International Journal of Hybrid Information Technology
Vol. 9, No.6 (2016), pp. 413-432
http://dx.doi.org/10.14257/ijhit.2016.9.6.37

Regional Drought Rainfall for Selangor River Basin in Malaysia Estimated Using L-Moments

Jer Lang Hong and Kee An Hong

Taylor's University, Hong and Associates
jerlanghong@hotmail.com, keeanhong@yahoo.co.uk

Abstract

Water from the Selangor river basin is the main source of supply for the federal capital of Kuala Lumpur and the various districts of Selangor State located in the centre of Peninsular Malaysia since 2000. As this region is the most populated area in the country, its rapid economic development and population growth have caused concern over the adequacy of the quantity and quality of water abstracted from the Selangor basin, both at present and in the future. A recent prolonged drought has caused a one-month water rationing in the Kuala Lumpur and adjacent areas which basically has seriously affected the everyday life of the people and the industrial and agricultural sectors. A rainfall drought analysis is therefore required for assessing the severity of the drought events in the study area. In this study, L-moments have been used to compute the rainfall quantile values for 9 probabilities, 6 durations, 12 starting months for the 2 regions across the Selangor basin. The choice of the Pearson Type III and Wakeby distributions for fitting the Selangor rainfall data is presented. Quantile values are expressed as a percentage of mean rainfall of the particular duration and presented as drought indication maps. Rainfall of specific return period can be calculated easily using these maps and the mean rainfall for the particular station. The return period (severity) of a historical rainfall event from the station can be known when the magnitude of the historical event is compared with that of the rainfall event of a particular return period. This helps in monitoring and management of droughts.

***Keywords**: Drought, L-Moment*

1. Introduction

Drought rainfall analysis gives valuable information for efficient water resources management. Extensive regional drought rainfall analysis for different time scales and probabilities have been carried out by Guttman (1993), Nunez et al. (2011), Blain (2011), Giovannettone et al. (2013). In particular, the L moment method has been used for these studies. As droughts are normally assessed in terms of probabilities and durations, this study uses the L-moments to derive drought quintiles for durations of 1, 3, 6, 12, 24, and 36 months for probabilities of 0.01, 0.02, 0.05, 0.1, 0.2, and 0.5.

2. The Study Area

A location map of Selangor basin up to the water intake point is shown in Figure 1. Selangor basin has an area of 1450 km² and the maximum length and width of the basin are 48 km and 39 km respectively. About 30% of the basin is steep mountainous country above 600 m, 38% is in hilly country and the remainder undulating low terrain. Two-thirds of the basin is under jungle and the remainder under rubber and oil palm. Fine to coarse granite and other allied rocks occupy the eastern half of the basin and sandstone is found in the western half. Wet seasons dominate in April and May in the south west monsoon season and October to

ISSN: 1738-9968 IJHIT

December in the north east monsoon season. Dry periods generally occur in January to March and June to September. The rainfall stations with long term records are shown in Figure 1. There are also some short term rainfall stations which are usually located within 10 km of the respective stations shown in Figure 1. Figure 2 shows the mean annual rainfall map for the study area created based on the data of long term rainfall stations.

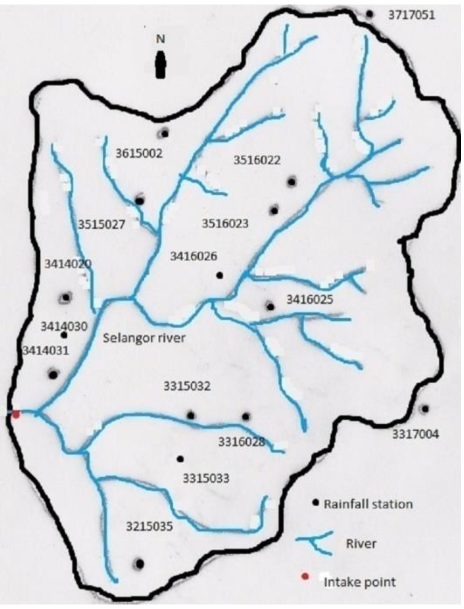

Figure 1. Rainfall Stations of Selangor Basin

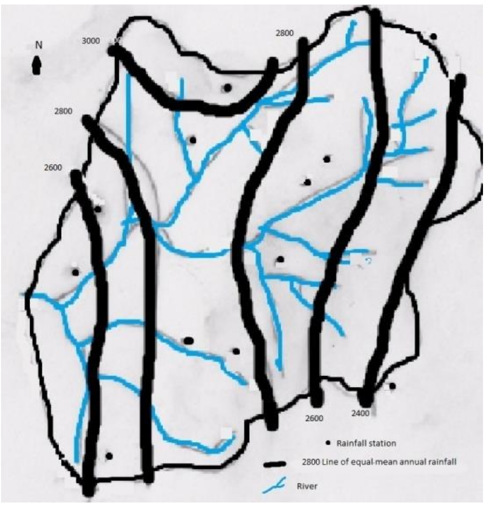

Figure 2. Mean Annual Rainfall Map of Selangor Basin

3. Rainfall Data

Rainfall data are collected by various agencies such as Drainage and Irrigation Department, Public Works Department, Estates, Hospital and the data are processed and published by the Drainage and Irrigation Department. Rainfall data and the periods of record available are shown in Table 1.The records in the study area are not available for most of the stations for the period 1940-1947 due to the interruption of Second World War.

3.1. Infiiling of Missing Data

In order to preserve continuity of the monthly rainfall for this study, estimates of missing data were made. As usual, not all the stations in the study area are with complete data and stations where only a few months of data are missing were infilled using the records of nearby stations to increase the record length. Nearby stations are defined as stations located within 10 km with consideration to elevation difference (Perica et al. 2009).No attempts were made to infill missing data more than a year. It is possible to estimate the missing data of most of the stations in this way as there are available nearby stations near the key rainfall stations listed in Table 1.

Table 1. Rainfall Records of Selangor Basin

Station no	Station name	Period of record	Remarks
3215035	Ladang Strathairlie	1922-1940,1947-1997	Gap in record due to world war II
3316028	Ldg. Sg. Gapi	1970-2000	
3416030	Ldg. Hopeful	1931-1939, 1947-2000	Gap in record due to world war II
3315033	JKR reservoir Rawang	1951-1996	
3317004	Genting Sempah	1974-2000	
3416026	Rasa Estate	1928-1941,1947-2000	Gap in record due to world war II
3416025	Ldg Batang Kali	1935-1939,1950-1993	Gap due to world war II
3414029	Sg Tinggi Estate	1927-1940,1945-2000	Gap in record due to world war II
3414031	Selangor Tin Dredging	1913-1939,1947-1993	Gap in record due to world war II
3516023	Kuala Kubu Hospital	1893-1940,1947-1996	Gap due to world war II
3516022	Loji Air Kuala Kubu	1943-2000	
3717051	Bt Fraser	1922-1940,1947-2000	Gap due to world war II
3615022	Ldg Sg Gumut	1947-2000	
3515027	Ldg Sg Beleta	1947-2000	

3.2. Initial Data Screening

Quality control data are needed for the frequency analysis model. In this study, the annual rainfall data are tested for consistency by checking the presence of outliers, trends, and data independence.

3.3. OutlierS

Tests for outliers for the station records were carried out using annual rainfall data and the generalized Extreme Studentized Deviate (ESD) test. The ESD is a generalization of the Grubb's test (Zaiontz 2014). Results are presented in Table 2.The identified outliers are not excluded for further analysis unless there are strong hydrological and statistical evidences that they are real outliers.

3.4. Graphical Checks

Consistency of the annual rainfall data sometimes can be investigated using graphical plots. The annual rainfall of the adjacent stations are grouped and plotted as shown in Figures 3 to 5.Figures 3 to 5 show that stations 3414030,3315033 and 3414026 recorded low rainfall consistently for certain periods compared to adjacent stations. Station 3414029 has extremely low rainfall for 1991 to 1993.

Table 2. Results of Outlier Test

Station no	No of outliers
3215035	1
3315032	1
3315033	1
3316028	0
3317004	0
3414030	1
3414031	1
3416025	1
3416026	0
3515027	1
3615022	0
3516023	0
3516022	0
3717051	0
3414029	3

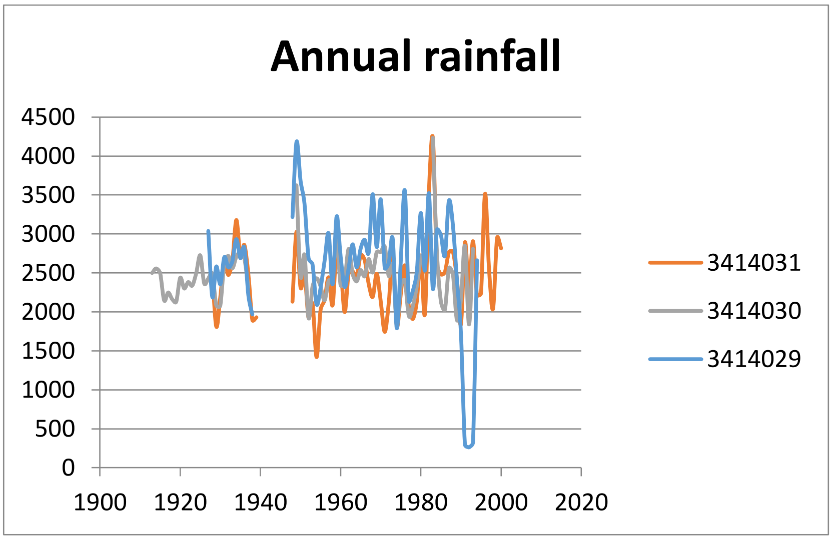

Figure 1. Graphical Plot of Annual Rainfall (A)

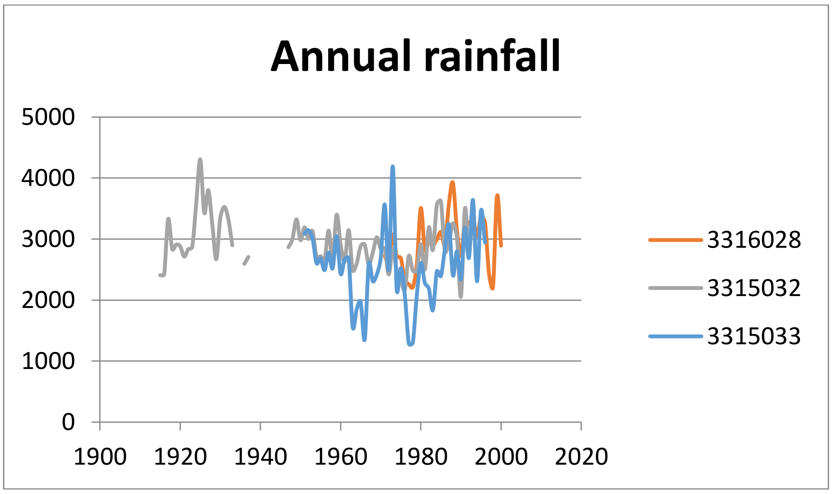

Figure 2. Graphical Plot of Annual Rainfall (B)

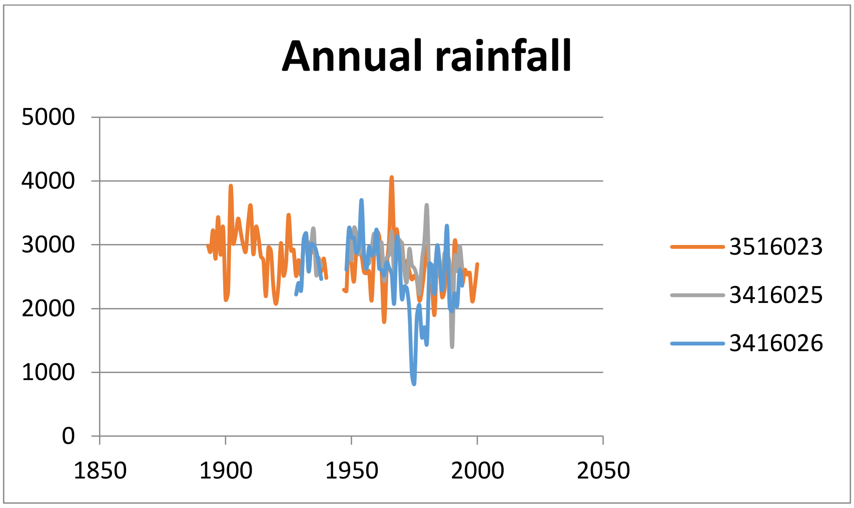

Figure 3. Graphical Plot of Annual Rainfall (C)

2.2.3. Trend and Independence of Annual Rainfall

A typical assumption in frequency analysis is that the observed records are stationary, *i.e*. the year to year variations in the data are expected, but the time series does not exhibit major temporal fluctuations or long term trend. For this study, we use the trend /change detection software prepared for CRC for catchment Hydrology (Chiew and Siriwardena 2005) to test the trend for the post world war II continuous records for stations with more than 45 years. The software uses 11 different statistical tests to detect trend for time series samples. This includes parametric and non-parametric methods. Lag one autocorrelation is also included to test randomness of a sample. The 11 methods used are: Mann Kendall test, Sperman's Rho test Linear regression test, Distribution free CUSUM test, Worsley Likelihood ratio test, Rank sum test, Student's t test, Median crossing test, Turning point test, Rank difference test. Results at 5% significance level for the test are summarized in Table 3. As 10 out of 13 stations passed at least 6 out of 11 tests, the regional increase or decrease in rainfall is not marked. Among these 10 out of 13 stations also show no significant autocorrelation. Annual rainfall trends in this region are considered not significant as for most stations, it occurs that only marginally failures were noted for the 5% level. Figures 6 and 7 show the trend test for two stations in the study area.

Table 3. Trend Test Results at 5% Significance Level

Station	Monn Kendall	Sperman's Rho	Linear regression	CUSUM	Cumulative variation	Worsley	Rank sum	Student's t test	Median crossing	Turning point	Rank difference	Auto-correlation
3215035	NS	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS
3315032	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
3315033	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
3414030	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS	NS
3416025	S	S	S	NS	NS	NS	S	S	NS	NS	NS	NS
3416026	S	NS	NS	NS	S	S	S	S	S	NS	S	S
3515027	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
3515022	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
3515023	NS	NS	NS	NS	S	NS	S	S	NS	NS	S	S
3615002	NS	S	S	S	S	S	S	S	NS	NS	NS	NS
3717051	NS	NS	NS	S	S	S	S	S	NS	S	S	S
3414029	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS	NS
3414031	NS	NS	NS	NS	NS	S	NS	NS	NS	NS	NS	NS

S denotes significant
NS not significant

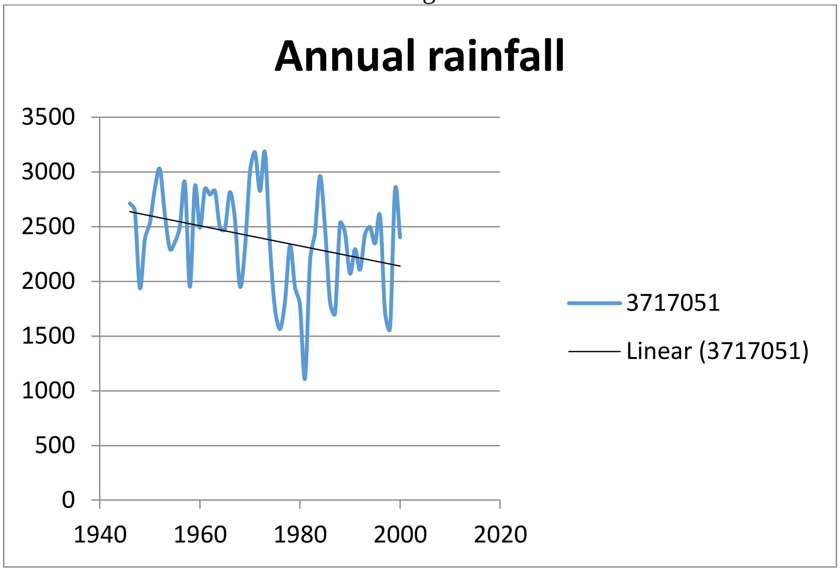

Figure 4. Annual Rainfall for Station 3717051 with Trend Line for 1946-2000

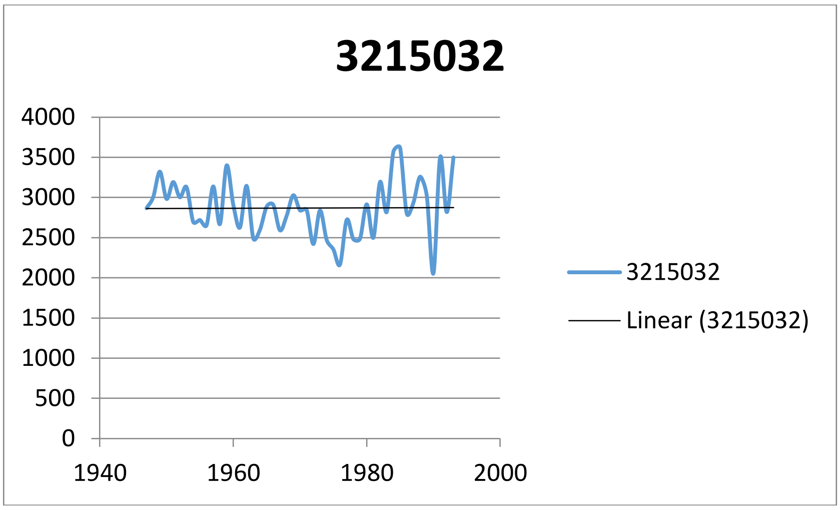

Figure 5. Annual Rainfall for Station 3215032 with Trend Line for 1946-1993

4. Regional Frequency Analysis

In this paper, L-moments were used for the regional rainfall frequency analysis. The ICI –RAFT software developed by Jason Giovannettone and Michael Wright (2013) was used for the computation of rainfall quantiles for this study. ICI –RAFT is a software designed to estimate regional rainfall frequency based on L-moment approach .Details of L moments are given by Hosking and Wallis (1997).

4.1. L-Moments

Hosking and Wallis (1997) found that some linear combinations of probability weighted moments, which they termed “L-moments” could be considered as measures of probability distributions which in turn form a basis for the estimation of these distributions. The L-moment statistics (Hosking and Wallis 1997) are similar to the conventional moments but can be estimated by linear combinations of the elements of an ordered sample. L-moment approach has advantages over conventional moments as it is able to characterize a wider range of distributions. L –moments do not raise data to powers of 2, 3, 4 as required for variance, skewness, and kurtosis, and they give better parameter estimates for data containing outlying values. L-moment ratio estimators of location, scale and shape are almost unbiased irrespective of the probability distribution from which the observations arise. L-moment ratio estimators such as L-coefficient of variation, L-skewness, and L-kurtosis exhibit lower bias than conventional product moment ratios, particularly for highly skewed samples. L-coefficient of variation and L-skewness do not have bounds which depend on sample size as do the ordinary product moment ratio estimators of coefficient of variation and skewness. L-moment ratio diagrams are particularly efficient in identifying the distributional properties of highly skewed sample data.

4.2. L-Moment Analysis

L-moment estimators are linear combinations of the observations. These techniques provide an alternative of describing the shape of probability distributions. The L-moments are derived by modifying the function of probability weighted moments .Probability weighted moments are calculated from the ranked observations $x_{j:n}$. For ordered statistics of $x_{j:n}$, the unbiased estimator of probability weighted moments is:

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)...(j-r)}{(n-1)(n-2)...(n-r)} x_{j:n} \quad (1)$$

Where n is the sample size

and let $x_{1:n} >= x_{2:n} >= \ldots\ldots\ldots >= x_{n:n}$ be the ordered sample

r=0, 1, 2, 3, …

The sample L-Moments are defined as:

$$l_1 = b_0 \quad (2)$$

$$l_2 = 2b_1 + b_0 \quad (3)$$

$$l_3 = 6b_2 - 6b_1 + b_0 \quad (4)$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \quad (5)$$

Hosking defined L-moment ratios as:

L-coefficient of variation L-CV $(t) = \frac{l_2}{l_1}$

L-coefficient of skewness $(t_3) = \frac{l_3}{l_2}$

L –coefficient of kurtosis $(t_4) = \frac{l_4}{l_2}$

4.3. Application of L-Moments

Hosking and Wallis (1997) suggested four steps in the application of L –Moments for regional frequency analysis. These are: (1) Screening of data by calculating the discordancy measure, D_i ,which is used to identify unusual sites (2) Identification of regional homogeneity using the heterogeneity measure, H (3) Choice of a regional frequency distribution by means of the goodness of fit test ,the z_i^{dist} measure, which decides whether a distribution is an adequate fit for the data (4) Estimation of the regional hydrological extreme events such as floods and low flows using the selected distribution.

4.4. Diiscordancy Measure Test

The objective of this test is to check the consistency of data used for regional analysis. Screening of data can be carried out using the L-moment based discordancy measure (D_i) as proposed by Hosking and Wallis (1997).Let $\mu_i = \{ t^i \quad t_3^i \quad t_4^i \}^T$ be a vector containing the t, t_3, t_4 values for site i, with a total of n sites, the superscript T denotes transposition of a vector or matrix. Let

$$\bar{\mu} = n^{-1} \sum_{i=1}^{n} \mu_i \quad (6)$$

be the unweighted group average and define the matrix of sums of squares and cross product as:

$$A = \sum_{i=1}^{n} (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})^T \quad (7)$$

Then the discordancy measure for site i is:

$$D_i = \frac{1}{3} n(\mu_i - \bar{\mu})^T A^{-1} ((\mu_i - \bar{\mu}) \quad (8)$$

Sites with large values of D_i indicate that there may be data errors for these sites and further investigations may be required.

Site i is discordant if D_i is greater than the critical value which depends on the number of sites(Hosking and Wallis,1997).The possibility of large D_i value of a site may be an error in the data of the site or the station may probably belong to another region or no region at all.

4.5. Test of Regional Homogeniety

Hosking and Wallis (1997) also developed a homogeneity test which allows us to assess whether a group of sites are reasonably homogeneous. The measure (H1) compares the inter-site variations in sample L-Moments for a group of sites with what would be expected of a homogeneous region. The inter-site variation of L-moment ratio is measured as the standard deviation (V) of the at site L coefficient of variation (L-CV) weighted proportionally to the record length at each site. Simulations are used (Hosking and Wallis 1997) to establish what would be expected of a homogeneous region. A number of say 500 data regions are generated based on the regional weighted average statistics using the four parameter distribution such as kappa distribution(Hosking and Wallis ,1997).The inter-site variation of each generated region is calculated and the mean (μ_v) and standard deviation (σ_v) of the computed inter-site variation are obtained. The heterogeneity measure (H1) is derived as:

$$H1 = \frac{V - \mu_v}{\sigma_v} \quad (9)$$

The region is declared acceptably homogeneous if H1<1.The region is possibly heterogeneous if 1<=H1<=2 and if H1>2, the region is definitely heterogeneous.
However, in a recent study (Nunez et al. 2014) a region was accepted as homogeneous if H1<2, possibly heterogeneous if 2<H1<3,heterogeneous if H1>3.These criteria were also

adopted for developing the NOAA Atlas (Perica et al. 2009).We adopt the same criteria for this study.

Hosking and Wallis (1997) suggested that if the region tested is not acceptably homogeneous, some redefinition of the region can be considered. The region can be subdivided into two or more regions; some sites can be removed from the region, or reassign sites to different regions.

4.6. Goodness of Fit Test

Comparing the moments of the distribution to the average moment statistics from regional data is a way to choose the suitable frequency distribution for a homogeneous region. This is achieved by seeing the match between the L-kurtosis and L - skewness of the fitted distribution and the regional average L- skewness and L-kurtosis of the observed data. The goodness of fit measure, z^{dist} is defined (Hosking and Wallis, 1997) as

$$z^{dist} = \frac{\tau_i^R - \tau_i^{dist}}{\sigma_i^{dist}} \quad (10)$$

Where τ_i^R is the weighted regional average of L-moment statistics

τ_i^{dist} and σ_i^{dist} are simulated regional average and standard deviation of L-moment statistics for the given distribution

The fit is considered adequate if $| z_i^{dist} |$ is close to zero, with a reasonable criterion that $| z_i^{dist} |$ is less than 1.64

4.7. Index Flood

The estimation of rainfall quantile values using L-moments involves regional analysis using “index flood” procedures (Hosking and Wallis 1997), it is assumed that throughout a region the distribution of rainfall is the same at all sites except for a scale factor that might vary from site to site. A weighted regional dimensionless growth curve is calculated, and the site quantile values are obtained by multiplying the mean of the specific site by the regional growth factor.

5. Formation of Homogeneous Regions

In principle regions could be defined separately for each combination of rainfall duration and starting month, this would result in producing a drought rainfall map which is excessive large and difficult to use. It is therefore considered appropriate to construct a set of regions, based on data for annual rainfall totals, and to use these regions when fitting regional frequency distributions to the data for all durations and starting month as proposed by Hosking and Wallis (1997).

The first step in applying L-moments is to delineate homogeneous regions. No attempts are made to delineate homogeneous regions using cluster analysis as the river basin is small with not many rainfall stations. As a start, all the rainfall stations are used to obtain the H1 value, considering the whole basin as one region. As the H1 value is high, the offending stations 3414030, 3416026, and 3215033 are excluded based on the data quality check, discordancy test, L –moment ratio diagrams and the L-statistics. The H1 value reduced but the region is still heterogeneous. The basin is divided into two regions and a test using ICI-RAFT shows that the regions are homogeneous. Results are shown in Table 4. Figure 8 shows the two regions and the rainfall stations included in the regions.

Table 4. Homogeneous Measure for Selangor Basin and the Two Homogeneous Regions

Region	Number of stations	H1
All	15	5.19
All	12	4.32
West	7	0.82
East	5	0.74

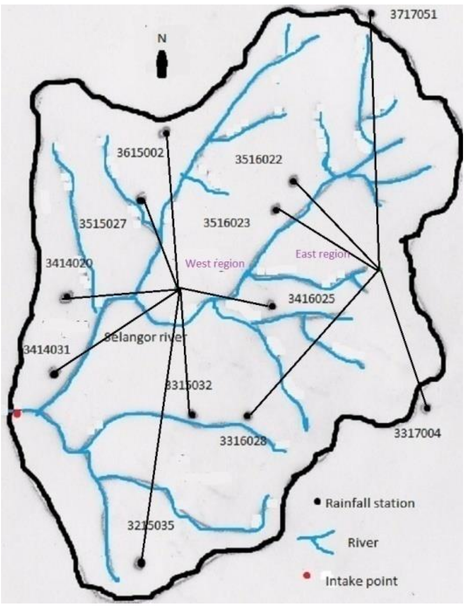

Figure 6. Rainfall Regions, Lines Radiating from the Centre End at the Locations of Stations at the Region

6. Distribution Fitting

The station data of each region are used to calculate the regional L-moments to fit the 3 parameter probability distributions, namely Pearson type III, generalised normal distribution, generalised logistic distribution and the generalized extreme value distribution. The L-moment relationships are shown in Figures 9 to 10 for the two regions for the annual rainfall data. The goodness of fit is measured using Equation (10). Tables 5 and 6 show the distributions accepted by the goodness of fit test by duration and starting month for the two regions.

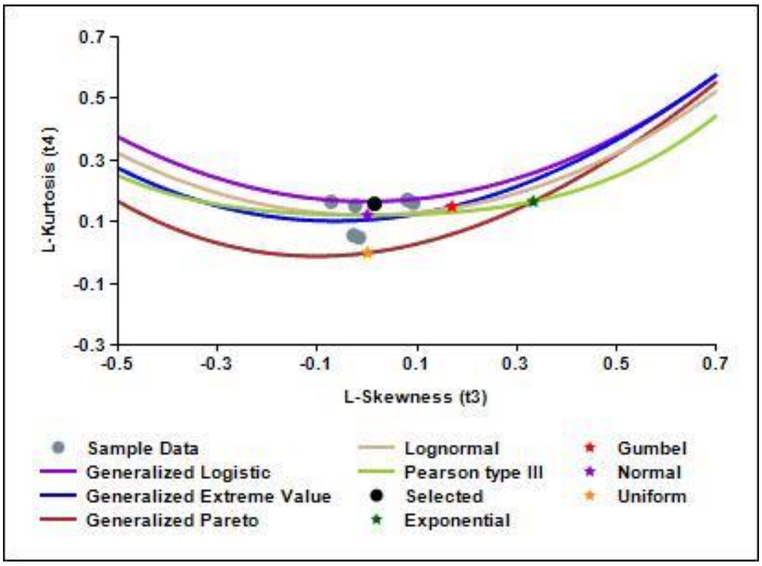

Figure 7. L Moment Diagram For Western Region

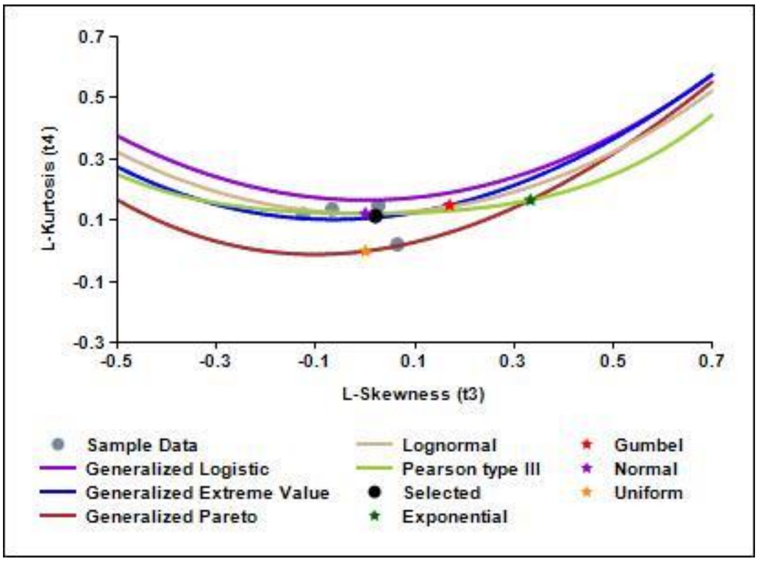

Figure 8. L-Moment Diagram for Eastern Region

Table 5. Distributions Accepted By the Goodness of Fit Test by Duration and Starting Month for Theeastern Region

Duration (Months)	Distribution	Starting month											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	Pearson III	A	A	A	A	A	-	A	A	-	A	A	A
	Lognormal	A	A	A	A	A	-	A	A	-	A	A	A
	Gen Ev	A	A	A	A	A	-	A	A	-	A	A	A
	Gen Log	NA	NA	NA	NA	NA	-	NA	NA	-	NA	NA	NA
3	Pearson III	A	A	A	A	A	A	-	A	A	A	A	A
	Lognormal	A	A	A	A	NA	A	-	NA	A	A	A	A
	Gen Ev	A	A	A	A	A	NA	-	NA	A	NA	A	A
	Gen Log	A	NA	NA	NA	A	NA	-	A	NA	A	NA	NA
6	Pearson III	A	A	A	A	A	A	A	A	A	A	A	-
	Lognormal	A	A	NA	A	NA	NA	A	NA	A	A	A	-
	Gen Ev	NA	NA	NA	A	NA	NA	NA	A	A	A	A	-
	Gen Log	NA	A	A	NA	A	A	A	A	A	A	NA	-
12	Pearson III	A	A	A	A	A	A	A	A	A	A	A	A
	Lognormal	A	A	A	NA	A	A	A	A	A	A	A	A
	Gen Ev	NA	NA	NA	NA	NA	NA	A	A	A	A	NA	A
	Gen Log	NA	A	A	A	A	A	NA	NA	NA	NA	NA	NA
24	Pearson III	A	A	A	A	A	A	A	A	A	A	A	-
	Lognormal	A	A	A	A	A	A	A	A	A	NA	A	-
	Gen Ev	NA	NA	A	A	A	A	A	A	A	A	A	-
	Gen Log	A	A	NA	NA	NA	NA	NA	NA	A	A	A	-
36	Pearson III	-	-	-	-	-	-	-	-	-	-	-	-
	Lognormal	-	-	-	-	-	-	-	-	-	-	-	-
	Gen Ev	-	-	-	-	-	-	-	-	-	-	-	-
	Gen Log	-	-	-	-	-	-	-	-	-	-	-	-

Note A denotes accepted
NA not accepted
-denotes region is heterogeneous

Table 6. Distributions Accepted By the Goodness of Fit Test by Duration and Starting Month for the Western Region

Duration (Months)	Distribution	Starting month											
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	Pearson	A	A	A	-	A	A	A	A	A	A	A	A
	III	A	A	A	-	A	A	A	A	A	A	A	A
	Lognormal	A	A	A	-	A	A	A	A	A	A	A	A
	Gen Ev	N	N	A	-	A	N	N	NA	N	N	NA	NA
	Gen Log	A	A				A	A		A	A		
3	Pearson	A	-	A	A	A	A	-	-	-	A	A	A
	III	A	-	A	A	NA	N	-	-	-	A	A	A
	Lognormal	A	-	A	A	NA	A	-	-	-	A	A	A
	Gen Ev	A	-	NA	NA	A	A	-	-	-	N	NA	NA
	Gen Log						A				A		
6	Pearson	A	A	A	A	A	A	A	A	A	A	A	A
	III	A	A	NA	A	NA	N	A	A	A	A	A	A
	Lognormal	N	N	NA	A	NA	A	N	NA	A	A	A	A
	Gen Ev	A	A	A	NA	A	N	A	A	A	A	NA	NA
	Gen Log	N	N				A	A					
		A	A				A						
12	Pearson	A	A	A	A	A	A	A	A	A	A	A	A
	III	A	A	A	NA	A	A	A	A	A	A	A	A
	Lognormal	N	N	NA	NA	NA	N	A	A	A	A	A	A
	Gen Ev	A	A	A	A	A	A	N	NA	N	N	NA	NA
	Gen Log	N	A				A	A		A	A		
		A											
24	Pearson	A	A	A	A	A	A	A	A	A	A	-	A
	III	A	A	A	A	A	A	A	A	A	A	-	A
	Lognormal	N	N	A	A	A	A	A	A	A	A	-	A
	Gen Ev	A	A	NA	NA	NA	N	N	NA	N	N	-	NA
	Gen Log	N	A				A	A		A	A		
		A											
36	Pearson	A	A	A	A	A	A	-	-	-	-	-	-
	III	A	A	A	A	A	A	-	-	-	-	-	-
	Lognormal	A	A	A	A	A	A	-	-	-	-	-	-
	Gen Ev	N	N	NA	NA	NA	N	-	-	-	-	-	-
	Gen Log	A	A				A						

The Pearson type III distribution is found to be acceptable for rainfall totals for all durations. The lognormal distribution is the second best fit distribution as it is acceptable for most of the rainfall totals for various durations. The generalized extreme value distribution is more acceptable for longer rainfall durations. The generalized logistic distribution is accepted least often. When more than one distribution is accepted by the goodness of fit, the estimated quantiles may be expected to be very similar except in the extreme tails of the distribution. As we are interested in quantiles in the range 0.01 t0 0.5, it is considered appropriate to use distributions that pass the goodness of fit test for quantile estimations. To produce a drought rainfall map which is easy and simple to use, it is reasonable to adopt only one distribution for quantile estimation. For our case, the Pearson type III is a suitable candidate. However, the Pearson type III distribution is not acceptable for all time periods when the region becomes heterogeneous. The number of heterogeneous regions increased as the duration and the starting month of rainfall changed. For such cases, the Wakeby distribution was chosen as the single fitted distribution for a heterogeneous region (Hosking and Wallis 1997).

7. Quantile Estimates

Values are computed for 6 quantiles (0.01 to 0.5) for the two regions for 6 durations (1 to 36 months) for the 12 time scales (January to December).Figure 10 shows the 24 month regional growth curve of Western region with a starting month of November. The difference in growth factors between Pearson type III and Wakeby is not marked.

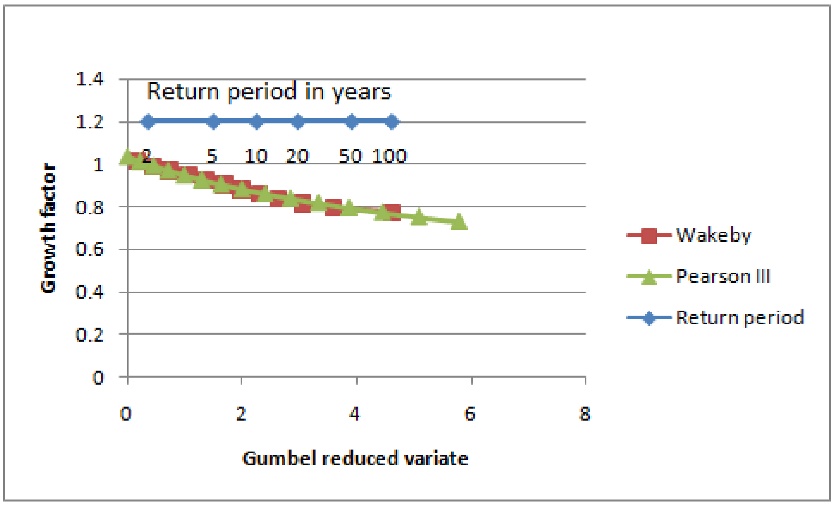

Figure 9. 24 Month Regional Growth Curves of Western Region, Starting November

Since the concern is to characterize drought, the discussion is focused on patterns of non – exceedence probability of 0.01 to 0.5.For rainfall duration of one year or longer, the spatial patterns of a quantile value expressed as a percentage of the mean of the rainfall period show very little change as the starting month of the duration series varies from January through December. The 12 through 60 – month rainfall totals, starting from January, can therefore be considered representative of the totals for other starting months.
The quantile values are presented on “starburst” maps (Figures 12 to 17). Lines radiate from the centroid of a region to the location of the rainfall sites make up the region. Regional statistics are plotted on the respective region. Heterogeneous regions are indicated by parentheses around the regional statistics. The map shows the geographic location and the number of sites in the region.

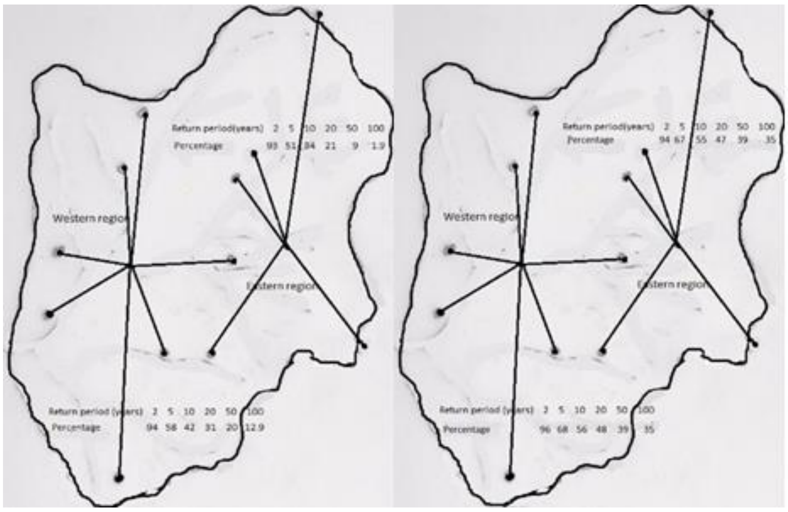

Figure 12. Left: 1 Month Rainfall Starting February as A Percentage of Mean February Rainfall. The Return Period and Respective Percentage are Shown for Eastern and Western Regions Right: 1 Month Rainfall Starting May as a Percentage of Mean May Rainfall. The Return Period and Respective Percentage are shown for Eastern and Western Regions

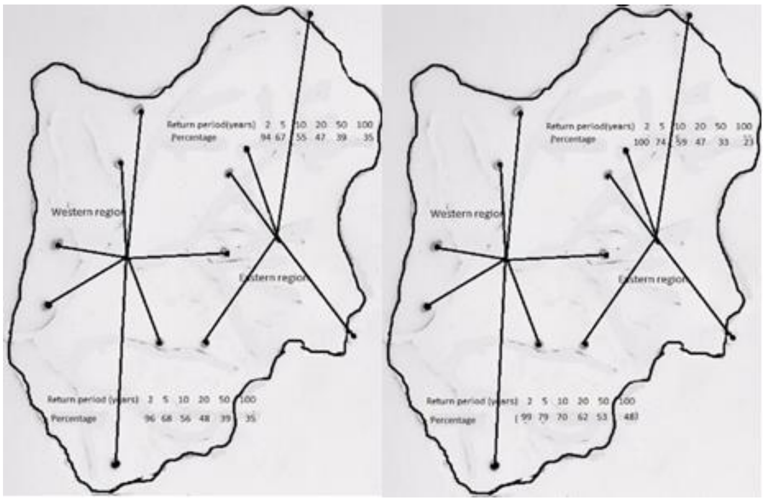

Figure 13. Left: 1 Month Rainfall Starting November as a Percentage of Mean November Rainfall. The Return Period and Respective Percentage are Shown for Eastern and Western Regions Right: 3 Month Rainfall Starting February as A Percentage of Mean February-April Rainfall. The Return Period and Respective Percentage Mean Total Are Shown For Eastern and Western Regions

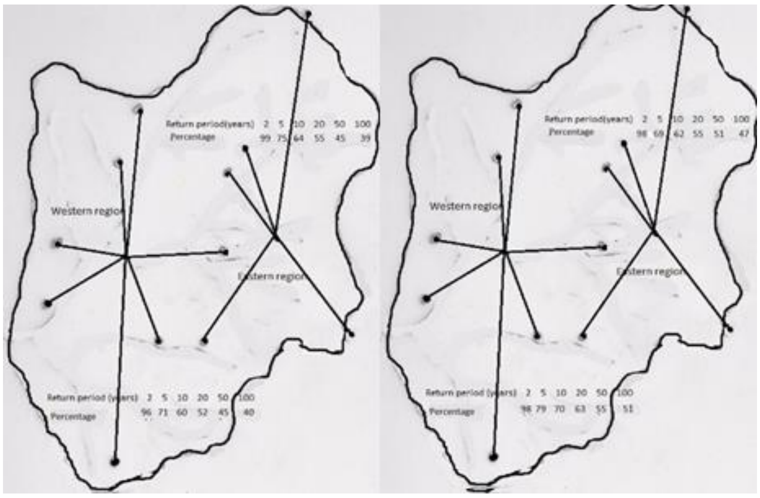

Figure 14. Left: 3 Month Rainfall Starting June as a Percentage of Mean June-August Rainfall. The Return Period and Respective Percentage of Mean Total are Shown for Eastern and Western Regions Right: 3 Month Rainfall Starting October as a Percentage of Mean October-December Rainfall. The Return Period and Respective Percentage of Mean Total are shown for Eastern and Western Regions

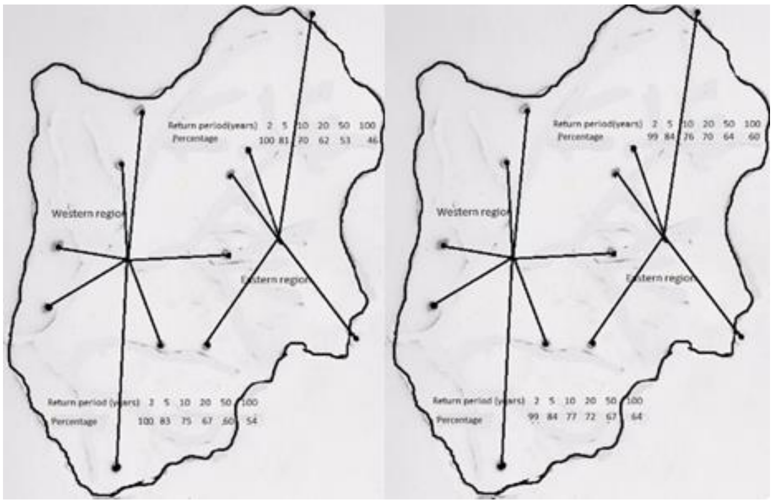

Figure 15. Left: 6 Month Rainfall Starting January as a Percentage of Mean January - June Rainfall. The Return Period and Respective Percentage of Mean Total are shown for Eastern and Western Regions Right: 6 Month Rainfall Starting July as a Percentage of Mean July-December Rainfall. The Return Period and Respective Percentage are Shown for Eastern and Western Regions

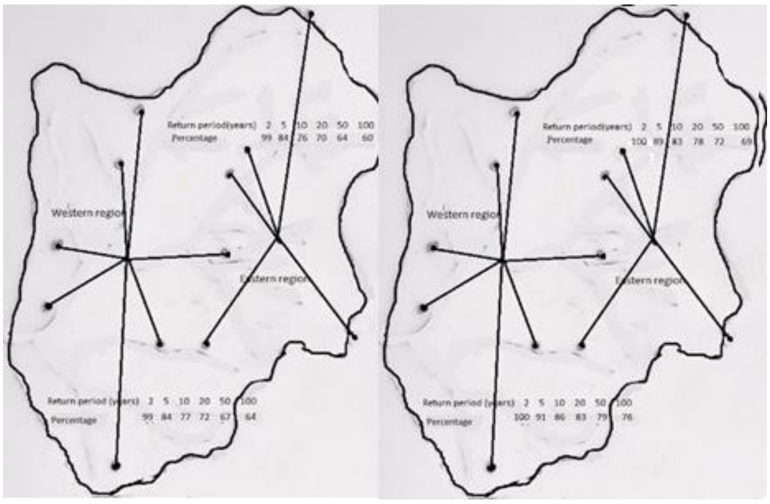

Figure 16. Left: 12 Month Rainfall as a Percentage of Mean 12 Month Rainfall Totals for Period Starting any Month From January Through December Right: 24 Month Rainfall as a Percentage of Mean 24 Month Totals for Period Starting Any Month From January Through December

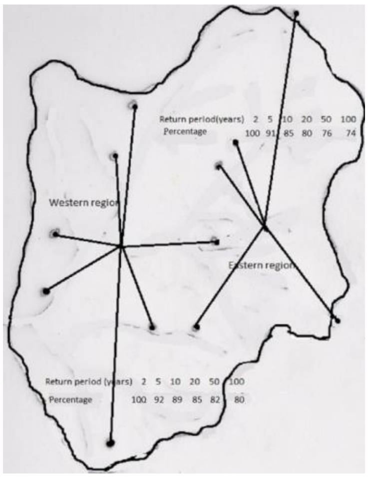

Figure 17. 36 Month Rainfall as A Percentage of Mean 36 Month Totals for Period Starting any Month From January Through December

It is observed that for shorter durations, the rainfall totals expressed as a percentage of mean of the periods vary with the starting month, showing deviation from symmetry of the rainfall for dry seasons. Temporal and spatial patterns in the quantile values expressed as a percentage of mean total becomes evident as the duration decreases from 12 months to 1 month. For durations shorter or longer than 12 months, some of the regions become heterogeneous resulting from the fact that regions were determined from the annual pattern of rainfall rather than from duration and starting time of year rainfall characteristics.

8. Accuracy of the Regional Estimates

The performance of the regional model is assessed using the jack-knife cross - validation (see e.g. Brath et al. 2003).The cross validation procedure allows us to compare the regional and resampled estimates of the annual quantile values (January to December) at all rainfall stations for the two regions taken as an example. Through the comparison the uncertainty of the regional estimates when applied to ungauged sites in the regions can be assessed. The jack-knife resampling method is briefly presented as follows:

1. One of the N rainfallstation, say i, and its corresponding data were removed from the set of N stations;

2. Regional growth factor (expressed as percentage of mean) for the annual rainfall (January to December) was estimated using data from the remaining N-1 stations;

3. Jack-knife quantile value for site i was calculated using the recalculated growth factor from step 2 and the jackknife mean annual rainfall(January to December) estimated from the mean annual rainfall map constructed without station i,i. e. assuming i is an ungauged site

4. Steps 1- 3 were repeated N-1 times, removing in turn one of the remaining stations

The N quantile values estimated using the jack-knife method were then compared with values estimated using the regional procedure. The comparison allows us to assess the robustness and reliability of the regional model through the following measures:

relative bias $$\text{BIAS}=\mathbf{1/N}\sum_{\mathbf{1}}^{\mathbf{N}}\frac{(\mathbf{R}_{\varrho i}-\mathbf{R}_i)}{\mathbf{R}_i} \quad (11)$$

relatve root mean square error $$\text{RMSE}=\sqrt{1/N\sum_1^N\left[\frac{(R_{\varrho i}-R_i)}{R_i}\right]^2} \quad (12)$$

Where $R_{\varrho i}$ is the jack-knife quantile value for station i

R_i is the the quantile value estimated from the regional model for station i

The relative bias and root mean square error were calculated for the Eastern and Western regions for 2 year and 50 year return periods and results are shown in Table 7.

Table 7 Relative Bias and Root Mean Square for Annual Rainfall

Region	Relative bias		Relative root mean square error	
	2 year	50 year	2 year	50 year
Eastern	0.010	-0.03	0.055	0.075
Western	-0.016	-0.017	0.036	0.037

The BIAS is negligible and the RMSE is less than 0.1. The results are considered satisfactory since the RMSE is small.

9. Conclusions

The L-moment method has been used to estimate the quantile values for Selangor rainfall data for 6 durations at 6 probability levels with starting months from January through December. The river basin is divided into 2 homogeneous regions and using the long term rainfall records, satisfactory results were obtained. Using quantile values expressed as a percentage of period mean rainfall presented as drought rainfall indication maps, the rainfall values for a specific return level, duration and starting month at any stations of the region can be estimated. Dimensionless regional quantile values presented in the drought rainfall indication maps will help water managers to assess and monitor historical droughts, and the allocation of water resources in a drought period.

Acknowledgements

We are grateful for being able to use publicly available softwares such as ICI-RAFT, TREND, and the REAL STATISTICS USING EXCEL so that work can be carry out more efficiently.

References

[1] G. C. Blain, "Standardised Precipitation Index Based On Pearson Type Iii Distribution", Rev.Bras. Meteorol Sao Paulo June 2011, vol 26, no 2, **(2011)**.
[2] A. Brath, A.Casterallin and A.Montanari, "Assessing The Reliability Of Regional Depth Duration Frequency Equations For Gauged And Ungauged Sites Water Resources Research", vol. 39, no. 12, **(2003)**.
[3] F, Chiew and L. Siriwardena, Trend User Guide Www.Toolkit.Net.Au/Trend, **(2005).**
[4] J. Giovannettone, F. Paredes, R. Seiler and A.Revelo, "Effectiveness Of A New Method (Software) To Evaluate The Impacts Of The Madden-Julian Oscillation On Rainfall In Central And South America" ,Aqua- Lac-vol 5, -no. 2, **(2013)**.
[5] J. Giovannettone and M. Wright, "Ici-Warm Regional Analysis Frequency Tool(Ici-Raft)", Institute Of Water Resources (Iwr), **(2013)**.
[6] N. B. Guttman, "Regional Precipitation Quantile Values For The Continental United States Computed From L-Moments Journal Of Climate", vol 6, **(1993)**.
[7] J.R.M. Hosking and J.R., "Wallis Regional Frequency Analysis",An Approach Based On L-Moments ,Cambridge University Press, **(1997)**.
[8] J. H. Nunez, K.V. erbist, J.R. Wallis, M.G. Schaefer, L. Morales and W.M. Cornelis, "Regional Frequency Analysis For Mapping Drought Events In North Central Chile", Journal Of Hydrology, **(2011)**.
[9] S. Perica, D. Martin, B. Z. Lin, T. Parzybok, D Riley, M. Yekta, L. Hinder, L.C. Chen., F. Brewer, F. L. Yan, K. Maitaria, C. Trypaluk and G. M Bonnin, "Precipitation Frequency Atlas Of The United States" , Noaa Volume 14 Version 2 Hawaiian Islands, **(2009)**.
[10] C. Zaiontz, Real Statistics In Excel , Http//Www.Real-Statistics.Com, **(2014)**.