The New Regional Flood Frequency Estimation Model for Australia: RFFE Model 2015

Ataur Rahman [1]
Khaled Haddad [1]
Md M. Haque [1]
George Kuczera [2]
Erwin Weinmann [3]
Peter Stensmyr [4]
Mark Babister [4]
William Weeks [3]

[1]Western Sydney University, Penrith (Kingswood) Campus, NSW 2751, Australia

[2] The University of Newcastle, NSW, Australia

[3] Private Consultant

[4] WMA Water, Sydney, NSW

Abstract

A new regional flood frequency estimation (RFFE) model for Australia has been developed as a part of 4th edition of Australian Rainfall and Runoff, which is referred to as 'RFFE Model 2015'. To develop and test this model, flood data from 853 gauged catchments have been utilised, which includes 798 gauged catchments from the humid coastal areas and 55 catchments from the arid and semi-arid areas. The model allows the derivation of design flood estimates for annual exceedance probabilities (AEP) of 50% to 1% for small to medium catchments anywhere in Australia. In the RFFE Model 2015, the humid coastal and arid/semi-arid areas of Australia have been divided into five and two different regions, respectively. The boundaries between the arid and humid coastal regions have been drawn approximately based on the 500 mm mean annual rainfall contour line. To reduce the effects of sharp variation in flood estimates for the ungauged catchments located close to these regional boundaries, seven fringe zones have been delineated. In the humid coastal regions, a region-of-influence approach has been adopted to derive design flood estimates for ungauged catchments. In developing the prediction model for the regionalised Log Pearson Type 3 distribution, a Bayesian generalised least squares regression technique has been applied, which considers the inter-station correlation and variation in record lengths of the annual maximum flood series across different sites. For the arid/semi-arid regions, a simple index type regional method has been adopted. For easy application by the industry, an application tool has been developed, which automates the application of the RFFE Model 2015. The user is required to provide simple input data (e.g. catchment area and catchment location) to obtain design flood quantiles and associated uncertainty. This paper provides essential technical information, which will assist the user to apply the RFFE Model 2015 in practice with confidence. Further details can be found in the ARR draft chapter on regional flood methods and technical reports.

1. INTRODUCTION

Flood is one of the worst natural disasters that cause significant economic damage. To manage flood risk, one of the primary steps is to estimate a design flood, which is the peak flood flow associated with a specified annual exceedance probability (AEP). To estimate design floods at a given location, one ideally needs a sufficiently long period of recorded streamflow data. However, at many locations of Australia, there are limited or no recorded streamflow data available, representing 'poorly gauged' or 'ungauged catchments', respectively. In these poorly gauged and ungauged catchments, regional flood frequency estimation (RFFE) models are the most commonly used method to estimate design floods. RFFE transfers flood characteristics information from a group of gauged catchments to the catchment location of interest. Even in cases where there is recorded streamflow data it is beneficial to pool the information in the gauged record with the RFFE information.

Design flood estimates at ungauged locations have many applications, e.g. design of culverts, small to medium sized bridges, farm dams, soil conservation works and for many other water management tasks. This can also serve as a 'benchmark' for checking the consistency of rainfall-based design flood estimates. Also, regional estimates of distributional parameters can enhance the accuracy of at-site flood frequency estimates using the Bayesian approach, such as ARR FLIKE (Kuczera, 1999; Kuczera and Frank, 2006). A RFFE model is based on simplified assumptions with respect to data and modeling and it is generally adopted when detailed rainfall runoff modeling is not warranted. It should also be noted that rainfall runoff modelling in ungauged catchments case involves regionalisation of model parameters, which introduces significant uncertainty in design flood estimates, which could be even higher than that of a RFFE technique.

In 1987, Australian Rainfall and Runoff (ARR) recommended various regional flood estimation techniques for different regions of Australia (I. E. Aust., 1987). In ARR 1987, regions were based on state boundaries, which resulted in the potential for inconsistency across boundaries, e.g. the border areas of Victoria and NSW can potentially generate quite different flood estimates based on methods recommended for the two states. In ARR 1987, the Probabilistic Rational Method (PRM) was recommended for Victoria and eastern New South Wales (NSW). The PRM was based on a map of the runoff coefficient (C10); in plotting and using the C10 map, a simple linear geographical interpolation method was adopted, involving an assumption that may not be satisfied at many locations.

As a part of the current revision of the ARR (4thedition), 'Project 5 Regional Flood Methods' for Australia focused on the development, testing and recommendation of new regional flood estimation methods for Australia by incorporating the latest data and regionalization techniques. This paper presents a summary of the RFFE technique, including database and model development and testing, thus providing useful background information to the users.

2. OVERVIEW OF THE RFFE DEVELOPMENT

The research on the development the RFFE technique (reported in this paper) started in 2006 by a team of Australian hydrologists, and has been continuing for the last 10 years. This research has produced a wealth of knowledge in the form of three technical reports (Rahman et al, 2009, 2012, 2015a, b) and over 30 refereed publications (e.g. Rahman et al, 2011; Haddad et al, 2012a; Haddad et al, 2012b; Palmen and Weeks, 2011; Zaman et al, 2012). The draft ARR chapter (Chapter 3, Book 3) for ARR (4th edition) and a RFFE Model in the form of a software have been released for testing. Figure 1 illustrates major steps of the development of the RFFE model. The project has been completed in three stages. Stage 1 involved collation of a national database and testing of various RFFE methods; Stage 2 involved testing of a region-of-influence (ROI) approach to form sub-regions to reduce regional heterogeneity and development of a Bayesian generalised least squares (GLS) regression method to estimate model coefficients (Haddad et al, 2011; Haddad and Rahman, 2012; Micevski et al, 2015). Stage 3 involved finalization of the ARR RFFE model with an enhanced national database and new IFD data, and preparation of the chapter on RFFE in ARR (4th edition) and development of application tool.

3. DATABASE USED IN DEVELOPING THE RFFE TECHNIQUE

In the new RFFE technique, Australia has been divided into humid coastal areas and arid areas. Humid coastal areas include most of the coastal belts of Australia where there are a good number of recorded streamflow stations (Figure 2). An upper limit of catchment size of 1,000 km^2 was adopted. However, in some states (e.g. the NT), a few larger catchments were included as the total number of catchments with areas less than 1,000 km^2 was too small. The cut-off record length was selected as 19 years to maximise the number of eligible stations, on the consideration that a higher cut-off would reduce that number. The selected streams are essentially unregulated since major regulation by storage affects the rainfall-runoff relationship significantly. Streams with minor regulation, such as small farm dams and diversion weirs, are not excluded because this type of regulation is unlikely to have a significant effect on large annual floods.
The data sets for the initially selected potential catchments were further examined, as detailed in Haddad et al (2010), Ishak et al (2013) and Rahman et al (2015): gaps in the annual maximum flood

series were filled as far as could be justified, outliers were detected using the multiple Grubbs-Beck test (Lamontagne et al, 2013), error associated with rating curve extrapolation was investigated and presence of trend was checked. From an initial number of about 1,200 catchments, a total of 798 catchments were finally adopted from all over Australia (excluding catchments in the arid regions, where some of the above criteria were relaxed). The record lengths of the annual maximum flood series of these 798 stations range from 19 to 102 years (median: 37 years). The catchment areas of the selected 798 catchments range from 0.5 km^2 to 4,325 km^2 (median: 178 km^2). The geographical distribution of the selected 798 catchments is shown in Figure 2 and Table 1 presents a summary of the selected catchments.

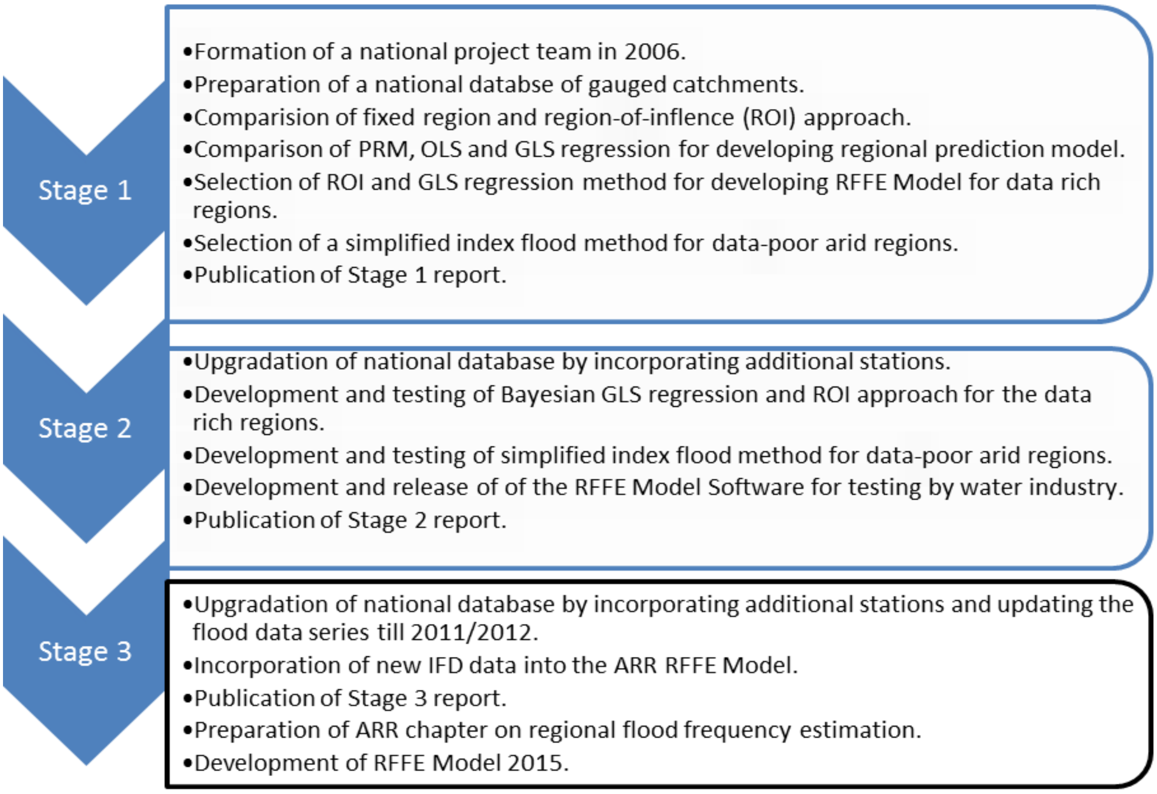

Figure 1 Overview of the development of RFFE Model 2015. (OLS refers to ordinary least squared regression and GLS refers to generalised least squares regression)

In ARR1987, only a few catchments were used from arid areas to develop RFFE methods. Since the publication of ARR1987, there has been little improvement in terms of streamflow data availability in most of the arid areas of Australia. In the preparation of the regional flood estimation database, only a handful of catchments from the arid areas satisfy the selection criteria. To increase the number of stations from the arid areas to develop a ‘reasonably meaningful’ RFFE method, the selection criteria were relaxed i.e. the threshold streamflow record length was reduced to 10 years and the limit of catchment size was increased from 1,000 km2 to 6,000 km2. These criteria resulted in the selection of 55 catchments from the arid areas of Australia (Figure 2 and Table 1). The selected catchments have average annual rainfall smaller than 500 mm. The catchment areas range from 0.1 to 5,975 km2 (median: 259 km2) and streamflow record lengths range from 10 to 46 years (median: 27 years).

The at-site flood frequency analyses (for both the humid coastal and arid areas were conducted using the FLIKE software (Kuczera, 1999). The potentially influential low flows (PILFs) were identified using the multiple Grubbs-Beck test (Lamontagne et al, 2013) and were censored in the flood frequency analysis. A Bayesian parameter estimation procedure with LP3 distribution was used to estimate flood quantiles for each of the gauged sites for AEPs of 50%, 20%, 10%, 5%, 2% and 1%.

Over ten candidate predictor variables were selected initially; however, it was found that the accuracy of a RFFE model was not increased with the number of adopted predictor variables. A total of five candidate predictor variables were finally adopted in the RFFE technique: (i) catchment area in km2 (area); (ii) design rainfall intensity at catchment centroid (in mm/h) for 6-hour duration and AEP of 50% (I6,50); (iii) design rainfall intensity at catchment centroid (in mm/h) for 6-hour duration and AEP of 2% (I6,2); (iv) ratio of design rainfall intensities of AEPs of 2% and 50% for duration of 6-hour (I6,2/I6,50); and (v) catchment shape factor, which is defined as the shortest distance between catchment outlet and centroid divided by the square root of catchment area. Design rainfall values were extracted using the new intensity-frequency-duration (IFD) data from Australian Bureau of Meteorology website.

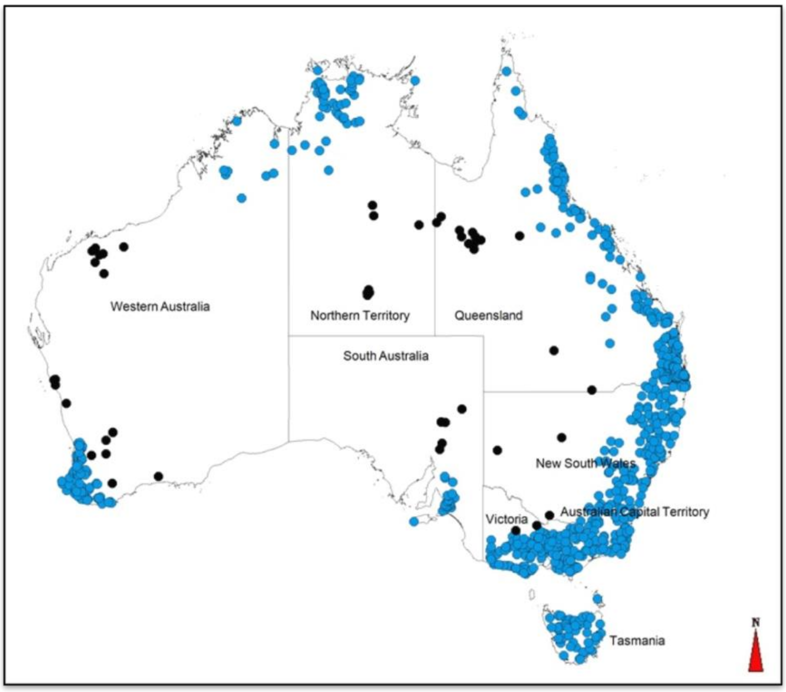

Figure 2 Location of selected 853 catchments: humid coastal (blue circles) and arid/semi-arid areas (black).

Table 1. Summary of the selected 853 catchments (humid coastal and arid/semi-arid areas)

State	No. of stations	Streamflow record length (years) (range and median)	Catchment size (km^2) (range and median)
New South Wales & Australian Capital Territory	176	20 – 82 (34)	1 – 1036 (204)
Victoria	186	20 – 60 (38)	3 – 997 (209)
South Australia	28	20 – 63 (37)	0.6 – 708 (62.6)
Tasmania	51	19 – 74 (28)	1.3 – 1900 (158.1)
Queensland	196	20 – 102 (42)	7 - 963 (227)
Western Australia	111	20 – 60 (30)	0.5 – 1049.8 (49.2)
Northern Territory	50	19 – 57 (42)	1.4 - 4325 (352)
Sub Total	**798**	19 – 102 (37)	0.5 – 4325 (178.5)
Arid areas	55	10 – 46 (27)	0.1 - 5975 (259)
TOTAL	**853**	10 – 102 (36)	0.1 – 5975 (181)

4. ADOPTED RFFE TECHNIQUE

In the new RFFE technique, Australia has been divided into seven regions. There are five regions in the humid coastal areas, as shown in Table 2 and Figure 3. The remaining two regions are located in the semi-arid and arid areas (referred to as arid areas) of Australia. The boundary of the fringe zone located near humid coastal region was approximately defined by the 500 mm mean annual rainfall isohyet, while the other side was defined by 400 mm mean annual rainfall isohyet to establish a fringe zone. There are seven fringe zones, as shown in Figure 3.

In the adopted RFFE technique for the humid coastal areas of Australia a Region of Influence (ROI) approach is applied, where information from a number of catchments surrounding the site of interest is pooled for regional flood estimation. The first three moments of the annual maximum flood series (i.e. the mean, standard deviation and skewness of the natural logarithms of the annual maxima) are extracted to estimate the flood quantiles from the LP3 distribution. This method is referred to as the parameter regression technique (PRT). The flood quantile estimates from the LP3 distribution are described by the following equation:

$$\ln Q_x = M + K_x S \tag{1}$$

where Q_x = the discharge having an AEP of x% (design flood or flood quantile);
M = mean of the natural logarithms of the annual maximum flood series;
S = standard deviation of the natural logarithms of the annual maximum flood series; and
K_x = frequency factor for the LP3 distribution for AEP of x%, which is a function of the AEP and the skewness (SK) of the natural logarithms of the annual maximum flood series.

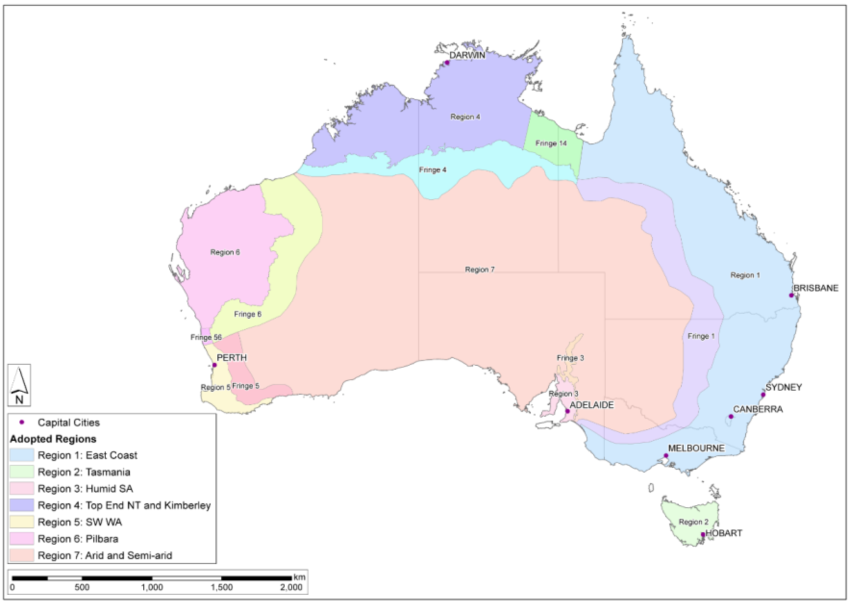

Figure 3 Adopted regions for RFFE technique in Australia

Table 2. Details of RFFE technique for humid coastal areas of Australia

Region	Method to form region	Number of stations	Estimation model
Region 1: East Coast	ROI (based on geographical proximity)	558	Bayesian GLS regression-PRT
Region 2: Tasmania		51	
Region 3: Humid SA		28	
Region 4: Top End NT and Kimberly		58	
Region 5: SW WA		103	

The prediction equations for M, S and SK have been developed for all the gauged catchment locations in the high rainfall areas using Bayesian GLS regression and model parameters noted. These model parameters are integrated with the RFFE Model 2015. The adopted estimation equations for M, S and SK for the regional LP3 model (equation 1) have the following general form:

$$M = b_0 + b_1(\ln(area)) + b_2(\ln(I_{6,50})) + b_3(\ln(shape\ factor)) \quad (2)$$
$$S = c_0 + c_1(\ln(I_{6,2}/I_{6,50})) \quad (3)$$
$$SK = d_0 + d_1(\ln(area)) + d_2(\ln(I_{6,2}/I_{6,50})) + d_3(\ln(I_{6,2})) \quad (4)$$

where, *area* = catchment area (km^2);
$I_{6,50}$ = design rainfall intensity (mm/h) at catchment centroid for 6-hour duration and AEP of 50%;
shape factor = shortest distance between catchment outlet and centroid/$area^{0.5}$; and
$I_{6,2}$ = design rainfall intensity (mm/h) at catchment centroid for 6-hour duration and AEP of 2%.

For Region 1 and Region 2, equations 3 and 4 have only the model intercepts, i.e. for these two regions, the weighted average values of *S* and *SK* are adopted, which are determined on the basis of record lengths at the stations within the ROI sub-region. The values of b_0, b_1, b_2, b_3, c_0, c_1, d_0, d_1, d_2 and d_3 at all the 798 individual gauged catchment locations (in the high rainfall regions) were estimated and embedded in the REEF Model 2015. To derive flood quantile estimate at an ungauged catchment of interest, RFFE Model 2015 takes the inverse distance weighted average value of flood quantile estimates from up to 15 of the nearest gauged catchment locations within a 300 km radius from the catchment of interest. This ensures a smooth variation of flood quantile estimates over space.

In developing the confidence limits for the estimated flood quantiles, a Monte Carlo simulation approach was adopted by assuming that the uncertainty in the first three moments (i.e. the mean, standard deviation and skewness of the logarithms of the annual maximum flood series) can be specified by a multivariate normal distribution. Here the correlations among the three moments for a given region were estimated from the residuals of the GLS regression models of the moments. The mean of the moment is given by its regional predicted value and the standard deviation of the moment is the square root of the average variance of prediction of the moment at the nearest gauged sites. Based on 10,000 simulated values of the moments from the multivariate normal distribution as defined above, 10,000 Q_x values are estimated, which are then used to develop the 90% confidence intervals.

The formation of regions in the arid areas in Australia is a difficult task, as there are only 55 catchments available from this vast part of Australia. There are two alternatives: (i) formation of one region with all the 55 stations; and (ii) formation of smaller sub-regions based on geographical proximity, noting that too small a region makes the developed RFFE model of little statistical significance. Examination of a number of alternative sub-regions led to the formation of two regions from the 55 arid catchments: Region 6 (11 catchments from the Pilbara area of WA) and Region 7 (44 catchments from all other arid areas). Application of the ROI and PRT methods for arid regions was deemed inappropriate as the ROI method requires a minimum number of gauging stations to form sub-regions and the number of gauging stations in the arid areas of Australia is insufficient for this purpose. Hence a simpler RFFE method is considered more appropriate for the two arid regions. Here, an index type approach as suggested by Farquharson *et al*. (1992) is adopted. The 10% AEP flood quantile (Q_{10}) is used as the index variable and a dimensionless growth factor for AEP of *x*% (GF_x) is used to estimate Q_x:

$$Q_x = Q_{10} \times GF_x \quad (5)$$

A prediction equation is developed for Q_{10} as a function of catchment characteristics, and regional growth factors were developed based on the observed annual maximum flood series data (where quantiles were estimated using LP3 distribution, similar to humid coastal areas). The Q_x/Q_{10} values were first estimated at individual stations; the weighted average of these values (weighting based on record length at individual sites) over all the stations in a region then defined the growth factors (GF_x) for the region.

The adopted prediction equation for the index variable Q_{10} has the following form:

$$\log_{10}(Q_{10}) = b_0 + b_1(\log_{10}(area)) + b_2(\log_{10}(I_{6,50})) \quad (6)$$

where b_0, b_1 and b_2 are regression coefficients, estimated using ordinary least squares regression; *area* represents catchment area in km^2, and $I_{6,50}$ is the design rainfall intensity (mm/h) at catchment centroid for 6-hour duration and 50% AEP. The values of b_0, b_1 and b_2 and the regional growth factors (GF_x) are embedded into the RFFE Model 2015. For the fringe zones, the flood estimate at an ungauged catchment location is taken as the inverse distance weighted average value of the flood estimates based on the two nearest regions. This interpolation method is embedded into the ARR RFFE Model 2015.

The RFFE technique described above has been incorporated into a software tool referred to as RFFE Model 2015. Figure 4 presents a screen shot of the software landing page. The RFFE Model 2015 is applicable to any catchment that has similar attributes and flood producing characteristics as the catchments used in the derivation of the flood estimation equations. Catchments which do not satisfy this requirement can be divided into three groups: (i) catchments which have been substantially modified from their natural characteristics and for which the RFFE Model 2015 is not applicable and should thus not be used, (ii) catchments for which flood estimates must be expected to have lower accuracy; and (iii) 'atypical catchments' where additional catchment attributes need to be considered and adjusted for. This issue is further detailed in the draft ARR chapter on regional flood methods (Rahman et al., 2015c).

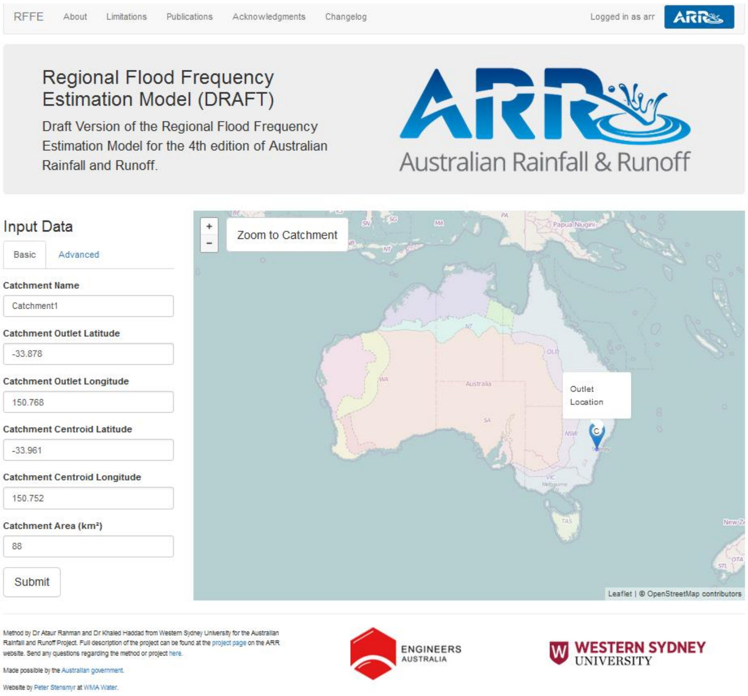

Figure 4 Screen shot of RFFE Model 2015 (landing page).

5. CONCLUSION

This paper presents an overview of the development of ARR RFFE Model 2015 for Australia to be included in Australian Rainfall and Runoff (ARR) - 4th Edition. The ARR RFFE Model will allow the derivation of design flood estimates for annual exceedance probabilities (AEP) of 50% to 1% for catchments of 0.1 to 1,000 km2 anywhere in Australia. In the development of the model, data from 798 catchments from the humid coastal areas and 55 catchments from the semi-arid and arid areas have been used. Australia has been divided into several regions to apply the RFFE Model. In the six humid coastal regions, a region-of-influence (ROI) approach has been adopted to form sub-regions. In developing the prediction equations, a Bayesian generalised least squares (GLS) regression technique has been adopted for the humid coastal regions, which considers the inter-station correlation and variation in record lengths from site to site in developing regional parameter prediction equations. A regionalised Log Pearson Type 3 (LP3) distribution is used to derive design flood estimates for ungauged catchments in the range of AEPs of 50% to 1%. For the arid regions, a simplified index type regional flood frequency method has been adopted. The RFFE Model 2015 may not be able to generate flood quantiles very accurately in few cases such as catchments having significant natural flood plain storage areas as explained in the ARR chapter on regional flood methods. The new ARR RFFE Model 2015 will have a wider application in estimating design floods for ungauged catchments, as well as providing prior information for at-site flood frequency analysis using

ARR-FLIKE and a useful means of benchmarking other flood estimation methods in Australia. However, the RFFE Model has limitations which need to be recognised and understood by the users.

6. ACKNOWLEDGEMENTS

ARR Revision Project 5 was made possible by funding from the Federal Government through the Department of Climate Change and Energy Efficiency and Geoscience Australia. Project 5 reports and the associated publications including this paper are the result of a significant amount of in kind hours provided by Engineers Australia Members. The authors would also like to acknowledge various agencies in Australia for supplying data: Australian Bureau of Meteorology, DSE (Victoria), Thiess Services (Victoria), DTM (Qld), DERM (Qld), ENTURA, DPIPWE (TAS), DWLBC (SA), NRETAS (NT), University of Western Sydney, University of Newcastle, University of South Australia, University of New South Wales, Department of Environment, Climate Change and Water (NSW), Department of Water (WA) and WMA Water (NSW).

7. REFERENCES

Haddad, K., Rahman, A., Weinmann, P.E., Kuczera, G., 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., Rahman, A., Kuczera, G. (2011). 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, 15, 2, 59-70.

Haddad, K., Rahman, A. (2012a). 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 (2012), 142-161.

Haddad, K., Rahman, A., Stedinger, J.R. (2012b). Regional Flood Frequency Analysis using Bayesian Generalized Least Squares: A Comparison between Quantile and Parameter Regression Techniques, Hydrological Processes, 26, 1008-1021.

Institution of Engineers Australia (I. E. Aust.) (1987). Australian Rainfall and Runoff: A Guide to Flood Estimation, Editor: Pilgrim, D.H., Engineers Australia, Canberra.

Ishak, E., Rahman, A., Westra, S., Sharma, A., Kuczera, G. (2013). Evaluating the Non-stationarity of Australian Annual Maximum Floods. Journal of Hydrology, 494, 134-145.

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

Kuczera G. and Frank S. (2006*). Australian Rainfall and Runoff, Book IV, Estimation of Peak Discharge–Draft*, Engineers Australia, downloaded 12 July 2012.

Lamontagne, J. R., Stedinger, J. R., Cohn, T. A., Barth, N.A. (2013*). Robust National Flood Frequency Guidelines: What is an Outlier?,* World Environmental and Water Resources Congress 2013, 2013, pp. 2454-2466.

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.

Palmen, L.B., Weeks, W.D. (2011). Regional flood frequency for Queensland using the quantile regression technique, Australian Journal of Water Resources, 15, 1, 47-57.

Rahman, A., Haddad, K., Kuczera, G., Weinmann, E. (2009). *Australian Rainfall and Runoff Revision Project 5: Regional Flood Methods. Stage 1 report*, Engineers Australia.

Rahman, A., Haddad, K., Zaman, M., Kuczera, G., Weinmann, P.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, 127-137.

Rahman, A., Haddad, K., Zaman, M., Ishak, E., Kuczera, G., Weinmann, P.E. (2012). Australian Rainfall and Runoff Revision Projects, Project 5 Regional flood methods, Stage 2 Report.

Rahman, A., Haddad, K., Rahman, A.S., Haque, M.M. (2015a). 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., Haque, M., Kuczera, G., Weinmann, P.E. (2015b). 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., Kuczera, G., Weinmann, P.E. (2015c). Regional flood methods, in Australian Rainfall & Runoff, Chapter 3, Book 3, edited by J. E. Ball, Engineers Australia, 78-114 (In press).

Zaman, M., Rahman, A., Haddad, K. (2012). Regional flood frequency analysis in arid regions: A case study for Australia. Journal of Hydrology, 475, 74-83.