# Extreme Value Analysis of Annual Maximum Rainfall Data in Colombo and Anuradhapura

P. P. Witharana[1] and S. S. Wickramasuriya[2]
*Department of Civil Engineering, University of Moratuwa, Sri Lanka*

ABSTRACT: Sri Lanka experiences various natural disasters like floods and landslides, where extreme rainfall is a major contributory factor. In relation to climate change, a crucial question also arises about the possible variability of these extremes. In this study, annual maximum daily rainfall data recorded at Colombo (1901- 2013) and Anuradhapura (1889-2005) meteorological stations from the wet and dry zones respectively are analysed. The major objective of this research is to obtain the best fitting probability distribution for the data, considering the Gumbel, Log normal and Log Pearson type III distributions. From the analysis, rainfall of 25, 50 and 100 year return periods have been estimated. These results have been compared with the values given by the Irrigation Department of Sri Lanka and used in design work. The results of this study are useful for future work related to flood control, disaster management and water resources planning.

## 1 INTRODUCTION

There is great concern that the magnitude and frequency of extreme rainfall events have increased during the recent years, and the resulting floods causing severe damages to life and property all over the world. Hence the analysis of extreme rainfall data using various probability distributions continues to draw the attention of researchers in order to mitigate flood hazards and for the improved design of hydraulic structures such as spillways. (Chu et al., 2012) ; (Koutsoyiannis & Papalexiou, 2013) ; (Kilsby & Serinaldi, 2014)

In relation to Sri Lanka, extreme rainfall pose a major challenge to urban infrastructure and dams, and ensuring their safety in a changing global climate is of critical importance to the country (Ministry of Finance and Planning, 2012). The major flood events of Nov 2010 ( Colombo ) , followed by the flood disasters of Jan- Feb 2011 (eastern and north central provinces ) and December 2014 where critical conditions were encountered at several major dams such as Nachchaduwa, and the Nalanda dam were overtopped once again, should be noted. (Krishantha & Wickramasuriya, 2015); (Wickramasuriya & Fernando, 2012)

## 2 OBJECTIVES OF THE PAPER

- Examine whether there is a change in the rainfall pattern at the given stations.
- Find the best fitting probability distribution curve for annual maximum daily rainfall.
- Examine the relative changes in results from different periods of data and compare with the values derived by the Irrigation Department and used in design.

## 3 METHODOLOGY

### 3.1 *Data*

Two data sets were used for the analysis. A set of 113 values of annual maximum daily precipitation from Colombo (Wet zone) covering year 1901 – 2013, and a set of 107 values from Anuradhapura (Dry zone) covering year 1889 – 2005 were selected. The rainfall data were divided into three sub-sets comprising the first 30 years, the recent 30 years and the remaining values were taken as the mid-value sample.

### 3.2 *Analysis*

The rainfall data were analysed using the Gumbel, Log Normal and Log Pearson type III probability distributions. The respective probability density functions and other properties of these distributions are given in (Kite, 1978) and (McCuen, 2005).

The annual maximum rainfall data in each sample were arranged in descending order and a rank (m) was assigned such that the largest value gets a rank of 1.

1 Former Student 2 Professor e-mail : sunilw@uom.lk

The Weibull plotting position which is widely used was considered in the analysis. Thus,

$$P = \frac{m}{n+1} \qquad (1)$$

and T = 1/P

where $P$ represents the probability of exceedence, T is the return period and $n$ is the sample size.

Essentially the analysis focuses on three aspects. Firstly, the time series plots of the data are subjected to a visual examination to detect the possible existence of temporal changes in the data. This is followed by the statistical tests for checking the stability of data for variance and mean. Secondly, the various return periods of extreme rainfall are estimated using the three probability distributions, followed by a good-of-fit check. Finally these results are compared with those specified by the Irrigation Department (Ponrajah, 1984).

# 4 PRESENTATION AND DISCUSSION OF RESULTS

The important results of the analysis are presented in this section. Figs 1 and 2 show the time series plots of the data.

## 4.1 *Time series plots of the two stations*

The time series plot of Colombo shows that the two largest values have occurred in recent times (1992 and 2010). Here a greater variability of recent extreme values is apparent and this possibly indicates a greater severity and higher frequency of such events in the future. As for Anuradhapura, the analyzed data set shows lesser temporal variation. However the analysis has not included the extreme values of 2011 and 2014 for Anuradhapura.

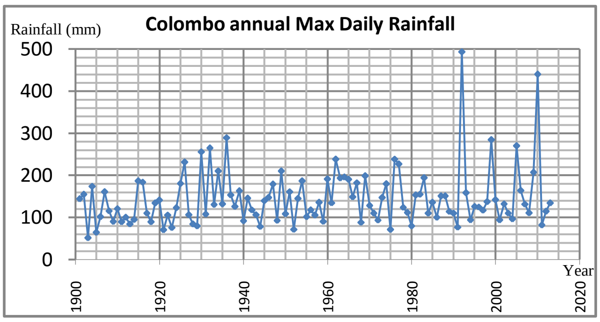


Fig.1 Time series plot of Colombo

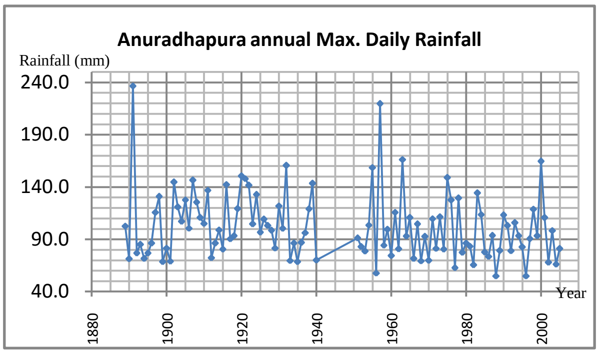


Fig.2 Time series plot of Anuradhapura

## 4.2 *Tests for stability of variance and mean*

These tests are carried out by comparing the first and last 30 year samples, to identify whether there are any statistically significant changes in variance and mean.

For the stability of the variance, the test statistic $\boldsymbol{F_t}$ should satisfy the following condition (Dahmen & Hall, 1990)

$$F_{2.5\%} < F_t < F_{97.5\%} \qquad (2)$$

where, $$F_t = \frac{\sigma_1^2}{\sigma_2^2} \qquad (3)$$

$\sigma_1, \sigma_2$ are the standard deviations of the first 30 year sample and recent 30 year sample and Table 1 shows the results.

Table 1. Results of stability of variance test

| | $F_{2.5\%}$ | $F_t$ | $F_{97.5\%}$ |
|---|---|---|---|
| Colombo | 2.07 | 0.26 | 0.48 |
| Anuradhapura | 2.07 | 1.60 | 0.48 |

Interestingly, these results confirm what was observed visually. As mentioned previously, the data for Colombo does not satisfy the criterion for stability of variance, although it is so in the case of Anuradhapura.

As for the stability of mean, the test statistic $\boldsymbol{t_t}$ should satisfy the following condition (Dahmen & Hall, 1990)

$$t_{2.5\%} < t_t < t_{97.5\%} \qquad (4)$$

$$t_t = \frac{\overline{x_1} - \overline{x_2}}{\left[\frac{(n_1-1)\sigma_1^2 + (n_2-1)\sigma_2^2}{n_1+n_2-2} * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]^{0.5}} \qquad (5)$$

where; $\overline{x_1}, \overline{x_2}$ are mean values of each sample, $n_1, n_2$ are the sample sizes and $\sigma_1, \sigma_2$ are the standard deviations of each sample. Table 2 shows the results of this test.

Table 2. Results of stability of mean test

| | $t_{2.5\%}$ | $t_t$ | $t_{97.5\%}$ |
|---|---|---|---|
| Colombo | -2.04 | -1.84 | 2.04 |
| Anuradhapura | -2.04 | 1.43 | 2.04 |

Hence the stability of mean test condition is satisfied for both data sets.

## 4.2 *Trend Analysis*

According to the trend analysis, for the absence of a significant trend, the test statistic $T_t$ should satisfy the following condition (Dahmen & Hall, 1990)

$$T_{2.5\%} < T_t < T_{97.5\%} \quad (6)$$

$$T_t = R_{sp} \frac{[n-2]^{0.5}}{[1 - R_{sp} * R_{sp}]^{0.5}} \quad (7)$$

where; $R_{sp} = 1 - \frac{6*\sum_{i=1}^{n}(D_i*D_i)}{n*(n*n-1)}$ (8)

and $D_i = Kx_i - Ky_i$ (9)

$Kx_i$ is the rank of variable (rainfall value) $x_i$ which is the chronological order number of the observation, $Ky_i$ is the rank value of the variable after sorting in the ascending order.

Table 3. Results of trend analysis

| | $T_{2.5\%}$ | $T_t$ | $T_{97.5\%}$ |
|---|---|---|---|
| Colombo | -2.04 | 0.29 | 2.04 |
| Anuradhapura | -2.04 | 0.198 | 2.04 |

Table 3 shows the results of the trend analysis and it can be seen that both Colombo and Anuradhapura data could be deemed as trend free.

## 4.3 *Goodness of fit test*

The standard error (***SE***) is calculated according to the following equation,

$$SE = \sqrt{\frac{\sum_{i=1}^{n} D_i^{\,2}}{n-m}} \quad (10)$$

where, $D_i$ is the difference between value derived from the distribution and the respective actual rainfall value, $n$ is the sample size and $m$ is number of estimated parameters. For Gumbel and log normal $m=2$; and for Log Pearson type III distribution $m=3$ (Kite, 1978).

According to the results of Table 4, the best fitting distribution for Colombo is the Log Pearson type III and for Anuradhapura it is the Gumbel distribution. As for Colombo, the considerably large deviation of the result for the last 30 year data set is due to the two extreme values in the sample. Improving the fit in this case would be an aspect for investigation in future research.

Table 4. Comparison of Standard Error values

| **Colombo** | | | |
|---|---|---|---|
| | 1st 30 | mid 53 | last 30 |
| Gumbel | 6.29 | 21.33 | 47.84 |
| Log normal | 8.41 | 4.98 | 50.53 |
| Log Pearson | 7.67 | 6.63 | 36.49 |
| **Anuradhapura** | | | |
| | 1st 30 | mid 47 | last 30 |
| Gumbel | 11.08 | 5.31 | 2.65 |
| Log normal | 12.80 | 7.05 | 4.09 |
| Log Pearson | 10.53 | 5.93 | 3.82 |

## 4.4 *Predicted annual maximum rainfall values*

Estimated annual maximum rainfalls of return periods 25, 50 and 100 years are made in accordance with the best fitting distribution for each station, using the three samples. These have then been compared with the values given by Irrigation Department (Ponrajah, 1984) (Figs 3-6).

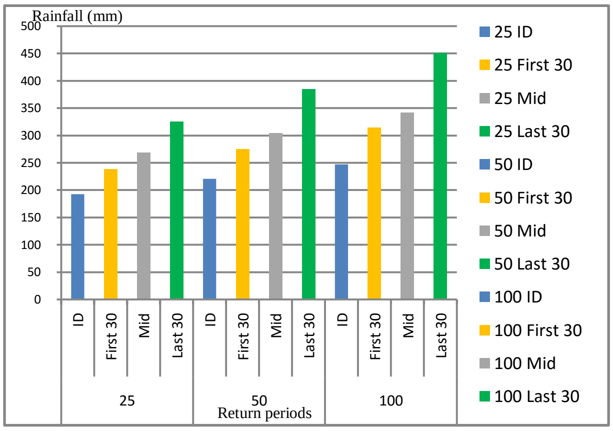


Fig. 3 Annual maximum daily rainfall value comparison for Colombo

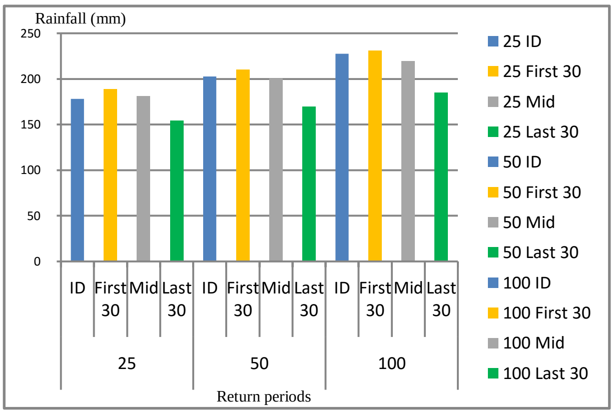


Fig.4 Annual maximum daily rainfall value comparison for Anuradhapura station

"ID" refers to the values derived by the Irrigation Department.

When observing the values for Colombo, it can clearly be seen that the last 30 year sample gives higher annual maximum rainfall values compared to the other two samples and the values predicted by the Irrigation Department. The estimated rainfall for the 100 year return period using the last 30 year sample is about 200mm larger than the ID estimate. (Fig 3,5) However recent studies done by the Irrigation Department confirm these findings. (Hettiarachchi, 2012)

As for Anuradhapura, the first 30 year and mid-value samples give higher estimates when compared to the last 30 year sample. Furthermore, there is also a lesser deviation when compared with values derived by the Irrigation Department.(Fig. 4,6)

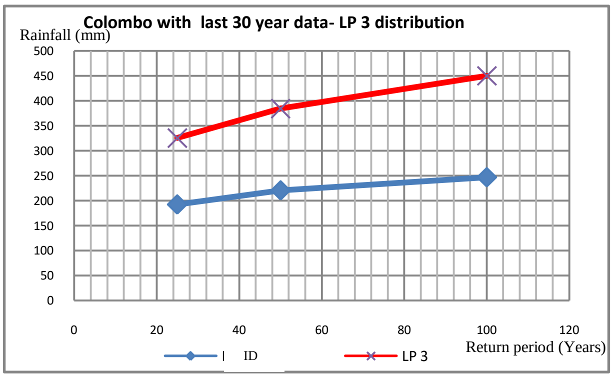


Fig.5 Comparison with Irrigation Department Estimates

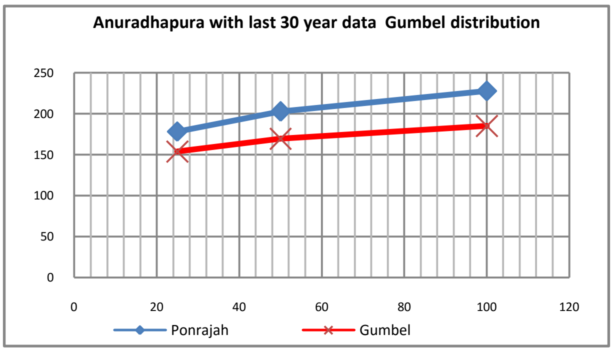


Fig.6 Comparison with Irrigation Department Estimates

## 5 CONCLUSION

- The best fitting distribution curve for Colombo (wet zone) annual maximum rainfall data is the Log Pearson type III and for Anuradhapura (dry zone), the Gumbel distribution gives the best fit.
- Examination of the time series plots show that extreme rainfalls of Colombo indicate a change in variability. This is most likely because of two extreme events which have occurred in the recent past. As for Anuradhapura, the data does not show such a significant change in pattern and demonstrates stability in mean and variance.
- The reason for changes in the rainfall pattern in Colombo could possibly be due to the effects of climate change.
- The estimated rainfalls for Colombo for 25, 50 and 100 year return periods are much larger than the values estimated according to the publication "Design of Irrigation Headworks for Small Catchments", which is published by the Irrigation Department. A re-evaluation of the various coefficients given in this publication, considering recent extreme events is essential. This is further confirmed by the findings reported in the recent Hydrological Annual 2011/2012.
- Considering the recent extreme rainfall events in Colombo, one can expect more frequent extreme values of precipitation and hence flood mitigation measures should be taken to reduce the negative impacts of flooding and minimize damage to property. In relation to dam safety continuous monitoring and analysis of extreme rainfalls is essential.

## REFERENCES

Anon., 2012. *Ministry of Finance and Planning "Annual Report"*. Annual Report. Ministry of Finance and Planning. n.d. http:/www.treasury.gov.lk.

Chu, L.F., McAleer, M. & Chang, C.-C., 2012. *Statistical Modelling of Extreme Rainfall in Taiwan*. Thaiwan: National Science and Technology Center for Disaster Reduction.

Dahmen, E.R. & Hall, M.J., 1990. *Screening of hydrological data*. Wageningen: International Institute for Land Reclamation and Improvement.

Hettiarachchi, P., 2012. *Hydrological Annual 2011/2012*. Hydrology Division, Irrigation Department.

Kilsby, C. & Serinaldi, F., 2014. Rainfall extremes: Toward reconciliation after the battle of distributions. *Water Resources Research*, 50(1), pp.336-52.

Kite, G.W., 1978. *Frequency and Risk Analyses in Hydrology*. 2nd ed. Colorado: Water Resources Publications.

Koutsoyiannis, D. & Papalexiou, S.M., 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49(1), pp.187–201.

Krishantha, R.A.R.V. & Wickramasuriya, S.S., 2015. Climate Change Adaptation: A Hydraulic Model Study to Improve the Spillway Discharge of Giritale Reservoir. In *MERCON, University of Moratuwa.*., 2015.

McCuen, R.H., 2005. *Hydrologic analysis and design*. 3rd ed. New Jersey: Pearson Prentice Hall.

Ponrajah, A.J.P., 1984. *Design of Irrigation Headworks for Small Catchments*. 2nd ed.

Wickramasuriya, S.S. & Fernando, W.C.D.K., 2012. Challenges in dam safety and extreme rainfall estimation in relation to Sri Lanka. *Engineer*, xxxxv(no 1), pp.39-49.