

Development and Validation of Artificial Intelligence Based Regional Flood Estimation Model for Eastern Australia

K. Aziz^a, F. Kader^b, A. Ahsan^{c,d} and A. Rahman^a

^a*School of Computing Engineering and Mathematics and Institute for Infrastructure Engineering, University of Western Sydney, Australia*

^b*Private Consultant*

^c*Department of Civil Engineering, and Institute of Advanced Technology, University Putra Malaysia, Serdang, Malaysia*

^d*Department of Civil & Environmental Engineering, University of Sharjah, PO Box 27272, Sharjah, UAE*

Email: faruk.kader718@gmail.com

Abstract: This paper presents the development and validation of an artificial intelligence based regional flood frequency analysis (RFFA) model for application in ungauged catchments of eastern Australia. The artificial intelligence based techniques with a flexible model structure and non-linear approach can overcome the limitations of the conventional RFFA models, which are generally based on linear relationship between flood statistics and catchment characteristics. Till now, there have been limited applications of artificial intelligence based techniques to RFFA problems in Australia. This study has developed four artificial intelligence based RFFA models for eastern Australia, using a comprehensive flood database available as a part of Australian Rainfall and Runoff (ARR) revision 'Project 5 Regional flood methods'. These four RFFA models are based on artificial neural network (ANN), genetic algorithm based artificial neural network (GAANN), gene-expression programming (GEP) and co-active neuron fuzzy inference system (CANFIS).

A total of 452 catchments from the states of New South Wales, Victoria, Queensland and Tasmania have been considered by this study. These dataset is divided into training and validation sets. Data of 362 catchments (training data set) have been used to train the model and the data from the remaining 90 catchments (validation data set) used to validate the model. The models have been trained/calibrated using the training data set that involved minimisation of the mean squared error between the observed and predicted flood quantiles by the model (being trained) for a given ARI for the training data set. Six average recurrence intervals (ARI) flood quantiles (2, 5, 10, 20, 50 and 100 years) were considered in this study. Four evaluation statistics are adopted to assess the model accuracy: median ratio of the predicted flood quantile (Q_{pred}) and observed flood quantile (Q_{obs}), denoted by Q_{pred}/Q_{obs} ratio, plots of Q_{obs} and Q_{pred} , median relative error and coefficient of efficiency. This is initially done for the training data set and then repeated for the validation data set.

The artificial intelligence based RFFA models have been ranked based on their relative performances in relation to the above criteria to identify the best trained/calibrated model. It has been found that none of the four models is superior across all the six ARIs against the adopted criteria. The ANN based RFFA model has a better ranking score in terms of training/calibration; therefore, it is suggested that the ANN-based RFFA model is the best calibrated model among the four artificial intelligence based RFFA models for eastern Australia.

Keywords: *Floods, genetic algorithm, regional flood frequency analysis, ungauged catchments, ANN*

1. INTRODUCTION

Regional flood frequency analysis (RFFA) is widely used for estimation of design floods in ungauged catchments (I. E. Aust., 1987). It is also used to enhance the accuracy of flood estimates in poorly gauged catchments (e.g. when there is a shortage of recorded flood data and poor data quality). The commonly used techniques for developing the regional flood prediction equations include the probability rational method (PRM), index flood method (IFM) and quantile regression technique (QRT). A significant number of studies on RFFA have been done in Australia recently (e.g. Bates *et al.*, 1998; Rahman *et al.*, 1999, Rahman *et al.*, 2002, Rahman, 2005; Rahman *et al.*, 2011; Haddad and Rahman, 2011; Haddad *et al.*, 2010; Haddad *et al.*, 2012; Haddad and Rahman, 2012; Haddad *et al.*, 2014, Caballero and Rahman, 2014a; Caballero and Rahman, 2014b; Rahman *et al.*, 2015a, Rahman *et al.*, 2015b). Recent studies have shown that the QRT can provide more accurate design flood estimates than the PRM (Rahman *et al.*, 2011). Quantile regression technique has also been compared with Parameter Regression Technique (PRT) (Haddad *et al.*, 2012; Haddad and Rahman, 2012; Micevski *et al.*, 2015).

The RFFA methods generally assume linear relationships between flood statistics and predictor variables in log domain for developing the regional prediction equations, although the hydrologic processes are inherently nonlinear and exhibit a high degree of spatial and temporal variability, as is the case with the catchments in Australia. Non-linear techniques such as artificial intelligence based models can be employed to overcome the limitations associated with the existing linear RFFA models. In the last two decades, artificial intelligence based techniques such as artificial neural networks (ANN), genetic algorithm based ANN (GAANN), gene expression programming (GEP) and co-active neuro-fuzzy inference system (CANFIS) have been applied in water resources engineering such as rainfall runoff modeling and hydrologic forecasting, (e.g. Daniell, 1991; Muttiah *et al.*, 1997; Jain and Srinivasulu, 2004; Dawson *et al.*, 2006; Shu and Ouarda, 2008, Gao *et al.*, 2010). A very limited number of artificial intelligence applications on RFFA has been done in Australia (Aziz *et al.*, 2011; Aziz *et al.*, 2014; Aziz *et al.*, 2015). To fill this knowledge gap, this study evaluates four artificial intelligence based RFFA models, which are ANN, GAANN, GEP and CANFIS.

2. DATA

A total of 452 small to medium sized gauged catchments have been selected from New South Wales (NSW), Victoria (VIC), Queensland (QLD) and Tasmania (TAS) states of eastern Australia for this study as shown in Figure 1. At site flood quantiles have been estimated using ARR FLIKE (Kuczera, 1999) software for average recurrence intervals (ARIs) of 2, 5, 10, 20, 50 and 100 years using Bayesian Log Pearson Type 3 (LP3) distribution. In this study, five catchment characteristics and eight model combinations are considered (Table 1). For this purpose, 80% (362 catchments) are used to train the model (training data set) and the remaining 20% (90 catchments) are used to validate the model (validation data set)

Table 1. Candidate models and catchment characteristics.

Model ID	Variables	Variable name
1	A, I_{tc_ARI}	<p>A: catchment area I_{tc_ARI}: design rainfall intensity S: slope E: evapo-transpiration R: mean annual rainfall</p>
2	A, I_{tc_ARI}, S	
3	A, I_{tc_ARI}, E	
4	A, I_{tc_ARI}, R	
5	A, I_{tc_ARI}, S, E	
6	A, I_{tc_ARI}, R, E	
7	A, I_{tc_ARI}, R, S	
8	A, I_{tc_ARI}, R, S, E	

3. TRAINING OF FOUR ARTIFICIAL INTELLIGENCE BASED RFFA MODELS

Each of the four artificial intelligence based RFFA models has been trained using MATLAB codes (developed as a part of this research) by minimising the mean squared error between the observed and predicted flood quantiles for each of six ARIs (2, 5, 10, 20, 50 and 100 years). This is done using the training data set consisting of 362 catchments. The artificial intelligence based RFFA models have been evaluated

against four criteria: median ratio of the predicted flood quantile (Q_{pred}) and observed flood quantile (Q_{obs}), denoted by Q_{pred}/Q_{obs} ratio, plots of Q_{obs} and Q_{pred} , median relative error (RE) and coefficient of efficiency (CE) (Nash and Sutcliffe, 1970). This is initially done for the training data set and then repeated for the validation data set. Models are ranked based on their relative performances in relation to these criteria to identify the best trained/calibrated model.

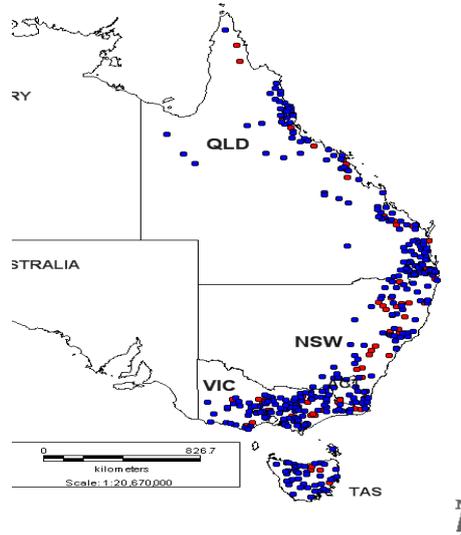


Figure 1. Study catchments (Red indicates test catchments).

Table 2 shows the CE values for the ANN, GANN, GEP and CANFIS based RFFA models. Among these four models, the GAANN is found have the highest CE values for ARIs of 2, 5, 10 and 20 years. For ARIs of 50 and 100 years, the ANN has the highest CE values. Considering all the six ARIs, GAANN has the highest CE value (0.71) and the three other models have similar CE values in the range of 0.67 to 0.66.

Table 3 shows the median Q_{pred}/Q_{obs} ratio values for the four artificial intelligence based RFFA models. The ANN based RFFA model shows the best performance (i.e. Q_{pred}/Q_{obs} ratio value is closest to 1.00 for ARIs of 20, 50 and 100 years). Considering all the six ARIs, the ANN based RFFA model performs better than any of the other three models with an overall Q_{pred}/Q_{obs} ratio value of 1.09. The second best performance is demonstrated by the GEP based RFFA model (1.19), while the GAANN and CANFIS perform similarly. In terms of consistency over the ARIs, GAANN, GEP and CANFIS show very high Q_{pred}/Q_{obs} ratio values for some ARIs as can be seen in Table 3. Here again, the ANN shows the best consistency over the ARIs.

Table 2. CE values of four artificial intelligence based RFFA models.

ARI (years)	ANN	GAANN	GEP	CANFIS
2	0.59	0.76	0.69	0.64
5	0.73	0.79	0.72	0.67
10	0.64	0.76	0.73	0.75
20	0.71	0.76	0.65	0.73
50	0.70	0.57	0.61	0.53
100	0.64	0.63	0.57	0.62
Overall	0.67	0.71	0.66	0.66

Table 3. Median Q_{pred}/Q_{obs} ratio values of four artificial intelligence based RFFA models.

ARI (years)	ANN	GAANN	GEP	CANFIS
2	1.03	1.22	0.99	1.76
5	1.12	1.20	1.08	0.99
10	1.06	1.02	1.08	0.87
20	1.10	1.11	1.17	1.26
50	1.08	1.52	1.45	1.04
100	1.15	1.18	1.39	1.36
Overall	1.09	1.21	1.19	1.21

Table 4 shows the median of the absolute relative error values for the ANN, GAANN, GEP and CANFIS based RFFA models. It shows that ANN based RFFA model performs better than the other models with an overall median RE value of 42.07% for the six ARIs. Both GEP and CANFIS have relatively higher RE values (GEP = 54.02%, CANFIS = 59.46%). Importantly, CANFIS shows very high RE values for 2 years ARI (94.02%) and 50 years ARI (71.94%). Overall, ANN performs best in terms of RE value, followed by GAANN, GEP and CANFIS. The predicted and observed flood quantiles for the four artificial intelligence based RFFA model for ARI 20 years (Q_{20}) are shown in Figure 2 to 5. The reason of adopting Q_{20} is that it is the frequently applied ARI in design.

Table 4. Median RE values of four artificial intelligence based RFFA models.

ARI (years)	ANN	GAANN	GEP	CANFIS
2	43.75	40.92	73.3	94.02
5	39.53	39.31	43.91	43.55
10	39.14	41.01	43.25	45.27
20	40.38	40.29	54.61	46.07
50	43.32	60.00	54.22	71.94
100	46.30	45.28	54.82	55.89
Overall	42.07	44.47	54.02	59.46

Figures 2 to 5 generally present a good agreement; however, there are some over estimations by the ANN-based RFFA model when the observed flood quantiles are smaller than 50 m³/s for all the ARIs except 50 years (not shown in this paper). Overall, the ANN based RFFA model shows better training results for higher discharges. The CE, median Q_{pred}/Q_{obs} ratio and median relative error values for ANN based RFFA model are compared in Table 5 for the training and validation datasets. The best agreement between the training and validation data sets for CE value is found for ARIs of 10, 20 and 50 years, while a relatively poor agreement is noticed for the ARI of 100 years. The median Q_{pred}/Q_{obs} ratio values agree quite well between the training and validation data sets. In case of the median relative error values, the best agreement between the training and validation data sets is found for ARIs of 5 and 100 years. The results suggest different degrees of agreement between the training and validation data sets for the ANN based RFFA model for different ARIs.

The training results indicate that none of the four models has explicit superiority above other models across the six ARIs and four statistical criteria. Therefore, a ranking system is adopted to score the performance results of the four models for the training and validation cases. Four different ranks are used, with a relative score ranging from 4 to 1. A model is ranked 1 for a criterion 1, if it scores 4. For ranks of 2, 3 and 4, scores of 3, 2 and 1, respectively are assigned. For training, the ranking is applied against the previous four different criteria as shown in Table 6. The table shows that ANN based model has the overall highest score of 15, while CANFIS has the lowest score of 7. In Table 7, the rankings of the four artificial intelligence based RFFA models with regard to the agreement between the training and validation results are provided. Figures 6 and 7 compare the boxplots of relative error values for the ANN and GEP based RFFA model, which shows that ANN has much smaller error bands compared with the GEP based RFFA model. Boxplots for the other methods (not shown here) also reveal the superiority of the ANN based RFFA model.

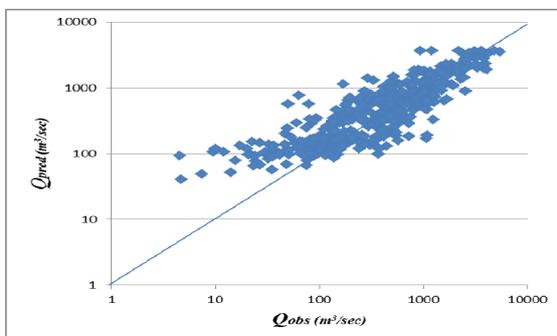


Figure 2. Observed and predicted flood quantiles for ANN based RFFA model for Q_{20} .

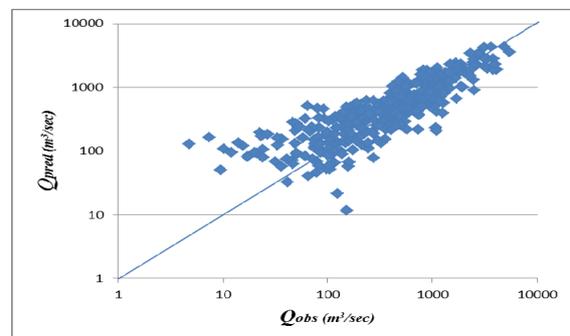


Figure 3. Observed and predicted flood quantiles for GAANN based RFFA model for Q_{20} .

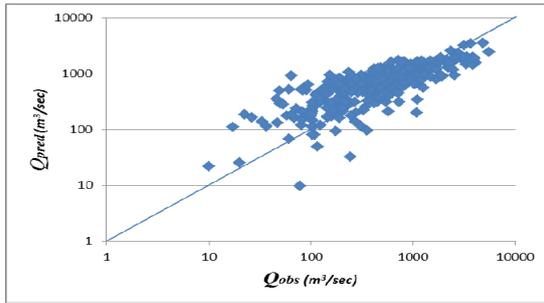


Figure 4. Observed and predicted flood quantiles for GEP based RFFA model for Q_{20} .

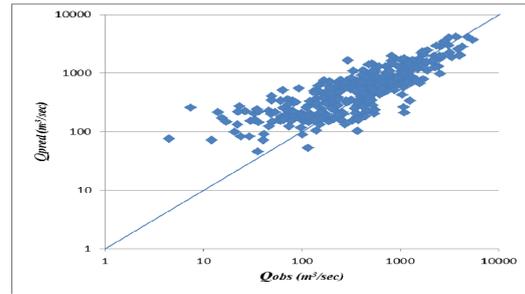


Figure 5. Observed and predicted flood quantiles for CANFIS based RFFA model for Q_{20} .

Table 5. Training and validation results for the ANN based RFFA model.

ARI (years)	Training			Validation		
	CE	Q_{pred}/Q_{obs} ratio (median)	RE (%) (median)	CE	Q_{pred}/Q_{obs} ratio (median)	RE (median)
2	0.59	1.03	43.75	0.69	1.04	37.56
5	0.73	1.12	39.53	0.59	0.99	40.39
10	0.64	1.06	39.14	0.63	1.02	44.63
20	0.71	1.10	40.38	0.69	1.04	35.62
50	0.70	1.08	43.32	0.68	1.14	39.09
100	0.64	1.15	46.30	0.40	1.10	44.53
Overall	0.67	1.09	42.07	0.61	1.06	40.30

Table 6. Ranking of the four artificial intelligence based RFFA models (with respect to training).

Criterion	Rank 1	Rank 2	Rank 3	Rank 4
Scatter plot of Q_{obs} Vs Q_{pred}	ANN	GANN	CANFIS	GEP
Median Q_{pred}/Q_{obs}	ANN	GEP	GAANN/CANFIS	#
Median RE	ANN	GAANN	GEP	CANFIS
Median CE	GAANN	ANN	GEP/CANFIS	#
Overall Score: ANN-15, GAANN-12, GEP-10, CANFIS-7				

Table 7. Ranking of four artificial intelligence based RFFA models (with respect to validation).

Criterion	Rank 1	Rank 2	Rank 3	Rank 4
Median Q_{pred}/Q_{obs}	GEP	ANN	CANFIS	GAANN
Median RE (%)	GEP	ANN	GAANN	CANFIS
Median CE	GAANN	ANN	CANFIS	GEP
Overall Score: ANN-9, GEP-9, GAANN-7, CANFIS-5				

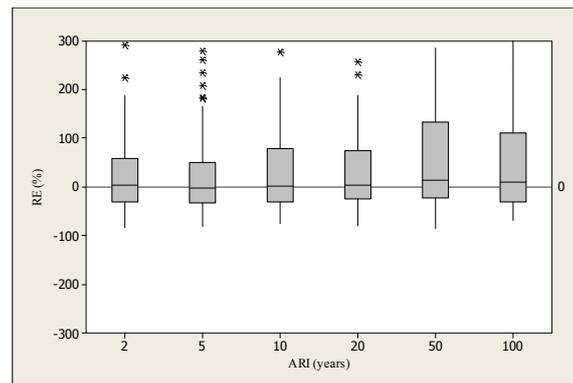


Figure 6. Boxplot of relative error (RE) values for ANN based RFFA model.

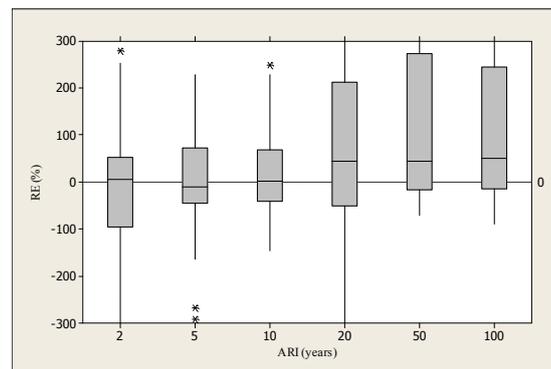


Figure 7. Boxplot of relative error (RE) values for GEP based RFFA model.

4. CONCLUSION

This study has carried out training and validation of four artificial intelligence based regional flood frequency analysis (RFFA) models to identify a suitable non-linear technique for application in eastern Australia. These four RFFA models are based on artificial neural network (ANN), genetic algorithm based artificial neural network (GAANN), gene-expression programming (GEP) and co-active neuron fuzzy inference system (CANFIS). Six average recurrence intervals (ARI) and a range of statistical criteria are adopted. The training results indicate that none of the four models has explicit superiority above other models across all the six ARIs and four statistical criteria. The validation results show that the ANN based RFFA model has the best performance among the four models considered for eastern Australia.

REFERENCES

- Aziz, K., Rahman, A., Shrestha, S. and Fang, G. (2011). Derivation of optimum regions for ANN based RFFA in Australia, 34th IAHR World Congress, 26 June – 1 July 2011, Brisbane, 17-24.
- Aziz, K., Rahman, A., Fang, G., Shrestha, S. (2014). Application of Artificial Neural Networks in regional flood frequency analysis: A case study for Australia, *Stoch Environ Res Risk Assess*, 28, 3, 541-554.
- Aziz, K., Rai, S., and 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.
- 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 Resour Res*, 34, 12, 3369 – 3381.
- Caballero, W. L., and Rahman, A. (2014a). Development of regionalized joint probability approach to flood estimation: A case study for New South Wales, Australia, *Hydrological Processes*, 28, 13, 4001-4010.

- Caballero, W.L., and Rahman, A. (2014b). Application of Monte Carlo simulation technique for flood estimation for two catchments in New South Wales, Australia, *Natural Hazards*, 74, 1475-1488.
- Daniell, T.M. (1991). Neural networks—applications in hydrology and water resources engineering. International Hydrology & Water Resources Symposium. Perth, 2-4 October, Engineers Australia.
- Dawson, C.W., Abrahart, R.J., Shamseldin, A.Y., and Wilby, R.L. (2006). Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, 319, 391–409.
- Gao, C., Gemmer, M., Zeng, X., Liu, B., Su, B., and Wen, Y. (2010). Projected streamflow in the Huaihe River Basin (2010-2100) using artificial neural network. *Stoch Environ Res Risk Assess*, 24, 685-697.
- 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. (2011). Regional flood estimation in New South Wales Australia using Generalised Least Squares Quantile Regression. *Journal of Hydrologic Engineering*, ASCE, 16, 11, 920-925, doi:10.1061/(ASCE)HE.1943-5584.0000395.
- 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., 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 (2012), 142-161.
- Haddad, K., Rahman, A., and Ling, F. (2014). Regional flood frequency analysis method for Tasmania, Australia: A case study on the comparison of fixed region and region-of-influence approaches, *Hydrological Sciences Journal*, DOI:10.1080/02626667.2014.950583.
- Institution of Engineers Australia (I. E. Aust.) (1987). Australian Rainfall and Runoff: A Guide to Flood Estimation. Editor: D.H. Pilgrim, Vol.1, I. E. Aust., Canberra.
- Jain, A., and Srinivasulu, S. (2004). Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms and artificial neural network techniques. *Water Resources Research*, 40, W04302, doi:10.1029/2003WR002355.
- Kuczera, G. (1999). Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian Inference, *Water Resources Research*, 35(5), 1551-1557.
- Muttiah, R.S., Srinivasan, R. and Allen, P.M. (1997). Prediction of two year peak stream discharges using neural networks. *Journal of the American Water Resources Association*, 33 (3), 625–630.
- 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.
- Nash, J. E., and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models, Part I - A discussion of principles. *Journal of Hydrology*, 10, 282–290.
- 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.
- Rahman, A., Weinmann, P.E., Hoang, T.M.T and Laurenson, E.M. (2002). Monte Carlo Simulation of flood frequency curves from rainfall. *Journal of Hydrology*, 256 (3-4), 196-210. ISSN 0022-1694.
- Rahman, A., (2005). A quantile regression technique to estimate design floods for ungauged catchments in South-east Australia. *Australian Journal of Water Resources* 9(1), 81-89.
- Rahman, A., Haddad, K., Zaman, M., Kuczera, G., and 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.
- Shu, C. and Ouarda, T.B.M.J. (2008). Regional flood frequency analysis at ungauged sites using the adaptive neuro-fuzzy inference system. *Journal of Hydrology*, 349, 31-43.