, suburban and urban river basins in tropical areas have shown an increase in their hydrological risk. In other words, the occurrence probability of an extreme hydrometeorological event that exceeds a specific value of social and economic damage has been increased (USAL, 2013). The aforementioned risk scenario is generated by the combination of both natural and anthropic factors. Among the natural factors, it is important to mention the role played by the climate change phenomenon in tropical countries. This generates an increase in rainfall in rainy seasons that exceed the buffering capacity of flows in river basins (Mendivelso et al., 2014). Anthropic factors to highlight are the high demographic concentration, deforestation of high areas in river basins, and increase of impervious coverage in urban areas. The latter represents a decrease in natural soil permeability that generates an increase in surface runoff and a decrease in the concentration time of river basins (Norambuena, 2009).

Therefore, it is important to adequately forecast the river flows for decision-making in engineering projects, territorial planning, water resources management, disaster and emergency prevention, and civil works design. CN method proposed by U.S. Soil Conservation Service in 1972 is a conceptual deterministic model of the relationship between rainfall and surface runoff (Alonso, 2001). This method is used in estimating maximum river flows in basins and has reached a high application due to the requirement of few physical and meteorological information (Weber and Jorquera, 2010). However, the reliability of estimated flows must be assessed in a temporal perspective, due to the change in environmental conditions under which the method is applied. These environmental conditions can affect the accuracy degree of the forecasts obtained (Deshmukh et al., 2013).

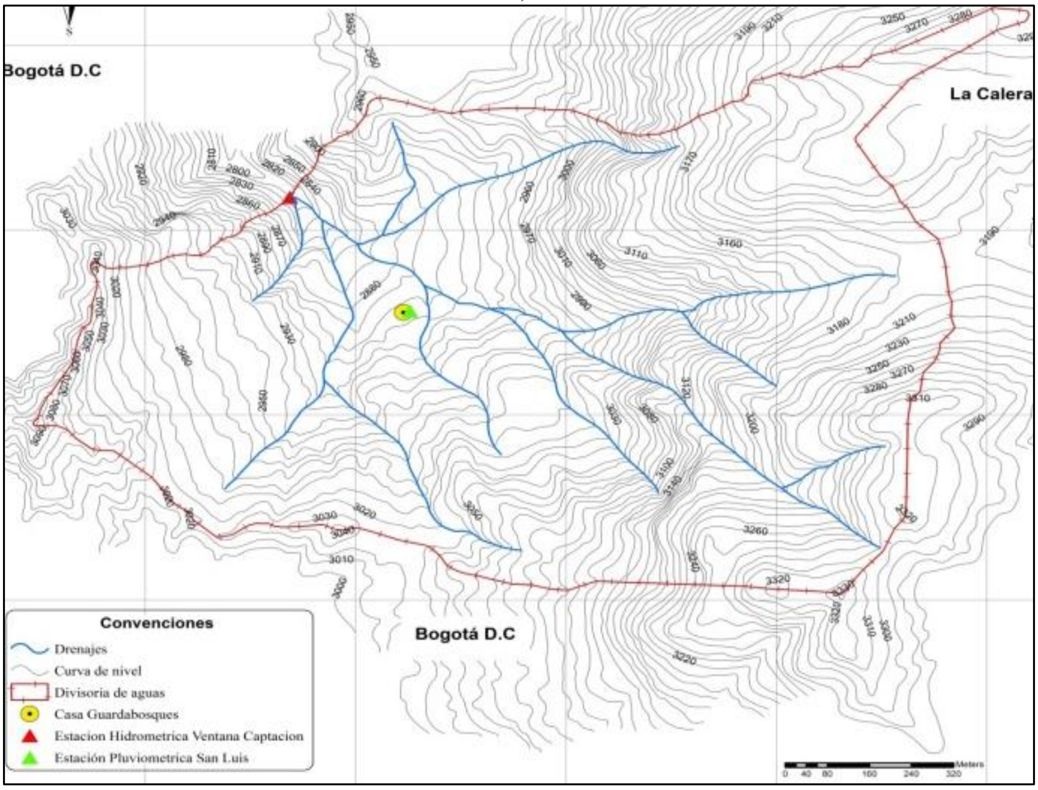
The main objective of this paper is to show a temporal analysis using ARIMA models of the river flows forecasted with CN method in a sub-urban basin of Bogota city (Colombia). It also compares the temporal structure of flow time series foretold with CN method and those observed in a hydrometric station located at the basin exit. This study will be useful for institutions responsible for managing water resources systems, specifically, to propose new methodologies for river flow analysis in projects of exploitation and risk management.

MATERIALS AND METHODS

Study site

The river basin (Quebrada La Vieja) was in the eastern hills of Bogota city (4°39’ N; 74°03’ W) and its waters were discharged in the middle basin of Bogota River (Figure 1). Rodríguez-Barrios and Ospina (2007) reported that the river basin showed physical characteristics of mountain (maximum elevation: 3300 masl, minimum: 2690 masl), good conservation status, and continuous forest coverage. Climatological information was obtained from a station located within the river basin (San Luis Station), which was operated by Empresa de Acueducto de Bogotá (EAB). In the study area, a bimodal rainfall regime was evidenced, with two annual maximums: March-May and October-December. Average daily and annual rainfalls were 3.6 mm and 1317 mm, respectively. The river basin displayed an average annual relative humidity of 77% and an average annual temperature of 16.1 °C. Average daily flow of the main river was 22 L/s. The river flow information was obtained from a hydrometric station installed at the basin exit (Quebrada la Vieja Station), which was operated by EAB.

a)



b)

Figure 1. Cartography of the river basin under study (Google Earth, 2017). (a) Satellite image. (b) Cartographic map.

Sampling system

The following secondary information was used to develop this study: (i) Basic cartography of the study area to scale 1:10000, (ii) thematic cartography to scale 1:100000 (soil map), (iii) air and satellite photographs, and (iv) time series of hydroclimatic information. Statistical (IBM-SPSS statistics 19.0) and geographic information systems (ArcGis 10.0) software were also used. A five-year period (01/01/2010-31/12/2014) was defined as a timescale of study. During this period, a total of 1826 days of daily rainfall and flow information were taken. Figure 2 shows the methodology used in this study.

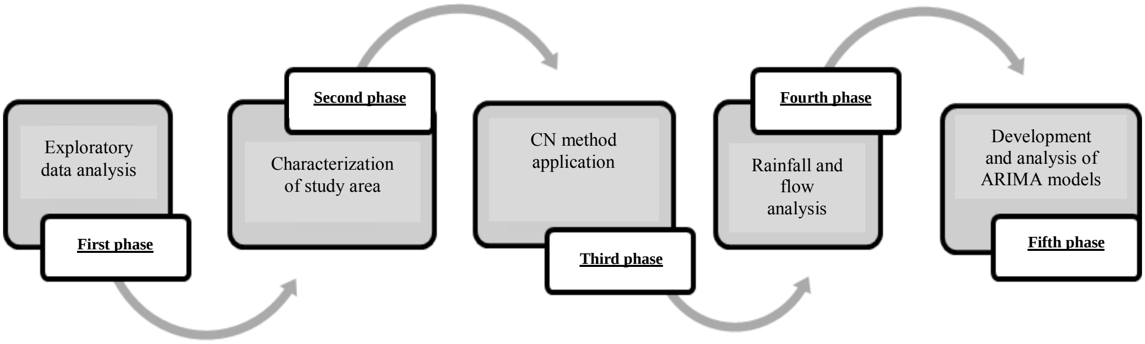


Figure 2. Methodology applied to develop this study.

Information analysis

In the first phase of information analysis was carried out the identification of missing information in time series of daily rainfall and daily flow measured in the stations pluviographic (San Luis Station, EAB) and hydrometric (Quebrada la Vieja Station, EAB), respectively. Valid information exceeded 75% of the data used in this study (rainfall: 99% and river flow: 98%). The missing information was completed by a preliminary ARIMA model (validated), generated from the original time series (Zafra et al., 2017). In other words, the missing information was generated with the temporal methodology incorporated by ARIMA models.

In the second phase of information analysis, the study area was digitally delimited by watershed. We also digitally locate the hydrometric station at the basin exit. The following analyses were also included in this phase: determination of morphological parameters, and identification of type, use, and treatment of soil in the study area. Subsequently, the base flow was calculated by means of flow hydrograph and applying the straight-line method for components separation of hydrographs (Chow et al., 1994).

In the third phase, the application of CN method was carried out to determine flow rates. The information analysis activities were as follows: Determination of average CN for the river basin, calculation of soil antecedent moisture conditions, determination of effective daily rainfall, and application of triangular synthetic hydrograph for generation of daily flows (Kottegoda et al., 2000). In order to verify the CN value calculated above, we proceeded in the fourth phase to calculate this value from the daily flows observed in the hydrometric station.

In the last phase, we identified the ARIMA models that adequately represented the behavior of maximum and average daily flows calculated through CN method, and the average daily flows observed in the study area. ARIMA model methodology developed by Box and Jenkins (1970) was used: (i) identification, (ii) estimation, and (iii) model verification (assumption compliance). During this phase, IBM-SPSS statistical software was used.

RESULTS AND DISCUSION

River basin characterization

Morphological parameters indicated that the river basin had an area of 1.42 km^2, and an elongated and oval shape (form factor: 0.73; compactness coefficient: 1.34). The river basin also showed a high drainage density (4.18 km/km^2), high slopes (main river: 35.4%; river basin: 52.5%), and a low concentration time (0.14 h). This suggested the presence of high surface runoff speeds and a tendency to the rapid generation of instantaneous flows. In other words, the study basin was probably characterized as torrential.

Additionally, in the study area, impervious soils (loamy clay texture) and open forests prevailed. This suggested a high surface runoff potential, and flow attenuation due to the rainfall interception and water regulation by the vegetable coverage. Pereira et al. (2016) suggested a similar trend in the interpretation of these variables in hydrological projects.

CN method application

On average, the results showed that 11.1% of total rainfall was probably transformed into surface runoff when applying the CN method under hydrologic conditions of the study area (CN = 75.4; timescale: daily). We displayed that the behavior of maximum flow time series foretold with CN method did not reflect the bimodal regime observed in the river basin for rainfall.

We made a comparison between the flow time series. On average, the results showed that the maximum daily flows foretold with CN method were 5.27 times higher than the daily flows observed in the hydrometric station. Findings also revealed that the daily flows foretold were 2.91 times higher than the daily flows observed in the hydrometric station (Figure 1). Tucci (1998) established a relationship between the maximum and average daily flows in a river basin. This researcher reported for regions less than 200 km^2 a relationship between 3 and 5. Therefore, in this study the variation in magnitude between the time series of maximum daily flow foretold and of average daily flow observed were those expected in a river basin with these physical characteristics (i.e., drainage area).

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 21 (2018) pp. 15118-15123

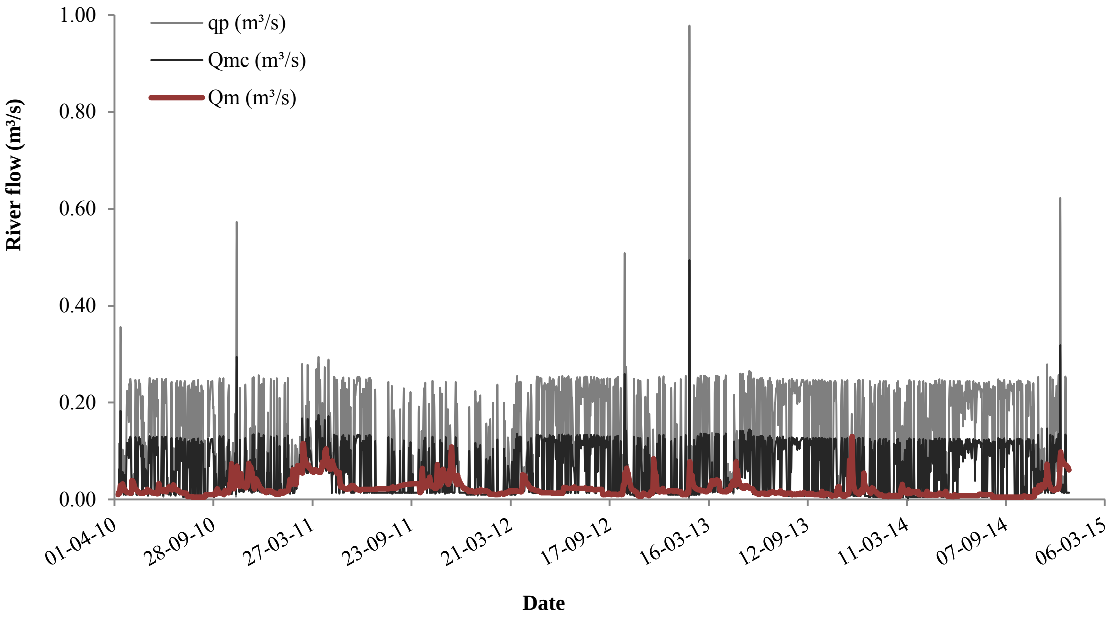


Figure 3. Graphic comparison between times series of maximum (qp) and average daily flow (Qmc) foretold with CN method, and average daily flow (Qm) observed in the hydrometric station.

ARIMA models

ARIMA models identified as satisfactory to represent the temporal behavior of time series evaluated in this study are shown in Table 1. The results showed that the maximum and average daily flows foretold with CN method exhibited a similar temporal structure: [0,1,3] and [1,0,3]. As expected, the results suggested that CN method foretold under the same temporal structure the maximum and average daily flows in the study basin. In other words, the temporal behavior was the same for these time series. ARIMA models for time series of average daily flow foretold with CN method and observed in the hydrometric station showed differences in its temporal structure. The results suggested that CN method could not adequately forecast the temporal variation of daily flows observed in the study basin. However, the polynomials that intervened in ARIMA models (p,d,q) were not comparatively different in relation to the autoregressive (p) and moving average (q) processes.

Therefore, it was observed that ARIMA models for maximum daily flows foretold with CN method were characterized by limited or null memory. Maximum daily flow depended or was influenced by the river flow of one previous day ($p = 1$). The third-order moving average term ($q = 3$) suggested random fluctuations for the maximum daily flows foretold with CN method around the time-series equilibrium point. In other words, there were random fluctuations (temporal variability) in the series of maximum daily flows during the three days preceding the flow foretold. It was also observed that ARIMA model for the time series of average daily flow observed in the hydrometric station had a short memory ($p = 2$). The average daily flow observed was influenced by the river flow of two previous days. The moving average term of second order ($q = 2$) suggested random fluctuations (temporal variability) in the time series of average daily flow for two days previous to the observed event.

As noted, the orders of ARIMA models (p and q) developed for the flow time series foretold with CN method and those observed in the hydrometric station were comparatively low (Table 1). Possibly this trend was due to morphological characteristics of the study basin, which were characterized by a low concentration time and the occurrence of high runoff speeds that conferred its classification of torrential basin. The results also showed that ARIMA models developed for CN method had a worse adjustment in relation to the model developed for river flows observed in the hydrometric station. The mean absolute percentage error (MAPE) for ARIMA forecasts of maximum and average daily flows generated with CN method were 117% and 120%, respectively. MAPE for average daily flows observed in the hydrometric station was 8.9% (Table 1).

Therefore, the results suggested a greater variability in time series foretold with CN method in relation to time series observed in the hydrometric station. It is important to mention that when rainfall events did not occur it was not possible to make forecasts with CN method. In these cases, were included the base flows calculated to maintain a continuous time series and being able to develop ARIMA models.

Table 1. ARIMA models developed for flow time series.

Time series	ARIMA models					
	(p,d,q)	Trans.[a]	Equation	R^2	MAPE[b] (%)	MaxAPE[c] (%)
Maximum daily flow forecasted (m^3/s)	(0,1,3)	P	$W_t = (1 - 0.824B - 0.065B^2 - 0.055B^3)a_t$	0.12	117.8	812
	(1,0,3)	P	$(1 - 0.987B)\tilde{Z}_t = (1 - 0.816B - 0.064B^2 - 0.052B^3)a_t$	0.13	119.9	764
Average daily flow forecasted (m^3/s)	(0,1,3)	P	$W_t = (1 - 0.821B - 0.066B^2 - 0.054B^3)a_t$	0.12	78.8	514
	(1,0,3)	P	$(1 - 0.986B)\tilde{Z}_t = (1 - 0.812B - 0.064B^2 - 0.051B^3)a_t$	0.13	80.3	484
Average daily flow observed (m^3/s)	(2,1,2)	Ln	$(1 + 0.283B + 0.650B^2)W_t = (1 + 0.348B + 0.613B^2)a_t$	0.88	8.90	142

Note. [a] Transformation: P = Potential and Ln = Natural logarithm. [b] MAPE = Mean absolute percentage error. [c] MaxAPE = Maximum absolute percentage error.

CONCLUSIONS

The development of this study allows exposing the following conclusions:

1) Forecasts with CN method tend to overestimate the flow rates observed in the hydrometric station. On average, the daily flows foretold are 2.91 times higher in relation to the daily flows observed.

2) CN method forecasts under the same temporal structure ARIMA the time series of maximum and average daily flows. Namely, the temporal structure of time series is conditioned by the rainfall regime rather than by CN method. CN method does not include in its development a temporal analysis of the generated time series (river flow).

3) The temporal analysis ARIMA allows to demonstrate in torrential basins a low order in the polynomials of models developed: [0,1,3], [1,0,3], and [2,1,2]. This is probably attributed to the low concentration time of the study basin (0.14 h).

4) Adjustment of the developed ARIMA models allows suggesting a greater variability in the flow time series foretold with CN method in relation to time series observed in the hydrometric station.

Finally, the analyses and conclusions presented in this study will be useful for institutions responsible for managing the water resource, to propose new methodologies for temporal analysis of river flows in projects for their use and risk management (disaster and emergencies prevention).

ACKNOWLEDGMENTS

The authors appreciate the support of Subdirección de Geografía y Cartografía del Instituto Geográfico Agustín Codazzi, Secretaría Distrital de Ambiente de Bogotá D.C., y Empresa de Acueducto y Alcantarillado de Bogotá D.C. (Colombia).

REFERENCES

[1] Alonso, R.L. (2001). Consideraciones acerca de los límites de aplicación del método de número de curva de Soil Conservation Sevice. Montes Ciencia y tecnología, 66, 92-97.

[2] Box, G., Jenkins, G. (1970). Time Series Analysis, Forecasting and Control. San Francisco: Holden-Day.

[3] Chow, V.T., Maidment, D., Mays L. (1994). Hidrología aplicada. Santafé de Bogotá, Colombia: McGraw-Hill.

[4] Deshmukh, D.S., Chaube, U.C., Hailu, H.E., Gudeta, D.A., Kassa, M.T. (2013). Estimation and comparision of curve numbers based on dynamic land use land cover change, observed rainfall-runoff data and land slope, Journal of Hydrology, 492, 89-101.

[5] Kottegoda, N.T., Natale, L., Raiteri, E. (2000). Statistical modelling of daily streamflows using rainfall input and curve number technique, Journal of Hydrology, 234(3–4), 170-186.

[6] Mendivelso, H.A., Camarero, J., Gutiérrez, E., Zuidema, P.A. (2014). Time-dependent effects of climate and drought on tree growth in a Neotropical dry forest: Short-term tolerance vs. long-term sensitivity. Agricultural and Forest Meteorology, 188, 13-23.

[7] Norambuena, P. (2009). Análisis y predicción de los efectos del crecimiento urbano sobre escorrentía e infiltración en la provincia de Chacabuco, región metropolitana de Santiago, Chile. Santiago. Recuperado el 22 de 02 de 2016. Disponible en: http://mgpa.forestaluchile.cl/Tesis/Normabuena,%20 Pablo.pdf

[8] Pereira, F.L., Valente, F., David, J.S., Jackson, N., Minunno, F., Gash, J.H. (2016). Rainfall interception modelling: Is the wet bulb approach adequate to estimate mean evaporation rate from wet/saturated canopies in all forest types?, Journal of Hydrology, 534, 606-615.

[9] Rodríguez-Barrios, J., Ospina, R. (2007). Retención de materia orgánica particulada gruesa en una montaña tropical Bogotá. Colombia. Acta Biológica Colombiana, 12(2), 33-46. Recuperado el 22 de 02 de 2016. Disponible en: http://www.bdigital.unal.edu.co/29187

[10] Tucci, C. (1998). Modelos Hidrológicos. Porto Alegre, Brasil: Editora da Universidade.

[11] USAL. (2013). Universidad del Salvador. Recuperado el 02 de Mayo de 2014, disponible en:

http://www.salvador.edu.ar/vrid/iiefgs/drenarieshidro.htm

[12] Weber, J., Jorquera, E. (2010). Desarrollo y calibración de un modelo hidrológico de simulación mixta. Mecánica computacional.

[13] Zafra, C., Ángel, Y., Torres, E. (2017). ARIMA analysis of the effect of land surface coverage on PM10 concentrations in a high-altitude megacity. Atmospheric Pollution Research, 8(4), 660-668.