DOI: 10.24850/j-tyca-15-01-09

Articles

Intensity-duration-frequency curves for Santa Clara City, Cuba

Curvas de intensidad-duración-frecuencia para la ciudad de Santa Clara, Cuba

Carlos Castillo-García[1], ORCID: https://www.orcid.org/0000-0002-6430-2775

Ismabel Domínguez-Hurtado[2], ORCID: https://www.orcid.org/0000-0002-7841-8031

Yoel Martínez-González[3], ORCID: https://www.orcid.org/0000-0002-8023-7897

Diego Abreu-Franco[4], ORCID: https://www.orcid.org/0000-0001-6161-2922

[1]Central University Marta Abreu of Las Villas, Santa Clara, Cuba, ccgarcia@uclv.cu

[2]Meteorological Center de Villa Clara, Santa Clara, Cuba, ismabel.dominguez@vcl.insmet.cu

[3]Higher Institute of Technologies and Applied Sciences, La Habana, Cuba, ymg@instec.cu

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

[4]Hydraulic Projects and Research Company of Villa Clara, iph.tecproy14@vc.giat.cu

Corresponding author: Carlos Castillo-García, ccgarcía@uclv.cu

Abstract

The intensity-duration-frequency (IDF) curves are a representation of extreme hydrometeorological phenomena of rainfall to be used in hydrological projects. In this article, an analysis of 243 convective rainy events of more than 25 mm that occurred at the Yabú Meteorological Station in Cuba, Villa Clara province, in the interim period from 1990 to 2019 was carried out with the objective of elaborating the IDF curves of the station aforementioned. A series of annual maximums was elaborated for the durations between 5 and 4 320 minutes, which was subjected to a missing data imputation process using the multiple imputation algorithm by linear regression, anomalous values were found, and their treatment was highlighted. The resulting series were tested in non-parametric tests to verify their independence, randomness and seasonality, with which they were adjusted to the Gumbel probabilistic distribution of extreme values and subsequently to a parametric equation of the Montana model. The results obtained showed that there is a point where the adjustment of the Montana model begins to obtain discordant results with the series adjusted to the Gumbel distribution, for which two families of IDF Curves are proposed: For durations ≤ 360 min and another for durations > 360 min, with which correlation coefficients greater than 0.99 are obtained.

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

Keywords: Extreme events, rainfall intensity, return period, precipitation.

Resumen

Las curvas de intensidad-duración-frecuencia (IDF) son una representación de fenómenos hidrometeorológicos extremos de la lluvia para su uso en proyectos hidrológicos. En el presente artículo se realizó un análisis de 243 eventos lluviosos convectivos de más de 25 mm ocurridos en la estación meteorológica Yabú de la provincia Villa Clara, Cuba, en el periodo comprendido desde 1990 hasta 2019, con el objetivo de elaborar las curvas IDF de dicha estación. Se elaboró una serie de máximos anuales para las duraciones comprendidas entre los 5 y 4 320 minutos, la cual se sometió a un proceso de imputación de datos faltantes usando el algoritmo de imputación múltiple por regresión lineal; se encontraron valores anómalos y se destacó su tratamiento. Las series resultantes se testearon en pruebas no paramétricas para comprobar su independencia, aleatoriedad y estacionalidad, con lo cual se procedió a ajustarlas a la distribución probabilística de valores extremos Gumbel y posteriormente a una ecuación paramétrica del modelo de Montana. Los resultados obtenidos demostraron que existe un punto donde el ajuste del modelo de Montana empieza a obtener resultados discordantes con la serie ajustada a la distribución Gumbel, por lo cual se proponen dos familias de Curvas IDF para duraciones ≤ 360 min y otra para duraciones > 360 min, con las que se obtienen coeficientes de correlación superiores a los 0.99.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Palabras clave: eventos extremos, intensidad de la lluvia, periodo de retorno, precipitación.

Received: 11/03/2022

Accepted: 05/07/2022

Published online: 19/07/2022

Introduction

The design storm usually needs Intensity-Duration-Frequency (IDF) or Sheet-Duration-Frequency (DDF) for its acronym in English, which contain the ratio of the probability of occurrence of the variables sheet and intensity with their duration. Both are used as the primary input to rainfall runoff models to estimate the magnitude of the design flood, particularly in catchments with no flow rate measurements. Singh (2017) argues that the derivations of these relationships require high-quality data handling, fitting them to a distribution of extreme values where they can then be used to extrapolate to an exceedance probability of interest. The storm duration and intensity parameters obtained from the IDF curves have great significance in the field of hydrology, and are basic elements for the study of large floods and the development of urban infrastructures (Yong, Ng, Huang, & Ang, 2021).

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

With the development of computational methodologies and studies based on experiences in other branches of engineering statistics, two classifications of IDF curves are identified; the first is according to the variability in the trends of the data series used. Authors such as Agilan and Umamahesh (2017b); Gregersen, Madsen, Rosbjerg and Arnbjerg-Nielsen (2017), and Soumya, Anjitha, Mohan, Adarsh and Gopakumar (2020) have proposed the non-seasonality of the data series as a condition of interest for a study of short-term rainfall, the author Agilan himself has made several contributions to this topic stating that the values of the data series can be increased or decrease according to a linear function, or linear trend, whose slope value can be added to the extreme value distribution function used in its position parameter (Agilan & Umamahesh, 2017a); This trend is contrasted with another part of the scientific community that is not yet fully convinced. It is worth mentioning that Agilan and Umamahesh (2017c) himself take a cautious position on the results in comparisons with periods of occurrence of up to 10 years, however other authors such as Ganguli and Coulibaly (2017), and Yilmaz and Perera (2014) despite verifying significant trends in the series of annual maximums (Noor, Ismail, Chung, Shahid, & Sung, 2018) show that non-stationary models with functions of generalized extreme values (GEV, for its acronym in English) do not yet see clear advantages over similar stationary models.

The second classification is made according to the type of data series chosen for the analysis, a preference of researchers at present is the use of partial duration series (SDP), Ben-Zvi (2009); Emmanouil, Langousis, Nikolopoulos and Anagnostou (2020), and Chang, Lai and

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Faridah (2013)have obtained satisfactory results using the Generalized Pareto (GP) distribution, this methodology consists of obtaining a series of data with values above a threshold, which would allow for each year to obtain peak values that with a series of annual maximums (SMA) would be excluded. This technique has already been used previously in similar services. Egea-Pérez, Cortés-Molina and Navarro-González (2021) carry out analyzes with rainfall in localities with scarce annual data; Masseran and Safari (2020) apply the SDPs to obtain the extreme air pollution risk assessment based on an IDF approximation. Despite this boom, Sane *et al*. (2018), and Olsson, Södling, Berg, Wern and Eronn (2019) maintain research with MAS in countries such as Senegal and Sweden, respectively.

There are comparative studies between the methodologies with SMA and SDP, Vrban, Wang, McBean-Edward, Binns and Gharabaghi (2018) show that the SDP for obtaining the design storm is more effective than the SMA since the rainfall exceeds 4 to 10 % and is therefore more conservative, in addition to the fact that greater results are obtained in return period between 2 to 5 years. . Van Campenhout, Houbrechts, Peeters and Petit (2020) use SDP to find a relationship with the SMA with respect to the return period, although the study is carried out for series of maximum runoff costs. Agilan and Umamahesh (2017b) compare their non-stationary models obtained with SMA and SDP obtaining similar results to Vrban *et al*. (2018) although they state that for short durations and small return periods the difference is higher than for long durations and longer return periods.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Singh (2017) compiles several works and presents an assessment between the relationship that exists between the return period of the SMA and SDP where the following expression is reached:

$$T_P = -\frac{1}{\ln\left(1-\frac{1}{T_A}\right)} \qquad (1)$$

Where T_P is the return period of the analysis with SDP and T_A is the return period obtained in an analysis with SMA. Even for return periods greater than 10 years, this expression can be reduced to $T_A = T_P + 1/2$, which makes both results, both for SDP and SMA, relatively equal.

Under the above criteria, the main objective of this contribution is to obtain the IDF Curves corresponding to the Yabú meteorological station near the city Santa Clara of through the study of 30 years of records between 1990 and 2019 of rainy events greater than 25 mm in 24 hours. The study is based on the best fit probability function and the parametric equation with similar conditions. In this sense, three distribution functions and four parametric equations will be evaluated for the best fit. Subsequently, the respective IDF curves will be obtained, characteristics of the station under study, to be used in a future regionalization study in future contributions.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Materials and methods

The Yabú agrometeorological station (Code 78343) is located in the province of Villa Clara, Cuba, located at 22º 26' N and 79º 59' W, at 116.44 m above mean sea level, with the presence of a flat relief (Figure 1), approximately 7 km from the center of the city of Santa Clara. It is framed on the east bank of the Sagua La Grande river basin, the largest river system in the province and where two of the most economically important reservoirs in the region are located, Palmarito and Alacranes, although it is not the only weather station of the basin, its privileged position in it allows to have an accurate behavior of the climatic variables that affect the place. It began operating on September 3, 1976, the date of its first measurement record of all its own variables. It has several instruments suitable for meteorological activity, among which is the pluviometer and pluviograph:

1. Rain gauge: Model (USWB), with measurement start date in 1976, measurements are made in millimeters (mm).
2. Rain gauge: Brand (Standard), Model (P-2), Series (281) manufactured in the former Union of Soviet Socialist Republics (USSR), with measurements beginning in 1976. However, records from the years 1976 to 1990 have intervals where information losses have occurred that compromise their analysis, therefore they will not be considered in this investigation.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

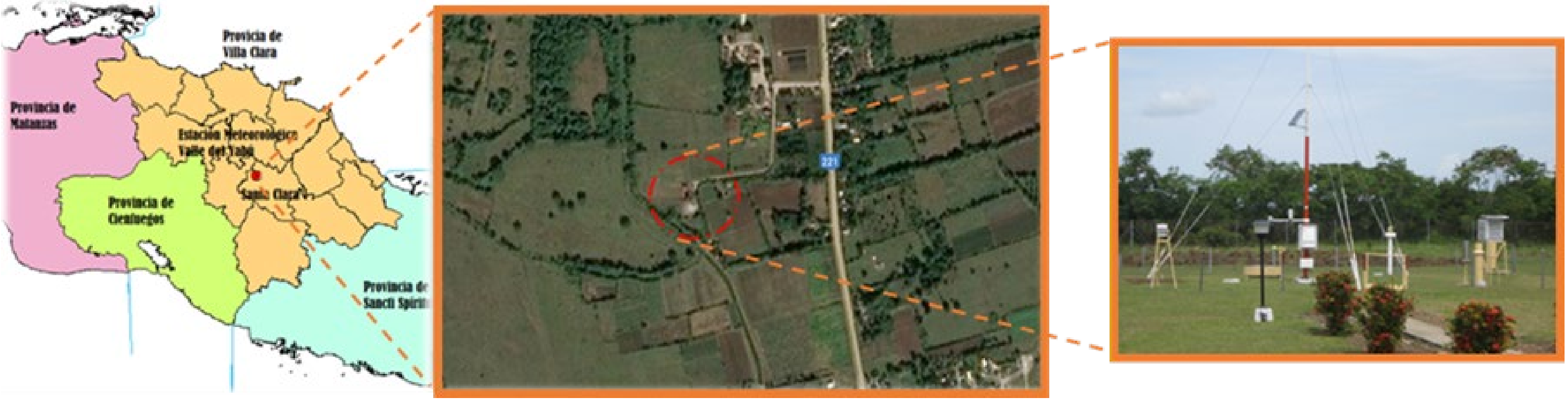

Figure 1. Geographical location of the Yabú weather station, Villa Clara province, Cuba.

An analysis of 30 years of pluviographic records is carried out from 1990 to 2019 with records only interrupted in periods that do not exceed three months, due to breakages, maintenance, malfunction or poor quality of the pluviogram for reading. Table 1 presents the data series obtained from the processing of the station charts.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table 1. Maximum intensities in mm/min for different durations recovered from the Yabú Station.

Year	1 hour	2 hours	4 hours	12 hours	24 hours
1990	-	-	-	-	-
1991	1.231	0.780	0.400	0.139	0.061
1992	0.930	0.500	0.250	0.084	0.043
1993	0.730	0.380	0.190	0.083	0.042
1994	0.939	0.563	0.291	0.097	0.051
1995	0.900	0.480	0.240	0.082	0.046
1996	0.720	0.482	0.263	0.136	0.115
1997	0.940	0.500	0.250	0.084	0.042
1998	0.910	0.470	0.255	0.158	0.194
1999	0.840	0.460	0.230	0.134	0.067
2000	1.080	0.570	0.280	0.106	0.053
2001	0.620	0.320	0.190	0.091	0.045
2002	0.870	0.520	0.260	0.094	0.047
2003	0.980	0.650	0.330	0.116	0.058
2004	1.100	0.620	0.310	0.115	0.058
2005	1.050	0.580	0.290	0.162	0.115
2006	0.920	0.480	0.240	0.101	0.051
2007	0.890	0.440	0.220	0.086	0.043
2008	1.200	0.600	0.300	0.153	0.076
2009	-	-	-	-	-
2010	1.000	0.540	0.270	0.114	0.057
2011	1.220	0.740	0.490	0.169	0.085
2012	1.000	0.530	0.260	0.093	0.075
2013	0.980	0.680	0.400	0.138	0.069
2014	1.799	1.035	0.520	0.184	0.092
2015	1.140	0.610	0.310	0.111	0.056
2016	1.020	0.550	0.280	0.095	0.048
2017	0.740	0.370	0.230	0.161	0.064
2018	1.240	0.620	0.310	0.152	0.081
2019	0.680	0.340	0.170	0.063	0.038

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

In the case of the year 1990, its records were lost and in the case of the year 2009 the amount of rain analyzed was insufficient due to problems in the data collection of the equipment (breakages and poor condition of the pluviogram). The pluviogram or pluviographic chart resulting from the analog measurement of the equipment has a minimum scale of 10 minutes for every 0.5 cm of paper on the horizontal scale, which allows the digitalization to witness one or two siphoning of the equipment within an interval of 10 min, which can be reduced to the minimum interval of 5 min that can be identified visually. The durations processed are 5, 10, 20, 40, 60, 90, 120, 150, 240, 300, 720, 1 440, 2 880 and 4 320 minutes, the last 3 correspond to 24, 48 and 72 hours of duration, typical of cyclonic events. Table1 shows the SMA for 1, 2, 4, 12 and 24 hours.

Missing and data analysis

Having missing data for various reasons in a data series is a common problem for any researcher, according to Molenberghs, Fitzmaurice, Kenward, Tsiatis and Verbeke (2015), and Little and Rubin (1987) there are three basic types of these data:

- Completely at random (MCAR, Missing Completely at random). Represents a situation for which the absence is independent of the variables of an investigation. It is a lack of data due exclusively to chance.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

- Missing at random (MAR, Missing at random). It refers to the fact that the absence of data is present in the independent variables of the study, but not in the dependent one. It is the most typical case of data loss and it is considered the type present in this investigation due to the absence of data in the primary analog source (pluviograph) due to loss of pluviographic charts and their deterioration, the intensities obtained from this source then the independent variables of the study are considered.
- A process that is not MCAR or MAR is non-random.

Singh (2017) proposes three paths to follow when there are missing data:

- Skip lost data: Only the existing data is analyzed without completing the missing records.
- Impute missing data: Using established techniques and methodologies, find a substitute for the non-existent data that propitiates the filling of the series. The Expectation-Maximization (EM) and Multiple Imputation (IM) algorithms will be analyzed in this contribution.
- Accommodate missing data: It is done with data filling techniques, but using the statistics of the data series.

Maximum likelihood methods, such as EM, can be applied to any estimation problem. In the analysis of missing data, and assuming that the missing data follow a MAR pattern, it is shown that the marginal distribution of the observed records is associated with a likelihood function

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

for an unknown parameter, under the assumption that the model is adequate for the complete data set (Mallol, 2017).

According to Little and Rubin (1987), cited in Mallol (2017), this function is known as the likelihood function, which ignores the mechanism that generated the missing data. The procedure for estimating the parameters of a model using a sample with missing data is summarized below:

- Estimate the parameters of the model with the complete data with the maximum likelihood function.
- Use the estimated parameters to predict the missing values.
- Substitute the data for the predictions, and obtain new values of the parameter, maximizing the likelihood of the complete sample.

An efficient procedure to maximize likelihood when there are missing data is the Expectation-Maximization algorithm (Miró, Caselles, & Estrela, 2017).

The multiple imputation method consists of making several imputations of the missing observations and then analyzing the completed data sets and combining the obtained results to obtain a final estimate. Multiple imputation analysis is divided into three phases: imputation phase, analysis phase and pooling phase (Mallol, 2017).

The imputation phase creates multiple copies of the data sets (m), each containing different estimates of missing values. Conceptually, this step is an iterative version of stochastic regression imputation, although its mathematical underpinnings are often based on Bayesian estimation principles (Mallol, 2017).

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

- The goal of the analysis phase, as its name suggests, is to analyze the populated data sets. This step applies the same statistical procedures that an individual would have used if he had all the data. The only difference is that we perform each analysis *m* times, once for each imputed data set.
- The analysis phase leads to *m* sets of parameter estimates and standard errors, so the purpose of the pooling phase is to combine everything into a single set of results. Little and Rubin (1987) outlined relatively simple formulas for pooling parameter estimates and standard errors. For example, the pooled parameter estimate is simply the arithmetic mean of the *m* estimates from the analysis phase. Combining the standard errors is slightly more complex, but follows the same logic. The process of analyzing multiple data sets and pooling the results seems laborious, but multiple imputation software packages such as SPSS, XLSTAT, R, which fully automate the procedure. The *m* estimates are combined into an ensemble estimate and a variance-covariance matrix using Rubin's rules, which are based on asymptotic theory in a Bayesian framework. The combined variance-covariance matrix incorporates the variability within the imputation (uncertainty about the results of imputed data sets) and the variability between the imputations (reflecting the uncertainty due to missing information).

To carry out this iterative process and speed up the imputation process to make the most convenient decisions, SPSS software version 22 is used. The procedure will be to obtain results through EM and three random imputations with IM (using the linear regression technique), and

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

the results of the largest maximum intensities obtained in any method will be chosen, always guaranteeing the following aspects:

- The mean of the series cannot vary significantly.
- The variance and covariances must remain constant with an error interval of less than 1 %.
- The results obtained in the years 1990 and 2009 that are intended to be imputed must have a logical order of obtaining intensities in relation to their duration, that is, the intensities of 5 min are greater than those of 10 min and these in turn are greater than those of 20 min and so on consecutively.
- The completed series will be analyzed consecutively to obtain outliers (values out of range); none of the imputed values can be found as an outlier, and if any non-imputed values are detected, the SMA is reconfigured and the imputation process is started again.

Outlier analysis

In geophysical sciences, such as hydrology, observations are regularly obtained for analysis and capture of changes in historical processes over a time interval. Hydrological data often contain extreme observations or anomalous data (outliers) due to real events or factors external to the measurement (Singh, 2017).

An anomalous data is one that appears far from the set of data. The presence of outliers in a data sample can create difficulties when fitting a

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

distribution to the sample. In a sample, there can be high or low value anomalous data, or both, which can influence the frequency analysis in different ways. Although the treatment of anomalous data is still a highly debated topic, certain procedures have been used in hydrology to identify and treat them, such as those described in the publication of the Water Resources Council (1981) of the United States for the analysis of flood frequency, or for extreme precipitation (OMM, 2011).

OMM (2011) and Naghettini (2017) recommend the use of the US-WRC (United States-Water Resources Council) method. To apply it, it will be necessary to assume that the logarithms or another function of the hydrological series are normally distributed, since the test is only applicable to samples obtained from a normal population. To carry out the US-WRC test, the following two expressions are calculated:

$$X_H = \exp(\bar{x} + K_N s) \tag{2}$$

$$X_L = \exp(\bar{x} - K_N s) \tag{3}$$

Where $\bar{x}$ and s in equations (2) and (3) are the mean and standard deviation of the natural logarithms of the sample, respectively, K_N is the Grubbs and Beck statistics tabulated for various sample sizes and levels of importance, and N the size of the sample, X_H is the upper limit of the test and X_L is the lower limit. For $5 \leq N \leq 150$, K_N can be calculated from Equation (4) (Stedinger *et al*., 1993, cited in OMM, 2011):

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$K_N = -0.9043 + 3.345\sqrt{\log(N)} - 0.4046\log(N) \qquad (4)$$

Data serie quality

For the results of frequency analysis to be theoretically valid, the data series must satisfy certain statistical criteria, such as randomness, independence, homogeneity and seasonality (OMM, 2011). This text also recommends the ideal tests to apply to test the hypotheses in each case, as shown in Table 2. The tests indicated are non-parametric, thus avoiding any assumption about the underlying parametric distribution of the data.

Table 2. Non-parametric tests for data quality analysis at the Yabú weather station.

Statistical Criteria	**Recommended test**	**Confidence interval (%)**
Randomness	Runs test	95
Independency (1)	Mann-Whitney test	
Independency (2)	Wald-Wolfowitz test	
Seasonality	Mann- Kendall test, Sen´s Slope	

Due to the importance observed in the analysis of the state of the art of this topic to the trend tests, to confirm the existence of seasonality

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

or not of a series of annual maximums, the following is a summary of the theory that accompanies it.

The Mann-Kendall test is a nonparametric test based on rank correlation that allows one to assess the significance of a trend. The null trend hypothesis H0 is that a time-ordered data sample is independent and identically distributed. The S statistic is defined as follows (Maity, 2018):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn\left(x_j - x_i\right) \tag{5}$$

Where:

$$sgn(x) = \begin{Bmatrix} 1\ si\ x > 0 \\ 0\ si\ x = 0 \\ -1\ si\ x < 0 \end{Bmatrix} \tag{6}$$

When *n* ≥ 40, the S statistic has an asymptotically normal distribution with mean 0 and variance given by the following equation:

$$Var\,\{S\} = \frac{1}{18[n(n-1)(2n+5)-\sum_t t(t-1)(2t+5)]} \tag{7}$$

Where *t* is the size of a given bound group and ∑ is the sum of the set of all bound groups in the data sample. The normalized test statistic *K* is calculated using the following equation:

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$K = \frac{S-1}{\sqrt{Var(S)}}; 0; \frac{S+1}{\sqrt{Var(S)}} \; para: S > 0, S = 0, S < 0, respectivamente \tag{8}$$

The normalized *K* statistic has a standard normal distribution with mean equal to 0 and variance equal to 1. The probability value *P* of the *K* statistic of the sample data can be estimated using the normal cumulative distribution function, in the form:

$$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt \tag{9}$$

For independent data samples with no trend, the *P* value should be equal to 0.5. When the sample data show a strong positive trend, the *P* value should be close to 1, while a strong negative trend should give a *P* value close to 0. If the sample data are serially correlated, whitening will be necessary, previously the data and apply a correction to calculate the variance (OMM, 2011).

For the linear trend, the slope is usually estimated by computing the least squares estimate using linear regression. However, it is only valid when there is no serial correlation and the method is very sensitive to outliers. Sen (1968), cited in OMM (2011), developed a more robust method (OMM, 2011).

The slope of a trend can be estimated as follows:

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$Q = mediana\left(\frac{x_i - x_j}{i-j}\right);\ \forall\, j < i \qquad (10)$$

Where Q is the estimated value of the slope of the trend and x_j is the 20th observation. The rising trends are represented by a positive value of Q, while downtrends are represented by a negative value of Q.

The slope estimator of Sen is the median of the *N*' values of Q. The same procedure is followed whether there are one or several observations per time period.

Sen (1968), cited in OMM (2011), provides a nonparametric method for obtaining a confidence interval for this slope, although a simple normal approximation method is more commonly used. For this we need the standard deviation of the Mann-Kendall statistic, *S* (OMM, 2011).

Statistical distributions functions

Probability distributions are used in a wide variety of hydrological studies, particularly in studies of extreme high and low flows, floods, reservoir volumes, rainfall amounts, and time series models. It should be noted that in the SMA study, the analysis distributions are well defined, and in recent years the Generalized Extreme Value Distribution (GVE) has been used more strongly in stationary and non-stationary models (Olsson *et al*., 2019; Yong *et al*., 2021; Agilan & Umamahesh, 2017a).

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table 3 shows the distributions of best fit obtained by various authors and their models and types of IDF relationships according to the classification explained above.

Table 3. Compilation of IDF studies and summary of their most relevant characteristics.

References	Stationary or non-stationary curves	Type of series used	Best Fit Distribution	Adjustment method used	Location
Olsson *et al*. (2019)	Stationary	AMS[1]	Generalized Extreme Value.	Neighbourhood	Sweden
Yong *et al*. (2021)	Stationary	AMS	Gumbel	L-Moments	Malaysia
Agilan and Umamahesh (2017a)	Non stationary	AMS	Generalized Extreme Value.	Neighbourhood	Wilmington (USA) Hyderabad (India)
Agilan and Umamahesh (2017b)	Non stationary	AMS[1]	Generalized Extreme Value.	Neighbourhood	Hyderabad (India)
Ganguli and Coulibaly (2017)	Both	AMS	Generalized Extreme Value.	Markov chain (DE-MC) Monte Carlo simulation (Vrban *et al*., 2018)	Ontario (Canada)
Ng *et al*. (2021)	Stationary	AMS[1]	Generalized Extreme Value.	Neighbourhood	Kelantan (Malaysia)
Sane *et al*. (2018)	Stationary	AMS	Generalized Extreme Value.	L-Moments	Senegal
Vrban *et al*. (2018)	Stationary	AMS[1]	Generalized Extreme Value.	L-Moments	Ontario (Canada)

[1]It also carries out a study with PDS with the Generalized Pareto Distribution.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

For this study, three statistical approximations are used to describe the short duration events of the rain according to the theory of extreme values; (i) generalized extreme value (GEV) distribution, (ii) bi-parametric and tri-parametric log-normal distribution (LN2 and LN3, respectively), (iii) Pearson Type III log-normal distribution (LP3).

The cumulative function GEV is expressed as follows:

$$F(x;\ \mu,\sigma,\xi) = \exp\left[-\left(1+\xi * \left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)\right] \tag{11}$$

Where μ is the location parameter, σ is the scale parameter, and ξ is the shape parameter. GEV represents a family of distributions depending on the value of ξ: Gumbel (ξ = 0), Fréchet (ξ > 0) and Weibull (ξ < 0) (Olsson *et al*., 2019). For this study, ξ = 0 will be used.

In general, flood distributions have a positive skewness and are not adequately described by a normal distribution. In many cases, the random variable corresponding to the logarithm of the flood flows will be adequately described by a normal distribution. The parametric log-normal distribution has a probability density function indicated in Equation (12). Frequently, the logarithms of a random variable X do not fit a normal distribution. In such cases, the problem can be solved by introducing a boundary parameter τ before calculating the logarithms, thus obtaining a three-parameter log-normal distribution (OMM, 2011):

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$F(x;\ \mu,\sigma) = \frac{1}{x\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right] \tag{12}$$

$$F(x;\ \mu,\sigma,\xi\ ) = \frac{1}{(x-\tau)\sqrt{2\pi\sigma^2}} exp\left[-\frac{1}{2}\left(\frac{\ln(x-\tau)-\mu}{\sigma}\right)^2\right] \tag{13}$$

The log-Pearson type III distribution (LP3) describes a variable x whose logarithm y = log x presents a Pearson type III distribution. This was recommended for the description of floods in the United States by the Water Resources Council of that country, initially in 1966, and later by the Interagency Advisory Committee on Water Data in 1982. It was also adopted in Canada, among other methods (OMM, 2011):

$$F(x;\ \beta,\xi,\alpha) = \frac{|\beta|\{\beta[\ln(x)-\varphi]\}^{\alpha-1} exp\{-\beta[\ln(x)-\varphi]\}}{x\Gamma(\alpha)} \tag{14}$$

In this opportunity α and β are scale and shape parameters, while φ is a location parameter.

Parameters estimation

Possibly the simplest approach is the method of moments, which allows parameter estimates to be obtained such that the theoretical moments of a distribution agree with the calculated sample moments. The recommended procedure for US federal agencies, references all cited in

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

OMM (2011), is based on the moments of the logarithms of the flood flows $X = \log Q$.

Teegavarapu, Salas and Stedinger (2019) propose the concept of L-moments as a linear combination of probability-weighted moments (PWM). Teegavarapu *et al*. (2019) explains through some previous studies the equation that describes the PWM and the examples of estimators for any probability distribution.

Goodness of fit

Several rigorous and useful statistical tests are available in hydrology to determine whether or not it is reasonable to conclude that a given set of observations has been derived from a particular family of distributions (OMM, 2011). The Kolmogorov-Smirnov test allows to be obtained bounds for each of the observations of a probability plot when the sample has been effectively obtained from the assumed distribution.

This procedure is a non-parametric test that allows testing whether two samples come from the same probabilistic model. Suppose we have two samples of total size $N = m + n$ composed of observations x_1, x_2, x_3,…, x_n and y_1, y_2, y_3,…, y_m. The test assumes that the variables *x, y* are mutually independent and that each x comes from the same continuous population P_1 and that the variables y come from another continuous population P_2. The null hypothesis is that both distributions are identical, that is, they are two samples from the same population. The test is based on calculating the J statistician defined as the maximum value of the

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

absolute difference between two cumulative distribution functions. Among the advantages of the Kolmogórov-Smirnov test is its superiority with respect to the Chi square (X^2) test, its ease of calculation and the fact that it does not use a grouping of data, in addition to the fact that the statistic is independent of the expected frequency distribution, it just depends on the sample size.

Intensity-frequency-duration models

The results of frequency analysis are usually expressed in terms of intensity-duration-frequency relationships at a given location, or in the form of precipitation frequency atlases, in which the accumulated heights of rainfall for different durations and return periods in the region of interest (OMM, 2011).

Based on the parameterization proposed by Sherman (1931), cited in Gutiérrez and Barragán (2019), the mathematical and graphical representation of the calculation of intensity (I)-duration (D)-frequency (F) curves is adopted throughout the world. This formulation is a rational equation of the type:

$$i_d^T = f(x) = \frac{P(T)}{Q(d)} \quad (15)$$

The numerator *P*(*T*) is a function of the return period (*T*) and indicates the cumulative frequency quantile (1-1/*T*) of a probability

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

distribution function of a random variable. For the denominator *Q*(*d*), which is also a function of the time or duration of the intensity, it is admitted that it can be expressed as a polynomial that allows factorization (Gutiérrez & Barragán, 2019). Leaving the following general parameterization:

$$i_d^T = \frac{kT^m}{(d^\theta + C)^n} \tag{16}$$

Where i_d^T is the maximum intensity of precipitation, expressed in mm/min or mm/h; *T* is the return period in years; *d* is the duration of precipitation in minutes and *k*, *m*, θ and *n* are the adjustment parameters to estimate Equation (16) is widely used and several authors have proposed different values of the parameters *k*, *m*, θ and *n*. In all cases, these parameters are estimated by numerical, analytical, linear numerical, non-linear, statistical and optimization procedures. Optimal values of can be calculated through a trial and error procedure as quoted by Gutiérrez and Barragán (2019). However, in no case is there evidence that any of these parameters have physical significance. That is, it has not been shown that the parameters of Equation (16) are related to any physiographic or climatological characteristic of the environment.

Equation (16) shows several adjustment models where the values of k, m, θ and n are simplified or reduced to 0. In this contribution the models to be adjusted will be Montana, Sherman, Bernard and Chow cited in Gutiérrez and Barragán (2019), from equations (17) to (20) respectively:

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$i_d^T = \frac{kT^m}{d^\theta + C} \tag{17}$$

$$i_d^T = \frac{kT^m}{(d + C)^n} \tag{18}$$

$$i_d^T = \frac{kT^m}{d^n} \tag{19}$$

$$i_d^T = \frac{kT^m}{d + C} \tag{20}$$

Results and discussion

As explained in previous sections, in the years 1990 and 2009 it was not possible to obtain the values of maximum annual rainfall intensities due to different factors; To make up for this lack of data, the Maximization-Expectation (Miró *et al*., 2017) and Multiple Imputation (Mallol, 2017) methodologies were used with five variants. Of these five data groups, the result with the highest number was chosen to complete the registry worth. Table 4 shows a summary of the results obtained and subsequently analyzes them.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table 4. Results of the imputation of values to the missing data of 1990 and 2009 for durations of 1, 2, 4, 12 h.

Hour	Imputed value 1990 mm/min	Imputed value 2009 mm/min	Mean mm/min		Standard deviation mm/min		Kurtosis		Skewness	
			S/I	C/I	S/I	C/I	S/I	C/I	S/I	C/I
1	1.008	0.856	0.99	0.99	0.23	0.22	4.66	4.91	1.48	1.45
2	0.572	0.488	0.55	0.55	0.15	0.14	3.45	3.58	1.29	1.24
4	0.287	0.298	0.29	0.29	0.08	0.08	2.36	2.26	1.46	1.38
12	0.118	0.107	0.12	0.12	0.03	0.03	0.98	0.83	0.37	0.36

C/I: With imputation; S/I: Without imputation.

The descriptive statistics of the sample before and after the imputation are summarized, as can be seen there is no significantly relevant difference between the results for the selected durations, thus fulfilling the conditions proposed above.

Data series processing

The application of the US-WRC method demonstrates the obtaining of four anomalous data, three of them occurred on August 2, 2014 with a convective rainy event of 124.6 mm in 120 minutes, the other value was presented on September 24, 1998 with the passage of Hurricane George which left a record of 274 mm in 24 hours. Figure 2 shows these results.

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

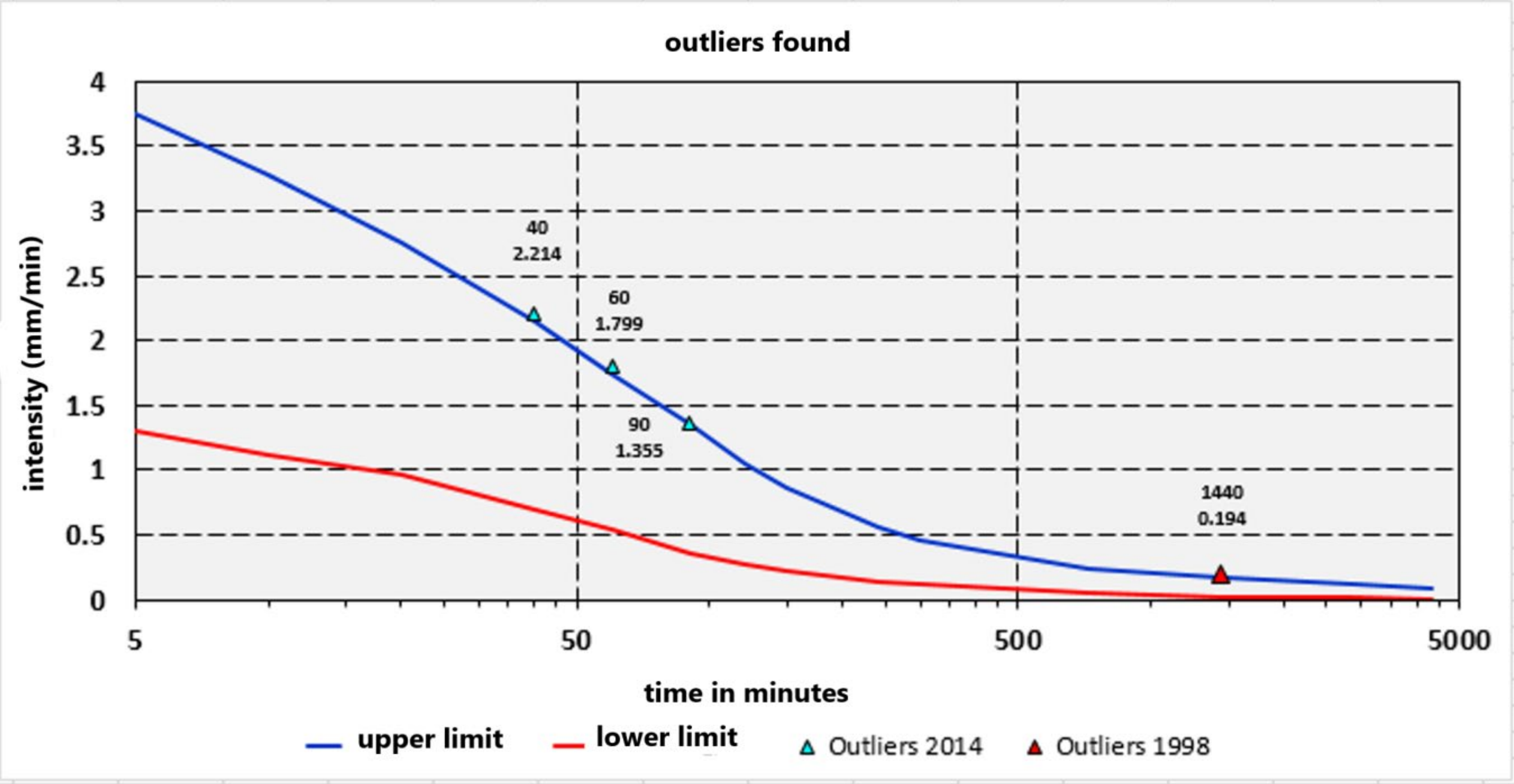

Figure 2. Anomalous data from the data series.

After review by specialists and considering criteria of experiences with other stations, the anomalous values presented by the series are accepted, since they do not exceed the upper limits of the US-WRC model by 10 % for a confidence level of 95 %.

The results of the quality tests applied to the series of annual maximums recommended in OMM (2011), Gusts, Mann-Whitney (M-W), Wald-Wolfowitz (W-W) and Mann-Kendall (M-K) are summarized in Table 5.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

The word Yes in Table 5 means that the null hypothesis is accepted that:

1. The series is random for a significance of 5 % (Runs Test).
2. The series is independent for a significance of 5 % (Tests M-W, W-W).
3. The series is seasonal for a significance of 5 % (M-K test).

Table 5. Quality tests results for all data series.

Serie	Gusts	M-W	W-W	M-K
5 min	Yes	Yes	Yes	Yes
10 min	Yes	Yes	Yes	Yes
20 min	Yes	Yes	Yes	Yes
40 min	Yes	Yes	Yes	Yes
60 min	Yes	Yes	Yes	Yes
90 min	Yes	Yes	Yes	Yes
120 min	Yes	Yes	Yes	Yes
150 min	Yes	Yes	Yes	Yes
240 min	Yes	Yes	Yes	Yes
300 min	Yes	Yes	Yes	Yes
720 min	Yes	Yes	Yes	Yes
1 440 min	Yes	Yes	Yes	Yes
2 880 min	Yes	Yes	Yes	Yes
4 320 min	Yes	Yes	Yes	Yes

When verifying that all the series are seasonal, the trends of said series are investigated, Figure 3 shows the result of the Mann-Kendall and

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

Sen's Slope tests for the 40-minute SMA, in which it can be seen that, although there is a tendency to increase by the linear estimator, it is still insufficient to consider it within the analysis.

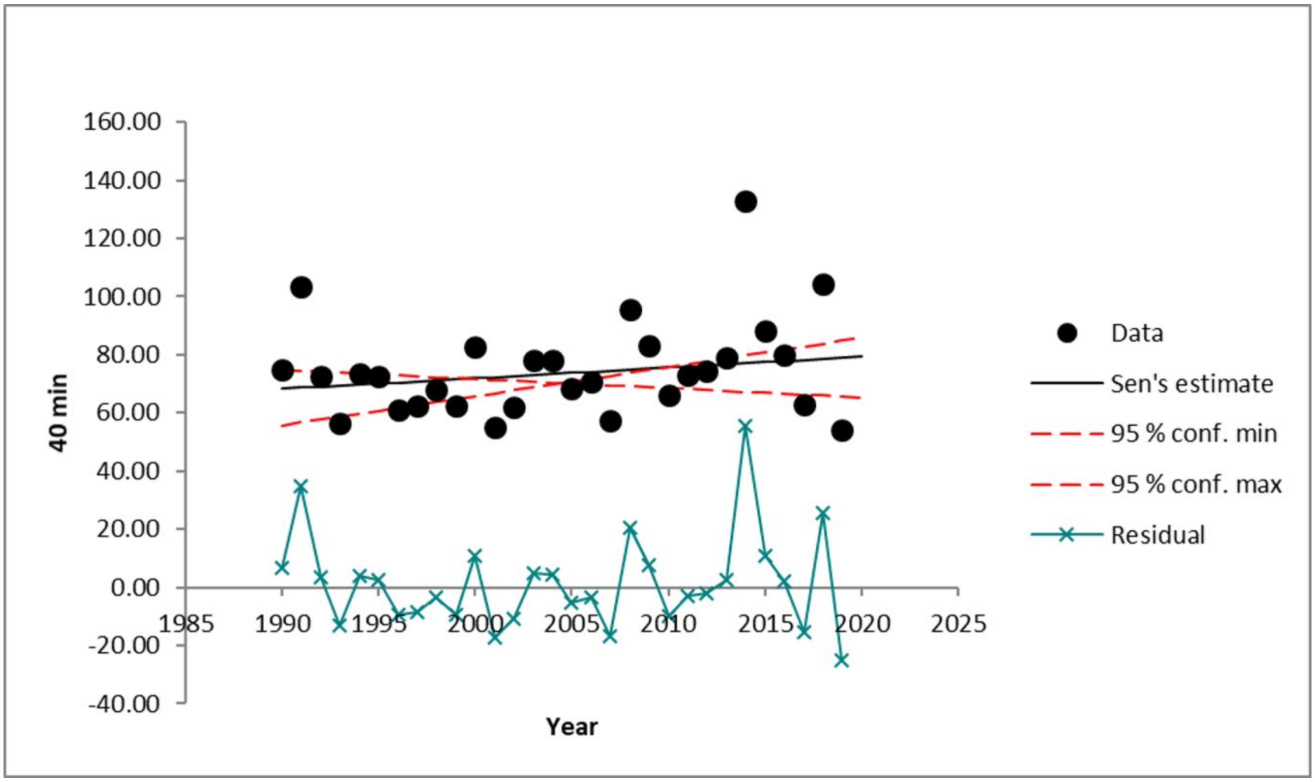

Figure 3. Trend analysis for the 40-minute series whose slope was one of the highest of the series studied.

The results obtained in the quality tests show that the data series for the elaboration of the IDF curves of the Yabú Meteorological Station are suitable for probabilistic processing, highlighting that stationary models can be used for their representation without the need to resort to models not stationary.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

The summary of the analysis with the Gumbel distribution (ξ = 0) with an adjustment for *L*-moments is shown in figure 4 for 1, 2, 4, 12 h and Table 6 shows the results of the position and scale parameters obtained.

Table 6. Parameters of the Gumbel probability distribution obtained for the series of 1, 2, 4, and 12 h.

Serie	**Location Parameter μ**	**Standard error**	**Scale Parameter σ**	**Standard error**
1 hour	1.11481	0.05835	0.29973	0.03425
2 hours	0.63042	0.03536	0.18169	0.02117
4 hours	0.33243	0.01902	0.09768	0.01192
12 hours	0.13387	0.00595	0.03069	0.00415

Figure 4 plots the fit for the aforementioned durations using the cumulative probability form.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

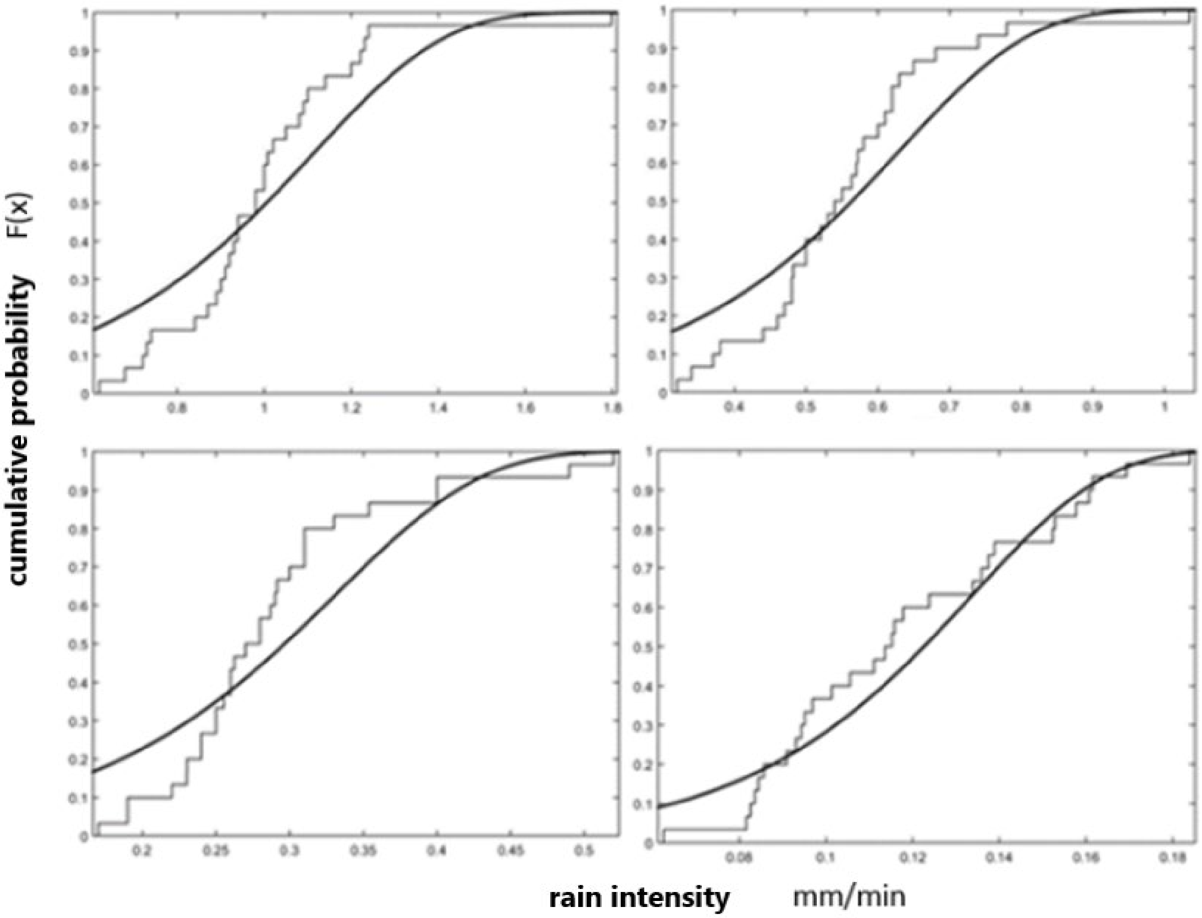

Figure 4. Adjustment to the Gumbel cumulative probability function.

To obtain the probability function of best fit, the Kolmorogov-Smirnov goodness-of-fit test is performed for the results of the functions; (a) distribution of extreme values ($\xi = 0$); (b) bi and tri parametric logarithmic normal distribution, which for a significance level of 5 %, showed that the Gumbel adjustment is the most effective.

The application of the Montana, Sherman, Bernard and Chow models for the adjustment of the results of the Gumbel probability

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

function, in the series of annual maximums is summarized in Table 7 where the values of *k*, *m*, θ, *C* and *n* for each model, in addition to using the Pearson correlation coefficient to find which of them is the one with the best fit.

Table 7. Parameters obtained and results of the Pearson correlation for the applied models.

Model	***k***	***m***	**θ**	***C***	***n***	**Pearson**
Montana	120.304	0.131	1.065	54.101	-	0.999194
Sherman	176.351	0.131	-	48.229	1.124	0.999192
Bernard	4.500	0.131	-	-	0.424	0.975301
Chow	89.042	0.131	-	38.757	-	0.999010

The analysis carried out shows that the Montana model is the one that best fits the results of the Gumbel probability distribution, so it is chosen for its analysis. Figure 5 illustrates the Montana fit function for the Yabú Weather Station.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

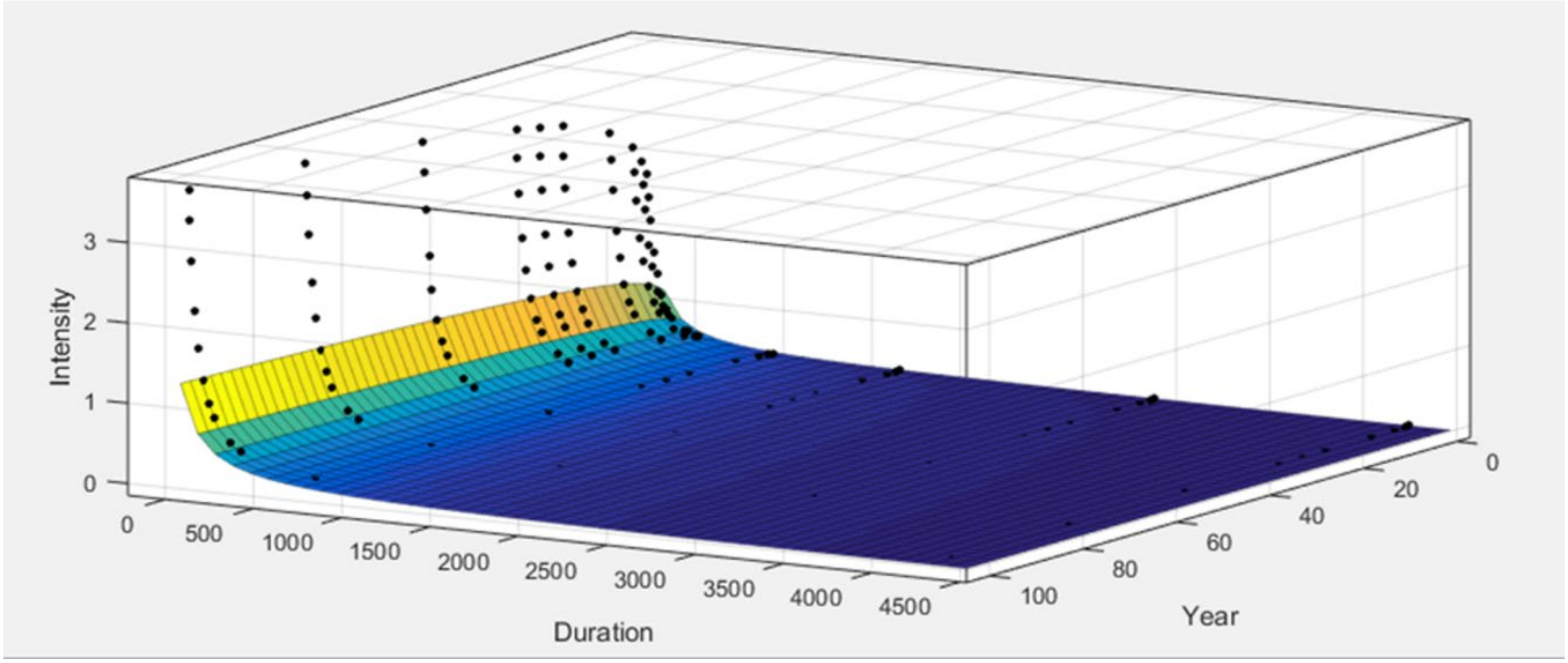

Figure 5. Montana model adjusted to the data obtained by the Gumbel probabilistic model.

To delve deeper into the behavior of the model, it is necessary to convert the IDF curves obtained with Montana into Precipitation-Duration-Frequency curves (PDF) with which the residual of the model can be obtained and the range of validity of the model can be clearly appreciated. same. Figure 6 shows a graph where the residual sheet in mm of precipitation is plotted against the duration in minutes of precipitation.

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

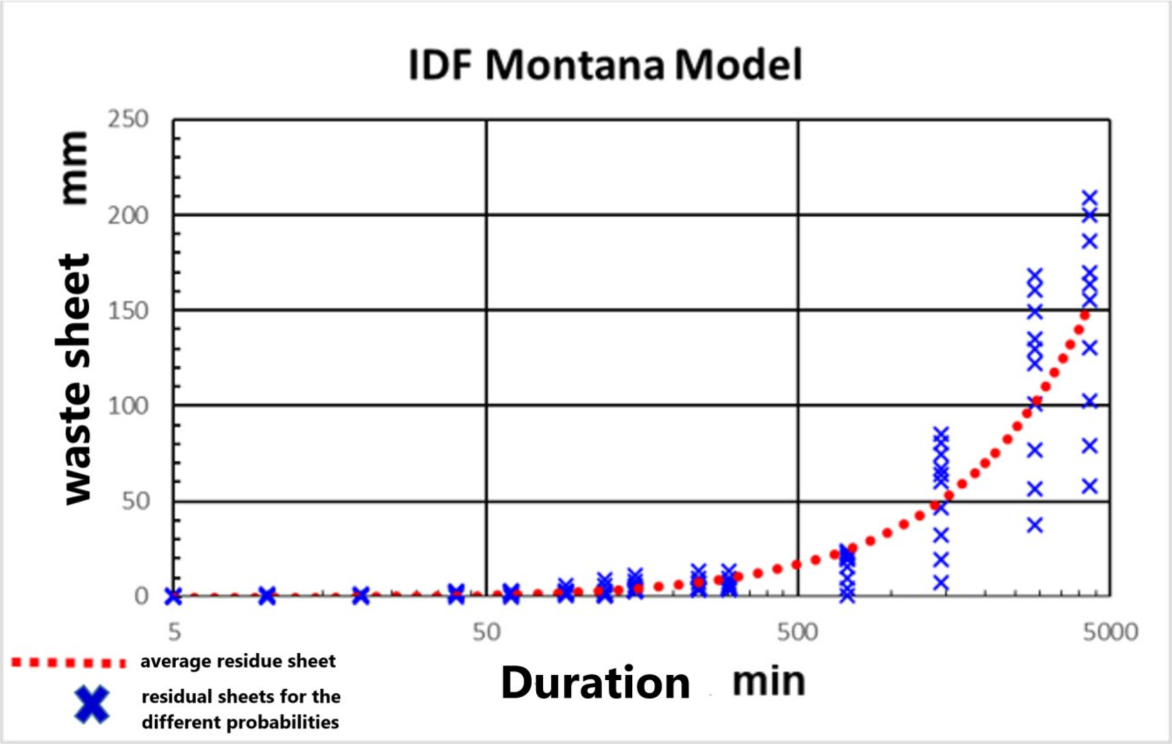

Figure 6. Residue sheet found for the different rainfall durations using the parameters in Table 7 in the Montana model.

The results obtained in figure 6 show that the Montana model presents a linear residual behaviour with an upward slope that is more clearly evident from 500 mm for all probabilities. This observation allows us to infer that there may be an IDF Montana model that describes the intensities for durations less than 360 min and another for durations greater than that value. Table 8 shows these new adjustments and Equation (21) the final result for the analysis station.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table 8. Parameters and correlation found for the Montana model with durations less than 360 min and greater than that value.

Model	***k***	***m***	**θ**	***C***	***n***	**Pearson**
Montana (-6 hours**)**	135.5	0.1309	1.092	61.58	-	0.9981
Montana (+6 hours**)**	4.348	0.1826	0.5816	-2.905	-	0.9911

$$I=\begin{Bmatrix} \frac{135.5T^{0.1309}}{(D^{1.092}+61.58)}\ para\ toda\ D\leq 360 \\ \frac{4.348T^{0.1826}}{(D^{0.5816}-2.905)}\ para\ toda\ D>360 \end{Bmatrix} \quad (21)$$

Where:

I: Intensity of average precipitation in mm/min.

T: Return period in years.

D: Duration of the storm in minutes

The values obtained with Equation (21) never exceed 20 % of the value obtained from the Gumbel probability function and the largest discrepancies (10 to 20 %) are the result of the lowest return periods (2 and 3 years) which It is also favorable from the point of view of engineering design, as it has a safety percentage for its use. Table 9 shows the values that are between 10 and 20 % higher than the value obtained in the probability function.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table 9. Comparison between the values obtained with Eq.19 and the values of the Gumbel probability distribution for the most unfavorable results.

Gumbel probability function (mm/min)	Return period (years)	Duration (minutes)	Montana model result (mm/min)
0.031	2	4 320	0.039
0.040	2	2 880	0.049
0.061	2	1 440	0.075
0.276	2	240	0.323
0.427	2	150	0.495
0.226	2	300	0.261
0.037	3	4 320	0.042
0.531	2	120	0.598
0.073	3	1 440	0.081

Conclusions

After the study and analysis of the results obtained during this investigation, the authors arrived at the following conclusions:

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

- Data from 30 years of rainfall records from the Yabú Meteorological Station were processed, obtaining the series of annual maximums for durations of 5, 10, 20, 40, 60, 90, 120, 150, 240, 300, 720, 1 440, 2 880 and 4 320 minutes, the last 3 correspond to 24, 48 and 72 hours of duration, typical of cyclonic events.
- There were problems in the measurements of the years 1990 and 2009; however, the multiple imputation algorithm was applied through linear regression and the missing data was obtained without showing significant changes in the measures of central tendency of the data series.
- Four values considered outliers by the US-WRC method were found, which were accepted as they come from well-documented real rainfall events, one of them a hurricane.
- The data series was analyzed to verify its randomness, independence and trend, with which the hypotheses were confirmed and it was concluded that an IDF stationary model would clearly represent the phenomenon.
- The Gumbel distribution was matched to the data series and it was verified using the Kolmogorov-Smirnov method that the distribution adjusts to them, also being the one with the best results compared to the logarithmic distributions tested.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

- The Montana model parameterized the data obtained from the probabilistic distribution with greater correlation; however, it was found that for durations greater than 500 minutes, there were significant residuals that led to the definition of truncating the series and elaborating a new adjustment.
- The proposed equations for durations less than 360 min and greater than that figure were represented by a Montana model of two equations with different parameters. It is to be expected that this situation will be repeated for the other meteorological stations in the province analysis of this contribution for its implementation.

Acknowledgments

The authors wish to convey their deepest gratitude to Villa Clara Provincial Meteorological Center, whose support in the development of the investigations has been commendable and noteworthy despite the harsh conditions imposed by the COVID-19 pandemic affecting the country. To the Hydraulic Engineering students of Las Villas Central University: José Solis Quintana, Alejandro Marrero, José Manuel Alba Bacallao and Roberto Fernández.

References

Agilan, V., & Umamahesh, N. V. (2017a). Modelling nonlinear trend for developing non-stationary rainfall intensity–duration–frequency curve. *International Journal of Climatology*, 37(3), 1265-1281. Recovered from https://doi.org/https://doi.org/10.1002/joc.4774

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Agilan, V., & Umamahesh, N. V. (2017b). Non-stationary rainfall intensity-duration-frequency relationship: A comparison between annual maximum and partial duration series. *Water Resources Management*, 31(6), 1825-1841. Recovered from https://doi.org/10.1007/s11269-017-1614-9

Agilan, V., & Umamahesh, N. V. (2017c). What are the best covariates for developing non-stationary rainfall Intensity-Duration-Frequency relationship? *Advances in Water Resources*, 101, 11-22. Recovered from https://doi.org/https://doi.org/10.1016/j.advwatres.2016.12.016

Ben-Zvi, A. (2009). Rainfall intensity–duration–frequency relationships derived from large partial duration series. *Journal of Hydrology*, 367(1), 104-114. Recovered from https://doi.org/https://doi.org/10.1016/j.jhydrol.2009.01.007

Chang, K. B., Lai, S. H., & Faridah, O. (2013). RainIDF: Automated derivation of rainfall intensity–duration–frequency relationship from annual maxima and partial duration series. *Journal of Hydroinformatics*, 15(4), 1224-1233. Recovered from https://doi.org/10.2166/hydro.2013.192

Egea-Pérez, R., Cortés-Molina, M., & Navarro-González, F. J. (2021). Analysis of rainfall time series with application to calculation of return periods. *Sustainability*, 13(14). Recovered from https://doi.org/10.3390/su13148051

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Emmanouil, S., Langousis, A., Nikolopoulos, E. I., & Anagnostou, E. N. (2020). Quantitative assessment of annual maxima, peaks-over-threshold and multifractal parametric approaches in estimating intensity-duration-frequency curves from short rainfall records. *Journal of Hydrology*, 589, 125151. Recovered from https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125151

Ganguli, P., & Coulibaly, P. (2017). Does non-stationarity in rainfall require nonstationary intensity–duration–frequency curves? *Hydrology and Earth System Sciences*, 21(12), 6461-6483. Recovered from https://doi.org/10.5194/hess-21-6461-2017

Gregersen, I. B., Madsen, H., Rosbjerg, D., & Arnbjerg-Nielsen, K. (2017). A regional and nonstationary model for partial duration series of extreme rainfall. *Water Resources Research*, 53(4), 2659-2678. Recovered from https://doi.org/https://doi.org/10.1002/2016WR019554

Gutiérrez, A., & Barragán, R. (2019). Ajuste de curvas IDF a partir de tormentas de corta duración. *Tecnologías y ciencias del agua*, 10, 1-24. Recovered from https://doi.org/10.24850/j-tyca-2019-06-01

Little, R. J., & Rubin, D. B. (1989). El análisis de datos de ciencias sociales con valores faltantes. *Métodos e Investigación Sociológicos*, 18(2-3), 292-326.

Maity, R. (2018). *Statistical methods in hydrology and hydroclimatology*. Berlin, Germany: Springer. Recovered from https://doi.org/https://doi.org/10.1007/978-981-10-8779-0

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Mallol, P. (2017). *Importancia del tratamiento de datos perdidos. Aplicación en estudios longitudinales pequeños*. Barcelona, España: Universitat Oberta de Catalunya. Recovered from http://openaccess.uoc.edu/webapps/o2/bitstream/10609/64105/6/pmallolrTFM0617memoria.pdf

Masseran, N., & Safari, M. A. M. (2020). Risk assessment of extreme air pollution based on partial duration series: IDF approach. *Stochastic Environmental Research and Risk Assessment*, 34(3), 545-559. Recovered from https://doi.org/10.1007/s00477-020-01784-2

Miró, J., Caselles, V., & Estrela, M. (2017). Multiple imputation of rainfall missing data in the Iberian Mediterranean context. *Atmospheric Research*, 197, 313-330. Recovered from https://ui.adsabs.harvard.edu/abs/2017AtmRe.197..313M/abstract

Molenberghs, G., Fitzmaurice, G., Kenward, M. G., Tsiatis, A., & Verbeke, G. (2015). *Handbook of missing data methodology*. London, UK: Chapman & Hall/CRC.

Naghettini, M. (2017). *Fundamentals of statistical hydrology*. Berlin, Germany: Springer. Recovered from https://doi.org/DOI 10.1007/978-3-319-43561-9

Ng, J. L., Tiang, S. K., Huang, Y. F., Noh, N. I. F. M., & Al-Mansob, R. A. (2021). Analysis of annual maximum and partial duration rainfall series. *IOP Conference Series: Earth and Environmental Science*, 646(1), 012039. Recovered from https://doi.org/10.1088/1755-1315/646/1/012039

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Noor, M., Ismail, T., Chung, E.-S., Shahid, S., & Sung, J. H. (2018). Uncertainty in rainfall intensity duration frequency curves of peninsular Malaysia under changing climate scenarios. *Water*, 10(12). Recovered from https://doi.org/10.3390/w10121750

Olsson, J., Södling, J., Berg, P., Wern, L., & Eronn, A. (2019). Short-duration rainfall extremes in Sweden: A regional analysis. *Hydrology Research*, 50(3), 945-960. Recovered from https://doi.org/10.2166/nh.2019.073

OMM, Organización Meteorológica Mundial. (2011). *Guía de prácticas hidrológicas: gestión de recursos hídricos y aplicación de prácticas hidrológicas* (6a. ed.). Vol. II. Ginebra, Suiza: Organización Meteorológica Mundial.

Sane, Y., Panthou, G., Bodian, A., Vischel, T., Lebel, T., Dacosta, H., Quantin, G., Wilcox, C., Ndiaye, O., Diongue-Niang, A., & Diop Kane, M. (2018). Intensity–duration–frequency (IDF) rainfall curves in Senegal. *Natural Hazards and Earth System Sciences*, 18(7), 1849-1866. Recovered from https://doi.org/10.5194/nhess-18-1849-2018

Singh, V. (2017). *Handbook of applied hydrology* (2nd ed. to replace the classic 1963 edition edited by Ven Te Chow). New York, USA: McGraw-Hill Education.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Soumya, R., Anjitha, U. G., Mohan, S., Adarsh, S., & Gopakumar, R. (2020). Incorporation of non-stationarity in precipitation intensity-duration-frequency curves for Kerala, India. *IOP Conference Series: Earth and Environmental Science*, 491, 012013. Recovered from https://doi.org/10.1088/1755-1315/491/1/012013

Teegavarapu, R., Salas, J., & Stedinger, J. (2019). *Statistical analysis of hydrologic variables. Methods and Applications*. Reston, USA: American Society of Civil Engineers. Recovered from https://doi.org/10.1061/9780784415177

Van Campenhout, J., Houbrechts, G., Peeters, A., & Petit, F. (2020). Return period of characteristic discharges from the comparison between partial duration and annual series, application to the Walloon Rivers (Belgium). *Water*, 12(3). Recovered from https://doi.org/10.3390/w12030792

Vrban, S., Wang, Y., McBean Edward, A., Binns, A., & Gharabaghi, B. (2018). Evaluation of stormwater infrastructure design storms developed using partial duration and annual maximum series models. *Journal of Hydrologic Engineering*, 23(12), 04018051. Recovered from https://doi.org/10.1061/(ASCE)HE.1943-5584.0001712

Yilmaz, A., & Perera, B. (2014). Extreme rainfall non-stationarity investigation and intensity–frequency–duration relationship. *Journal of Hydrologic Engineering*, 19, 1160-1172. Recovered from https://doi.org/10.1061/(ASCE)HE.1943-5584.0000878

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Yong, S. L. S., Ng, J. L., Huang, Y. F., & Ang, C. K. (2021). Assessment of the best probability distribution method in rainfall frequency analysis for a tropical region. *Malaysian Journal of Civil Engineering*, 33(1). Recovered from https://doi.org/10.11113/mjce.v33.16253

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)