

# Developing Intensity-Duration-Frequency Curves in Scarce Data Region: An Approach using Regional Analysis and Satellite Data

Ayman G. Awadallah<sup>1</sup>, Mohamed ElGamal<sup>2</sup>, Ashraf ElMostafa<sup>3</sup>, Hesham ElBadry<sup>3</sup>

<sup>1</sup>Civil Department, Fayoum University, Al Fayoum, Egypt

<sup>2</sup>Civil Department, Cairo University, Giza, Egypt,

<sup>3</sup>Civil Department, Ain Shams University, Cairo, Egypt

E-mail: [aawadallah@darcairo.com](mailto:aawadallah@darcairo.com)

Received December 21, 2010; revised January 11, 2011; accepted February 9, 2011

## Abstract

The availability of data is an important aspect in frequency analysis. This paper explores the joint use of limited data from ground rainfall stations and TRMM data to develop Intensity Duration Frequency (IDF) curves, where very limited ground station rainfall records are available. Homogeneity of the means and variances are first checked for both types of data. The study zone is assumed to be belonging to the same region and checked using the Wiltshire test. An Index Flood procedure is adopted to generate the theoretical regional distribution equation. Rainfall depths at various return periods are calculated for all stations and plotted spatially. Regional patterns are identified and discussed. TRMM data are used to develop ratios between 24-hr rainfall depth and shorter duration depths. The regional patterns along with the developed ratios are used to develop regional IDF curves. The methodology is applied on a region in the North-West of Angola.

**Keywords:** Rainfall Frequency Analysis, Regional Analysis, Intensity Duration Frequency, TRMM, Angola

## 1. Introduction and Methodology

The availability of data is an important aspect in frequency analysis. The estimation of probability of occurrence of extreme rainfall is an extrapolation based on limited data. Thus the larger the database, the more accurate the estimates will be. From a statistical point of view, estimation from small samples may give unreasonable or physically unrealistic parameter estimates, especially for distributions with a large number of parameters (three or more). Large variations associated with small sample sizes cause the estimates to be unrealistic. In practice, however, data may be limited or in some cases may not be available for a site. In such cases, regional analysis is most useful.

The main objective of this paper is to present the methodology and results aiming at developing intensity-duration-frequency (IDF) curves in a region where ground rainfall stations data is scarce. To complement the data, Tropical Rainfall Measuring Mission (TRMM) corrected satellite data was used. TRMM is a joint U.S.-Japan sat-

ellite mission to monitor tropical and subtropical (40° S–40° N) precipitation. TRMM satellite data are available from 1998 to 2008. Several studies have compared the TRMM data with ground station data ([1-4]); however, the aim of this study is to explore the joint use of TRMM data with ground station data to produce intensity-duration-frequency relations.

The first step is to assess if the TRMM annual maximum daily data have different average (compared to ground station data) using the Mann-Whitney U, the Moses extreme reactions, the Kolmogorov-Smirnov Z, the Wald-Wolfowitz runs tests. Furthermore, the Levene test was performed to check the equality of variance between the two data types.

Once the possibility of the joint use of maximum daily rainfall from ground stations and TRMM data is checked, the combined maximum daily records at locations of interest are verified if they belong to the same region via the Wiltshire test and the ordinary moment diagram. An Index Flood method is then applied to get the regional estimates of the maximum daily rainfall at different re-

turn periods.

The regional estimates are first compared with the at-site estimates, and the ones that better fit the raw data are selected. Subsequently, the adopted estimates for all locations are compared in order to establish geographically coherent regions. Since the main purpose is to develop intensity-duration-frequency curves, the lower (less conservative) values that might appear in certain locations and that are not consistent with the geographically coherent region are discarded.

Following the establishment of the geographically coherent regional average estimates of the daily rainfall at different return periods, the ratios between intensities of the 24-hr and those of the 12-, 6-, 3-, 2-, 1-Hr, 30-, 15-, and 5-min based on Bell [5] and SCS type II dimensionless rainfall curve [6] are used to derive the short duration rainfall values of the IDF. As such, regional robust IDFs are developed for scarce data regions.

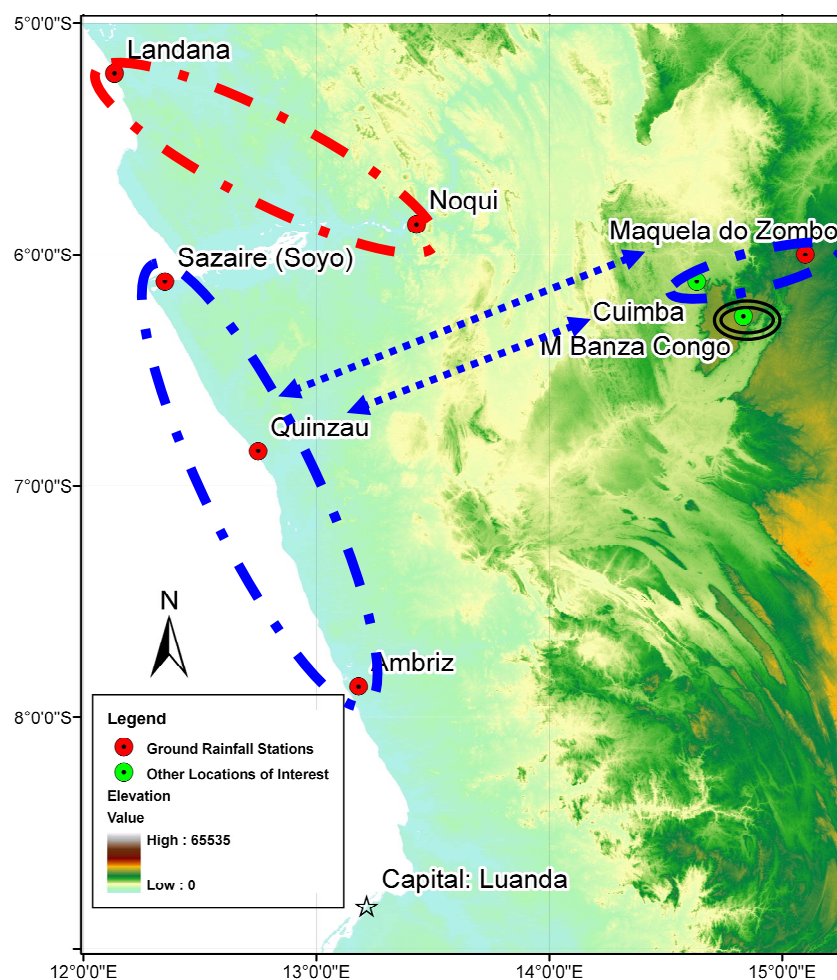
The application of the methodology is in a region of North-West Angola where IDF curves are needed for eight cities; namely, Sazaire (Soyo), Ambriz, Quinzau,

Landana, Noqui, Maquela do Zombo, M Banza Congo, and Cuimba, shown on **Figure 1**. Maximum annual rainfall data coming from ground rainfall stations for the eight cities are only available for 8 years at most.

The paper is organized as follows. After the current section presenting the paper objective, its methodology and region of application, the rainfall data available from the ground stations and TRMM are described. The following section presents the results of homogeneity tests verifying the possibility of the joint use of ground data and TRMM. The regionalization methodology is detailed in the fourth section illustrating the Wiltshire test and the index flood method. The results of the regionalization are illustrated in the subsequent section followed by the assumptions of the IDF development. Finally, conclusions are offered in the last section.

## 2. Rainfall Data Available for Ground Stations and TRMM

In the procedure of developing the Intensity Duration



**Figure 1.** Rainfall stations that will be used during the analysis.

Frequency (IDF) curves for the North-West of Angola, the data was collected from the ground rain gauging stations. The following **Table 1** shows the names and coordinates of ground rainfall stations used in this study. **Figure 1** shows these stations along with other locations of interest for which an IDF is required, on the Shuttle Radar Topography Mission (SRTM) elevation data (90 m resolution at ground) in the background. SRTM data is obtained on a near-global scale to generate the most complete high-resolution digital topographic database of Earth. It consisted of a modified radar system that flew onboard the Space Shuttle Endeavour during an 11-day mission in February of 2000.

Unfortunately, recent ground stations records were not available. Old data were retrieved from the National Oceanic and Atmospheric Agency in the USA, which keeps in its data rescue website, scanned reports of “Elementos Meteorológicos e Climatológicos” of the “Serviços de Marinha, Repartição Técnica de Estatística Geral”. The reports available on the NOAA website cover the period from 1943 to 1952. However, some stations were not functioning properly even during this short period. **Table 2** shows the maximum daily rainfall depth recorded for each ground station in millimeter.

As limited data are available, Tropical Rainfall Measuring Mission (TRMM) data was used. TRMM data give rainfall depths every 3 hours and are downloadable from <http://disc2.nascom.nasa.gov/Giovanni/tovas/site>. **Table 3** shows the maximum daily rainfall depth recorded for each location of interest in millimeter.

### 3. Homogeneity Check

The frequency analysis of rainfall data records is affected by the number of records for each station; therefore carrying the analysis for all the records available shall give higher confidence to the results.

However, the TRMM data should be tested first if they can be combined to the ground data in one set. Several tests are available to test the homogeneity of the mean and the variance. This is described in the below sub-sections.

#### 3.1. Homogeneity of the Mean

To test the homogeneity of the mean, several non parametric tests [7] are available, among them the Mann-Whitney U [8] and Wilcoxon W [9] tests which are equivalent. Non-parametric tests are preferred as the data dealt with are known a-priori to be non-normal. In the following we present the procedure of the Mann-Whitney test as a representative of the applied test and we present the results of both tests. Both of them confirm that there

is no statistical evidence- to a 5% level of significance — that there is a significant difference in the mean of the two samples (**Table 4**).

In this test, two samples of size  $p$  and  $q$  are compared. The combined dataset of size  $N = p + q$  is ranked in increasing order. The Mann-Whitney (M-W) test considers the quantities  $V$  and  $W$  in Equations (1) and (2)

$$V = R - \frac{p(p+1)}{2} \quad (1)$$

$$W = pq - V \quad (2)$$

$R$  is the sum of the ranks of the elements of the first sample (size  $p$ ) in the combined series (size  $N$ ), and  $V$  and  $W$  are calculated from  $R$ ,  $p$ , and  $q$ .  $V$  represents the number of times an item in sample 1 follows an item in sample 2 in the ranking. Similarly,  $W$  can be computed for sample 2 following sample 1. The M-W statistic  $U$  is defined by the smaller of  $V$  and  $W$ .

Other non-parametric tests such as the Moses extreme reactions [10], the Kolmogorov-Smirnov  $Z$ , the Wald-Wolfowitz runs tests [11] are reported in **Tables 5 to 7**. They all confirm that there is no statistical evidence-to a 5% level of significance-that there is a significant difference in the mean of the two samples. For details about these tests, the reader is referred to [7].

#### 3.2. Homogeneity of the Variance

Levene's test [12] is an inferential statistic used to assess the equality of variance in different samples. Levene's test assesses that variances of the populations, from which different samples are drawn, are equal. It tests the null hypothesis that the population variances are equal. If the resulting p-value of Levene's test is less than some critical value (typically .05), the obtained differences in sample variances are unlikely to have occurred based on random sampling. Thus, the null hypothesis of equal variances is rejected and it is concluded that there is a difference between the variances in the population. One advantage of Levene's test is that it does not require

**Table 1. Ground rainfall gauges coordinates.**

Station Name	Longitude (South)	Latitude (East)
Sazaire (Soyo).	6.1167	12.3500
Ambriz	7.8667	13.0833
Quinzau	6.8500	12.7500
Landana	5.2167	12.1333
Noqui	5.8667	13.4333
Maquela do Zombo	6.0500	15.1000

**Table 2. Maximum daily rainfall depths based on ground rainfall gauges.**

Station Name	Year	Sazaire (Soyo)	Ambriz	Quinzau	Landana	Noqui	Maquela do Zombo
Ground Data	1943	106.40	56.80		100.00	71.30	50.80
	1944	49.00	66.70	90.00	114.00	139.10	110.20
	1945	66.00	72.30	54.00	85.00	127.40	76.00
	1946	56.70	73.00	46.10	133.00	80.60	65.00
	1947	164.00	106.00	158.50	68.20	65.00	76.00
	1948	71.00		123.50	82.30	94.50	65.00
	1949	130.00		215.00		115.00	
	1950	110.00				85.00	

**Table 3. TRMM maximum daily rainfall depth recorded in years 1998 to 2008 for locations of interest.**

Station Name	Sazaire (Soyo)	Ambriz	Quinzau	Landana	Noqui	Maquela do Zombo	M banza Congo	Cuimba
TRMM Data	119.82	50.00	81.30	124.45	113.40	68.82	54.59	60.05
	129.19	74.63	112.62	200.91	61.55	64.10	84.32	84.14
	88.09	46.60	51.20	107.37	62.74	50.82	59.85	67.70
	66.08	151.38	53.80	84.40	62.47	69.74	68.37	56.60
	120.90	90.72	60.14	104.64	82.09	173.44	94.85	112.58
	102.97	63.72	63.50	146.09	81.20	56.89	152.69	62.14
	55.17	79.08	85.84	138.25	81.96	169.69	93.87	84.00
	75.14	49.94	56.81	151.94	69.90	109.05	108.21	83.95
	62.55	46.00	57.84	109.49	101.09	73.86	86.17	89.13
	69.42	59.08	71.33	74.63	276.04	59.11	119.40	131.18
	51.15	72.15	48.72	62.94	49.00	61.35	85.00	53.72

**Table 4. Mann-Whitney U and Wilcoxon W Statistics.**

	Sazaire (Soyo)	Ambriz	Quinzau	Landana	Noqui	Maquela do Zombo
Mann-Whitney U	40.000	21.000	20.000	22.000	28.000	31.000
Wilcoxon W	106.000	87.000	86.000	43.000	94.000	97.000
Z	-.330	-.736	-1.307	-1.106	-1.321	-.201
Asymp. Sig. (2-tailed)	.741	.462	.191	.269	.186	.840
Exact Sig. [2*(1-tailed Sig.)]	.778(a)	.510(a)	.216(a)	.301(a)	.206(a)	.884(a)
Exact Sig. (2-tailed)	.778	.510	.216	.301	.206	.864
Exact Sig. (1-tailed)	.389	.255	.108	.151	.103	.432
Point Probability	.031	.034	.018	.022	.014	.019

**Table 5. Moses test statistics.**

	Sazaire (Soyo)	Ambriz	Quinzau	Landana	Noqui	Maquela do Zombo	Pango Aluquem
Observed Control Group Span	19	11	17	12	14	15	18
Sig. (1-tailed)	1.000	.346	1.000	.261	.153	.728	1.000
Trimmed Control Group Span	15	4	12	8	11	6	11
Sig. (1-tailed)	.929	.214	.928	.484	.395	.205	.648
Outliers Trimmed from each End	1	1	1	1	1	1	1

**Table 6. Kolmogorov-Smirnov test statistics.**

		Sazaire (Soyo)	Ambriz	Quinzau	Landana	Noqui	Maquela do Zombo	Pango Aluquem
Most Extreme Differences	Absolute	.250	.364	.576	.394	.364	.288	.260
	Positive	.250	.364	.576	.091	.364	.288	.260
	Negative	-.125	-.164	-.167	-.394	-.091	-.182	-.143
Kolmogorov-Smirnov Z		.538	.674	1.134	.776	.783	.567	.537
Asymp. Sig. (2-tailed)		.934	.754	.152	.583	.573	.904	.935
Exact Sig. (2-tailed)		.865	.625	.092	.459	.467	.808	.870
Point Probability		.035	.087	.027	.065	.048	.075	.041

**Table 7. Wald-Wolfowitz test statistics.**

	Number of Runs	Z	Exact Sig. (1-tailed)	Point Probability
<b>Sazaire (Soyo)</b>	11(a)	.600	.722	.175
<b>Ambriz</b>	9(a)	.990	.846	.220
<b>Quinzau</b>	7(a)	-.698	.242	.133
<b>Landana</b>	11(a)	1.510	.942	.119
<b>Noqui</b>	9(a)	-.370	.352	.153
<b>Maquela do Zombo</b>	8(a)	-.146	.436	.194
<b>Pango Aluquem</b>	9(a)	-.028	.484	.189

normality of the underlying data. The application of Levene test (**Table 8**) shows that for all stations (except Quinzau) the variance of the ground data is not statistically significant (at 5% level of significance) from the variance of the TRMM data. Since a regionalization procedure will be used to undertake the frequency analysis, the variance of a single station is not a governing factor in the global regional analysis.

#### 4. Regionalization Methodology

Regionalization serves two purposes. For sites where data are not available, the analysis is based on regional data [13]. For sites with available data, the joint use of

data measured at a site, called at-site data, and regional data from a number of stations in a region provides sufficient information to enable a probability distribution to be used with greater reliability. This type of analysis represents a substitution of space for time where data from different locations in a region are used to compensate for short records at a single site [14] and [15].

Regional frequency analysis methods are based on the assumption that the standardized variable  $q_i = Q_T/\mu_i$  at each station ( $i$ ) has the same distribution at every site in the region under consideration, where  $Q_T$  is the quantile at  $t$  return period and  $\mu_i$  is the mean of the maximum daily rainfall record at location  $i$ . In particular  $C_v(q)$  and  $C_s(q)$ , the coefficient of variation and the coefficient of

**Table 8. Independent samples Levene test.**

	Sazaire (Soyo)	Ambriz	Quinzau	Landana	Noqui	Maquela do Zombo
F	1.698	0.574	11.648	1.391	0.712	3.541
Sig.	0.21	0.461	0.004	0.257	0.411	0.079

skewness of  $q$ , are considered to be constant across the region [13]. Departures from this assumption may lead to biased quantile estimates at some sites. Sites with  $C_v$  and  $C_s$  nearest to the regional average may not suffer from such bias, but large, biased quantile estimates are expected for sites whose  $C_v$  and  $C_s$  deviate from average. Good results may be obtained by regionalization, especially in cases of short records, provided that the degree of heterogeneity is not great. In such cases, the large number of sites contributing to parameter estimation compensates for regional heterogeneity. We will present hereafter the homogeneity test that checks if the stations used in the analysis could be grouped in one hydrological region and then the index flood method which is used as the regionalization method. The regionalization procedure is done using the Matlab code made available by Rao and Hamed and accompanying the book “Flood Frequency Analysis” [16].

#### 4.1. Regional Homogeneity Check

A method of assigning homogeneous regions is geographical similarity in soil types, climate and topography. However, geographically similar regions may not be similar from the rainfall frequency point of view [13]. On the other hand, two sites in different regions may prove to be similar with respect to frequency analysis, despite the fact that they are geographically different.

Wiltshire [17] and [18] used an approach to initially divide the entire group of catchments into two or more groups based on one or more chosen basin characteristic such as large and small, or wet and dry. The internal homogeneity and mutual heterogeneity of these groups are then expressed in terms of a flow statistic such as  $C_v$ . The process is then repeated by altering the partition points until an acceptable set of regions has been identified.

The Wiltshire  $C_v$ -based test involves the statistic  $S$  in Equation (4).

$$S = \sum_{j=1}^N \frac{(C_{vj} - C_{vo})^2}{U_j} \quad (4)$$

In Equation (4),  $N$  is the number of sites in the region,  $C_{vj}$  is the coefficient of variation at site  $j$  and  $C_{vo}$  and  $U_j$ , are given by Equations (5) and (6)

$$C_{vo} = \frac{\sum_{j=1}^N \frac{C_{vj}}{U_j}}{\sum_{j=1}^N \frac{1}{U_j}} \quad (5)$$

$$U_j = \frac{V}{n_j} \quad (6)$$

In Equation (6),  $n_j$  is the record length at site  $j$  and  $V$  is the regional variance defined in Equation (7).

$$V = \frac{1}{N} \sum_{j=1}^N n_j v_j \quad (7)$$

In Equation (7),  $v_j$  is given by

$$v_j = (n_j - 1) \cdot \sum_{i=1}^{n_j} \left[ C_{v_{n-1}}^{(i)} - \left( \sum_{k=1}^{n_j} C_{v_{n-1}}^{(k)} \right) / n_j \right]^2 / n_j \quad (8)$$

$C_{v_{n-1}}^{(k)}$  in Equation (8) is the  $C_v$  computed from a sample of size  $(n_j - 1)$  with the  $k^{th}$  observation removed. The statistic  $S$  in Equation (4) has the form of a  $\chi^2$  statistic.  $S$  is expected to be  $\chi^2$  distributed with  $(N - 1)$  degrees of freedom. If the value of  $S$  exceeds the critical value of  $\chi^2 (N - 1)$  at a particular significance level, then the hypothesis that the region is homogeneous is rejected and the region is regarded as heterogeneous. However, this test is likely to be effective only for large regions having large record lengths. Further developments using distribution-based tests are given by [17] and [18].

The results of the Wiltshire Chi-square statistic was found to be 4.81, which indicates (with a p-value = 0.778) that the eight rainfall stations used could be treated as one region.

Furthermore, the coefficient of skewness and kurtosis are plotted for the stations used in the analysis on the Moment ratio diagrams. For a given distribution, conventional moments can be expressed as functions of the parameters of distributions. It follows that the higher order moments can be expressed as functions of lower order moments. For two-parameter distributions, the moment  $\mu_3$  can be expressed as a unique function of  $\mu_2$ . For example, the coefficient of skewness ( $C_s$ )  $C_s = \mu_3 / \mu_2^{1.5}$  can be expressed as a unique function of  $C_v = \mu_2^{0.5} / \mu_1'$ . Similarly,  $C_k = \mu_4 / \mu_2^2$  of a three-parameter distribution can be expressed as a unique function of  $C_s$ . The  $C_s - C_k$  moment ratio relationship for some popular three pa-

parameter distributions are plotted in what is called the moment ratio diagrams.  $C_s$  and  $C_k$  moment ratio for the eight stations are also plotted on the same diagram (**Figure 2**). It shows that both the Generalized Pareto and the Pearson Type III distributions present adequate fits for the eight stations. The Pearson Type III was chosen since it is more known in rainfall frequency analysis. Furthermore, a comparison between the 100-year rainfall value using the Pareto distribution and the Pearson Type III distribution shows a difference of less than 5% in the value, which is negligible for this relatively high return period.

#### 4.2. Index Flood Method for Regionalization

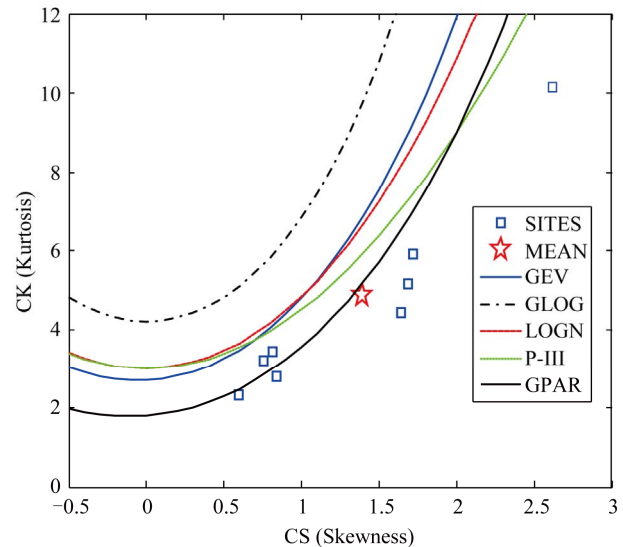
Many types of regionalization procedures are available [13] and [19]. One of the simplest procedures, which has been used for a long time, is the index flood method [20]. The key assumption in the index flood method is that the distribution of the variable of interest at different sites in a region is the same except for a scale or index flood parameter, which reflects rainfall and runoff characteristics of each region. The index flood parameter may be the mean of the maximum annual rainfall, although any location parameter of the frequency distribution may be used [21]. In this case, regional quantile estimates  $\hat{Q}_T$  at a given site for a given return period  $T$  can be obtained as in Equation (9), where  $q_T$  is the quantile estimate from the regional distribution for the given return period, and  $\mu_i$  is the mean of the maximum annual rainfall at the site.

$$\hat{Q}_T = \mu_i \cdot q_T \quad (9)$$

The regional distribution parameters are obtained by using the regional weighted average of dimensionless moments obtained by using the dimensionless rescaled data  $q_{ij} = Q_{ij} / \hat{\mu}_i$ . The joint use of at-site and regional data is advisable, provided that a reasonably homogeneous region can be identified. The data at a site may be used when the record at a station is exceptionally long, or when regional data are not available, or when this site departs somehow from the regional trend.

### 5. Regionalization Results and Discussion

The fitting of the Pearson Type III distribution to all data using both at site information as well as regional information is shown by **Figure 3**. It shows that only for three stations (Maquela de Zombo, Noqui and Quinzau) the at-site curve fits the observed data better. This is shown by the closer fit of the at-site blue curve compared to the regional red curve (**Figure 3**). However, for Noqui station, the 300 mm rainfall value in one day is departing from the global pattern of daily rainfall. Thus, the re-



**Figure 2. Moment Diagrams Showing the Eight (8) Stations Used in the Analysis.**

gional estimate, although lower, seems a more reasonable value. **Table 9** shows the regional and at-site Pearson type III distribution fittings for 5-, 10-, 25-, 50-, 100-, and 200-year return periods. The values in bold characters are the ones adopted for further analysis. A closer look at the bold values show that the 100-year estimates tend to group geographically in four regions as follows: Landana–Noqui, M Banza Congo, Sazaire (Soyo)–Ambriz–Quinzau, and Maquela de Zombo–Cuimba. This zoning—shown in **Figure 1**—is undertaken to add more robustness to the values adopted for design.

In the coastal region, The Landana–Noqui axis exhibits the highest rainfall with an average 100-year estimate of 267 mm. The Sazaire (Soyo)–Ambriz axis comes last, with low estimate at Ambriz and relatively higher at Quinzau and average at Sazaire (Soyo).

In the mountainous region, M Banza Congo location reveals higher rainfall than Cuimba and Maquela de Zombo stations. This could be explained by the digital elevation model which reveals that the elevations of Cuimba, Maquela de Zombo and M Banza Congo are 532, 877 and 1000 m, respectively. Furthermore, the Maquela de Zombo station is higher than Cuimba; however it is in the shade of the mountain while Cuimba is in the bottom of the mountain facing the prevailing winds. This might elucidate the reason why Cuimba and Maquela de Zombo experience similar extreme events although there is a difference of elevation between them. Based on the above M Banza Congo is considered in a separate region while as Maquela de Zombo and Cuimba are considered similar and nearly equal to the Sazaire (Soyo)–Quinzau axis. As such, all catchments discharging to the ocean in the stretch between Sazaire (Soyo)



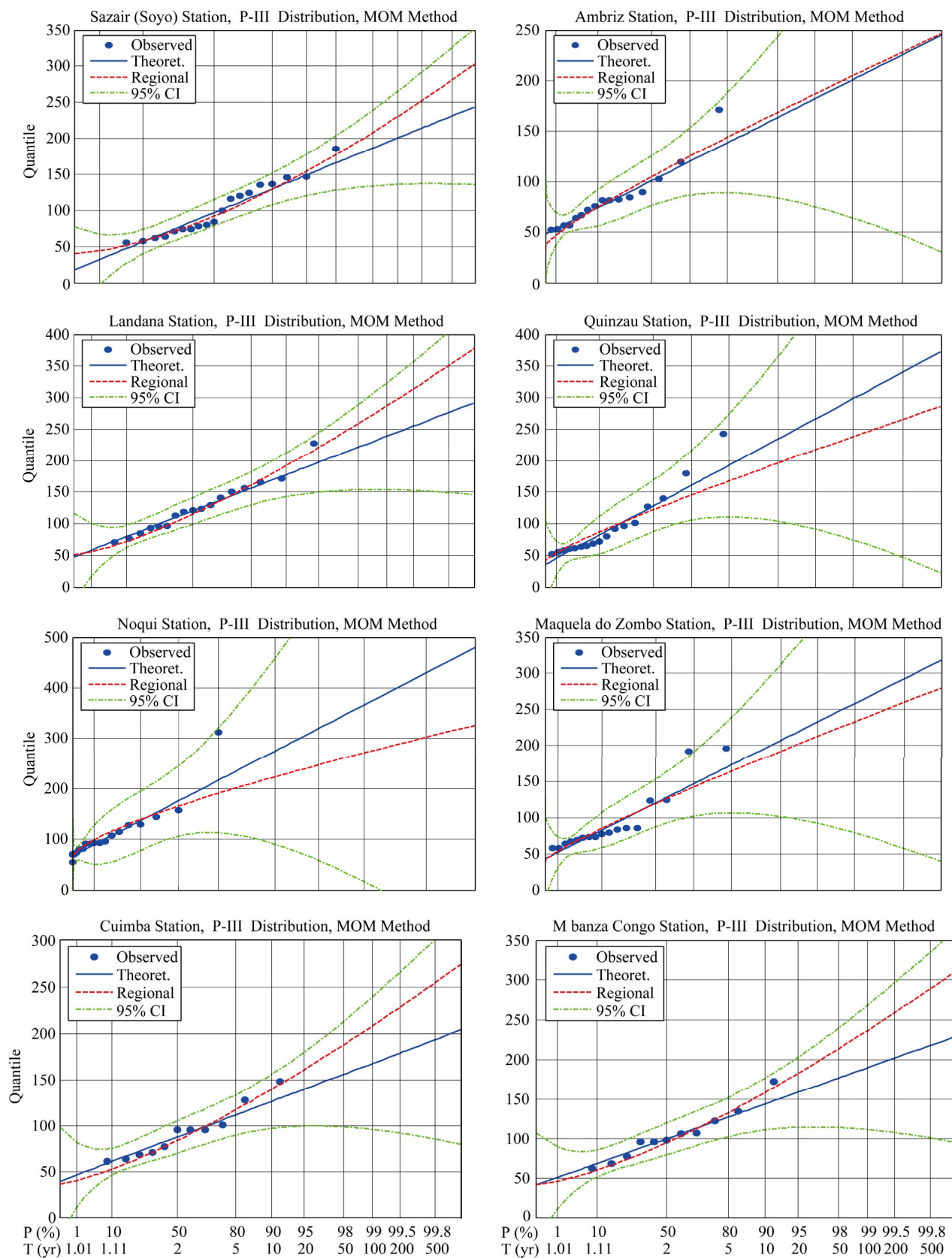


Figure 3. Regional and At-Site Pearson type III Distribution Fittings.



**Table 9. Regional and at-site Pearson type III distribution fittings.**

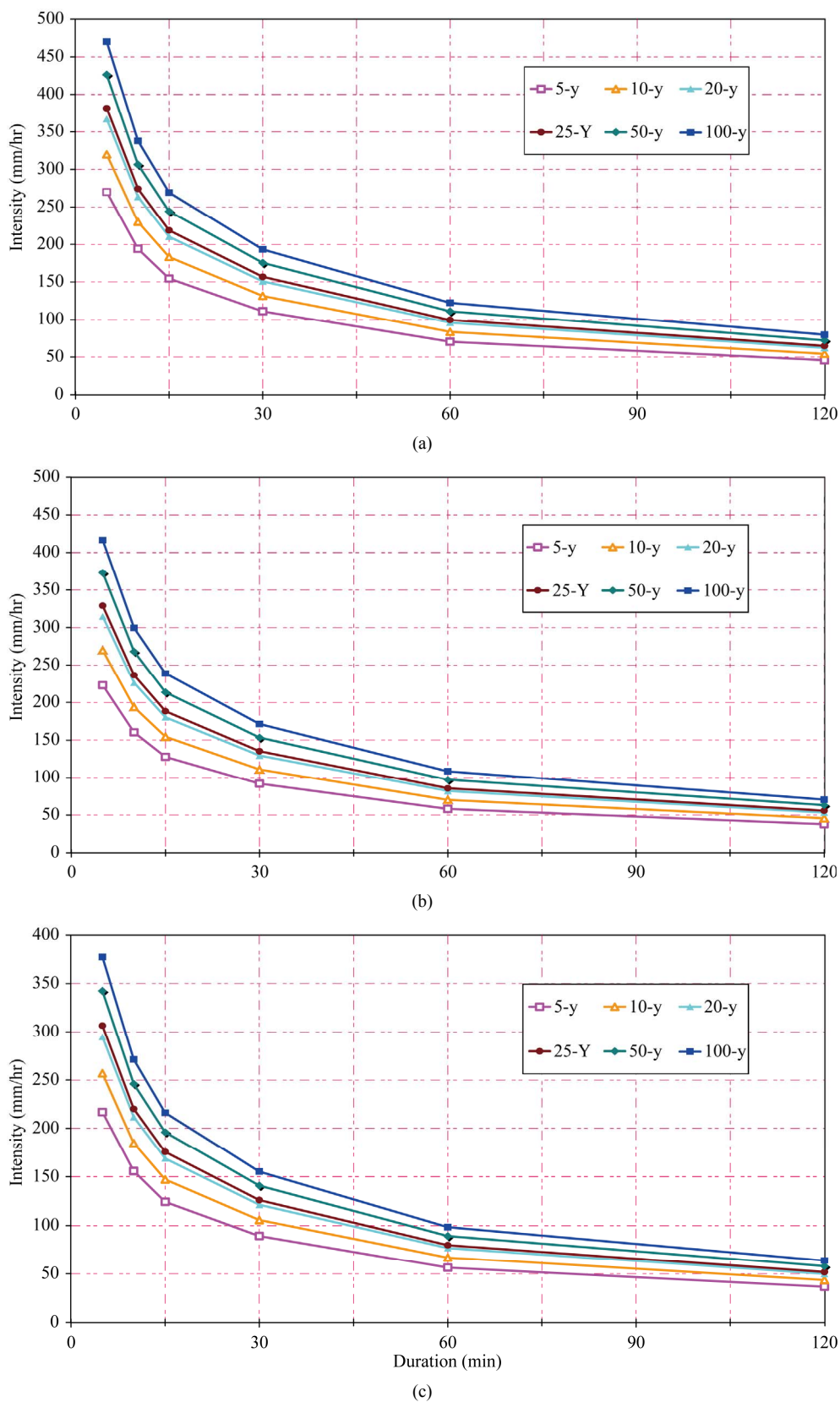
Return Period	Regional	At-site	Obs.	Regional	At-site	Obs.	Regional	At-site	Obs.
	Sazaire (Soyo)			Ambriz			Quinzau		
5	<b>129.31</b>	129.83	136.62	<b>105.01</b>	100.91	97.25	122.05	<b>128.09</b>	132.18
10	<b>154.11</b>	149.06	146.9	<b>125.14</b>	119.98	135.16	145.46	<b>160.65</b>	191.87
25	<b>185.16</b>	171.22	NaN	<b>150.36</b>	144.79	NaN	174.76	<b>202.88</b>	NaN
50	<b>207.92</b>	186.48	NaN	<b>168.84</b>	163.48	NaN	196.24	<b>234.64</b>	NaN
100	<b>230.32</b>	200.85	NaN	<b>187.02</b>	182.21	NaN	217.38	<b>266.4</b>	NaN
200	<b>252.47</b>	214.53	NaN	<b>205.01</b>	201.02	NaN	238.29	<b>298.27</b>	NaN
	Landana			Noqui			Maquela do Zombo		
5	<b>161.08</b>	155.82	159.77	<b>138.91</b>	136.41	129.95	119.46	<b>120.45</b>	123.75
10	<b>191.98</b>	177.61	182.76	<b>165.56</b>	175.62	157.18	142.37	<b>147.04</b>	192.6
25	<b>230.65</b>	203.37	NaN	<b>198.91</b>	230.7	NaN	171.05	<b>181.37</b>	NaN
50	<b>259.01</b>	221.43	NaN	<b>223.36</b>	274.45	NaN	192.08	<b>207.11</b>	NaN
100	<b>286.9</b>	238.65	NaN	<b>247.42</b>	319.79	NaN	212.77	<b>232.81</b>	NaN
200	<b>314.5</b>	255.23	NaN	<b>271.22</b>	366.59	NaN	233.24	<b>258.54</b>	NaN
	M Banza Congo			Cuimba					
5	<b>132.85</b>	127.14	129.86	<b>116.74</b>	111.27	116.62			
10	<b>158.33</b>	143.75	165.02	<b>139.13</b>	125.98	144.03			
25	<b>190.23</b>	163.27	NaN	<b>167.16</b>	143.41	NaN			
50	<b>213.61</b>	176.89	NaN	<b>187.71</b>	155.65	NaN			
100	<b>236.62</b>	189.84	NaN	<b>207.93</b>	167.33	NaN			
200	<b>259.38</b>	202.27	NaN	<b>227.93</b>	178.59	NaN			

**Table 10. Adopted daily rainfall values for each zone.**

Return Period	Landana-Cabinda-Noqui	M Banza Congo	Maquela de Zombo-Cuimba	Sazaire (Soyo)-Ambriz-Quinzau
5-y	161.75	133.65	124.5248	129.85
10-y	191.68	161.52	150.2393	153.88
25-y	228.61	197.03	182.9783	183.52
50-y	255.4	223.4	207.2805	205.03
100-y	281.56	249.56	231.3885	226.03

**Table 11. Ratios between 24-Hr duration and other storm duration intensities.**

Storm Duration (min)	5	10	20	30	60	120	180	360	720	1440
Bell's ratios	0.13	0.2	0.28	0.34	0.44	0.57	0.63	0.75	0.88	1



**Figure 4. Intensity Duration Frequency Curves for. (a) Landana-Cabinda-Noqui; (b) M Banza Congo; (c) Sazaire (Soyo)-Ambriz-Quinzau and Maquela de Zombo-Cuimba.**

and Ambriz are considered to follow the same rainfall pattern because of the similarity between the coastal zone (represented by Sazaire–Ambriz axis) and the mountainous region (represented by Maquela de Zombo–Cuimba zone). The high values depicted in M Banza Congo are only representative of the very high mountain where M Banza Congo is. To calculate the adopted values of daily rainfall at different return periods, a weighted average of the rainfall estimates of **Table 9** (based on the length of rainfall record at each location) are calculated for each region. The adopted values are shown in **Table 10**.

## 6. Intensity Duration Frequency Curves Development

A theoretical ratio of 1.13 to 1.14 is adopted between daily and 24-hr rainfall values [22]. In the absence of short duration records or any similar information, ratios that could be assumed between intensities of the 24-hr and those of the 12-, 6-, 3-, 2-, 1-Hr, 30-, 15-, and 5-min ratios (refer to **Table 11**) were first proposed from durations of 2-hr to 5-min by Bell [5] based on studies in the USA, and extended by the Soil Conservation Service of the USA through their SCS type II dimensionless rainfall curve [6]. Using The TRMM data from 3 hours and to 24 hours, we were able to confirm Bell/SCS type II ratios from 3 hours to 24 hours. It is well known that durations from 2 hours to 5 minutes are fairly constant in different climates because of the similarity of convective storms patterns ([5], [23] and references therein).

Based on **Tables 10** and **11** and the 24-hour rainfall depths at different return periods, the intensity-duration rainfall values are calculated. IDF curves for return periods of 100-, 50-, 25-, 20-, 10-, and 5-year are shown for the 4 above mentioned zones (**Figure 4**).

## 7. Conclusions

This research presents a methodology to overcome the lack of ground station rainfall data by the joint use of the available ground data with TRMM satellite data to develop Intensity Duration Frequency (IDF) curves. Homogeneity of the means and variances are first checked for both types of data. The study zone is verified as being a homogeneous region with respect to frequency analysis using the Wiltshire test. An Index Flood procedure is adopted to generate the theoretical regional distribution equation and the rainfall values at different return periods are calculated for all locations of interest. Regional coherence is identified and used to derive robust rainfall estimates. TRMM data from 3 hours to 24 hours are found consistent with Bell/SCS type II ratios and the

Bell/SCS type II are used to develop ratios between 24-hr rainfall depth and shorter duration depths. The regional patterns along with the developed ratios are used to develop regional IDF curves.

## 8. References

- [1] E. Habib, A. Henschke and R. F. Adler, "Evaluation of TMPA Satellite-Based Research and Real-Time Rainfall Estimates during Six Tropical-Related Heavy Rainfall Events over Louisiana, USA," *Atmospheric Research*, Vol. 94, No. 3, November 2009, pp. 373-388. [doi:10.1016/j.atmosres.2009.06.015](https://doi.org/10.1016/j.atmosres.2009.06.015)
- [2] M. Li and Q. Shao, "An Improved Statistical Approach to Merge Satellite Rainfall Estimates and Raingauge Data," *Journal of Hydrology*, Vol. 385, No. 1-4, May 2010, pp. 51-64. [doi:10.1016/j.jhydrol.2010.01.023](https://doi.org/10.1016/j.jhydrol.2010.01.023)
- [3] M. Almazroui, "Calibration of TRMM Rainfall Climatology over Saudi Arabia during 1998-2009," *Atmospheric Research*, Vol. 99, No. 3-4, March 2011, pp. 400-414.
- [4] L. Li, Y. Hong, J. Wang, R. F. Adler, F. S. Policelli, S. Habib, D. Irwn, T. Korme and L. Okello, "Evaluation of the Real-Time TRMM-Based Multi-Satellite Precipitation Analysis for an Operational Flood Prediction System in Nzoia Basin, Lake Victoria, Africa," *Natural Hazards*, Vol. 50, No. 1, 2009, pp. 109-123. [doi:10.1007/s11069-008-9324-5](https://doi.org/10.1007/s11069-008-9324-5)
- [5] F. C. Bell, "Generalized Rainfall-Duration-Frequency Relationship," *Journal of Hydraulic Division*, Vol. 95, No. HY1, 1969, pp. 311-327.
- [6] Soil Conservation Service, "Urban Hydrology for Small Watersheds, Technical Release 55," United States Department of Agriculture, 1986.
- [7] W. J. Conover, "Practical Nonparametric Statistics," 3rd Edition, John Wiley, Hoboken, 1980.
- [8] H. B. Mann and D. R. Whitney, "On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other," *Annals of Mathematical Statistics*, Vol. 18, No. 1, March 1947, pp. 50-60. [doi:10.1214/aoms/1177730491](https://doi.org/10.1214/aoms/1177730491)
- [9] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics*, Vol. 1, No. 6, December 1945, pp. 80-83. [doi:10.2307/3001968](https://doi.org/10.2307/3001968)
- [10] L. E. Moses, "A Two-Sample Test," *Psychometrika*, Vol. 17, No. 3, 1952, pp. 239-247. [doi:10.1007/BF02288755](https://doi.org/10.1007/BF02288755)
- [11] M. A. Stephens, "EDF Statistics for Goodness of Fit and Some Comparisons," *Journal of the American Statistical Association*, Vol. 69, No. 347, September 1974, pp. 730-737. [doi:10.2307/2286009](https://doi.org/10.2307/2286009)
- [12] H. Levene, "Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling," Stanford University Press, Palo Alto, 1960, pp. 278-292.
- [13] C. Cunneane, "Statistical Distributions for Flood Frequency Analysis," Word Meteorological Organization Operational Hydrology Report, No. 33, Geneva, 1989.
- [14] National Research Council, "Committee on Techniques

- for Estimating Probabilities of Extreme Floods, Estimating Probabilities of Extreme Floods, Methods and Recommended Research", National Academy Press, Washington D.C., 1988.
- [15] J. R. Stedinger, R. M. Vogel and E. Foufoula-Georgiou, "Frequency Analysis of Extreme Events," In: D. R. Maidment, Ed., *Handbook of Hydrology*, McGraw-Hill, New York, 1993.
- [16] A. R. Rao and K. H. Hamed, "Flood Frequency Analysis," CRC Press, Boca Raton, 2000.
- [17] S. E. Wiltshire, "Regional Flood Frequency Analysis I: Homogeneity Statistics," *Hydrological Sciences Journal*, Vol. 31, No. 3, 1986, pp. 321-333.  
[doi:10.1080/02626668609491051](https://doi.org/10.1080/02626668609491051)
- [18] S. E. Wiltshire, "Identification of Homogeneous Regions for Flood Frequency Analysis," *Journal of Hydraulics*, Vol. 84, No. 3-4, May 1986, pp. 287-302.  
[doi:10.1016/0022-1694\(86\)90128-9](https://doi.org/10.1016/0022-1694(86)90128-9)
- [19] C. Cunnane, "Methods and Merits of Regional Flood Frequency Analysis," *Journal Hydrology*, Vol. 100, No. 1-3, July 1988, pp. 269-290.  
[doi:10.1016/0022-1694\(88\)90188-6](https://doi.org/10.1016/0022-1694(88)90188-6)
- [20] T. Dalrymple, "Flood Frequency Analysis, Manual of Hydrology, Part 3, Flood-Flow Techniques: U.S. Geological Survey Water-Supply Paper 1543-A," 1960.
- [21] J. R. M. Hosking, and J. R. Wallis, "Some Statistics Useful in Regional Frequency Analysis, Res. Rep. RC 17096," IBM Research Division, New York, 1991.
- [22] D. M. Hershfield, "Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years," US Weather Bureau Technical Paper 40, Washington DC, 1962.
- [23] A. G. Awadallah, and N. S. Younan, "Impact of Design Storm Distributions on Watershed Response in Arid Regions," *8th International Association of Hydro-logical Sciences (IAHS) Scientific Assembly*, 6-12 September 2009, Hyderabad.