

A New Probabilistic Rational Method for design flood estimation in ungauged catchments for the State of New South Wales in Australia

O. Alsuwaidi^a, M.M. Haque^b, A. Rahman^b and K. Haddad^a

^a School of Computing, Engineering and Mathematics, University of Western Sydney

^b Institute for Infrastructure Engineering, University of Western Sydney

Email: m.haque@uws.edu.au

Abstract: Design flood estimates for ungauged catchments are needed in the planning and design of bridges, culverts and many other water infrastructure projects. The most commonly used methods of flood estimation in practice for the ungauged catchments include the Index Flood Method, quantile regression technique and Probabilistic Rational Method (PRM). In Australia, the regional flood frequency estimation (RFFE) methods recommended in Australian Rainfall and Runoff (ARR) 1987 include PRM for eastern New South Wales (NSW) and Victoria States. In the upcoming ARR 2016, the recommended RFFE method is based on regression and region-of-influence approaches for NSW, which is referred to as RFFE Model 2015. In this study, a new PRM is developed and tested for NSW.

A total of 106 catchments are used to develop and test the new PRM. These data are obtained from ARR Project 5 Regional Flood Methods. The catchment areas for the selected catchments range from 1 to 1010 km². The mean and median catchment areas are 273 km² and 169 km², respectively. The streamflow record lengths range from 20 to 82 years, with a mean of 38 years and median of 35 years.

Twelve different forms of PRM are examined, which are formed based on different combinations of observed runoff coefficients (C_{10}) and frequency factors (FF_Y). The C_{10} values at a test catchment site are estimated using inverse distance weighted method based on different combinations of the at-site C_{10} values i.e. (i) nearest one site; (ii) two nearest sites; (iii) three nearest sites; and (iv) five nearest sites. The FF_Y values are estimated using three different approaches. A leave-one-out (LOO) validation approach is adopted to compare various forms of the PRM. At-site flood frequency analysis (based on the log Pearson Type 3- Bayesian method) is used as the 'benchmark' for this LOO validation.

It is found that the best PRM case is achieved when C_{10} value is estimated from the three nearest sites and FF_Y is taken as the median value over all the model catchments. The results of this study reveal that the new PRM can be used to accurately predict the peak flow rates for small-to-medium sized catchments in the NSW region. For the new PRM, the typical median relative error values based on the leave-one-out validation is 45% to 55% for 20% to 1% annual exceedance probabilities. However, for the 50% AEP flood, the median relative error value is 77%. These relative error values are comparable to ARR RFFE Model 2015. The main advantage of the new PRM against the ARR 1987 is that it does not need a contour map to estimate the runoff coefficient and it can be automated in an application tool like the ARR RFFE Model 2015. The new PRM is being extended for other states of Australia.

Keywords: ARR, RFFE 2015, PRM, floods, regional flood frequency estimation, ungauged catchments

1. INTRODUCTION

Reliable estimation of design floods is needed in the hydrological design of water infrastructures, flood plain management tasks and various regulatory and environmental purposes. The design flood can be estimated more accurately for catchments where relatively long streamflow data is available; however, for ungauged catchments (where recorded streamflow data is unavailable or of limited length (less than 10 years) or of poor quality), accurate prediction of design floods remains a challenging task. Moreover, design flood estimates for ungauged catchments are generally associated with a large degree of uncertainty (Haque *et al.*, 2014; Haddad *et al.*, 2012). Regional flood frequency analysis (RFFA), which is based on the principle of pooling flood data within a homogeneous region, is commonly adopted to estimate design floods in the ungauged catchments. Various forms of RFFA were recommended in the Australian Rainfall and Runoff (ARR) 1987 for different states of Australia (IEAust, 1987). RFFA in essence attempts to transfer flood characteristics from gauged to ungauged catchments and hence can be used both for gauged and ungauged catchments; for gauged catchments, it enhances the accuracy of design flood estimates, in particular when record length is too short.

Various RFFA techniques have been adopted in the past such as Rational Method, Probabilistic Rational Method (PRM), Index Flood Method, Quantile Regression Technique, parameter regression technique, and Artificial intelligence-based methods (Bates *et al.*, 1998; Rahman *et al.*, 1999; Rahman *et al.*, 2011; Aziz *et al.*, 2014; Aziz *et al.*, 2015). The rational method was first introduced by Mulvaney (1851) to estimate peak discharge, which is generally regarded as a deterministic model. However, ARR 1987 recommended a probabilistic form of the Rational Method, known as Probabilistic Rational Method (PRM), for Victoria and Eastern New South Wales. The PRM in ARR 1987 was based on the studies by Pilgrim (1982), Pilgrim and McDermott (1982) and Adams (1984). The application of the PRM in ARR 1987 requires a contour map of runoff coefficient. The runoff coefficient is assumed to vary smoothly over geographic space; however, a sharp variation in the runoff coefficients has been found even within a close proximity indicating discontinuities at catchment boundaries (e.g. Rahman and Hollerback, 2003; Rahman *et al.*, 2008; Pirozzi *et al.*, 2009).

Since 1987, there have been many new developments in the statistical methods to develop RFFA techniques. Moreover, there has been the availability of additional flood data of 20-years of length at many gauged sites in Australia. To incorporate these new data and new statistical methods in RFFA, in 2006, the National Committee on Water Engineering (Engineers Australia) initiated a new RFFA project in Australia. In this regard, as a part of 'Project 5 Regional Flood Methods', a significant research on RFFA has been conducted (e.g. Rahman *et al.*, 2009, 2011, 2015a; Haddad and Rahman, 2012). The aim of this paper is to develop and test a new form of PRM, which does not require the use of a paper-based runoff coefficient map as in the ARR 1987 so that its application can be automated via software.

2. STUDY AREA

This study focuses on NSW state in Australia. Data from 106 gauged catchments are used (Figure 1). These catchments have not undergone any major land use change and are not affected by any major regulation during the period of streamflow data availability. These catchments are a subset of ARR Project 5 database (Rahman *et al.*, 2015b). The annual maximum flood data of these catchments were prepared following a stringent quality control as reported in Haddad *et al.* (2010). The catchment areas range from 1 to 1010 km², with a mean of 273 km² and median of 169 km². The annual maximum flood record lengths range from 20 to 82 years, with a mean of 38 years and median of 35 years.

3. ADOPTED METHODOLOGY

To estimate flood quantiles at the selected 106 gauged sites, at-site flood frequency analysis is carried out using log Pearson Type 3 (LP3) distribution and Bayesian parameter estimation procedure. This is done using FLIKE software (Kuczera, 1999). It should be mentioned that the at-site flood frequency analysis was conducted under the assumption of stationarity, i.e. the annual maximum flood series data do not show significant trends. This assumption was tested in selecting the stations for this study.

To make an independent assessment of the PRM, a leave-one-out (LOO) validation approach is adopted (Haddad *et al.*, 2013). A gauged catchment (i) is left out (called test catchment) from n gauged catchments (here $n = 106$) and the PRM is developed based on the $n-1$ gauged catchments; the developed PRM is then applied to the test catchment i and the procedure is repeated for n times. This in essence allows considering each of the n gauged catchments as 'ungauged test catchment'.

The PRM is regarded as a RFFA technique with two independent variables, namely catchment area (A) and design rainfall intensity ($I_{t_c,Y}$) (Rahman *et al.*, 2011). The dimensionless run-off coefficient (C_Y) is the central

component of the PRM, which was derived in the ARR 1987 based on gauged catchments in the study area. Some of the gauged sites used in ARR 1987 had as low as 10 years of streamflow data, which is likely to result in a massive sampling error in the estimated at-site flood quantiles.

The governing equation in the PRM is given by:

$$Q_Y = 0.278 C_Y A I_{t_c, Y} \quad (1)$$

where Q_Y is the peak flow rate in m^3/s for an AEP of 1 in Y , C_Y is the dimensionless runoff coefficient for AEP of 1 in Y ; A is the catchment area in km^2 ; and $I_{t_c, Y}$ is the average rainfall intensity (mm/h) for a time of concentration of t_c (hours) and AEP of 1 in Y .

From Eq. 1, the dimensionless runoff co-efficient for 10% AEP is calculated as:

$$C_{10} = \frac{Q_{10}}{0.278 A I_{t_c, 10}} \quad (2)$$

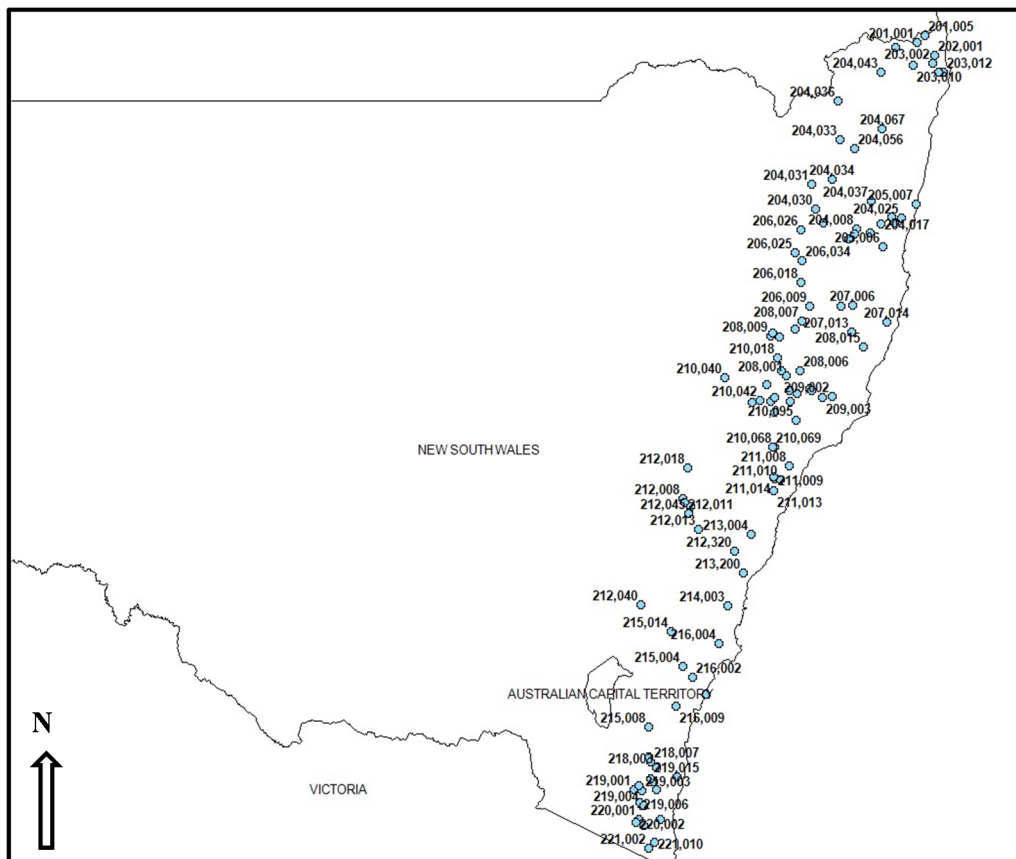


Figure 1. Location of selected 106 gauged catchments in New South Wales, Australia.

The values of Q_{10} and A are known for each of the study gauged catchments, and $I_{t_c, 10}$ is obtained from Australian Bureau of Meteorology website (2013 IFD values) as below:

$$I_{t_c, 10} \left(\frac{mm}{hour} \right) = \frac{I_{10} (mm)}{t_c (hour)} \quad (3)$$

t_c is calculated as below (IE Aust., 1987):

$$t_c = 0.76 A^{0.38} \quad (4)$$

where t_c is in hours and A is in km^2 .

The value of C_{10} at test catchment i is obtained using four different approaches as noted below.

- (a) The value of C_{10} for test catchment i is assumed to be equal to the value of C_{10} for the gauged site nearest to test catchment i .
- (b) The value of C_{10} for test catchment i is assumed to be the weighted average of C_{10} values at the two nearest gauged sites.
- (c) The value of C_{10} for test catchment i is assumed to be the weighted average of C_{10} values at the three nearest gauged sites.
- (d) The value of C_{10} for test catchment i is assumed to be the weighted average of C_{10} values at the five nearest gauged sites.

The inverse distance weighted average method is adopted in cases (b) to (d), which assigns higher weights to the gauged site located closer to the test catchment of interest relative to the distant sites.

Once C_{10} for test catchment i is known, the Q_{10} values are predicted using the following equation:

$$Q_{10}(\text{predicted}) = 0.278 C_{10}(\text{predicted}) A I_{t_c, 10} \quad (5)$$

Based on the above procedure, there are four different estimates of Q_{10} for test catchment i .

The peak flow rates for AEPs other than 10% are calculated using the following equation:

$$Q_Y(\text{predicted}) = FF_Y \times Q_{10}(\text{predicted}) \quad (6)$$

where, $Q_{10}(\text{predicted})$ is obtained from Eq. 5; and, FF_Y is the frequency factor obtained as below:

$$FF_Y = \frac{Q_Y(\text{observed})}{Q_{10}(\text{observed})} \quad (7)$$

Using Eq. 7, FF_Y values for each of the 106 gauged catchments are obtained. The design FF_Y is taken as the mean, median, and weighted average of FF_Y values, referred to $FF_Y(\text{mean})$, $FF_Y(\text{median})$, and $FF_Y(\text{weighted average})$, respectively.

Three different design FF_Y and four different C_{10} values give twelve different forms of PRM (see Table 1, column 1). Three performance statistics are adopted to evaluate the PRM: root means squared error (RMS error), percentage relative error (RE) and estimation ratio (ratio of the predicted (by each of the PRM techniques) and observed flood quantiles (obtained from at-site flood frequency analysis using FLIKE). In this comparison, it is assumed that at-site flood frequency analysis estimates are free of error, which is not the case since at-site flood frequency estimates are subject to various forms of sampling and modelling errors (Micevski *et al.*, 2015).

4. RESULTS

Graphical comparison between the observed (at-site FFA estimates) and predicted flood quantiles by the 12 different forms of PRM is presented in Figure 2 for 10% AEP. Graphs for other AEPs are not presented here as they have produced similar results. It has been observed that the predicted and observed flood quantiles for most of the catchments for all the 12 PRM cases and for the all the six AEPs lie close to the ideal line; however, there are some discrepancies for few larger catchments where the predicted flood quantiles are either notably overestimated or underestimated. This implies that the PRM cases are more accurate for smaller catchments compared with larger ones.

It has also been observed that the predicted flood quantiles are relatively more independent of the way design FF_Y is calculated, but are more influenced by the way design C_{10} is calculated (i.e. based on nearest site, nearest two sites, nearest three sites, and nearest five sites). For example, the PRM cases 1, 5, and 9 are similar (in Figure 2) since these three cases correspond to different FF_Y values but similar C_{10} (based on the nearest site).

Table 1 presents the median percentage relative error values (ignoring the sign of the individual site's relative error) corresponding to each of the twelve different PRM cases. It can be seen that Case 7 has the lowest median relative error value for all the AEPs except for 20% AEP. However, the relative error for 20% AEP for Case 7 (53.44%) is not much different to the lowest value for case 11 (52.7%). It can also be seen that the median relative error is much higher for 50% and 1% AEPs irrespective of the PRM cases. Relative errors are the smallest for 10% AEPs, followed by 5% AEPs for Case 7.

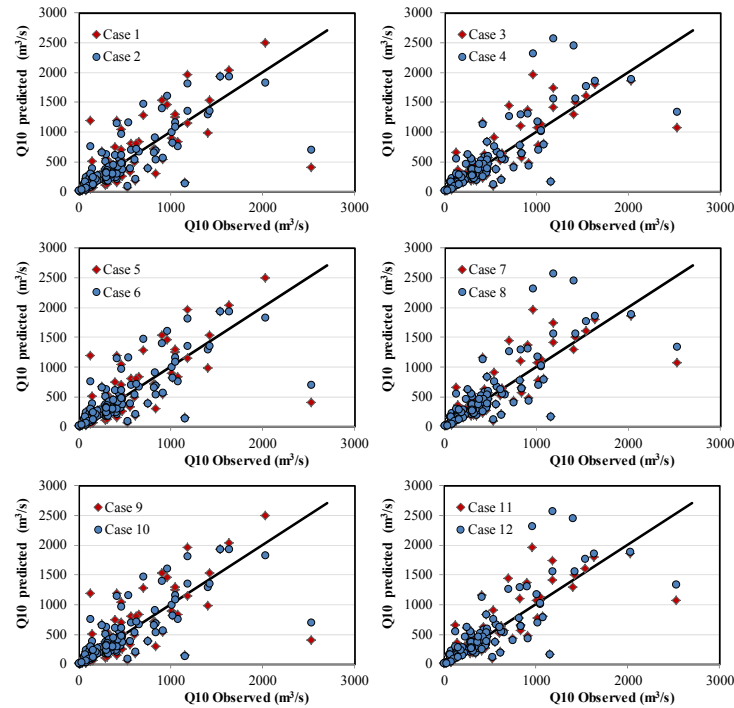


Figure 2. Observed (at-site FFA) versus predicted Q_{10} by the PRM for the twelve cases.

Table 1. Median percentage relative error for twelve PRM cases (Bold marked values represent the lowest median relative error value for a given AEP).

Case	FF_Y	C_{10}	50% AEP	20% AEP	10% AEP	5% AEP	2% AEP	1% AEP
1	Wt.Avg	Nearest site	103.86	68.63	56.10	53.58	61.46	76.82
2	Wt.Avg	2 nearest sites	88.53	57.76	49.48	49.68	59.53	77.32
3	Wt.Avg	3 nearest sites	81.64	52.78	45.48	46.60	56.93	75.15
4	Wt.Avg	5 nearest sites	82.39	53.91	47.82	49.64	61.91	81.52
5	Median	Nearest site	98.46	69.42	56.10	51.94	53.56	57.35
6	Median	2 nearest sites	83.59	58.49	49.48	48.06	51.92	56.58
7	Median	3 nearest sites	76.95	53.44	45.48	45.18	49.94	54.69
8	Median	5 nearest sites	77.47	54.80	47.82	47.85	53.09	58.36
9	Mean	Nearest site	103.85	68.53	56.10	53.82	62.51	79.25
10	Mean	2 nearest sites	88.52	57.67	49.48	49.93	60.56	80.03
11	Mean	3 nearest sites	81.63	52.70	45.48	46.81	57.97	77.80
12	Mean	5 nearest sites	82.38	53.80	47.82	49.91	63.07	84.36

The calculated RMSE ratios (RMSE/observed flood quantile by at-site FFA) for the twelve PRM cases are presented in Table 2. It can be seen that Case 7 has the lowest RMSE ratio values compared with the other PRM cases. Moreover, Cases 3, 6, and 11 perform consistently well in predicting the flood quantiles. It can be seen that that Case 7 consistently performs better than any other case. Thus, Case 7 where C_{10} is calculated

based on three nearest sites and FF_Y is estimated by taking the median of the FF_Y over all the gauged sites present the best form of the PRM in NSW.

Table 2. RMSE ratios for twelve PRM cases (bold marked values represent the lowest RMSE ratio for a given AEP).

Case	FF_Y	C_{10}	50% AEP	20% AEP	10% AEP	5% AEP	2% AEP	1% AEP
1	Wt.Avg	Nearest site	6.21	4.15	3.73	3.63	3.95	4.84
2	Wt.Avg	2 nearest sites	5.43	3.62	3.30	3.26	3.67	4.69
3	Wt.Avg	3 nearest sites	5.13	3.31	3.02	3.03	3.52	4.62
4	Wt.Avg	5 nearest sites	5.48	3.70	3.50	3.62	4.28	5.54
5	Median	Nearest site	6.05	4.18	3.73	3.55	3.59	4.06
6	Median	2 nearest sites	5.32	3.65	3.30	3.20	3.37	4.03
7	Median	3 nearest sites	5.04	3.33	3.02	2.97	3.23	3.97
8	Median	5 nearest sites	5.26	3.75	3.50	3.50	3.79	4.47
9	Mean	Nearest site	6.21	4.15	3.73	3.64	4.00	4.96
10	Mean	2 nearest sites	5.43	3.62	3.30	3.27	3.72	4.80
11	Mean	3 nearest sites	5.13	3.31	3.02	3.04	3.57	4.72
12	Mean	5 nearest sites	5.48	3.69	3.50	3.63	4.35	5.69

5. CONCLUSION

This study develops a new Probabilistic Rational Method (PRM) for NSW State in Australia to predict design floods for ungauged catchments. A total of 106 gauged catchments from NSW are used to develop the new PRM. It has been found that the new PRM more accurately predicts the flood quantiles for smaller to medium sized catchments compared with larger ones. It has also been found that the predicted flood quantiles by the new PRM are relatively independent of the way frequency factors (FF_Y) are calculated, while the predicted flood quantiles are more influenced by the way C_{10} are calculated. It has been found that among the 12 different forms of the PRM examined in this study, the median percentage relative error values are the smallest for the PRM Case 7 (i.e. when C_{10} is calculated based on three nearest sites, and FF_Y is taken as the median values of FF_Y over all the gauged sites). The results of the RMSE ratio values also reveal that case 7 PRM is the best performing case. For the new PRM, the typical median relative error values based on the leave-one-out validation is 45% to 55% for 20% to 1% AEPs. However, for the 50% AEP flood, the median relative error value is 77%. These relative error values are comparable to those of ARR RFFE Model 2015. The main advantage of the new PRM against the ARR 1987 PRM is that it does not need a contour map to estimate the runoff coefficient and hence it can be automated in an application tool like the ARR RFFE Model 2015. The new PRM is being extended for other states of Australia.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the ARR Revision Project 5 team and Australian Bureau of Meteorology for providing data for this study.

REFERENCES

- Adams, C. A. (1984). Regional flood estimation for ungauged rural catchments in Victoria. MEngSc Thesis, The University of Melbourne.
- Aziz, K., Rai, S., Rahman, A. (2015). Design flood estimation in ungauged catchments using genetic algorithm based artificial neural network (GAANN) technique for Australia, *Natural Hazards*, 77, 2, 805-821.
- Aziz, K., Rahman, A., Fang, G., Shreshtha, S. (2014). Application of Artificial Neural Networks in Regional Flood Frequency Analysis: A Case Study for Australia, *Stochastic Environment Research & Risk Assessment*, 28, 3, 541-554.
- Bates, B.C., Rahman, A., Mein, R.G. and Weinmann, P.E. (1998). Climatic and physical factors that influence the homogeneity of regional floods in south-eastern Australia. *Water Resources Research*, 34, 12, 3369 – 3381.

- Haddad, K., Rahman, A., Weinmann, P.E., Kuczera, G. and Ball, J.E. (2010). Streamflow data preparation for regional flood frequency analysis: Lessons from south-east Australia. *Australian Journal of Water Resources*, 14, 1, 17-32.
- Haddad, K., and Rahman, A. (2012). Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework – Quantile Regression vs. Parameter Regression Technique. *Journal of Hydrology*, 430-431, 142-161.
- Haddad, K., Rahman, A., and Stedinger, J.R. (2012). Regional Flood Frequency Analysis using Bayesian Generalized Least Squares: A Comparison between Quantile and Parameter Regression Techniques, *Hydrological Processes*, 26, 1008-1021.
- Haddad, K., Rahman, A., Zaman, M., Shrestha, S. (2013). Applicability of Monte Carlo cross validation technique for model development and validation using generalised least squares regression, *Journal of Hydrology*, 482, 119-128.
- Haque, M.M., Rahman, A., and Haddad, K. (2014). Rating curve uncertainty in flood frequency analysis: A quantitative assessment. *Journal of Hydrology and Environment Research*, 2(1), 50-58.
- Institution of Engineers Australia (IE Aust.) (1987). Australian Rainfall and Runoff: A Guide to Flood Estimation, Editor: D.H. Pilgrim, Vol.1, Institution of Engineers Australia, Canberra.
- Kuczera, G. (1999). Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *Water Resources, Res*, 35(5):1551–1557.
- Micevski, T., Hackelbusch, A., Haddad, K., Kuczera, G., Rahman, A. (2015). Regionalisation of the parameters of the log-Pearson 3 distribution: a case study for New South Wales, Australia, *Hydrological Processes*, 29, 2, 250-260.
- Mulvaney, T.J. (1851). On the use of self-registering rain and flood gauges in making observations of rainfall and flood discharges in a given catchment. *Transactions of the Institution of Civil Engineers of Ireland*, IV (II), 19–33.
- Pilgrim, D. H. (1982). Assessment of derived rural and urban runoff coefficients. *Civil Engineering Transactions*, IEAust, Vol. CE24, 235-241.
- Pilgrim, D. H., and McDermott, G. (1982). Design floods for small rural catchments in eastern New South Wales. *Civil Engineering Transactions*, IEAust, Vol. CE24, 226-234.
- 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, 805-816.
- Rahman, A., Haddad, K., Kuczera, G., and Weinmann, P.E. (2009). Regional flood methods for Australia: data preparation and exploratory analysis. ARR Revision Project 5 Stage I Report, Engineers Australia, Report No. P5/S1/003, 179.
- Rahman, A., Haddad, K., Haque, M., Kuczera, G., Weinmann, P.E. (2015a). Australian Rainfall and Runoff Project 5: Regional flood methods: Stage 3 Report, Technical Report, No. P5/S3/025, Engineers Australia, Water Engineering, 134pp.
- Rahman, A., Haddad, K., Rahman, A.S., Haque, M.M. (2015b). Australian Rainfall and Runoff Project 5: Regional flood methods: Database used to develop ARR REEF Technique 2015, Technical Report, No. P5/S3/026, Engineers Australia, Water Engineering, 68pp.
- Rahman, A., Haddad, K., Zaman, M., Kuczera, G., and Weinmann, E. (2011). Design flood estimation in ungauged catchments: A comparison between the Probabilistic Rational Method and Quantile Regression Technique for NSW. *Australian Journal of Water Resources*, 14(2), 127 - 139.
- 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, 199-203.
- Rahman, A., Bates, B.C., Mein, R.G. and Weinmann, P.E. (1999). Regional flood frequency analysis for ungauged basins in south-eastern Australia. *Australian Journal of Water Resources*. 3(2): 199-207. ISSN 1324-1583.
- 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, 329-340.