# Regional flood modelling in Western Australia: Application of regression based methods using ordinary least squares

**M. Taylor[a], K. Haddad[a], M. Zaman[a] and A. Rahman[a]**

[a] *School of Engineering, University of Western Sydney, Australia*

*Email: mel.taylor@uws.edu.au*

**Abstract:** Flooding is one of Australia's costliest natural disasters, which on average costs around \$AUD400 million annually. In 2010-2011 alone, the cost of flood damage has exceeded \$AUD20 billion. Flood estimation is therefore crucial in assessing and managing flood risks. Flood Frequency Analysis (FFA) is one of the most commonly adopted techniques used to estimate floods with an associated frequency, known as design floods. These design floods are used in the planning and design of water infrastructure projects, along with various floodplain planning regulations.

FFA largely relies on the existence of long recorded streamflow data. Due to the sheer size of Australia and the substantial costs involved, streamflow gauges are generally restricted to the highly populous and coastal regions of Australia. As design flood estimates are quite often needed in ungauged catchments, which have insufficient, unreliable or no streamflow data, an alternative method known as Regional Flood Frequency Analysis (RFFA) is generally adopted for design flood estimation.

RFFA attempts to substitute the lack of temporal data with spatial data, to make more accurate flood estimates at ungauged sites. The most commonly adopted RFFA techniques include the Probabilistic Rational Method (PRM), the Index Flood Method (IFM) and the Quantile Regression Technique (QRT). The national guideline for design flow estimation, Australian Rainfall and Runoff (ARR) 1987, currently recommends a particular form of the PRM and the IFM for use in Western Australia (WA). The PRM has been widely criticised due to the simplistic assumption involved in plotting and interpolating the dimensionless runoff coefficients. At the same time, identifying homogeneous regions in connection with the IFM, has also proven to be problematic in Australia.

Regression-based RFFA techniques have, however, shown promise in a number of recent studies across Australia and are the standard method in some other countries, such as the United States. The aim of the paper is to develop and compare two regression-based techniques: the QRT and the Parameter Regression Technique (PRT). In the QRT, individual flood quantiles are regressed against the catchment characteristics; while in the PRT, the parameters of a probability distribution (here the Log Pearson type 3 distribution is considered) are regressed against the catchment characteristics. An ordinary least squares regression method is used in this study for developing the prediction equations.

The study uses streamflow and catchment characteristic data from 206 catchments across WA and part of the Northern Territory to develop regional prediction equations. The dataset was divided into three regions, according to the Australian Drainage Divisions (VI, VII and VIII), with 125, 12 and 69 catchments respectively. Independent testing of the QRT and the PRT was carried out on independent test catchments, which were randomly selected and were not used in the development of the prediction equations.

It has been found that both the QRT and PRT provide reasonable predictions. The developed prediction equations are also relatively easy to apply as they only contain two predictor variables, being catchment area and design rainfall intensity. It should be noted that results from GLS regression for WA presented in Haddad et al. (2011b) should be more reliable than this study. The findings from this study and other relevant RFFA studies would form the basis of recommendation of a new RFFA for WA in the upcoming edition of ARR.

*Keywords:*     *Regional flood frequency analysis, Western Australia, regression technique, ordinary least squares.*

## 1 INTRODUCTION

Flooding is one of Australia's costliest natural disasters, which on average costs around $AUD400 million annually. In 2010-2011 alone, the cost of flood damage has exceeded $AUD20 billion. To assess flood risk and adopt appropriate control measures, one needs to estimate flood discharge of a given Average Recurrence Interval (ARI), which is commonly referred to as 'design floods'. This design flood is widely used in flood risk management e.g. determining the habitable floor level, design of bridges, culverts, flood control levees, sluices, weirs and spill-ways. Design floods can be estimated using Flood Frequency Analysis (FFA), but this requires a long period of recorded streamflow data. Australia has an abundance of streams, many of which are isolated due to the sheer size of the continent. It is for this reason and the fact that streamflow gauging is relatively expensive that there are little to no streamflow records at many locations of interest. Under these circumstances, Regional Flood Frequency Analysis (RFFA) is commonly adopted to estimate design floods, which attempts to use data from nearby gauged catchments to make flood estimation at ungauged sites of interest.

There have been notable researches on flow prediction in ungauged basins (PUB) (e.g. Sivapalan, 2003). Many different RFFA techniques have been proposed for design peak flow estimation, including the Probabilistic Rational Method (PRM), the Index Flood Method (IFM) and the Quantile Regression Technique (QRT). The Australian national guideline for flow estimation, Australian Rainfall and Runoff (ARR), recommended several RFFA methods (I. E. Aust., 1987), e.g. the PRM and IFM were recommended for use in Western Australia (WA). The PRM is widely used due to its simplicity and ease of application; nevertheless it is based on a dimensionless runoff coefficient which is assumed to vary smoothly over geographical space (Hodgkins et al., 2007; Young et al., 2009). It has been found that the runoff coefficient may show sharp variation within a close proximity, reflecting discontinuities at catchment boundaries. French (2002) noted that the isopleths of the runoff coefficient in ARR1987 ignore the existence of watercourses. Pirozzi et al. (2009), Rahman et al. (2008) and Rahman and Hollerbach (2003) investigated the PRM and attempted to link the runoff coefficient with catchment characteristics, but they had little success. There is a lack of independent testing of the PRM and users have little knowledge about the uncertainty in design flood estimates derived from the PRM. Another commonly used technique, the IFM, is also limited in that it requires the hydrologic region to be acceptably homogeneous (Hosking and Wallis, 1993); however, in the case of Australia, identifying acceptably homogeneous regions has been found to be difficult (e.g. Bates et al., 1998).

Since the release of the 3rd edition of ARR in 1987 there have been notable advances in RFFA. Moreover, there is an additional 24 years of hydrological data available (including streamflow data) at many locations across Australia, which would increase the accuracy of new RFFA methods. Recently, the National Committee on Water Engineering (NCWE) has put together 21 projects to revise ARR. Project 5 focuses on the development of new RFFA methods for Australia. This paper focuses on Western Australia and in particular compares two regression based RFFA methods, the QRT and the Parameter Regression Technique (PRT). The PRT regionalises the first three moments of the Log-Pearson Type 3 (LP3) distribution (explained in Section 3). An Ordinary Least Squares (OLS) regression method is adopted in this study as a first step to develop regional prediction equations. In a complementary paper (Haddad et al. 2011b) an advanced method employing the Generalised Least Squares Regression (GLS) along with Region-Of-Influence (ROI) approach (Burn, 1990) for WA has been adopted. The results from Haddad et al. (2011b) are likely to be more reliable than the results of this preliminary study since the GLS regression considers the inter-station correlation of the annual maximum flood series and ROI better deals with spatial heterogeneity in RFFA.

## 2 THE QUANTILE AND PARAMETER REGRESSION TECHNIQUES

The quantile regression technique (QRT) was originally developed by the United States Geological Survey (USGS), and due to the ongoing development it has remained the leading technique of the USGS to estimate flood quantiles in ungauged catchments (Brutsaert, 2005). The method estimates flood quantiles through multiple regression between recorded streamflow data and a set of climatic/catchment characteristics within a region (Pandey & Nguyen, 1999). The quantile regression technique is expressed as follows:

$$Q_T = aB^b C^c D^d \cdots \qquad (1)$$

where $B, C, D, \cdots$ are catchment/climatic characteristics; $a, b, c, d, \cdots$ are the regression coefficients and $Q_T$ is the flood magnitude with $T$ year ARI (flood quantile). There are a number of methods available to determine the regression coefficients, OLS, weighted least squares (WLS) and GLS. The OLS method is the simplest of the three and finds the best fitting regression parameters by minimising the sum of squared residuals. For the OLS estimator to be accurate and efficient the peak flow at each site must be uncorrelated, every site should have equal record lengths and at-site flood quantiles should have equal variance, which is unlikely to be satisfied in hydrological regression (Tasker & Stedinger, 1989). The GLS regression method is further explained in Haddad et al. (2011a).

The Parameter Regression Technique (PRT) is similar to the QRT, but instead of quantiles, here the first three moments of the LP3 distribution are taken as the dependent variables in the regression analysis. Let $Q$ be the annual maximum flood series at a site and $X = \ln(Q)$, then the mean ($M$), standard deviation ($S$) and skew ($g$) of the $X$ series are taken as dependent variables.

$$lnQ_T = M + K_T S \tag{2}$$

where $Q_T$ is flood quantile of $T$ years ARI and $K_T$ is the standardised LP3 frequency factor (which is a function of $g$) and can be obtained from ARR1987 or can be approximated using the Cornish-Fisher transformation (Fisher and Cornish, 1960).

## 3  STUDY CATCHMENTS

Several factors were considered when selecting suitable catchments for use in the WA; including, the catchment size, record length, any regulation (such as dams etc.) and land-use changes. It is suggested in ARR (I.E. Aust., 1987) that an arbitrary upper limit of 1000 km² be adopted for small to medium sized catchments; therefore, the initial dataset adopted 1000 km² as the upper limit. The imposed limit lead to a lack of suitable catchments in the northern parts of WA, therefore, additional stations with larger catchment areas were also included in the dataset. Also, part of NT, situated in Drainage Division VIII, was lumped with the WA as shown in Figure 1.

Streamflow records are required to be long enough to accurately characterise the underlying flood probability distribution. However, there must also be a suitable number of gauged catchments available in RFFA for greater spatial coverage. For this reason, a balancing act is required to ensure streamflow records are long enough, whilst also ensuring that the spatial coverage of stations is adequate. The ARR revision team suggested a lower limit of 25 years of streamflow record length, which was initially adopted; however, this resulted in insufficient spatial coverage. The minimum record length was therefore taken as 19 years. It is also crucial that each catchment does not contain any major regulation (such as dams etc.) and that no major land-use changes have occurred throughout the period of the streamflow record.

Based on the aforementioned criteria, a total of 206 catchments were selected across WA and the NT (Figure 1). The dataset was then divided into three regions, according to the Drainage Divisions VI, VII and VIII. These drainage divisions are defined by major topographical features and key climatic zones; therefore it seems appropriate to utilise these divisions. As central WA has a lack of data, drainage division XII was left out of the study.

In order to test the accuracy of prediction equations, a number of test catchments were then selected from the dataset and put aside. The split-sample validation technique was applied to Drainage Divisions VI and VIII, with approximately 20% of sites were left out for independent testing. Alternatively, due to the lack of catchments available in Drainage Division VII, a bootstrapping technique was adopted where one catchment was left out and the regression equation was developed, the developed equation is then applied to the catchment that was left out. The procedure was repeated for all the sites in Division VII. The catchment information is summarised in **Error! Reference source not found.**.
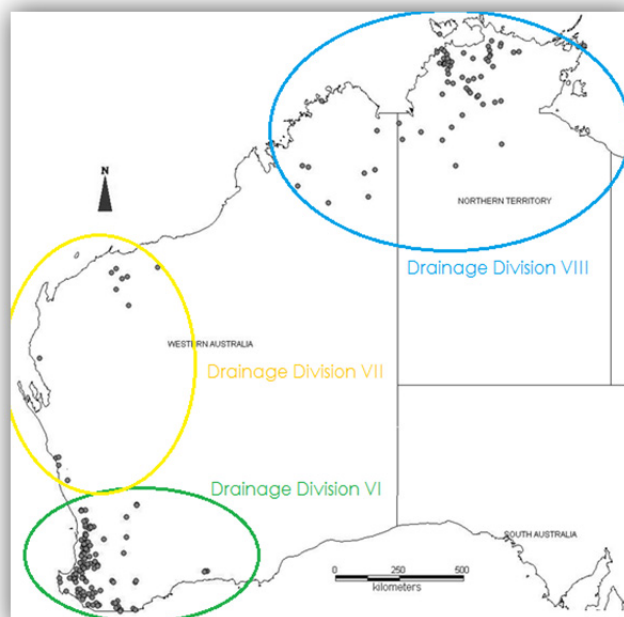


**Figure 1: Distribution of catchments between drainage divisions VI, VII, and VIII.**

**Table 1 Summary of catchment information for each drainage division from WA**

| Drainage Division | No. of Sites | | Record Length (yrs) | | Mean Record Length (yrs) | Catchment Area (km²) | | Mean Catchment Area (km²) |
|---|---|---|---|---|---|---|---|---|
| | Total | Validation | Min | Max | | Min | Max | |
| VI | 125 | 26 | 20 | 56 | 31 | 0.2 | 983.1 | 156.7 |
| VII | 12 | 12 | 20 | 35 | 29 | 0.1 | 1000 | 346.7 |

| VIII | 69 | 14 | 19 | 54 | 34 | 1.4 | 4325 | 645.3 |
|------|----|----|----|----|----|----|------|-------|

## 4 CATCHMENT DATA

Seven climatic and catchment characteristics were initially selected for use in this study, following the findings of previous Australian studies (Rahman, 2005). The main criteria which was used in the study involved the relevance of the characteristic to the generation of peak flows (i.e. hydrologically meaningful), ability to obtain the relevant data (i.e. easily obtainable) and that each variable is largely uncorrelated with other selected variables. The climatic and physical catchment characteristics which were selected are listed below:

- Catchment area (*area*, km$^2$) – the area of each catchment was obtained from the Department of Water.
- Design rainfall intensity (mm/h) – the Bureau of Meteorology's (BOM's) IFD calculator was used to obtain each of the design rainfall intensities, along with the AUSIFD software. The design rainfall intensities ($^{ARI}I_{duration}$) used in this study include $^{2}I_{1hr}$, $^{2}I_{12hr}$, $^{2}I_{72hr}$, $^{50}I_{1hr}$, $^{50}I_{12hr}$, $^{50}I_{72hr}$, $^{2}I_{t_c}$, $^{5}I_{t_c}$, $^{10}I_{t_c}$, $^{20}I_{t_c}$, $^{50}I_{t_c}$, and $^{100}I_{t_c}$,
- Slope of the central 75% of the mainstream (*S1085*, m/km) – S1085 was adopted as it excludes the extremes found at each end of the mainstream. The slope was determined using 1:100,000 topographic maps and is equal to: $S1085 = \frac{E_{85}-E_{10}}{0.75L}$ where $E_{85}$ and $E_{10}$ are the elevations of the streambed at 85% and 10% of the stream length, respectively.
- Stream density (*sden*, km/km$^2$) – it is the channel length per unit area of the catchment.
- Fraction forest (*forest*) – it is the fraction of the basin covered by medium to dense forest, measured using 1:100,000 topographic maps.
- Mean annual rainfall (*rain*, mm) – was obtained at the catchment outlet using the BOM's data CD of mean annual rainfall.
- Mean annual areal potential evapotranspiration (*evap*, mm) – was obtained at the catchment outlet using the BOM's data CD of evaporation.

## 5 MEASURES OF MODEL ADEQUACY

Each of the developed prediction equations were evaluated using a number of goodness-of-fit statistics. The coefficient of multiple determination ($R^2$) and adjusted $R^2$ determine how accurately the model fits the observed data. Note that $0 \leq R^2 \leq 1$, where 1 indicates the model is a perfect fit. Standard error of the estimate (*SEE*) indicates the spread of predicted values about the prediction line, where a lower value indicates a better fitting model, i.e. less variance.

Each regression equation was also checked to ensure that underlying assumptions of the model are met. The Durbin-Watson statistic (DW) is a measure of the correlation of the residuals where a value of 2 indicated that there is no serial correlation between residuals. The *t*-value for the coefficient and two-tailed significance level of *t* (*P*-value) indicates the significance of each predictor variable. Higher *t*-values and a low significance level (<0.10) indicates that the predictor is significant. The tolerance and variance inflation factor (VIF) measures the collinearity of a variable. A tolerance and VIF close to 1 indicate the predictor variables are not highly correlated. For highly correlated variables tolerance approaches zero, and VIF moves toward infinity. The preceding statistics provide useful information on the accuracy of prediction equations; however, the equations may have been affected by points outside the normal range. Outliers in the response variables were detected using standardised residuals. Values larger than an absolute value of three are classed as outliers. Influential points were identified using Cook's distance.

## 6 DEVELOPMENT OF PREDICTION EQUATIONS

At-site flood frequency analysis is the first step in developing RFFA methods. The annual maximum series data was adopted in this study, which consists of the largest flood in each year of record. The at-site flood frequency analysis was performed using the FLIKE software (Kuczera, 1999). The program facilitates the use of a number of distributions and fitting procedures, including the LP3, Log-Normal (LN), Gumbel, Generalised Extreme Value and Generalised Pareto distribution, using either the Bayesian or L moments parameter fitting procedures, which were tested in this study. The LP3-Bayesian parameter fitting procedure provided the most consistent results and hence was adopted in this study.

Each of the dependent variables, including the selected flood quantiles, for the QRT, and the LP3 moments/parameters, for the PRT, was regressed against each of the seven predictor variables. A natural log transformation was considered in the regression analysis for both the predictor and dependent variables. The Backward Multiple Regression technique in statistical package Predictive Analytical Software (PASW) was initially used; however, in most cases it produced impractical equations. Therefore, the Enter Multiple

Regression technique was adopted to allow for a trial-and-error approach, based on the most appropriate DW, *t*-value, *P*-value, tolerance and VIF statistic values. The models with the highest $R^2$, lowest SEE and that satisfied the underlying model assumptions in the best possible manner, were then taken as the best fitting models.

There has been much debate about the inclusion or exclusion of outliers and influential points, as such in the first set of prediction equations outliers and influential points were determined using the standardised residuals, Cook's distance and centred leverage values, but kept in the dataset. Similarly in the second set of prediction equations, outliers and influential points were determined and removed from the dataset (with a maximum limit of 10% of the dataset). This was carried out for Drainage Divisions VI and VIII; however due to the lack of sites in Drainage Division VII, it was not feasible to remove any outliers. The original models in both cases were found to be superior; it is thought this could be due to the decrease in variability that occurs when these points are removed. The selected prediction equations for each drainage division are shown in **Error! Reference source not found.**.

**Table 2: Selected prediction equations for each drainage division**

| | | | |
|---|---|---|---|
| **Drainage Division VI** | QRT | $ln(Q_2) = -7.464 + 0.830 \times ln(area) + 1.691 \times \ln(^{2}I_{1hr})$ | (3) |
| | | $ln(Q_5) = -3.239 + 0.829 \times ln(area) + 0.763 \times \ln(^{50}I_{12hr})$ | (4) |
| | | $ln(Q_{10}) = -1.515 + 0.816 \times ln(area) + 0.116 \times \ln(^{50}I_{12hr})$ | (5) |
| | | $ln(Q_{20}) = 1.202 + 0.811 \times ln(area) - 0.723 \times \ln(^{2}I_{1hr})$ | (6) |
| | | $ln(Q_{50}) = 1.622 + 0.792 \times ln(area) - 1.072 \times \ln(^{50}I_{12hr})$ | (7) |
| | | $ln(Q_{100}) = 5.693 + 0.788 \times ln(area) - 1.646 \times \ln(^{50}I_{1hr})$ | (8) |
| | PRT | $M = -7.462 + 0.833 \times ln(area) + 1.680 \times \ln(^{2}I_{1hr})$ | (9) |
| | | $S = C_1 \,;\; g = C_2$, where $C_1$ and $C_2$ are regional average values | (10) |
| **Drainage Division VII** | QRT | $ln(Q_2) = -11.366 + 0.521 \times ln(area) + 7.858 \times \ln(^{2}I_{12hr})$ | (11) |
| | | $ln(Q_5) = -15.913 + 0.486 \times ln(area) + 5.336 \times \ln(^{2}I_{1hr})$ | (12) |
| | | $ln(Q_{10}) = -14.285 + 0.465 \times ln(area) + 5.055 \times \ln(^{2}I_{1hr})$ | (13) |
| | | $ln(Q_{20}) = -12.949 + 0.445 \times ln(area) + 4.824 \times \ln(^{2}I_{1hr})$ | (14) |
| | | $ln(Q_{50}) = -4.914 + 0.431 \times ln(area) + 5.705 \times \ln(^{2}I_{12hr})$ | (15) |
| | | $ln(Q_{100}) = -4.072 + 0.413 \times ln(area) + 5.412 \times \ln(^{2}I_{12hr})$ | (16) |
| | PRT | $M = -11.411 + 0.527 \times ln(area) + 7.765 \times \ln(^{2}I_{12hr})$ | (17) |
| | | $S = C_1 \,;\; g = C_2$, where $C_1$ and $C_2$ are regional average values | (18) |
| **Drainage Division VIII** | QRT | $ln(Q_2) = 5.620 + 0.531 \times ln(area) - 1.391 \times \ln(^{50}I_{12hr})$ | (19) |
| | | $ln(Q_5) = 6.172 + 0.519 \times ln(area) - 1.284 \times \ln(^{50}I_{12hr})$ | (20) |
| | | $ln(Q_{10}) = 6.585 + 0.513 \times ln(area) - 1.287 \times \ln(^{50}I_{12hr})$ | (21) |
| | | $ln(Q_{20}) = 6.988 + 0.508 \times ln(area) - 1.317 \times \ln(^{50}I_{12hr})$ | (22) |
| | | $ln(Q_{50}) = 7.506 + 0.503 \times ln(area) - 1.381 \times \ln(^{50}I_{12hr})$ | (23) |
| | | $ln(Q_{100}) = 7.886 + 0.500 \times ln(area) - 1.440 \times \ln(^{50}I_{12hr})$ | (24) |
| | PRT | $M = 5.699 + 0.549 \times ln(area) - 1.512 \times \ln(^{50}I_{12hr})$ | (25) |
| | | $S = C_1 \,;\; g = C_2$, where $C_1$ and $C_2$ are regional average values | (26) |

Prediction equations for the QRT method for each region included *area* and design rainfall intensity. *Area* was found to be the most significant predictor variable across all flood quantiles and for the mean flood. Most prediction equations performed reasonably well, as can be seen in **Error! Reference source not found.**. Each of the developed prediction equations were examined for underlying assumptions and found to satisfy these

Taylor *et al.,* Regional flood modelling in Western Australia: Application of regression based methods…

assumptions quite well. No meaningful prediction equations could be developed for the standard deviation (*S*) and skew (*g*), therefore the regional average values were used.

**Table 3: Average goodness-of-fit statistics for each drainage division**

| Drainage Division | Method | $R^2$ | $R^2_{Adj}$ | SEE | DW |
|---|---|---|---|---|---|
| VI | QRT | 0.801 | 0.797 | 0.920 | 1.867 |
| | PRT | 0.785 | 0.781 | 1.007 | 1.750 |
| VII | QRT | 0.907 | 0.883 | 0.562 | 2.266 |
| | PRT | 0.963 | 0.954 | 0.421 | 2.160 |
| VIII | QRT | 0.514 | 0.495 | 0.892 | 1.969 |
| | PRT | 0.568 | 0.551 | 0.860 | 1.886 |

## 7    MEASURES OF PREDICTION ERROR

The following statistics were adopted to compare the QRT and PRT:

$$RMSE = \sqrt{\frac{\Sigma(Q_{pred}-Q_{obs})^2}{n}} \qquad (27)$$

$$RMSE\ Ratio = \frac{RMSE}{\bar{Q}_{obs}} \times 100\% \qquad (28)$$

$$RE = \frac{Q_{pred}-Q_{obs}}{Q_{obs}} \times 100\% \qquad (29)$$

$$Ratio = \frac{Q_{pred}}{Q_{obs}} \qquad (30)$$

The Root Mean Squared Error (RMSE) ratio provides an indication as to the overall accuracy of the model, a value of zero indicates the model is perfect; however, large values are undesirable. Relative Error (RE) indicates the accuracy of each prediction in relation to the observed value, with a value of zero being perfect and can also be either positive or negative. The $Q_{pred}/Q_{obs}$ ratio indicates the amount of bias and accuracy of the model. It should be noted that although the RMSE ratio and RE are not true measures of the error associated with a model, both values give a reasonable indication of the error associated with a model for practical purposes, as both $Q_{obs}$ and $Q_{pred}$ have reasonably large uncertainties associated with them.

## 8    SPLIT-SAMPLE VALIDATION

Prediction equations for each region were validated using a number of independent test catchments. Statistics that were used to evaluate the performance of each model included the RMSE ratio, median RE and $Q_{pred}/Q_{obs}$ ratio. The median RE and RMSE ratio results, for the three drainage divisions, are shown in **Error! Reference source not found.**. It can be seen that the QRT produces a lower RE for drainage divisions VI and VII, with RE values of 60% vs. 67% and 31% vs. 41% for the QRT and PRT, respectively. In contrast, the PRT produces a lower RE for drainage division VIII, with RE values of 56% vs. 61% for the PRT and QRT, respectively. The RMSE ratio for each of the drainage divisions can be seen to be lower for the PRT; however, there is only a 0.3% difference between the PRT and QRT in drainage divisions VI and VII.

**Table 4: Relative error and Root mean squared error for Drainage Divisions VI, VII, and VII**

| Drainage Division | | VI | | VII | | VIII | |
|---|---|---|---|---|---|---|---|
| | | *Range* | Average | *Range* | Average | *Range* | Average |
| Absolute Median RE % | PRT | 58 – 74 | 68 | 29 – 50 | 42 | 47 – 67 | 57 |
| | QRT | 53 – 73 | 61 | 28 – 37 | 31 | 52 – 79 | 62 |
| RMSE Ratio % | PRT | 10 – 13 | 11 | 4 – 10 | 7 | 6 – 13 | 9 |
| | QRT | 11 – 13 | 12 | 4 – 11 | 7 | 11 – 16 | 12 |

## 9    CONCLUSIONS

Prediction equations for the Quantile Regression Technique (QRT) and the Parameter Regression Technique (PRT) were developed for six flood quantiles ($Q_2$ to $Q_{100}$) in each of the three regions across WA. The final prediction equations for both methods contained two predictor variables, including catchment area and rainfall intensity of a given duration and ARI, which are easily obtainable from published data. Each of the developed

prediction equations satisfied the underlying assumptions of the OLS regression well, as the residuals are approximately normally distributed. Region 2 should be used with care, as the lack of sites made the statistics hard to interpret. Each model was found to produce reasonably high $R^2$ and $R^2_{Adj}$, with smaller *SEE* values. Both the QRT and PRT performed relatively well, based on the independent testing, however a small number of catchments produced estimates which were grossly above or below the at-site flood frequency estimates. However, it should be noted that at-site flood frequency estimates are not free of error and they have a large uncertainty range, for higher ARIs in particular. It should be noted that results from GLS regression for WA presented in Haddad et al. (2011a) are likely to be more reliable than this study. The findings from this study and other relevant RFFA studies would form the basis of recommendation of a new RFFA for WA in the upcoming edition of ARR.

## 10 REFERENCES

Bates, B.C., A. Rahman, R.G. Mein, and P.E. Weinmann (1998). Climatic and physical factors that influence the homogeneity of regional floods in south-eastern Australia. *Water Resources Research*, 34(12), 3369-3382.

Brutsaert, W. (2005). *Hydrology: An Introduction*, Cambridge University Press, Cambridge

Burn, D.H. (1990). An appraisal of the "region of influence" approach to flood frequency analysis. *Journal of Hydrological Sciences,* 35(2), 149-165.

Fisher, R.A. and E.A. Cornish, (1960). The Percentile Points of Distributions Having Known Cumulants, *Technometrics*, 2(2), 209-225.

French, R. (2002). Flaws in the rational method, 27th National Hydrology and Water Resources Symposium, Melbourne, 20-23 May 2002.

Haddad, K., Rahman, A. And Kuczera, G. (2011a). Comparison of Ordinary and Generalised Least Squares Regression Models in Regional Flood Frequency Analysis: A Case Study for New South Wales, *Australian Journal of Water Resources* (In press)

Haddad, K., Rahman, A., Weeks, W., Kuczera, G. and Weinmann, P.E. (2011b). Towards a New Regional Flood Frequency Analysis Method for Western Australia, MODSIM 2011, Perth, 12-16 December 2011.

Hodgkins, G.A., Hebson, C. Lombard, P.J. and Mann, A. (2007). Comparison of peak-flow estimation methods for small drainage basins in Maine, U.S. Geological Survey, Scientific Investigations Report 2007-5170.

Hosking, J.R.M., and J.R. Wallis (1993). Some statistics useful in regional frequency analysis. *Water Resources Research*, 29(2), 271-281.

Institution of Engineers Australia (1987). *Australian rainfall and runoff: A guide to flood estimation,* vol. 1, I. E. Aust., Canberra

Kuczera, G. (1999). Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian Inference, Water Resources Research, 35(5), 1551-1557.

Pandey, G.R. and Nguyen, V.T.V. (1999). A comparative study of regression based methods in regional flood frequency analysis, *Journal of Hydrology*, 225, 92-101.

Pirozzi, J., Ashraf, M., Rahman, A., and Haddad, K. (2009). Design flood estimation for ungauged catchments in Eastern NSW: Evaluation of the Probabilistic Rational Method, 32nd Hydrology and Water Resources Symposium*,* Newcastle, 30 November - 3 December, pp. 805-816.

Rahman, A. (2005). A quantile regression technique to estimate design floods in ungauged catchments in south-east Australia, *Australian Journal of Water Resources*, 9(1), 81-89.

Rahman, A. and Hollerbach, D. (2003). Study of Runoff Coefficients Associated with the Probabilistic Rational Method for Flood Estimation in South-east Australia, 28th Hydrology and Water Resources Symposium, Wollongong, 10-13 November 2003, pp. 199-203.

Rahman, A., Rima, K. and Weeks, W. (2008). Development of regional flood estimation methods using quantile regression technique: A case study for North-eastern part of Queensland, 31st Hydrology and Water Resources Symposium, Adelaide, 15-17 April, pp. 329-340.

Tasker, G.D. and Stedinger, J.R. (1989). An operational GLS model for hydrologic regression, *Journal of Hydrology*, 111, 361-375.

Young, C.B., McEnroe, B.M. and Rome, A.C. (2009). Empirical determination of rational method runoff coefficients, *Journal of Hydrologic Engineering*, 14(12), 1283-1289.