4th International Conference on Water Resources and Arid Environments (ICWRAE 4): 113-121
5-8 December 2010, Riyadh, Saudi Arabia

Design Rainfall Estimation for Short Storm Durations Using L-Moments and Generalised Least Squares Regression-Application to Australian Data

Khaled Haddad[1], *Ataur Rahman*[1], *Janice Green*[2] and *George Kuczera*[3]

[1]University of Western Sydney, Australia
[2]Australian Government Bureau of Meteorology
[3]University of Newcastle, Australia

Abstract: Short duration design rainfall intensity data is a basic input to many water related projects, particularly in urban drainage design. The main problem in deriving short duration design rainfall data is the lack of continuous pluviograph rainfall data. The use of regionalisation techniques is common in design rainfall estimation, which pools data from a large region. This paper presents a regional rainfall estimation method in Australia involving Generalised Least Squares Regression (GLSR) and *L* moments methods. It has been found that the GLSR and *L* moments based regional approach is a practical means for estimating short duration design rainfalls and that this provides an efficient statistical framework for assessing uncertainties in design rainfall estimates.

Key words: Design rainfalls • *L* moments • Generalised least squares • Australia

INTRODUCTION

Design rainfall data in the form of intensity-frequency-duration (IFD) curves are an important input in many water related projects. Short duration design rainfall estimates are particularly useful for urban drainage design tasks where catchment response time is quite short. The main problem in estimating short duration design rainfall data is the poor density of pluviograph stations as well as short record length at individual sites. Problems in estimating short duration design rainfall data have been discussed by many researchers such as [1-4]. The most common approach to deal with insufficient data in flood and rainfall analyses is the application of regional frequency analysis which pools data from many stations in the region. An index rainfall method is a widely adopted regional approach which is based on a regional growth factor and prediction equations for moments of various orders (such as coefficient of variation and skewness) and index rainfall. Use of Ordinary Least Squares (OLS) and Partial Least Squares (PLS) regression is common with the index rainfall approach. Some of the problems however with the OLS and PLS regressions are that they ignore the inter-station correlation and variation in record lengths from site to site. In contrast, use of Generalised Least Squares Regression (GLSR) avoids some of these problems [5].

This paper presents a framework for regional analysis to estimate short duration design rainfalls (6 minutes and 1 hour durations). This regional framework makes use of GLSR to relate *L* moments and index rainfall to climatic and physiographic characteristics. To identify a regional parent distribution, two goodness of fit tests are applied, the Akaike information criterion and Bayesian information criterion [6]. The study uses data from Australia where design rainfall estimates were obtained about 20 years ago [7]. There has been a recent initiative to derive new design rainfall estimates in Australia based on a national database. This study is aimed to explore the potential of the GLSR and *L* moments based index rainfall method in the derivation new design rainfalls for Australia.

Study Area and Data: The study uses data from 203 rainfall stations across Australia as described in [2]. For the purpose of this study, the data set was divided into a number of subsets as summarised below:

Corresponding Author: Khaled Haddad, University of Western Sydney, Australia.

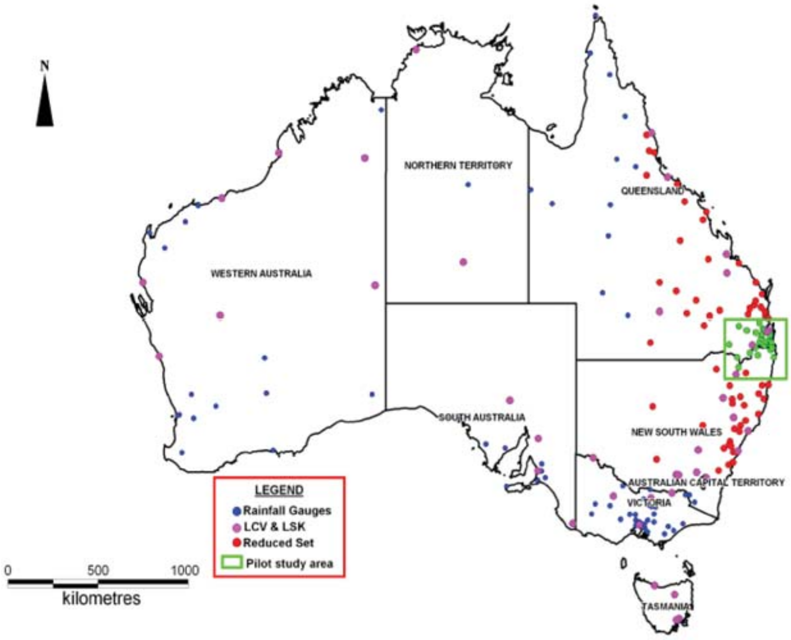


Fig. 1: Location of stations (pink) used in developing GLSR models to predict L-CV and L-SK at 1 hour and 6-minutes durations

- Data subset 1: 98 stations with minimum record length of 30 years (reduced set);
- Data subset 2: 40 stations having at least 45 years of data;
- Data subset 3: 30 stations from pilot study area (green box area in Figure 1);
- Data subset 4: 20 stations randomly selected from data subset 1 for model testing; and
- Data subset 5: 5 stations selected randomly from data subset 4 for model testing.

To enhance the stability of the *L* coefficient of variation (L-CV) and *L* coefficient of skewness (L-SK) estimates (from the annual maximum rainfall depth event series for a given duration), stations with a minimum record length of 45 years were adopted to estimate these statistics (data subset 2 containing 40 stations). The rainfall event durations considered in this study were: (a) 6 minutes (sub-hourly duration) and (b) 1 hour (sub-daily duration). The predictor variables considered in the GLSR were: (a) latitude (*lat*); (b) longitude (*long*); (c) distance from coast (*dcoast*); and (d) rainfall statistics such as 24 and 1 hours duration mean rainfall event depth (*mean_24hr*, *mean_1* hour) and the L-CV and L-SK of the 24 and 1 hours duration rainfall event depth.

MATERIALS AND METHODS

Description of GLSR: The regression model based on GLSR c an explicitly account for sampling uncertainty and intersite dependence. A brief description of the GLSR model is presented below.

Let $\hat{y}_i$ by an estimate of annual maximum series (AMS) parameter at station *i*. The following linear relationship is considered:

$$\hat{y}_i = \beta_0 + \sum_{k=1}^{p} \beta_k X_{ik} + \varepsilon_i + \delta_i \tag{1}$$

Where X_{ik} are predictor variables (climatic and physiographic characteristics), β_k are the regression coefficients, ε_i is the random sampling error and δ_i is the residual model error. To evaluate Eq. (1), the covariance structure of the sampling error must be known. The sampling error variances of the AMS parameters can be estimated by Monte Carlo simulations [8]. Estimates can be derived for the sampling error variances (diagonal of error covariance matrix) by substituting the population parameters by the sample estimates. It must be noted though, to solve the GLSR equations, the error covariance estimator should be independent, or nearly so, of the AMS parameter estimate $\hat{y}_i$ [5]. Following a similar approach as outlined by [8], estimates of the sampling error variance that is nearly independent of the three AMS parameters is summarised below.

For the index rainfall estimation (this was derived as the average of the annual maximum series at a site for a given duration), the sampling error variance is given by $\sigma_{\varepsilon i}^2 = \sigma_i^2 / n_i$ where σ_i^2 is the population variance. A reasonable estimate of $\sigma_{\varepsilon i}^2$ can be obtained from the average of the variance values. For estimation of the sampling error variance of the L-CV and L-SK estimators, a simple Monte Carlo simulation is carried out in the following way. For each of the 40 stations used in deriving L-CV and L-SK, the population parameter values are estimated for the selected distribution (GEV or GP); using these parameters and for each of the stations, 10,000 sets of values of AMS data equal to the observed record lengths are simulated. From these simulations, the variances of the L-CV and L-SK estimates are calculated. The GEV distribution was considered inappropriate for use with the 6 minutes duration as noted in [3]. As examined in Section 4 the GP distribution has been found to fit the observed data for 6 minutes duration reasonably well. A discussion on the applicability of the GP distribution is given in the results section of the paper. For the estimation of inter-site correlation for various parameters, we consider concurrent records of annual maximum rainfall series across all the sites within a selected region. The inter-site correlation between the sample mean values $\rho_{\mu ij}$ is equal to the correlation coefficient between concurrent rainfall events of sites i and j. The correlation between higher order sample moments depends on the order of the moment [9,8]. For the L-CV and L-SK estimates, the inter-site correlation coefficient is approximated by $\rho_{\tau_2 ij} = \rho_{ij}^2$ and $\rho_{\tau_3 ij} = \rho_{ij}^3$ respectively. The estimated cross correlation coefficients have reasonably large sampling uncertainties associated with them, especially if the concurrent record length is small. Relatively better estimates of cross correlation can be found when the sample cross correlation coefficients are smoothed by relating them to the distance between sites. In this study an exponential correlation function is used:

$$\rho_{ij} = \psi^{\left[\frac{d_{ij}}{\omega d_{ij} + 1}\right]} \quad (2)$$

Where d_{ij} is the distance between sites i and j and Ψ and ω are parameters estimated from the data.

Goodness of Fit Tests: For the 1 hour duration, the GEV distribution was adopted based on the findings of previous studies [2]. It was found that the L-SK model did not provide a reasonable overall fit to the 6 minutes duration data and hence it was decided to search among the one and two-parameter distributions for a suitable parent distribution. The Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) are used to evaluate the goodness-of-fit of four candidate two-parameter distributions.

The AIC has been used in hydrological applications to select the flood frequency model (e.g. [10,11,6]).; However in this study the second order variant of AIC, called AIC_c is used and is given by Eq. (3), where n is the sample size and P is the number of parameters of the desired probability distribution. AIC_c accounts for the biases in smaller sample sizes. As reported, AIC_c should be used when $n/p < 40$ to avoid bias [11].

$$AIC_c = -2\Pi(Y) + 2P \frac{n}{(n - P - 1)} \quad (3)$$

In practice, after the computation of the *AIC*, for all of the operating models, one selects the model with the minimum AIC value, AIC_{min}. The Bayesian Information Criterion (BIC) is very similar to the AIC, but is developed in a Bayesian framework:

$$BIC = -2\Pi(Y) + \ln(n)P \quad (4)$$

The BIC penalizes more heavily small sample sizes and models with high values of P. Since $\Pi(Y)$ depends on the sample, the candidate models can be compared using AIC and BIC only if fitted on the same sample. In this study the competing models are fitted to the same samples; the candidate distributions (models) are the generalised pareto (GP), gamma (GAM), extreme value type 1 (EV1) and exponential (EXP). The above criteria were used to find the best fitting two parameter frequency distributions for the 6 minutes duration design rainfall.

Estimation of Regional Growth Curves: Using the mean value as the index-rainfall parameter, the regional T-year event estimator can be written:

$$\hat{I}_{T_i} = \hat{\mu}_i \hat{z}_T \quad (5)$$

$$\hat{z}_T = 1 + \frac{\hat{\tau}_2^R}{1 - 2^{-\hat{k}^R}} \left[1 - \frac{\exp(-\hat{k}^R y_T)}{\Gamma(1 + \hat{k}^R)} \right]$$

Where,

$\hat{I}_{Ti}$ = Design rainfall intensity for a given duration *i* and *T*;

$\hat{\mu}_i$ = Regional mean rainfall (index rainfall) for a given duration *i;* and

$\hat{z}_T$ = Regional growth factor, regional inverse CDF (by either GEV, or GP) and

y_T = -ln[-ln(1-1/T)] is the gumbel reduced variate. For the two parameter GP distribution the regional $\hat{\tau}_2^R$ value is usually evaluated as the weighted regional value and the regional shape parameter $\hat{k}^R$ is estimated. However in this study $\hat{\tau}_2^R$ is found by GLSR and $\hat{k}^R$ is estimated and used in Eq. (5). For the three parameter GEV distribution [12] used regional weighted averages of *L*-moment ratios where $\hat{\tau}_3^R$ is usually estimated as a weighted regional value. In this study $\hat{\tau}_3^R$ is found by GLSR, c^R is estimated by Eq. (6) and the regional shape parameter $\hat{k}^R$ is found by Eq. (7), $\hat{\tau}_2^R$ is then estimated from GLSR and Eq. (5) is evaluated.

$$c^R = \frac{2}{\hat{\tau}_3^R + 3} - \frac{\ln(2)}{\ln(3)} \quad (6)$$

$$\hat{k} = 7.8590c^R + 2.9554c^{2R} \quad (7)$$

Model Performance: In developing the GLSR models, the final choice of equation was based on pseudo R^2_{GLS} values [12] and a range of graphical statistical diagnostics. It should be noted that the R^2_{GLS} differs from the R^2 of OLSR. In GLSR, the sampling error variance is partitioned from the total error variance. In this study a framework is presented where the new estimates at 6 minutes and 1 hour durations (for *T* of 5, 20 and 100 years) for the five stations chosen randomly are compared to previous Australian estimates (ARR87 estimates) and at-site estimates. This is considered as an approximate guide to validate the new estimates. At-site estimates were derived by fitting a GEV distribution (1 hour duration) and GP distribution (6 minute duration) to the annual maximum data series. To assess relative differences Eq. (8) was applied.

$$RE\ (\%) = \left(\frac{I_{inew} - I_{iat-site}}{I_{iat-site}} \right) \times 100 \quad (8)$$

The "median" relative error %" represents the percentage error between the at-site and new quantiles, standardised by the at-site estimate. To verify whether the new estimates were over or under estimating, the median relative errors were standardised to a normal score by Eq. (9).

$$Z = \left(\frac{M_{RE\%} - M_{aveRE\%}}{M_{stdevRE\%}} \right) \quad (9)$$

$M_{RE\%}$ represents the median relative error at a specific duration and specific *T*, $M_{aveRE\%}$ is the average of the median relative errors over all the cases and $M_{stdevRE\%}$ is the standard deviation of the median relative errors over all the cases.

The advantage of using the GLSR is that it provides an estimate of the AMS parameters and their associated variances. The variance reflects the uncertainty related to regional heterogeneity as well as sampling uncertainty corrected for inter-site correlation. In the final part of the validation we make use of the variance estimated for each site to examine whether the differences in the new and ARR87 estimates are statistically significant. The differences between the two estimates are compared with the estimation uncertainties. The following statistic as shown in [4] Eq. (10) is calculated.

$$S = \frac{\hat{I}^N_{ARI} - \hat{I}^{ARR87}_{ARI}}{\sqrt{0.5(\text{var}\{\hat{I}^N_{ARI}\} + \text{var}\{\hat{I}^{ARR87}_{ARI}\}}} \quad (10)$$

Where $\hat{I}^{ARR87}_{ARI}$ and $\hat{I}^N_{ARI}$ are the regional *T* year event estimates based on respectively, the ARR87 and the new estimates derived in this study and $\text{var}\{\hat{I}^N_{ARR}\}$ is the corresponding estimated variance from GLSR and $\text{var}\{\hat{I}^{ARR87}_{ARI}\}$ is the corresponding estimated variance for the ARR87 estimates found by Monte Carlo simulation. The *S*-statistic can be interpreted statistically by comparison with the quantiles in a standard normal distribution.

RESULTS

Intersite Correlation Analysis: For the index rainfall, the intersite correlation decreases for increasing distance between stations. Also the correlation structure was seen to depend on the considered duration, the correlation being higher for the 1 hour as compared to the 6 minute duration. The differences maybe due to different causative meteorological mechanisms between the two durations.

Table 1: GLSR prediction equations for index rainfall, L-CV and L-SK (1 hour and 6 minutes durations)

Duration	Index rainfall (mean)	$\hat{\sigma}_{\delta}^{2}$
1 hour	log(*mean_1hr*)=1.54+0.016(*lat*)+0.019(*long*)+0.0011(*mean_24hr*) L-CV	0
1 hour	L-CV_1hr = 0.23+0.860(*L-CV_24hr*)-0.09(*L-SK_24hr*)-0.004(*lat*)+0.011(*dcoast*) L-SK	0
1 hour	L-SK_1hr = 0.23+0.746(*L-CV_24hr*)+0.057(*L-SK_24hr*)-0.0093(*lat*)+0.0003(*dcoast*)	2.83E-03
Duration	Index rainfall (mean)	$\hat{\sigma}_{\delta}^{2}$
6 min	log(*mean_6min*)=1.05+0.005(*lat*)+0.011(*long*)+0.005(*mean_1hr*) L-CV	0
6 min	L-CV_6min = 0.095+0.614(*L-CV_1hr*)-0.074(*L-SK_1hr*)-0.0028(*lat*) -0.00082(*dcoast*) L-SK	0
6 min	L-SK_6min = 0.22+0.040(*L-CV_1hr*)+0.410(*L-SK_1hr*)-0.0034(*lat*)-0.0021(*dcoast*)	0

Table 2: R^2_{GLS} values for index rainfall, L-CV and L-SK models (1 hour and 6 minutes)

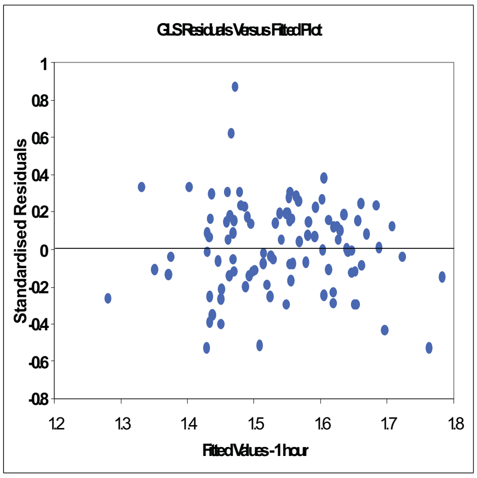
Duration	1 Hour	6 Minutes
Index rainfall	86%	75%
L-CV	68%	55%
L-SK	36%	34%

Development of Prediction Equations: The dependant variable (rainfall depth for a given duration) was transformed using a logarithm base 10 transformation and the independent variables were centered to mean zero. In the GLSR to select the predictor variables, a method similar to 'backward stepwise regression' was adopted where variables were entered/removed stepwise to achieve minimum model error variance. For the index rainfall models, data subset 1 was used that consisted of 98 stations. For the L-CV and L-SK models, data subset 2 was used that consisted of 40 stations. Table 1 shows the derived prediction equations for the index rainfall for the 1 hour duration along with the associated residual variances. The 1 hour estimates (from GLSR) were used as a predictor variable in developing prediction equation for the 6 minute duration event. Table 1 also presents the prediction equations for index rainfall, L-CV and L-SK for the 6 minutes duration. The R^2_{GLS} values of index rainfall, L-CV and L-SK models are shown in Table 2. As observed in Table 2, the 6 minutes duration index rainfall model is performing reasonably well as indicated by reasonably high R^2_{GLS} (75%) which is 11% lower than that of the 1 hour duration R^2_{GLS} (86%). The L-CV model has only provided a R^2_{GLS} of 55% which is relatively lower than the 1 hour estimate (R^2_{GLS} of 68%). The L-SK model shows the poorest fit with a R^2_{GLS} value of 34%. Table 1 also shows that the residual variances of the models are either zero or close to zero, which indicates that the sampling error has most probably dominated the total error in the estimation of the index rainfall, L-CV and L-SK.

The behaviour of the residual errors associated with the prediction equations were examined for the AMS parameters. These plots are important in assessing the adequacy of the GLSR models. Often in hydrological regression the normality of residuals and homogeneity of variance are assumed (homoskedastic errors). However, this assumption is often violated. If the prediction equations developed using GLSR are to perform well, the residuals should follow an approximate normal distribution, with no major outliers being present. The standardised theoretical estimates (derived from GLSR models) should closely match the standardised sample estimates; on the QQ plot this should show as an approximate slope of 1 and an intercept of 0. (i.e. mean zero and variance 1). The standardised residual vs. fitted values plots did not show any particular pattern nor did it show any true outlier site having undue influence on the regression (Figure 2 shows an example plot). The *QQ* plots showed that the residuals were approximately normally distributed with intercept close to 0 and slope of approximately 1 (Figure 3 shows a sample plot). Furthermore, the R^2 values between the sample and theoretical estimates were quite high (0.93 to 0.98). This suggests that the GLSR models have performed reasonably well in the estimation of the index rainfall, L-CV and L-SK for the 1 hour duration. For the 6 minute duration index rainfall, the standardised residuals vs. fitted values plots (a sample plot in Figure 4) revealed no true outlier site.

The QQ plots (a sample plot in Figure 5) showed that the slopes were close to 1. It can be observed from Figures 3 and 5 that the slopes of the 6 minute and 1 hour duration were the same, which show that there are no gross errors in the estimation of the 6 minute index rainfall using the 1 hour estimates. The standardised residuals vs. fitted values plot for L-CV did not show any outliers and the QQ plot had a slope approximately 0.97.

For L-SK, the standardised residuals showed both low and high outliers, with standardised residual values sightly above 2 and below - 2. In the QQ plot, the slope was approximately 0.97, similar to L-CV model. The QQ plot however revealed both high and low outliers, which suggests that GLSR has not provided a good fit



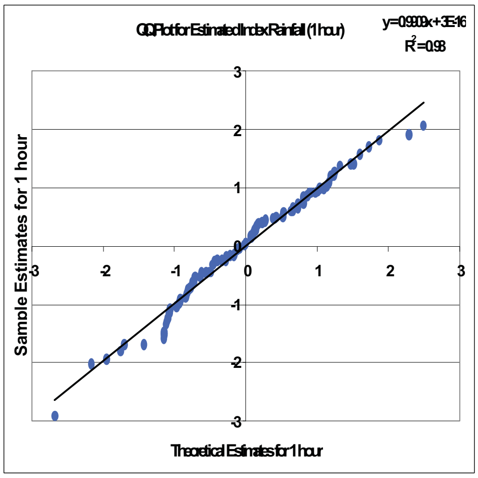
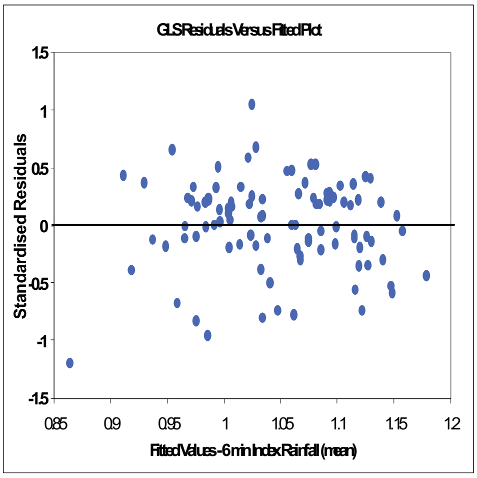


Fig. 2 (left): Standardised residuals vs. fitted values for the index rainfall – 1 hour duration
Fig. 3(right): Standardised theoretical vs. sample estimates for the index rainfall – 1 hour duration



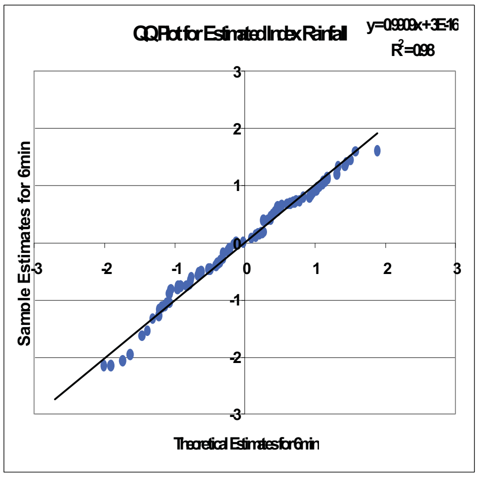


Fig. 4 (left): Standardised residuals vs. fitted values for index rainfall – 6 minutes duration
Fig. 5(right): Standardised theoretical vs. sample estimates for index rainfall – 6 minutes duration

for L-SK model. This result thus provided the motivation to explore the one- and two-parameter distributions to fit the index rainfall and L-CV for 6 minutes duration. Overall the important properties discussed in Section 4.2 enable the derivation of the parameters of all the shorter durations (i.e. 12, 18 and 30 minutes), utilising the longer durations. Such an approach however is feasible only if the statistical diagnostics show no major gross errors.

Selection of a Suitable Distribution for 6 Minute Rainfall Duration: The values of AICc and BIC were examined and the distribution that achieved the best minimum scores, considering all the tests, were taken as the parent distribution. The AIC_c test favoured the GP distribution for 49 stations out of 98 (i.e. 50% cases), the EV1 distribution for 18 stations (18%), the GAM distribution for 18 stations (18%) and the EXP distribution for only 13 stations (13%). The BIC test favoured the GP distribution for 42 stations out of 98 (43%), the GAM distribution for 27 stations (28%), EXP distribution for 17 stations (17%) and EV1 distribution for 12 stations (12%). Considering all the tests (294 cases), the EXP distribution shows the poorest fit while the GP distribution shows the best fit. The results are summarised as a plot in Figure 6. After tallying the results each distribution was ranked from 1 to 4. The GP was ranked as one, while the EXP was ranked 4. Since the GP distribution showed the best fit among the distributions considered here, it is was adopted as a regional parent for the 6 minutes duration rainfall.

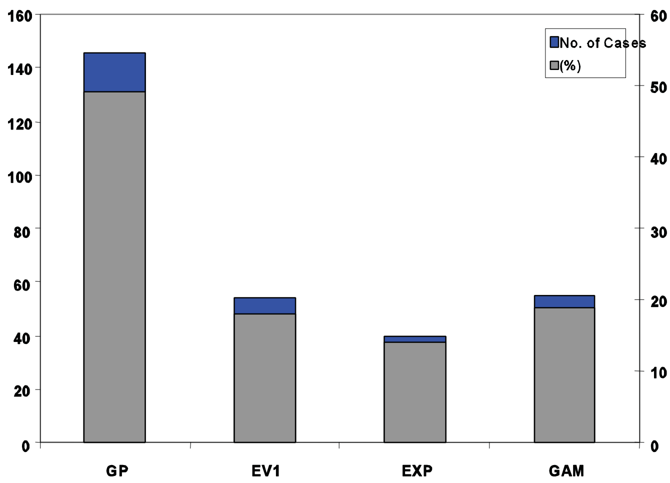


Fig. 6: Summary of goodness of fit test results for the 6 minutes duration event

Estimation of Regional Growth Factors: Differences in regional growth curves as compared to at- site growth curves for both the 1 hour and 6 minute durations were undertaken for T of 10 and 20 years. It should be noted though that the at-site estimates are considered useful approximations against which to make comparisons and should not be considered as a benchmark in validating any regional estimates, as the at-site estimates are subject to sampling uncertainties.

Typical growth factors for the 1 hour duration and T of 10 and 20 years for sub-daily durations are 1.6 and 1.9 respectively. The average differences between the regional and the at-site growth factor estimates for T of 10 years for 1 hour duration were very small (approximately 0.078). Typical growth factors for T of 10 and 20 years for the 6 minutes duration ranged from 1.4 to 1.6. Average differences between at-site and regional growth factor estimates for T 10 and 20 years based on 30 stations in the pilot study area (see Figure 1) were found to be 0.31 and 0.21, respectively. These differences cannot be considered to be extremely high but are typical for this sort of regionalisation approach. Figure 7 shows the boxplots of the differences in regional growth factors as compared to at-site growth factors for the 6 minutes duration and for T of 10 and 20 years.

Quantile Estimates: To derive new rainfall quantile estimates Eq. (5) was utilised. In applying this equation, $\hat{\mu}_i$ (the mean rainfall intensity value for a given duration at site i) was obtained from the developed prediction equations (Table 1) and $\hat{z}_T$ (the regional growth factor) was obtained from the fitted regional distribution (GEV for the 1 hour duration and GP for 6 minutes duration). Quantile estimates obtained from the proposed method were compared with the ARR87 and at-site estimates. Five stations were selected randomly for the comparison (data subset 5).

The selected T's and durations for comparison were respectively 5, 20 and 100 years and 6 minutes and 1 hour. In deriving the ARR87 estimates, for achieving consistency with ARR87, the Log Pearson Type 3 (LP3) distribution was fitted using the method of product moments (MOPM) to the at-site data up to the year 1983 for durations of 6 minutes and 1 hour. The relative differences were then plotted as boxplots to verify if the new estimates were over or under estimating the at-site estimates.

The developed boxplots revealed a pattern where for the Ts of 5 and 20 years, there tended to be an increase in the new estimates for the 1 hour duration and a decrease in the 1 hour duration for the 100 year event. It was also observed for the 6 minutes duration that the new method underestimates the at-site quantile, in particular for Ts of 5 and 20 years. However, it is interesting to note that the 6 minutes duration quantile showed less uncertainty at T of 100 years. The uncertainty was more pronounced for the 5 year event. It was also observed that the higher the T, the smaller the median relative error. Considering both the 1 hour and 6 minute durations over all the ARIs, the new estimates are slightly higher (based on median difference) as compared to the at-site estimates. This analysis concluded that there is definitely an over estimation of the quantile for the new method, as there were more positive Z scores than negative ones. This however was more pronounced for the 5 years ARI and appeared to become lower as the ARI increases.

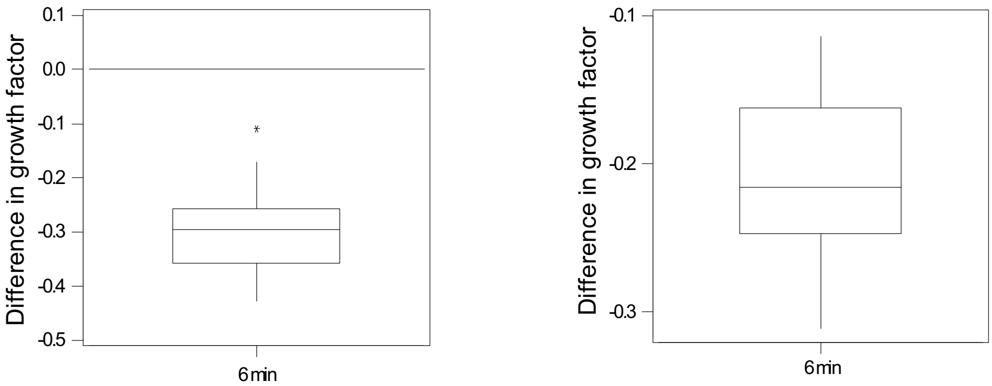


Fig. 7: Differences between regional and at-site growth factor estimates for 6 minutes

The new design rainfall estimates for all five stations were compared with ARR87 estimates for 3 different *T*'s being 5, 20 and 100 years. The comparison is made by dividing the new estimates with estimates from ARR87. Taking a ratio of 1 to depict no over/under estimation (perfect model), while values lower than 1 depict under estimation and values greater than 1 depict over estimation. Utilising this framework, in general, it was found there is a small increase in the new estimates for the 6 minutes duration for all the *T*s considered here. For the 1 hour duration, there is a slight under estimation for the 5 year event; however it is not greatly pronounced. The calculated *S*-statistics for the 1 hour and 6 minute, durations for three *T*'s of 5, 20 and 100 years revealed no significant differences in considering all the ARIs. Thus the null hypothesis that the two estimates (new and ARR87) are identical on average over the five stations cannot be rejected at a 5% level of significance.

CONCLUSIONS

The Generalised Least Squares Regression (GLSR) method has been applied to develop regional prediction equations for the index rainfall, *L* coefficient of variation (L-CV) and *L* coefficient of skewness (L-SK) for 6 minutes and 1 hour rainfall event durations. The GLSR-based design rainfall estimates have been compared with the ARR87 and at-site estimates. The GLSR-based prediction equations perform relatively well for both the 6 minutes and 1 hour durations, while also satisfying the underlying model assumptions adequately. Also there was no true outlier site in the model diagnostic plots. In the GLSR modelling, the sampling error has dominated the analysis masking the true uncertainty associated with the model error, which needs further investigation.

Goodness of fit tests were carried out to find a regional parent distribution for the 6 minute duration rainfall. It has been found that the two-parameter generalised pareto distribution approximates the at-site data reasonably well as indicated by two goodness of fit tests based on the Akaike Information Criterion and Bayesian Information Criterion. Regional growth factors were derived for the 6 minutes duration based on the GLSR estimates of L-CV. The regional growth factors as compared to at-site growth factors showed no major differences. Based on the regional growth factors, design rainfall estimates were calculated at 5 randomly selected stations for average recurrence intervals of 5, 20 and 100 years. Comparisons were made between the new design rainfall estimates, at-site and ARR87 estimates and the uncertainty associated with the new estimates was examined.

Based on the results from 5 test stations, the new design rainfall estimates were found to be slightly higher than the at-site estimates. The new design rainfall estimates showed some differences from the ARR87 estimates; statistical tests however revealed that these differences, on average, are not statistically significant at the 5% significance level. The GLSR modelling framework can be utilised to quantify uncertainty with the design rainfall estimates in a more efficient manner than traditional approaches. Hence, the GLSR modelling framework has the potential to provide more accurate and consistent design rainfall estimates plus associated uncertainty as compared to the current practice in ARR87. However, further testing on a larger data set should be carried out to confirm this.

REFERENCES

1. Nguyen, V.T.V., T.D. Nguyen and F. Ashkar, 2002. Regional frequency analysis of extreme rainfalls. Water Science and Technol., 45(2): 75-81.

2. Jakob, D., B. Taylor and K. Xuereb, 2005. A pilot study to explore methods for deriving design rainfalls for Australia – Part 1, Hydrology Report Series, HRS Report No, 10.
3. Jakob, D., J. Meighen, B. Taylor and K. Xuereb, 2008. Methods for deriving design rainfall estimates at sub-daily durations. Conference proceedings of Water Down Under 2008, 14-17 April 2008, Adelaide, Australia, pp: 11.
4. Madsen, H., K. Arnbjerg-Neilsen and P.S. Mikkelsen, 2009. Update of regional intensity-duration-frequency curves in Denmark: Tendency towards increased storm intensities. Atmospheric Res., 92: 343-349.
5. Stedinger, J.R. and G.D. Tasker, 1985. Regional hydrologic analysis, 1.Ordinary, weighted and generalised least squares compared, Water Resources Res., 2209: 1421-1432.
6. Laio, F., G. Di Baldassarre and A. Montanari, 2009. Model selection techniques for the frequency analysis of hydrological extremes. Water Resources Res., Vol. 45: W07416, doi:10.1029/2007WR006666.
7. Institution of Engineers Australia (I.E. Aust.), 1987, 1999. Australian Rainfall and Runoff: A Guide to Flood Estimation. Vol. 1, I.E. Aust., Canberra.
8. Madsen, H., P.S. Mikkelsen, D. Rosbjerg and P. Harremoes, 2002. Regional estimation of rainfall intensity duration curves using generalized least squares regression of partial duration series statistics. Water Resources Res., 38(11): 1-11.
9. Stedinger, J.R., 1983. Estimating a regional flood frequency distribution. Water Resources Res., 19: 503-510.
10. Strupczewski, W.G., V.P. Singh and S. Weglarczyk, 2002. Asymptotic bias of estimated methods caused by the assumption of false probability distribution. J. Hydrol., 258: 122-148.
11. Calenda, G., C.P. Mancini and E. Volpi, 2009. Selection of the probabilistic model of extreme floods: The case of the River Tiber in Rome. J. Hydrol., 27: 1-11.
12. Hosking, J.R.M. and J.R. Wallis, 1993. Some statistics useful in regional frequency analysis. Water Resources Res., 29(2): 271-281.
13. Reis, D.S., Jr., J.R. Stedinger and E.S. Martins, 2005. Bayesian generalized least squares regression with application to log Pearson type 3 regional skew estimation. Water Resources Res., Vol. 41: W10419, doi:10.1029/2004WR003445.