JOURNAL OF
ENVIRONMENTAL HYDROLOGY

The Electronic Journal of the International Association for Environmental Hydrology
On the World Wide Web at http://www.hydroweb.com

VOLUME 17 2009

MAXIMAL, MINIMAL AND MEAN SURFACE RUNOFF IN COLOMBIA: HOW IS IT DISTRIBUTED?

Efraín Domínguez[1,2]
Alexander Hassidoff[1,2]
Juan León[3]
Yulia Ivanova[4]
Hebert Rivera[5]

[1]Pontificia Universidad Javeriana, Bogotá, Colombia
[2]CeiBA-Complejidad, Bogotá, Colombia
[3]Universidad Nacional, Palmira, Colombia
[4]Universidad Central, Bogotá, Colombia
[5]Universidad Libre, Bogotá, Colombia

For 418 hydrological stations, the time series of maximal, minimal and mean annual runoff were built and analyzed to establish a more suitable theoretical probability density function (PDF) or cumulative density function (CDF) to describe the annual runoff in Colombia. For each time series the empirical CDF was compared to normal, lognormal, Gamma and Weibull theoretical CDFs. The Kolmogorov, Smirnov and Pearson criteria were used to test the goodness of fit at a significance level α=0.10. Our results show that the Gamma CDF is the best model to describe maximal, minimal and mean surface runoff. This work is preliminary research that establishes the baseline for building hydrological and climate change scenarios in a probabilistic manner. Further research will concentrate on how probabilistic runoff patterns will evolve under climate change conditions.

INTRODUCTION

Understanding surface runoff is crucial from several points of view. First, water is a vital resource for men and all ecosystems, and, therefore, measures are to be taken to guarantee its appropriate distribution. Second, surface runoff that is out of control can be a big threat for settlements neighboring rivers. Third, water has a number of uses in society, from agriculture to electric power generation. All these aspects make understanding the behavior of surface runoff of vital interest for governments and industry.

Surface runoff is a complex phenomenon involving rainfall, its timing, surface characteristics, subsurface runoff, and atmospheric processes, including evapotranspiration and others. Surface runoff properties can be usefully characterized by means of statistics.

A statistical approach to the study of surface runoff requires a series of historical data that is as long as possible from which the behavior of hydrological variables such as extreme flows values can be explored using statistical methods (Haan, 2002; Rozhdenstvenskiy and Chevotariov, 1974).

This paper studies the hydrological regime of surface runoff in Colombia by means of probability functions. Three surface runoff variables were studied: minimum, average and maximum annual flows. The results will be used as the baseline (initial conditions) for the stochastic modeling of hydrological scenarios under climate change conditions through the Fokker – Planck – Kolmogorov equation (Domínguez, 2007; Kovalenko et al., 2005).

METHODS AND DATASETS

Datasets

Data from 418 flow monitoring stations distributed in 9 main hydrological regions of the country were provided by the IDEAM, the organization responsible for hydrologic measurements in Colombia. These data consisted of daily average flow series for an interval of about 30 years, from 1970 to 2000. It has to be noted that not all the stations have the same monitoring interval; 1970 to 2000 is the most common, but some stations have shorter intervals with 1984 to 1998 being the shortest. It is also important to note that the series directly resulting from the monitoring process included, as is usual in hydrological measurements, blank intervals where no data where recorded. To fill these blank intervals, the IDEAM used the methods suggested by Martínez (2001) and Martínez and Ruíz (1998). The geographic distribution of the stations was so that the data represents the different climatic regions of the country (see Table 1).

Method (defining runoff probabilistic patterns)

- Runoff: a Random Variable

Extreme hydrological events such as maximum and minimum flows, as well as long term average flow values - annual average flow for instance - can be treated as random variables. In fact, in hydrology, there is a long tradition using the methods of the theory of probability when assessing

Table 1. Distributions of stations by geographic zones.

CLIMATE	REGION	NUMBER OF STATIONS
Tropical Rain Forest	Medio Magdalena, Pacífico, Catatumbo	112
Tropical Dry Forest	Medio Cauca, Caribe	74
Tropical Savanna	Llanos	88
Humid Tropical Mountain Forest	Alto Cauca, Alto Magdalena	79
Unclassified Climates	*	65

water resources (Lvovitch, 1970; WMO, 1994) or supporting hydraulic design in civil engineering (Foster, 1923; Fujita and Kudo, 1995; Haan, 1977; Klemes, 1995a; Klemes, 1995b; Rozhdenstvenskiy and Chevotariov, 1974).

- Typifying a catchment runoff by means of cumulative density functions for maximum, mean and minimum flows

To describe the behavior of a variable it is useful to assess measures of its central tendency as well as measures of its dispersion. For this, a number of parameters are evaluated from a sample of data: the mean, the median and the mode for describing the central tendency, and the standard deviation or the variance for describing the dispersion. A second step would be to study the moments of the sample – skewness and kurtosis essentially - in order to typify the shape of the data distribution. The next step to characterize the data distribution is to construct its frequency distribution, dividing its range into class intervals and counting the number of occurrences that fall in each class. The set of parameters that can be obtained following this method - parameters that partially describe the characteristics of the variable - are useful for making decisions related to the variable. For example, the mean flow assessed from historical data of a river is valuable information for a water management authority interested in solving water supply problems of a village as it gives an idea of the amount of flow that could be expected to be used in the future. In the same manner, the standard deviation reveals information about the dispersion of the flows that can be expected and, therefore, it gives an indication of the size required for storage of water during times of drought and allows assessment of hydrological risk. However, this procedure leads to a partial description of the behavior of the variable. For a more precise and useful description of the variable, probability density functions (PDFs) or cumulative density functions (CDFs) are used. Experience has demonstrated that, fortunately, there is no need to construct a function for each set of data. Instead of this, it is better to fit existing well known theoretical functions to the data. Briefly, the procedure consists of taking a number of these well known functions and fitting each one of them to the data. The fitting process consists of searching for the combination of parameters of the function that make it best represent or fit the data (Haan, 2002 ; Rozhdenstvenskiy and Chevotariov, 1974). Two main methods can be applied: the method of the maximum verisimilitude and the method of the moments. A third method would be to perform an optimization process where a measure of the difference between the theoretical function and the data is to be minimized; in this case it is very useful to begin the optimization process by using the set of parameters that can be obtained by one of the two methods cited above. This last optimization method has been used in this work. A more detailed explanation of the method can be found in Akai (1994) and Zwillinger (1997). Once the functions are fitted, the goodness of fit of each one of them is tested in order to identify the best one.

- Probability functions used

Four probability functions were used:

a) Normal distribution

A random variable X is said to have a normal distribution with mean μ and variance σ if it has the density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)[(x-\mu)/\sigma]^2} \quad -\infty < x < \infty \tag{1}$$

As the central limit theorem states that the sum of n independent random variables is approximately normally distributed, this distribution is particularly useful when working with variables resulting from the sum of others such as the average annual flow.

b) Lognormal Distribution

A random variable X is said to have a lognormal distribution if its logarithm follows the normal distribution. Calling Y the logarithm of X we have $Y = \ln X$ and Y is normally distributed with mean μ_Y and standard deviation $\sigma_{Y;}$ the density function of X is:

$$\begin{aligned} f(x) &= \frac{1}{x\sigma_Y\sqrt{2\pi}} e^{-(1/2)\left[(\ln x-\mu_Y)/\sigma_Y\right]^2}; x > 0 \\ &= 0; \quad \text{otherwise.} \end{aligned} \tag{2}$$

The mean of a random variable that has a lognormal distribution is $E(X) = \mu_X = e^{\mu_Y + (1/2)\sigma_Y^2}$ and its variance is $V(X) = \sigma_X^2 = e^{2\mu_Y + \sigma_Y^2}(e^{\sigma_Y^2} - 1)$. This distribution presents two advantages in relation to the normal distribution. First, it has positive constrained values; second, it does not have a symmetrical shape. Both of these characteristics are closer to the features of the majority of hydrological variables.

c) Gamma Distribution

A random variable X is said to have a gamma distribution if it has the density function:

$$\begin{aligned} f(x) &= \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1} e^{-\lambda x}; x > 0 \\ &= 0; \quad \text{otherwise.} \end{aligned} \tag{3}$$

Here: $\Gamma(n)$ is the gamma function, defined as $\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx$.

The parameters of this function are λ, which is called the shape parameter, and r, called the scale parameter. The mean of a random variable that has a gamma distribution is $E(X) = r/\lambda$ and its variance is $V(X) = r/\lambda^2$. This distribution works particularly well for variables related to Poisson processes. Extreme hydrological phenomena such as minimum or maximum flows can be seen as Poisson processes as they occur instantly and are independent from other extreme phenomena (Montgomery and Runger, 2003).

d) Weibull Distribution

A random variable X is said to have a Weibull distribution if it has the density function:

$$\begin{aligned} f(x) &= \frac{\beta}{\delta}\left(\frac{x-\gamma}{\delta}\right)^{\beta-1} \exp\left[-\left(\frac{x-\gamma}{\delta}\right)^{\beta}\right]; x > \gamma \\ &= 0; \quad \text{otherwise.} \end{aligned} \tag{4}$$

The parameters of this distribution are γ, the location parameter, δ, the scale parameter (positive) and β, the shape parameter (positive). The mean of a Weibull distributed random variable is $E(X) = \gamma + \delta\Gamma(1 + 1/\beta)$ and the variance is $V(X) = \delta^2 \{\Gamma(1 + 2/\beta) - [\Gamma(1 + 1/\beta)]^2\}$. In this particular work a simplified version of the Weibull distribution with the location parameter value equal to zero ($\gamma = 0$) has been used as this is the one provided by MS Excel, which is the program that has been used for all the statistical calculations. The Weibull distribution is an extreme value distribution and works very well for variables representing extreme hydrological phenomena such as maximum and minimum flows. As it has three parameters, it is generally easy to make it fit well to the data. It has been widely used for maximum flow analysis (Montgomery and Runger, 2003).

- Fitting process (algorithm, optimization method)

The first task to accomplish was to create the maximum, average and minimum annual flow series from the available series of daily average flows, for each flow monitoring station.

The maximum annual flows series and the average annual flow series were relatively easy to construct: for the first one the maximum flow value for each year of data was extracted from the original series, for the second one, the mean of the daily flow values was assessed for each year of the data. The construction of minimum annual flows series required a little longer procedure as local monitoring experience has showed that flow measurements during dry events are not as reliable as during average conditions or extremely high flow events. Minimum annual flow values were calculated by assessing the mean of runoff values that are exceeded 90 % of time within each year. The maximum, average and minimum annual flow series were arranged in a descending order and the probability for each value to be exceeded was assessed by the Weibull equation:

$$P(X \geq x_m) = \frac{m}{n+1} \tag{5}$$

where m is the position of the value for which probability to be exceeded is to be assessed and n is the total number of values in the series. Using Equation (5) we built the empiric cumulative frequency histogram of the random variable (maximum, minimum or mean annual runoff), the pattern to which the theoretic probability density functions are to be fitted.

The fitting process included two steps. First, the method of moments was used to find a set of parameters of the function that guarantee that its moments are equal to the moments of the sample. Secondly, the conjugate gradient method was implemented by means of the Excel Solver tool in order to optimize the set of parameters for each promoted theoretic CDF; the optimizing criterion was to reduce to a minimum the mean of the differences of the values of the theoretic function from the corresponding empirical values. To accomplish the first step, the following relations between the parameters of each function and the moments of the sample were used (Table 2, Montgomery and Runger, 2003).

For the Weibull distribution there are no explicit expressions relating the parameters and the moments, therefore, an arbitrary first set of parameters was defined to begin the optimizing process.

Goodness of Fit

In order to assess the goodness of fit between each proposed theoretic distribution and the empirical one, three non parametric tests were performed for each flow monitoring station. The tests that were performed are the Kolmogorov test, the Pearson (Chi-Square) test and the Omega Square test (or Cramer von Mises test). It should be noted that Kolmogorov test is oriented to

Table 2. Relationship between PDF parameters and time series statistical moments.

Distribution	Relation between parameters and moments
Normal	$\mu = \bar{x}$, $\sigma = S_x$
Log Normal	$\mu_y = \bar{y}$, $\sigma_y = S_y$
Gamma	$\lambda = \frac{\bar{x}}{S_x^2}$, $r = \frac{\bar{x}^2}{S_x^2}$

measure the biggest difference between individual data pairs while the Chi-Square and the Omega-Square tests are oriented to measure the difference between the sums of all the individual differences between data pairs. This choice of the tests to perform was made so that a wider view and understanding of the fitting process could be obtained. For all three tests a level of significance of 0.10 was used. The level of significance is the probability related to the risk of rejecting the hypothesis when it should be accepted.

Kolmogorov test

The Kolmogorov test compares observed and expected frequencies by means of the statistic $\sqrt{n}D_n$ where n is the number of data pairs and Dn is given by:

$$D_n = \sup_x |F_n(x) - F(x)| \tag{6}$$

The statistic D_n follows the Kolmogorov distribution, which corresponding cumulative distribution function is:

$$\Pr(K \le x) = 1 - 2\sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2x^2} \tag{7}$$

The procedure of the test consists of defining a level of significance and its corresponding K^2 value. This value will be called the critical value and denoted by $K_{0.05}{}^2$ if a level of significance of 0.05 is being used or by $K_{0.10}{}^2$ for a level of significance of 0.10. When $\sqrt{n}D_n$ is larger than the critical value the hypothesis that the flow values follow the theoretical distribution will be rejected, otherwise it will be accepted.

Pearson Test (Chi-Square)

The Chi-Square test compares observed and expected frequencies of a distribution by means of the statistic $\chi_0{}^2$ given by:

$$\chi_0{}^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{8}$$

where k is the number of intervals in the frequency histograms.

It should be noted that $\chi_0{}^2$ is positive. Small values of $\chi_0{}^2$ denote good agreement between the empirical and the theoretical distribution, large values of $\chi_0{}^2$ denote discrepancy between the distributions. The statistic $\chi_0{}^2$ approximately follows the chi square distribution with k-p-1 degrees of freedom, where p represents the number of parameters of the theoretical distribution estimated by sample statistics. In such a condition, the procedure of the test consists of defining a level of significance and its corresponding χ_0^2 value. This value will be called the critical value and is denoted by $\chi_{0.05}^2$ if a level of significance of 0.05 is being used or by $\chi_{0.10}^2$ for a level of significance of 0.10. When χ_0^2 is larger than the critical value the hypothesis that the flow values follow the theoretical distribution will be rejected, otherwise it will be accepted

It is also important to note that by using the procedure explained above the boundaries that have been chosen for the cells are such that the expected frequencies are equal for all the cells.

Omega-Square Test (Cramer Von Mises Test)

The Omega-Square test compares observed and expected frequencies by means of the statistic T given by:

$$T = nW^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left[\frac{2i-1}{2n} - F(x_i)\right]^2 \quad (9)$$

Once a level of significance is defined and a critical value of T can be defined by using the Cramer von Mises table, then, if T is larger than the critical value the hypothesis that the flow values follow the theoretical distribution will be rejected, otherwise it will be accepted.

RESULTS

Results for maximum annual flows appear in Table 3. For a total of 420 stations, as it can be seen that the hypothesis that the maximum annual flow follows a gamma distribution was the best. In fact, the gamma distributions obtained the highest acceptance percentages for the three tests performed: 79.8, 93.1 and 96.4 % for Kolmogorov, Chi Square and Omega Square tests respectively. The second best fitting results were obtained by the lognormal distribution for the Kolmogorov test (75.5 % of acceptance); for the Chi Square and the Omega Square test the second best fitting distribution was the Weibull distribution (91.2 and 90.5 % respectively). The Normal distribution fitted the worst for all three tests (63.8, 82.1 and 85.2 % for Kolmogorov, Chi Square and Omega Square respectively). The analysis of the Mean Absolute Relative error (MARE) of fit showed the minimum MARE for the gamma distributions to be 11.8% with a standard deviation of 4.7%. For the normal, lognormal and Weibull distributions the MARE was 13.4, 14.9% with standard deviations of 5.0, 13.1 and 4.1% respectively.

Results for minimum annual flows are shown in Table 4 for a total of 423 stations. The best results were the for the Gamma and Weibull distributions. For the Kolmogorov and Omega tests the Gamma distribution obtained the best results (73.0 and 86.8 % of the test were accepted respectively) while the Weibull distribution was second (57.9 and 78 % of the tests were accepted respectively). For the Chi Square test the Weibull distribution the one that obtained the best results (80.9 of the tests were accepted) while the Gamma distribution was second (76.8 of the tests were accepted). For all three tests the normal distribution was third best according to the number of accepted tests while the lognormal distribution gave the worst fit. For the minimum annual discharges the Gamma distributions showed the minimum MARE of fit with an average value of 11.9% with a standard deviation of 4.1%. In this case the normal, lognormal and Weibull distributions have shown average values for the MARE of fit of 14.0, 24.4 and 13.0% with standard deviations of 5.4, 31.1, and 5.7% respectively.

Table 3. Results of the tests for maximum annual flows performed on 420 stations.

Test	KOLMOGOROV			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	335	268	317	282
%	79.8	63.8	75.5	67.1
Classification of the fitting according to the percentage of acceptance	1	4	2	3
Test	CHI SQUARE			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	391	345	370	383
%	93.1	82.1	88.1	91.2
Classification of the fitting according to the percentage of acceptance	1	4	3	2
Test	OMEGA SQUARE			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	405	358	372	380
%	96.4	85.2	88.6	90.5
Classification of the fitting according to the percentage of acceptance	1	4	3	2

Table 4. Results of the tests for minimum annual flows performed on 423 stations.

Test	KOLMOGOROV			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	309	223	230	245
%	73.0	52.7	54.4	57.9
Classification of the fitting according to the percentage of acceptance	1	3	4	2
Test	CHI SQUARE			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	325	315	280	342
%	76.8	74.5	66.2	80.9
Classification of the fitting according to the percentage of acceptance	2	3	4	1
Test	OMEGA SQUARE			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	367	320	279	330
%	86.8	75.7	66.0	78.0
Classification of the fitting according to the percentage of acceptance	1	3	4	2

Results for average annual flows for a total of 433 stations are shown in Table 5. The hypothesis that the maximum annual flow follows a gamma distribution was the most accepted. In fact, the gamma distributions obtained the highest acceptance percentages for two of the three tests performed: 84.4, 85.0 and 91.2 % for Kolmogorov, Chi Square and Omega Square tests respectively. The second best fitting results were obtained by the lognormal distribution with 87.8, 79,7 and 86.8 % of acceptance for the Chi Square, Kolmogorov and Omega Square tests respectively. The worst fit results were for the Weibull distribution with 69.7, 85.7 and 88.2 % of acceptance for Chi Square, Kolmogorov and Omega Square tests respectively. The MARE of fit for the mean annual discharges showed 11.1, 12.0, 12.4 and 12.4% average values with standard

Table 5. Results of the tests for average annual flows performed on 433 stations.

Test	KOLMOGOROV			
Distribution	Gamma	Normal	Log N	Weibull
Number of tests where hypothesis was accepted	361	321	345	302
%	84.4	74.1	79.7	69.7
Classification of the fitting according to the percentage of acceptance	1	3	2	4
Test	CHI SQUARE			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	368	371	380	371
%	85.0	85.7	87.8	85.7
Classification of the fitting according to the percentage of acceptance	4	2	1	3
Test	OMEGA SQUARE			
Distribution	Gamma	Normal	LogN	Weibull
Number of tests where hypothesis was accepted	395	385	376	382
%	91.2	88.9	86.8	88.2
Classification of the fitting according to the percentage of acceptance	1	2	4	3

deviations of 4.1, 5.4, 31.1 and 5.7% for Gamma, normal, lognormal and Weibull distributions respectively.

Analysis by regions

As can be seen in the following tables (Tables 6, 7 and 8), results are quite uniform geographically with respect to the best fit distribution. In the majority of regions the best fit results were obtained for the Gamma distribution. In spite of this it has to be noted that uniformity of results for Kolmogorov and Omega Square tests was stronger than for the Chi Square Test. For the latter, best fitting results were obtained for the lognormal in two regions (Medio Cauca and Caribe) and for the Weibull distribution in two other regions (Otras and Pacífico).

CONCLUSIONS

It is recommended that the Gamma distribution be used to characterize the hydrological regime of runoff in Colombian catchments. It is valid for series of maximal, minimal and mean annual runoff. For all types of discharges (maximal, minimal and mean annual runoff) the Kolmogorov test showed a higher rejection of the null hypothesis for the lognormal, normal and Weibull distributions. The Chi Square and Omega Square tests have shown similar levels of rejection for all theoretic PDFs instead. Nevertheless, the MARE of fit assessment showed lower average values and standard deviations for the fit with Gamma distributions. The lognormal theoretic distribution has poorer results in this sense, having high average values and standard deviation for the MARE of fit.

The Gamma theoretic distribution constitutes a subfamily of CDF that belongs to Pearson III family type (Rozhdenstvenskiy and Chevotariov, 1974). At the same time a series of work has proposed the use of the Fokker – Planck – Kolmogorov equation as an approach to build hydrological scenarios under non-stationary conditions (Dolgonosov and Korchagin, 2007;

Table 6. Maximum Annual Flow - Results by Regions.

	KOLMOGOROV TEST					CHI SQUARE TEST					OMEGA SQUARE TEST				
Theoretic Distribution type	Gamma	Normal	Log Normal	Weibull	TOTAL	Gamma	Normal	Log Normal	Weibull	TOTAL	Gamma	Normal	Log Normal	Weibull	TOTAL
% of accepted hypothesis	79,8	63,8	75,5	67,1	100,0	93,1	82,1	88,1	91,2	100,0	96,4	85,2	88,6	90,5	100,0
Classification of the fitting	1	4	2	3		1	4	3	2		1	4	3	2	
MEDIO CAUCA															
% of accepted hypothesis	86,2	69,0	86,2	69,0	100,0	93,1	82,8	100,0	86,2	100,0	100,0	96,6	100,0	96,6	100,0
Classification of the fitting	1	2	1	2		2	4	1	3		1	2	1	2	
MEDIO MAGDALENA															
% of accepted hypothesis	77,8	63,0	74,1	61,1	100,0	92,6	79,6	75,9	92,6	100,0	100,0	81,5	77,8	92,6	100,0
Classification of the fitting	1	3	2	4		1	2	3	1		1	3	4	2	
OTRAS															
% of accepted hypothesis	81,5	78,5	73,8	73,8	100,0	87,7	87,7	84,6	90,8	100,0	92,3	90,8	86,2	90,8	100,0
Classification of the fitting	1	2	3	3		2	2	3	1		1	2	3	2	
PACÍFICO															
% of accepted hypothesis	84,2	68,4	73,7	60,5	100,0	92,1	86,8	73,7	94,7	100,0	89,5	81,6	73,7	86,8	100,0
Classification of the fitting	1	3	2	4		2	3	4	1		1	3	4	2	
ALTO CAUCA															
% of accepted hypothesis	80,0	60,0	84,0	60,0	100,0	96,0	76,0	96,0	96,0	100,0	100,0	92,0	100,0	84,0	100,0
Classification of the fitting	2	3	1	3		1	2	1	1		1	2	1	3	
ALTO MAGDALENA															
% of accepted hypothesis	74,1	35,2	72,2	59,3	100,0	96,3	63,0	87,0	88,9	100,0	98,1	63,0	87,0	83,3	100,0
Classification of the fitting	1	4	2	3		1	4	3	2		1	4	2	3	
CARIBE															
% of accepted hypothesis	79,2	62,5	75,0	75,0	100,0	91,7	81,3	97,9	87,5	100,0	97,9	89,6	95,8	91,7	100,0
Classification of the fitting	1	3	2	2		2	4	1	3		1	4	2	3	
CATATUMBO															
% of accepted hypothesis	89,5	63,2	68,4	89,5	100,0	89,5	73,7	78,9	89,5	100,0	100,0	89,5	78,9	94,7	100,0
Classification of the fitting	1	3	2	1		1	3	2	1		1	3	4	2	
LLANOS															
% of accepted hypothesis	77,3	69,3	76,1	65,9	100,0	96,6	93,2	95,5	93,2	100,0	95,5	89,8	95,5	93,2	100,0
Classification of the fitting	1	4	2	3		1	3	2	3		1	3	1	2	

Table 7. Minimum Annual Flow - Results by Regions.

	KOLMOGOROV TEST					CHI SQUARE TEST					OMEGA SQUARE TEST				
	Normal	Log Normal	Gamma	Weibull	TOTAL	Normal	Log Normal	Gamma	Weibull	TOTAL	Normal	Log Normal	Gamma	Weibull	TOTAL
%	52,7	54,4	73,0	57,9	100,0	74,5	66,2	76,8	80,9	100,0	75,7	66,0	86,8	78,0	100,0
Classification of the fitting	3	4	1	2		3	4	2	1		3	4	1	2	
MEDIO CAUCA															
%	48,3	86,2	82,8	62,1	100,0	79,3	100,0	86,2	93,1	100,0	75,9	96,6	100,0	79,3	100,0
Classification of the fitting	4	1	2	3		4	1	3	2		4	2	1	3	
MEDIO MAGDALENA															
%	55,6	50,0	79,6	61,1	100,0	88,9	63,0	88,9	94,4	100,0	87,0	59,3	96,3	87,0	100,0
Classification of the fitting	3	4	1	2		2	3	2	1		2	3	1	2	
OTRAS															
%	62,1	74,2	81,8	60,6	100,0	89,4	86,4	92,4	89,4	100,0	83,3	86,4	92,4	86,4	100,0
Classification of the fitting	3	2	1	4		2	3	1	2		3	2	1	2	
PACÍFICO															
%	61,0	58,5	75,6	58,5	100,0	82,9	73,2	82,9	90,2	100,0	82,9	73,2	95,1	87,8	100,0
Classification of the fitting	2	3	1	3		2	3	2	1		3	4	1	2	
ALTO CAUCA															
%	64,0	52,0	80,0	68,0	100,0	72,0	84,0	92,0	88,0	100,0	88,0	76,0	96,0	80,0	100,0
Classification of the fitting	3	4	1	2		4	2	1	3		2	4	1	3	
ALTO MAGDALENA															
%	64,8	55,6	79,6	70,4	100,0	87,0	70,4	77,8	85,2	100,0	94,4	68,5	98,1	96,3	100,0
Classification of the fitting	3	4	1	2		1	4	3	2		3	4	1	2	
CARIBE															
%	52,1	54,2	83,3	62,5	100,0	75,0	60,4	93,8	93,8	100,0	72,9	62,5	95,8	83,3	100,0
Classification of the fitting	4	3	1	2		2	3	1	1		3	4	1	2	
CATATUMBO															
%	78,9	57,9	89,5	78,9	100,0	73,7	57,9	57,9	73,7	100,0	94,7	68,4	100,0	84,2	100,0
Classification of the fitting	2	3	1	2		1	2	2	1		2	4	1	3	
LLANOS															
%	25,3	28,7	42,5	34,5	100,0	41,4	35,6	41,4	47,1	100,0	41,4	37,9	50,6	44,8	100,0
Classification of the fitting	4	3	1	2		2	3	2	1		3	4	1	2	

Domínguez, 2004a; Domínguez, 2004b; Domínguez, 2007; Kovalenko et al., 2005). It can be noted that the Pearson III family of CDF type is a particular solution of the Fokker – Planck – Kolmogorov equation and the use of the Gamma distribution as the starting point is recommended for stochastic models intended to predict behavior of variables such as maximum, mean and minimum annual flow under climate change conditions. Finally, any regionalization of probabilistic

Table 8. Average Annual Flow - Results by Regions.

	KOLMOGOROV TEST					CHI SQUARE TEST					OMEGA SQUARE TEST				
	Normal	Log Normal	Gamma	Weibull	TOTAL	Normal	Log Normal	Gamma	Weibull	TOTAL	Normal	Log Normal	Gamma	Weibull	TOTAL
%	74,1	79,7	83,4	69,7	100,0	85,7	87,8	85,0	85,7	100,0	88,9	86,8	91,2	88,2	100,0
Classification of the fitting	3	2	1	4		3	1	4	2		2	4	1	3	
	MEDIO CAUCA														
%	75,9	86,2	89,7	75,9	100,0	79,3	93,1	93,1	82,8	100,0	89,7	93,1	93,1	89,7	100,0
Classification of the fitting	3	2	1	3		3	1	1	2		2	1	1	2	
	MEDIO MAGDALENA														
%	80,0	83,6	90,9	78,2	100,0	94,5	90,9	90,9	96,4	100,0	96,4	90,9	94,5	94,5	100,0
Classification of the fitting	3	2	1	4		2	3	3	1		1	3	2	2	
	OTRAS														
%	83,6	80,8	89,0	78,1	100,0	90,4	82,2	89,0	91,8	100,0	91,8	82,2	91,8	89,0	100,0
Classification of the fitting	2	3	1	4		2	4	3	1		1	3	1	2	
	PACÍFICO														
%	39,5	44,2	44,2	30,2	100,0	48,8	53,5	53,5	51,2	100,0	51,2	53,5	48,8	48,8	100,0
Classification of the fitting	2	1	1	3		3	1	1	2		2	1	3	3	
	ALTO CAUCA														
%	72,0	88,0	92,0	56,0	100,0	96,0	96,0	96,0	84,0	100,0	92,0	88,0	100,0	92,0	100,0
Classification of the fitting	3	2	1	4		2	1	1	3		2	3	1	2	
	ALTO MAGDALENA														
%	77,8	77,8	87,0	68,5	100,0	85,2	85,2	81,5	81,5	100,0	92,6	88,9	98,1	92,6	100,0
Classification of the fitting	2	2	1	3		1	1	2	2		2	3	1	2	
	CARIBE														
%	59,6	89,4	83,0	59,6	100,0	85,1	97,9	93,6	87,2	100,0	91,5	95,7	97,9	89,4	100,0
Classification of the fitting	3	2	1	3		4	1	2	3		3	2	1	4	
	CATATUMBO														
%	84,2	78,9	84,2	89,5	100,0	94,7	100,0	94,7	94,7	100,0	100,0	94,7	100,0	100,0	100,0
Classification of the fitting	2	3	2	1		2	1	2	2		1	2	1	1	
	LLANOS														
%	83,0	85,2	86,4	80,7	100,0	92,0	96,6	83,0	92,0	100,0	93,2	94,3	96,6	95,5	100,0
Classification of the fitting	3	2	1	4		2	1	3	2		4	3	1	2	

characteristics of Colombian runoff must to consider the use of the Gamma distribution as the base model for estimations in un-gauged basins.

ACKNOWLEDGMENTS

This research was supported by Univerisidad Javeriana (project ID: 02185). The authors are grateful for the comments of Professors Vitaly Khaustov of the Department of Hydrophysics and Hydrological Forecasting at the Russian State Hydrometeorological University, and Miguel Ontiveros of the Centro de Levantamientos Espaciales at Universidad Mayor de San Simón in Bolivia.

REFERENCES

Akai, T.J. 1994. Applied numerical methods for engineers John Wiley & Sons.

Dolgonosov, B.M., and K.A. Korchagin. 2007. A non-linear stochastic model describing the formation of daily and mean monthly water flow in river basins. Water Resources, 34(6), pp. 624-634.

Domínguez, E. 2004a. Aplicación de la ecuación de Fokker–Planck–Kolmogorov para el pronóstico de afluencias a embalses hidroeléctricos (caso práctico de la represa de Betania). Meteorología Colombiana, 8, pp. 17-26.

Domínguez, E. 2004b. Stochastic forecasting of streamflow to Colombian hydropower reservoirs. PhD. Thesis, Russian State Hydrometeorological University, San Petersburg, 235 pp.

Domínguez, E. 2007. Hydrological scenarios modelling for climate change conditions using the Fokker-Planck-Kolmogorov equation. In: D. Schertzer, P. Hubert, S. Koide and K. Takeuchi (Editors), Proceedings of the PUB Kick-off meeting held in Brasilia, 20–22 November 2002. Red Book. IAHS Press, Brasilia.

Foster, A. 1923. Theoretical frequency curves and their application to engineering problems. Trans. Am. Soc. C.E, 87, pp. 142.

Fujita, M., and M. Kudo. 1995. Stochastic response of a storage function model for flood runoff estimation of higher-order moments. Environment International, 21(5), pp. 523-531.

Haan, C.T. 1977. Statistical Methods in Hydrology, Iowa State , Iowa.

Haan, C.T. 2002. Statistical methods in Hydrology, Iowa state University Press, Iowa.

Klemes, V. 1995a. The stochastic reservoir - Comment. Journal of Hydrology, 165, pp. 371-373.

Klemes, V. 1995b. Two early contributions to the theory of the stochastic reservoir. Journal of Hydrology, 172, pp. 351-354.

Kovalenko, V., N.V. Viktorova, and E. Gaidukova. 2005. Modelling of hydrological processes. Guidrometeoizdat, Saint Petersburg.

Lvovitch, M.I. 1970. World water balance (General Report). Symposium on the world water balance IAHS Press, Wallingford.

Martínez, L.F. 2001. Control Estadístico de Calidad, Información Hidrológica Nivel Mensual. Automatización SPSS y DEMETRA, Programa de Complementación de Series. Instituto de Hidrología, Meteorología y Estudios Ambientales - IDEAM, Bogotá, pp. 130.

Martínez, L.F., and L.F. Ruiz. 1998. Metodología para la estimación de datos faltantes en series temporales diarias. IDEAM, Subdirección de Hidrología, Bogotá, pp. 230.

Montgomery, D.C., and G.C. Runger. 2003. Applied statistics and probability for engineers John Wiley & Sons, United States of America.

WMO. 1994. Guide to hydrological practices - data acquisition and processing, analysis, forecasting and other applications World Meteorological Organization, Geneva.

Rozhdenstvenskiy, A.V., and A.I. Chevotariov. 1974. Statistical methods in hydrology. Guidrometeoizdat, Leningrad.

Zwillinger, D. 1997. Handbook of differential equations. Academic press, Boston.

ADDRESS FOR CORRESPONDENCE
Efraín Domínguez
Pontificia Universidad Javeriana
Bogotá
Colombia

Email: e.dominguez@javeriana.edu.co