

Development of Artificial Intelligence Based Regional Flood Estimation Techniques for Eastern Australia

Kashif Aziz, Ataur Rahman and Asaad Shamseldin

Abstract This chapter focuses on the development of artificial intelligence based regional flood frequency analysis (RFFA) techniques for Eastern Australia. The techniques considered in this study include artificial neural network (ANN), genetic algorithm based artificial neural network (GAANN), gene-expression programming (GEP) and co-active neuro fuzzy inference system (CANFIS). This uses data from 452 small to medium sized catchments from Eastern Australia. In the development/training of the artificial intelligence based RFFA models, the selected 452 catchments are divided into two groups: (i) training data set, consisting of 362 catchments; and (ii) validation data set, consisting of 90 catchments. It has been shown that in the training of the four artificial intelligence based RFFA models, no model performs the best across all the considered six average recurrence intervals (ARIs) for all the adopted statistical criteria. Overall, the ANN based RFFA model is found to outperform the other three models in the training. Based on an independent validation, the median relative error values for the ANN based RFFA model are found to be in the range of 35–44 % for eastern Australia. The results show that ANN based RFFA model is applicable to eastern Australia.

1 Introduction

Flood is one of the worst natural disasters that cause significant impacts on economy. About 951 people lost their lives and another 1326 were injured in Australia by floods during 1852–2011 [1]. During 2010–11, over 70 % of Queensland state

K. Aziz · A. Rahman (✉)

School of Computing, Engineering and Mathematics, Western Sydney University Australia,
Building XB, Kingswood, Locked Bag 1797, Penrith, NSW 2751, Australia
e-mail: a.rahman@westernsydney.edu.au

A. Rahman

Institute for Infrastructure Engineering, Western Sydney University, Penrith, NSW, Australia

A. Shamseldin

University of Auckland, Auckland, New Zealand

was affected by severe flooding, with the total damage to public infrastructure was estimated to be \$5 billion [2]. Design flood estimate, a probabilistic flood magnitude, is used in engineering design to safeguard water infrastructures and to minimise the overall flood damage. To estimate design floods, at-site flood frequency analysis is the most commonly adopted technique; however, at the locations where streamflow record is unavailable or is of limited length or of poor quality, regional flood frequency analysis (RFFA) is generally adopted to estimate design floods. A RFFA technique attempts to use flood data from a group of homogeneous donor gauged catchments to make flood estimation at ungauged location of interest.

For developing the regional flood prediction equations, the commonly used techniques include the rational method, index flood method and quantile regression technique. These techniques generally adopt a linear method of transforming inputs to outputs. Since hydrologic systems are often non-linear, RFFA techniques based on non-linear methods can be a better alternative to linear methods. Among the non-linear methods, artificial intelligence based techniques have been widely adopted in various water resources engineering problems. However, their application to RFFA problems is quite limited.

This chapter presents the development of artificial intelligence based RFFA methods for eastern Australia. The non-linear techniques considered in this chapter are artificial neural network (ANN), genetic algorithm based artificial neural network (GAANN), gene-expression programming (GEP) and co-active neuro fuzzy inference system (CANFIS).

2 Review of Artificial Intelligence Based Estimation Methods in Hydrology

RFFA essentially consists of two principal steps: (i) formation of regions; and (ii) development of prediction equations. Regions have traditionally been formed based on geographic, political, administrative or physiographic boundaries [3, 4]. Regions have also been formed in catchment characteristics data space using multivariate statistical techniques [5, 6]. Regions can also be formed using a region-of-influence approach where a certain number of catchments based on proximity in geographic or catchment attributes space are pooled together based on an objective function to form an optimum region [7–9].

For developing the regional flood prediction equations, the commonly used techniques include the rational method, index flood method and quantile regression technique (QRT). The rational method has widely been adopted in estimating design floods for small ungauged catchments [4, 10–12]. The index flood method has widely been adopted in many countries, which relies on the identification of homogeneous regions [13–18]. The QRT, proposed by the United States Geological Survey (USGS), has been applied by many researchers using either an Ordinary Least Square (OLS) or Generalized Least Square (GLS) regression techniques [19–27].

Most of the above RFFA methods assume linear relationship between flood statistics and predictor variables in log domain while developing the regional prediction equations. However, most of the hydrologic processes are nonlinear and exhibit a high degree of spatial and temporal variability and a simple log transformation may not guarantee achievement of linearity in modeling. Therefore, there have been applications of artificial intelligence based methods such as artificial neural networks (ANN), genetic algorithm based ANN (GAANN), gene expression programming (GEP) and co-active neuro-fuzzy inference system (CANFIS) in water resources engineering such as rainfall runoff modeling and hydrologic forecasting, but there have been relatively few studies till to-date involving the application of these techniques to RFFA [28–33]. Application of these techniques may help developing new improved RFFA techniques for Australia, which experiences a highly variable rainfall and hydrologic conditions. Unlike regression based approach, the artificial intelligence based techniques do not impose any fixed model structure on the data rather the data itself identifies the model form through use of artificial intelligence.

3 Methodology

Since the first neural model by McCulloch and Pitts [34], there have been developments of hundreds of different models which are considered ANN. The differences in them might be the functions, the accepted values, the topology, the learning algorithms, and the like. Since the function of ANN is to process information, they are used mainly in fields related to information processing. There are a wide variety of ANN that are used to model real neural networks, and study behaviour and control in animals and machines, but also there are ANN which are used for engineering purposes such as pattern recognition, flood forecasting, and data compression.

In this study, the adopted ANN modelling, Lavenberg-Marquardt method was used as the training algorithm to minimize the mean squared error (MSE) between the observed and predicted flood quantiles. The purpose of training an ANN with a set of input and output data is to adjust the weights in the ANN to minimize the MSE between the desired flood quantile and the ANN predicted flood quantile. Three hidden-layered neural networks were selected with 7, 3 and 1 neurons to each of these three layers. Two inputs, catchment area (A) and rainfall intensity with duration equal to time of concentration (t_c) and a given average recurrence interval (ARI) were used in one input layer and one output layer with one output called predicted flood quantile (Q_{pred}). The transfer function used for the hidden layers and the output layer was all hyperbolic tangent sigmoid function. Transfer functions calculate a layer's output from its net input. A maximum training iteration of 20,000 was adopted. All dependent and independent variables were standardized to the range of (0.05, 0.95), so that extreme flood events, which exceeded the range of the training data set could be modelled between the boundaries (0, 1) during testing.

A learning rate of 0.05 was used together with a momentum constant of 0.95. MATLAB was used to perform the ANN training. To select the best performing model the different combinations of hidden layers, algorithm, and number of neurons were observed against the MSE value. In order to obtain the best ANN-based model, the MSE values between the observed and predicted flood quantiles were calculated and the training was undertaken to minimize this error.

The major difference between GA and the classical optimization search techniques is that the GA works with a population of possible solutions; whereas, the classical optimization techniques work with a single solution [35]. GA is based on the Darwinian-type survival of the fittest strategy, whereby potential solutions to a problem compete and mate with each other in order to produce increasingly stronger individuals. Each individual in the population represents a potential solution to the problem that is to be solved and is referred to as a chromosome [36]. An initial population of individuals (also called chromosomes) is created and according to an objective function in focus the fitness values of all chromosomes is evaluated. From this initial population parents are selected who mate together to produce off springs (also called children). The genes of parents and children are mutated. The fittest among parents and children are sent to a new pool. The whole procedure is carried over until any of the two stopping criteria is met i.e. the required number of generations has been reached or convergence has been achieved. Chromosomes are the basic unit of population and represent the possible solution vector; they are assembled from a set of genes that are generally binary digits, integers or real numbers [37, 38].

An initial population is crowded with “ n ” number of chromosomes where “ n ” is referred to as the population size. An objective function comprising of feed forward ANN model with complete description of its architecture is defined. It reads training patterns once at the start of model and stores them in memory for applying to each chromosome. The total number of genes l of each chromosome represents the total synaptic weights of ANN model.

$$\{g_1, g_2, \dots, g_l\} = \{w_{(if \rightarrow hr)}, w_{(ib \rightarrow hr)}, w_{(hf \rightarrow or)}, w_{(hb \rightarrow or)}\} \quad (1)$$

where ‘ w ’ represents the value of a synaptic weight, subscript ‘ i ’ represents a node of input layer, ‘ h ’ is a node of hidden layer and ‘ o ’ represents the output layer node, ‘ f ’ is serial number of node which forwards the information (i.e. $f = 1, 2, 3, \dots$), ‘ r ’ is serial number of node which receives information (i.e. $r = 1, 2, 3, \dots$), ‘ ib ’ represent the bias node of input layer and ‘ hb ’ is bias node of hidden layer.

At the start of model, the fitness values of all the chromosomes of population are evaluated by ANN function. The real values stored in the genes of chromosome are read as the respective weights of ANN model. The ANN performs feed forward calculations with the weights read from genes of forwarded chromosome, and calculates MSE. The inverse of MSE is regarded as the fitness value of chromosome. By this way, the fitness values of all chromosomes of initial population are calculated by ANN function.

The selection operator selects two parent chromosomes randomly. The roulette wheel operator with elitism is used in this model. Elitism is a scheme in which the best chromosome of each generation is carried over to the next generation in order to ensure that the best chromosome does not lost during the calculations. The selected parents are mated to produce two children having the same number of genes. The uniform crossover operator is used with a crossover rate of $p_c = 1.0$. In uniform crossover, a toss is done at each gene position of an offspring and depending upon the result of toss, the gene value of 1st parent or 2nd parent is copied to the offspring. The genes of children are then mutated with the swap mutation operator with a mutation rate of $p_m = 0.8$. The mutated children are then evaluated by ANN function to know their fitness values. The fitness values of all the four chromosomes (2 parents & 2 children) are compared and the two chromosomes of highest fitness values are then sent to a new population and the other two are abolished. The evolutionary operators continue this loop of selection, crossover, mutation and replacement until the population size of new pool is same as old pool. One generation cycle completes at this stage and process is repeated until any of two stopping criteria is fulfilled i.e. maximum number of generations are reached or the convergence has been achieved. And the best chromosome which is tracked so far through the number of generations is sent to the ANN function. The genes of best chromosome are read as weights of ANN model and represent the optimised weights of ANN model. With these weights, the model is said to be fully trained. Finally, the train and test sets are simulated by using these weights.

The GAANN is coded in C language and some sub routines of LibGA package [39] for evolutionary operators of GA has been used with alterations to read and process the negative real values.

Gene-expression Programming (GEP) is used to perform a non-parametric symbolic regression. Symbolic regression although is very similar to traditional parametric regression, does not start with a known function relating dependent and independent variables as the latter. GEP programs are encoded as linear strings of fixed length (the genome or chromosomes), which are afterwards expressed as nonlinear entities of different sizes and shapes [40–42].

GEP automatically generates algorithms and expressions for the solution of problems, which are coded as a tree structure with its leaves (terminals) and nodes (functions). The generated candidates (programs) are evaluated against a “fitness function” and the candidates with higher performance are then modified and re-evaluated. This modification evaluation cycle is repeated until an optimum solution is achieved. In GEP a population of individual combined model solutions is created initially in which each individual solution is described by genes (sub-models) which are linked together using a predefined mathematical operation (e.g. addition). In order to create the next generation of model solutions, individual solutions from the current generation are selected according to fitness which is based on the pre-chosen objective function. These selected individual solutions are allowed to evolve using evolutionary dynamics to create the individual solutions of the next generation. This process of creating new generations is repeated until a certain stopping criterion is met [43].

Two important components of the GEP include the chromosomes and the expression trees (ETs). The ETs are the expression of the genetic information encoded in the chromosomes. The process of information decoding from chromosomes to the ETs is called translation, which is based on a kind of code and a set of rules. There exist very simple one to one relationships between the symbols of the chromosome and the functions or terminals they represent in the genetic code. To predict the flood quantiles the set of independent variables (predictor variables) to be used in the individual prediction equation are to be identified. Then a set of functions (e.g. e^x , x^a , $\sin(x)$, $\cos(x)$, $\ln(x)$, $\log(x)$, 10^x , etc.) and arithmetic operations (+, -, /, *) are defined. The terminals and the functions form the junctions in the tree of a program.

In GEP, k-expressions (from Karva notation) which are fixed length list of symbols are used to represent an ET. These symbols are called chromosomes, and the list is a gene. The Gene “*sqrt*, \times , \pm , a , b , c and d ” can be represented as ET. The GEP gene contains head and a tail. The symbols that represent both functions and terminals are present in the head while tail only contains terminals. The length of the head of the gene h is selected for each problem while the length of the tail is a function of length of the head of the gene.

In order to obtain the best GEP model, the mean squared error was used as ‘fitness function’, which was based on the observed and predicted flood quantiles; the training was undertaken to minimize this error.

For the CANFIS model development, model catchments were clustered based on model variables (A , I_{IC_ARI}) into several class values in layer 1 to build up fuzzy rules, and each fuzzy rule was constructed through several parameters of membership function in layer 2. A fuzzy inference system structure was generated from the data using subtractive clustering. This was used in order to establish the rule base relationship between the inputs.

In order to obtain the best CANFIS models, the MSE was used as the ‘fitness function’, which was based on the observed and predicted flood quantiles; the training was undertaken to minimise this error. Lavenberg-Marquardt (LM) method was used as the training algorithm to minimize the MSE. CANFIS model was trained with a set of input and output data to adjust the weights and to minimize the MSE between the desired outputs and the model outputs. The testing data set was selected randomly to produce a reasonable sample of different catchment types and sizes. Two inputs (A , I_{IC_ARI}) were used in one input layer and one output layer with one output (Q_{pred}).

In the case of CANFIS, the bell membership function and the TSK neuro fuzzy model were used, as this type of fuzzy model best fits the multi-input, single output system [44]. LM algorithm was used for the training of CANFIS model. The stopping criteria for the training of the CANFIS network was set to be a maximum of 1000 epochs and training was set to terminate when the MSE drops to 0.01 threshold value.

The following statistical measures were used to compare various RFFA models [45, 46]:

- Ratio between predicted and observed flood quantiles:

$$\text{Ratio of predicted and observed flood quantile} = \frac{Q_{pred}}{Q_{obs}} \quad (2)$$

- Relative error (RE):

$$\text{RE (\%)} = \text{Abs} \left[\frac{(Q_{pred} - Q_{obs})}{Q_{obs}} \times 100 \right] \quad (3)$$

- Coefficient of efficiency (CE):

$$\text{CE} = 1 - \frac{\sum_{i=1}^n (Q_{obs} - Q_{pred})^2}{\sum_{i=1}^n (Q_{pred} - \bar{Q})^2} \quad (4)$$

where Q_{pred} is the flood quantile estimate from the ANNs-based or GEP based RFFA model, Q_{obs} is the at-site flood frequency estimate obtained from LP3 distribution using a Bayesian parameter fitting procedure [47, 48] and \bar{Q} is the mean of Q_{obs} . The median relative error and median ratio values were used to measure the relative accuracy of a model. A Q_{pred}/Q_{obs} ratio closer to 1 indicates a perfect match between the observed and predicted value and a smaller median relative error is desirable for a model. A CE value closer to 1 is the best; however a value greater than 0.5 is acceptable.

4 Data Selection

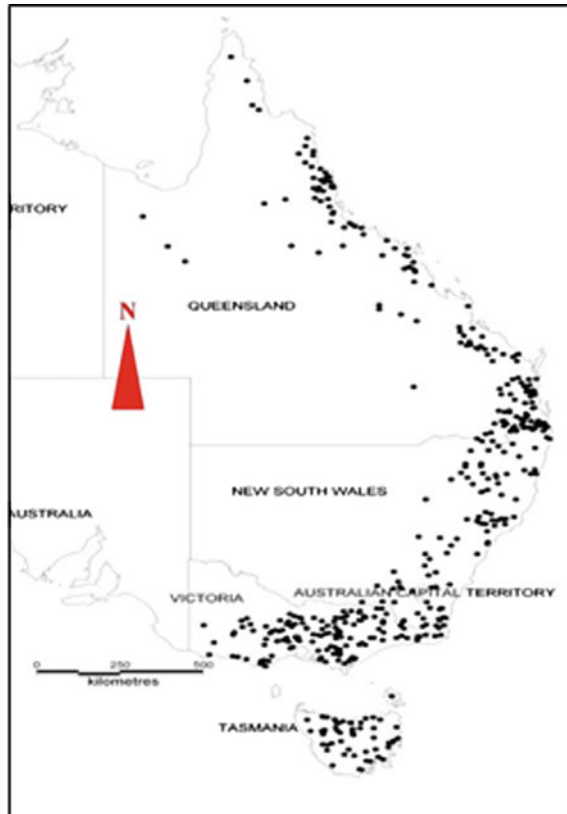
Eastern Australia was selected as the study area, which includes states of New South Wales (NSW), Victoria (VIC), Queensland (QLD) and Tasmania (TAS). This part of Australia was selected since this has the highest density of streamflow gauging stations with good quality data as compared to other parts of Australia. For RFFA study, streamflow data (annual maximum series) and climatic and catchment characteristics data are needed. A total of 452 catchments, which are rural and are not affected my major regulation were selected for this study. The locations of the selected catchments are shown in Fig. 1. A total of 96, 131, 172 and 53 stations from NSW, VIC, QLD and TAS, respectively were selected.

The catchment sizes of the selected 452 stations range from 1.3 to 1900 km² with the median value of 256 km². For the stations of NSW, VIC and QLD, the upper limit of catchment size was 1000 km², however for Tasmania; there were 4 catchments in the range of 1000 to 1900 km². Overall, there are about 12 % catchments in the range of 1 to 50 km², about 11 % in the range of 50 to 100 km², 53 % in the range of 100 to 500 km² and 24 % greater than 1000 km².

The annual maximum flood record lengths of the selected stations range from 25 to 75 years (mean: 33 years). The Grubbs and Beck [47] method was adopted in detecting high and low outliers (at the 10 % level of significance) in the annual maximum flood series data. The detected low outliers were treated as censored flows in flood frequency analysis. Only a few stations had a high outlier, which was not removed from the data set as no data error was detected for these high flows.

In estimating the flood quantiles for each of the selected stations, log-Pearson III (LP3) distribution was fitted to the annual maximum flood series using Bayesian method as implemented in FLIKE software [48]. According to Australian Rainfall

Fig. 1 Locations of the selected catchments



and Runoff (ARR), LP3 is the recommended distribution for at-site flood frequency analysis in Australia [4], and hence it was adopted. In previous applications [27, 49, 50], it has been found that LP3 distribution provide an adequate fit to the Australian annual maximum flood data.

In developing the prediction equations for flood quantiles, initially a total of five explanatory variables were adopted as part of study [45] using the same 452 catchments. These variables are: (i) catchment area expressed in km² (*A*); (ii) design rainfall intensity values in mm/h $I_{t_c, ARI}$ (where ARI = 2, 5, 10, 20, 50 and 100 years and t_c = time of concentration (hour), estimated from $t_c = 0.76A^{0.38}$) (iii) mean annual rainfall expressed in mm/y (*R*); (iv) mean annual areal evapo-transpiration expressed in mm/y (*E*); (v) main stream slope expressed in m/km (*S*). In both of these studies where the ANN and Gene Expression programming techniques were used to develop the prediction equation, this was found that two variables (*A* and $I_{t_c, ARI}$) model outperformed other models. Based on this finding, two predictor variables i.e., *A* and $I_{t_c, ARI}$ are selected for this study.

5 Results

At the beginning each of the four artificial intelligence based RFFA models is 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.

Table 1 shows the median of the absolute relative error values for the ANN, GAANN, GEP and CANFIS based RFFA models. It can be seen that ANN based RFFA model outperforms the other models with a median RE value of 42.1 % over all the six ARIs. In some cases, the GAANN based RFFA model performs better or equal to the ANN based model i.e. for ARIs of 2, 5, 20 and 100 years; however, for 50 years ARI it shows a very high RE (60 %). In terms of consistency over the ARIs, ANN outperforms the other three models. Both GEP and CANFIS have quite high 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 %).

Table 1 Median RE (%) values of four artificial intelligence based RFFA models (training)

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

Table 2 Comparison of training and validation results for the ANN based RFFA model

ARI (years)	Training			Validation		
	CE	Q_{pred}/Q_{obs} (median)	RE (%) (median)	CE	Q_{pred}/Q_{obs} (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

Overall, in terms of RE value, the ANN is the best performer, followed by the GAANN, GEP and CANFIS.

The CE, median Q_{pred}/Q_{obs} ratio and median relative error values are compared in Table 2 for the training and validation datasets for the ANN based RFFA model. In terms of CE value, the best agreement between the training and validation data sets is found for ARIs of 10, 20 and 50 years, a reasonable degree of agreement is found for ARIs of 2 and 5 years and relatively poor agreement is found for the ARI of 100 years where the CE value for the validation data set is remarkably small. With respect to median Q_{pred}/Q_{obs} ratio value, the best agreement between the training and validation data sets is found for 2 years ARI, a moderate agreement is noticed for 10, 20, 50 and 100 years ARIs and a poor agreement is found for 5 years ARI. However, for 5 years ARI validation data set gives a very good Q_{pred}/Q_{obs} ratio value (0.99). In relation to the median relative error values, the best agreement between the training and validation data sets is found for ARIs of 5 and 100 years, a moderate agreement for ARI of 50 years and poor agreement for ARIs of 2 and 10 years. From these results, it is noted that the ANN based RFFA model shows different degrees of agreement between the training and validation data sets for different ARIs across the three criteria adopted here.

From the results of the training of the four artificial intelligence based RFFA models, it has been found that none of the four models perform the best in all the adopted assessment criteria over the six ARIs. Based on the four different criteria as shown in Table 3, the performances of the four models are assessed in a heuristic

Table 3 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	#
<i>Overall score ANN-15, GAANN-12, GEP-10, CANFIS-7</i>				

manner. In this assessment, a model is ranked based on four different criteria as shown in Table 3. Four different ranks are used, with a relative score ranging from 4 to 1. If a model is ranked 1 for a criterion, it scores 4. For ranks of 2, 3 and 4, scores of 3, 2 and 1, respectively are assigned.

Table 3 shows that the ANN based RFFA model has the highest score of 15, followed by the GANN with a score of 12. The GEP receives a score of 10, while the CANFIS receives only 7 making it the least favourable model in terms of its performance during training. The ANN based model is placed at rank 1 in the 3 out of 4 criteria. Hence, it is decided that the ANN based RFFA model is the best performing artificial intelligence based model in terms of training/calibration of the model.

Table 4 shows the ranking of the four artificial intelligence based RFFA models based on the agreement between the training and validation using three criteria. Four different ranks are used with a relative score ranging from 4 to 1 as mentioned

Table 4 Ranking of the four artificial intelligence based RFFA models with respect to agreement between training and validation

Criterion	Rank 1	Rank 2	Rank 3	Rank 4
Median Q_{pred} Q_{obs}	GEP (Best agreement: Q_2, Q_5, Q_{10}, Q_{20} Moderate agreement: Q_{50}, Q_{100} Poor agreement: none)	ANN (Best agreement: Q_2, Q_{10}, Q_{100} Moderate agreement: Q_{20}, Q_{50} Poor agreement: Q_5)	CANFIS (Best agreement: $Q_5, Q_{10}, Q_{20}, Q_{50}, Q_{100}$ Moderate agreement: none Very poor agreement: Q_2)	GAANN (Best agreement: $Q_2, Q_{10}, Q_{20}, Q_{100}$ Moderate agreement: Q_5 Very poor agreement: Q_{50})
Median RE (%)	GEP (Best agreement: $Q_2, Q_5, Q_{10}, Q_{20}, Q_{100}$ Moderate agreement: Q_{50} Poor agreement: none)	ANN (Best agreement: Q_5, Q_{100} Moderate agreement: Q_{50}, Q_{20} Poor agreement: Q_2, Q_{10})	GAANN (Best agreement: Q_{50}, Q_{100} Moderate agreement: Q_{20} Very poor agreement: Q_2, Q_5, Q_{10})	CANFIS (Best agreement: $Q_5, Q_{10}, Q_{20}, Q_{50}, Q_{100}$ Moderate agreement: none Significantly poor agreement: Q_2)
Median CE	GAANN (Best agreement: $Q_2, Q_5, Q_{20}, Q_{100}$ Moderate agreement: Q_{10} Poor agreement: Q_{50})	ANN (Best agreement: Q_{10}, Q_{20}, Q_{50} Moderate agreement: Q_2, Q_5 Poor agreement: Q_{100})	CANFIS (Best agreement: $Q_{10}, Q_{20}, Q_{50}, Q_{100}$ Moderate agreement: Q_5 Poor agreement: Q_2)	GEP (Best agreement: Q_5, Q_{20}, Q_{50} Moderate agreement: Q_{10}, Q_{100} Poor agreement: Q_2)

Overall score ANN-9, GEP-9, GAANN-7, CANFIS-5

earlier. It is found that the ANN and GEP based RFFA models both score 9, followed by the GAAANN and CANFIS. Overall, ANN based RFFA model shows the best training/calibration and the CANFIS the least favourable one.

Figure 2 compares the predicted flood quantiles for the selected 90 test catchments from the ANN based RFFA model with the observed flood quantiles for 20 years ARI (Q_{20}). The observed flood quantiles are estimated using an LP3 distribution and Bayesian parameter estimation procedure [48]. It should be noted here that the observed flood quantiles are not free from error; these are subject to data error (such as rating curve extrapolation error), sampling error (due to limited record length of annual maximum flood series data), error due to choice of flood frequency distribution and error due to selection of parameter estimation method. This error undermines the usefulness of the validation statistics (e.g. RE); however, this provides an indication of possible error of the developed RFFA model as far as practical application of the RFFA model is concerned. The ratio $Q_{\text{pred}}/Q_{\text{obs}}$ and RE values are used for the assessment of models; however, the CE value is not very useful here as the mean of observed flood quantile is not known.

Figure 2 shows a good agreement overall between the predicted and observed flood quantiles; however, there is some over-estimations by the ANN based RFFA model when the observed flood quantiles are smaller than about 50 m^3/s . Most of the test catchments are within a narrow range of variability from the 45-degree line except for a few outliers. The plots of predicted and observed flood quantiles for other ARIs showed very similar for ARIs of 2, 5, 10 and 20 years. Results for ARIs of 50 and 100 years exhibited some overestimation by the ANN based RFFA model for smaller to medium discharges.

Figure 3 shows the boxplot of relative error (RE) values of the selected test catchments for ANN based RFFA model for different flood quantiles. It can be seen from Fig. 3 that the median RE values (represented by the thick black lines within the boxes) are located very close to the zero RE line (indicated by 0–0 horizontal

Fig. 2 Comparison of observed and predicted flood quantiles for ANN based RFFA model for Q_{20}

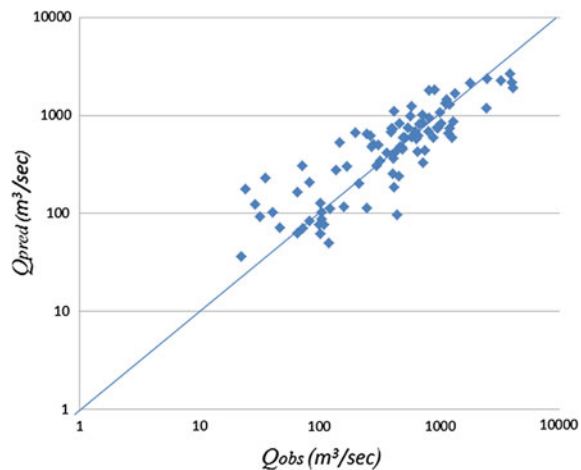
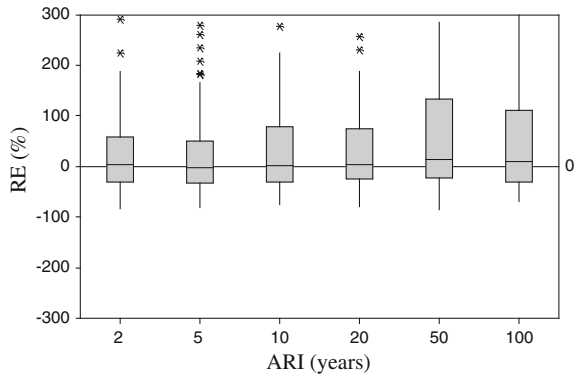


Fig. 3 Boxplot of relative error (RE) values for ANN based RFFA model



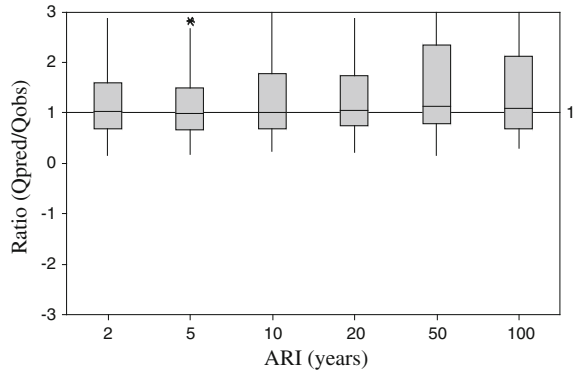
line in Fig. 3), in particular for ARIs of 2, 5, 10 and 20 years. However, for ARIs of 50 and 100 years, the median RE values are located above the zero line with ARI of 50 years showing the highest departure, which indicates an overestimation by the ANN based RFFA model. Overall, the ANN based RFFA model produces nearly unbiased estimates of flood quantiles as the median RE values match with the zero RE line quite closely as can be seen in Fig. 3.

In terms of the spread of the RE (represented by the width of the box), ARI of 50 and 100 years present the highest RE band and ARIs of 2 and 5 years present the smallest RE band, followed by ARI of 20 and 10 years. The RE bands for 50 and 100 years ARIs are almost double to RE bands of 2 and 5 years ARIs. This implies that ANN based RFFA model provides the most accurate flood quantile estimates for 2 and 5 years ARIs, and the least accurate flood quantiles for ARIs of 50 and 100 years. Overall, the boxplot in Fig. 3 shows that better results in terms of RE values are achieved for the smaller ARIs (i.e. 2, 5, 10 and 20 years ARIs) as compared to higher ARIs for the ANN based RFFA model. Some outliers (evidenced by notable overestimation with a positive RE) can be seen for all the ARIs, which may need to be examined more closely for data errors or issues regarding the hydrology and physical characteristics of these catchments; if these catchments are deemed to be genuine outliers they should be removed to enhance the ANN based RFFA model; however, this has not been undertaken in this chapter.

Figure 4 shows the boxplot of the Q_{obs}/Q_{pred} ratio values of the selected 90 test catchments for ANN based RFFA model for different ARIs. The median Q_{obs}/Q_{pred} ratio values (represented by the thick black lines within the boxes) are located closer to 1–1 line (the horizontal line in Fig. 4), in particular for ARIs of 2, 5, 10 and 20 years. However, for ARI of 50 years (and to a lesser degree for ARI of 100 years), the median Q_{obs}/Q_{pred} ratio value is clearly located above the 1–1 line.

These results indicate that the ANN based RFFA model generally provides reasonably accurate flood quantiles with the expected Q_{obs}/Q_{pred} ratio value very close to 1.00, although there is a noticeable overestimation for ARI of 50 and 100 years. In terms of the spread of the Q_{obs}/Q_{pred} ratio values, ARI of 2 and 5 years provide the lowest spread followed by ARIs of 20, 10, 100 and 50 years.

Fig. 4 Boxplot of Q_{pred}/Q_{obs} ratio values for ANN based RFFA model



Considering, the RE and Q_{obs}/Q_{pred} ratio values as discussed above, it can be concluded that ANN based RFFA model generally provide unbiased flood estimates for smaller to medium ARIs (2 to 20 years); however, the model slightly overestimates the observed flood quantiles for higher ARIs (50 to 100 years).

6 Conclusion

This chapter presents development and testing of non-linear artificial intelligence based regional flood frequency analysis (RFFA) models. For this purpose, a database of 452 small to medium sized catchments from eastern Australia has been used. Four different artificial intelligence based RFFA models have been considered in this research. It has been found that non-linear artificial intelligence based RFFA techniques can be applied successfully to eastern Australian catchments. Among the four artificial intelligence based models, the ANN based RFFA model has been found to be the best performing model, followed by the GAANN based RFFA model. It has been shown that in the training of the four artificial intelligence based RFFA models, no model performs the best for all the six ARIs over all the adopted criteria. Overall, the ANN based RFFA model is found to outperform the three other models in the training/calibration. Based on independent validation, the median relative error values for the ANN based RFFA model are found to be in the range of 35 to 44 % for eastern Australia.

References

1. D. Carbone, J. Hanson, Floods: 10 of the deadliest in Australian history. Australian Geographic, <http://www.australiangeographic.com.au/journal/the-worst-floods-in-australian-history.htm>. Accessed 18 July 2013
2. PWC, *Economic impact of Queensland's natural disasters*. (Price Waterhouse Cooper (PWC), Australia, 2011)

3. NERC, *Flood studies report*. (Natural Environment Research Centre (NERC), London, 1975)
4. Institution of Engineers Australia (I.E. Aust.), *Australian rainfall and runoff: A guide to flood estimation*, vol. 1, ed by D.H. Pilgrim (I. E. Aust., Canberra, 1987, 2001)
5. M.C. Acreman, C.D. Sinclair, Classification of drainage basins according to their physical characteristics and application for flood frequency analysis in Scotland. *J. Hydrol.* **84**(3), 365–380 (1986)
6. R.J. Nathan, T.A. McMahon, Identification of homogeneous regions for the purpose of regionalisation. *J. Hydrol.* **121**, 217–238 (1990)
7. D.H. Burn, Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resour. Res.* **26**(10), 2257–2265 (1990)
8. T.R. Kjeldsen, D. Jones, An exploratory analysis of error components in hydrological regression modelling. *Water Resour. Res.* **45**, W02407 (2009)
9. K. Haddad, A. Rahman, Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework—Quantile Regression vs Parameter Regression Technique. *J. Hydrol.* **430–431**(2012), 142–161 (2012)
10. T.J. Mulvany, On the use of self-registering rain and flood gauges. *Inst. Civ. Eng. (Ireland) Trans.* **4**(2), 1–8 (1851)
11. G.G.S. Pegram, M. Parak, A review of the regional maximum flood and rational formula using geomorphological information and observed floods, ISSN 0378-4738. *Water South Africa* **30**(3), 377–392 (2004)
12. A. Rahman, K. Haddad, M. Zaman, G. Kuczera, P.E. Weinmann, Design flood estimation in ungauged catchments: a comparison between the probabilistic rational method and quantile regression technique for NSW. *Aust. J. Water Resour.* **14**(2), 127–137 (2011)
13. T. Dalrymple, Flood frequency analyses. U.S. Geological Survey Water Supply Paper, 1543-A, 11–51, (1960)
14. J.R.M. Hosking, J.R. Wallis, Some statistics useful in regional frequency analysis. *Water Resour. Res.* **29**(2), 271–281 (1993)
15. B.C. Bates, A. Rahman, R.G. Mein, P.E. Weinmann, Climatic and physical factors that influence the homogeneity of regional floods in south-eastern Australia. *Water Resour. Res.* **34**(12), 3369–3382 (1998)
16. A. Rahman, B.C. Bates, R.G. Mein, P.E. Weinmann, Regional flood frequency analysis for ungauged basins in south-eastern Australia. *Aust. J. Water Res.* **3**(2), 199–207, 1324–1583 (1999)
17. T.R. Kjeldsen, D.A. Jones, Predicting the index flood in ungauged UK catchments: on the link between data-transfer and spatial model error structure. *J. Hydrol.* **387**(1–2), 1–9 (2010)
18. E. Ishak, K. Haddad, M. Zaman, A. Rahman, Scaling property of regional floods in New South Wales Australia. *Nat. Hazards* **58**, 1155–1167 (2011)
19. M.A. Benson, Evolution of methods for evaluating the occurrence of floods, U.S. Geological Surveying Water Supply Paper, 30, 1580-A (1962)
20. D.M. Thomas, M.A. Benson, Generalization of streamflow characteristics from drainage-basin characteristics, U.S. Geological Survey Water Supply Paper 1975, US Governmental Printing Office, 1970
21. J.R. Stedinger, G.D. Tasker, Regional hydrologic analysis - 1. Ordinary, weighted and generalized least squares compared. *Water Resour. Res.* **21**, 1421–1432 (1985)
22. G.R. Pandey, V.T.V. Nguyen, A comparative study of regression based methods in regional flood frequency analysis. *J. Hydrol.* **225**, 92–101 (1999)
23. A. Rahman, A quantile regression technique to estimate design floods for ungauged catchments in South-east Australia. *Aust. J. Water Resour.* **9**(1), 81–89 (2005)
24. V.W. Griffis, J.R. Stedinger, The use of GLS regression in regional hydrologic analyses. *J. Hydrol.* **344**, 82–95 (2007)
25. T.B.M.J. Ouarda, K.M. Bâ, C. Diaz-Delgado, C. Cârsteanu, K. Chokmani, H. Gingras, E. Quentin, E. Trujillo, B. Bobée, Intercomparison of regional flood frequency estimation methods at ungauged sites for a Mexican case study. *J. Hydrol.* **348**, 40–58 (2008)

26. K. Haddad, A. Rahman, Regional flood estimation in New South Wales Australia using generalised least squares quantile regression. *J Hydrol Eng ASCE* **16**(11), 920–925 (2011)
27. K. Haddad, A. Rahman, J.R. Stedinger, Regional flood frequency analysis using bayesian generalized least squares: a comparison between quantile and parameter regression techniques. *Hydrol. Process.* **25**, 1–14 (2011)
28. T.M. Daniell, Neural networks—applications in hydrology and water resources engineering. Paper presented at the international hydrology and water resources symposium. Perth, Australia, 2–4 October 1991
29. R.S. Muttiah, R. Srinivasan, P.M. Allen, Prediction of two year peak stream discharges using neural networks. *J. Am. Water Resour. Assoc.* **33**(3), 625–630 (1997)
30. C. Shu, D.H. Burn, Artificial neural network ensembles and their application in pooled flood frequency analysis. *Water Resour. Res.* **40**(9), W09301 (2004). doi:[10.1029/2003WR002816](https://doi.org/10.1029/2003WR002816)
31. U.C. Kothyari, Estimation of mean annual flood from ungauged catchments using artificial neural networks, in *Hydrology: Science and Practice for the 21st Century*, vol. 1 (British Hydrological Society, 2004)
32. C.W. Dawson, R.J. Abrahart, A.Y. Shamseldin, R.L. Wilby, Flood estimation at ungauged sites using artificial neural networks. *J. Hydrol.* **319**, 391–409 (2006)
33. C. Shu, T.B.M.J. Ouarda, Flood frequency analysis at ungauged sites using artificial neural networks in canonical correlation analysis physiographic space. *Water Resour. Res.* **43**, W07438 (2007). doi:[10.1029/2006WR005142](https://doi.org/10.1029/2006WR005142)
34. W.S. McCulloch, W. Pitts, A logic calculus of the ideas immanent in nervous activity. *Bull Math Biophys* **5**, 115–133 (1943)
35. A. Jain, S. Srinivasalu, R.K. Bhattacharjya, Determination of an optimal unit pulse response function using real-coded genetic algorithm. *J. Hydrol.* **303**, 199–214 (2005)
36. A.J.F.V. Rooij, L.C. Jain, R.P. Johnson, *Neural network training using genetic algorithms* (World Scientific Publishing Co. Pty. Ltd., River Edge, 1996), p. 130
37. M. Mitchell, *An Introduction to Genetic Algorithms* (MIT Press, Cambridge, 1996)
38. K. Aziz, A. Rahman, A.Y. Shamseldin, M. Shoaib, Co-Active neuro fuzzy inference system for regional flood estimation in Australia. *J Hydrol Environ Res* **1**(1), 11–20 (2013)
39. L.C. Arthur, L.W. Roger, in *LibGA for solving combinatorial optimization problems*, ed. by Chambers. Practical Handbook of Genetic Algorithms (CRC Press, Boca Raton, 1995)
40. C. Ferreira, Gene expression programming in problem solving, in *proceedings of the 6th Online World Conference on Soft Computing in Industrial Applications* (invited tutorial), 2001
41. C. Ferreira, Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst.* **13**(2), 87–129 (2001)
42. C. Ferreira, *Gene-expression programming; mathematical modeling by an artificial intelligence* (Springer, Berlin, 2006)
43. D.A.K. Fernando, A.Y. Shamseldin, R.J. Abrahart, Using gene expression programming to develop a combined runoff estimate model from conventional rainfall-runoff model outputs. Paper presented at the 18th world IMACS / MODSIM Congress, Cairns, Australia 13–17 July 2009
44. A. Aytel, Co-Active neuro-fuzzy inference system for evapotranspiration modelling. *Soft. Comput.* **13**(7), 691–700 (2009)
45. K. Aziz, A. Rahman, G. Fang, S. Shreshtha, Application of artificial neural networks in regional flood frequency analysis: a case study for Australia. *Stochast Environ. Res. Risk Assess.* **28**(3), 541–554 (2013)
46. K. Aziz, A. Rahman, G. Fang, K. Haddad, S. Shreshtha Design flood estimation for ungauged catchments: application of artificial neural networks for eastern Australia. In: *World Environment and Water Resources Congress*, ASCE, Providence, Rhodes Island, USA, 2010
47. F.E. Grubbs, G. Beck, Extension of sample sizes and percentage points for significance tests of outlying observations. *Technometrics* **14**, 847–854 (1972)
48. G. Kuczera, Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *Water Resour. Res.* **35**(5), 1551–1557 (1999)

49. K. Haddad, A. Rahman, P.E. Weinmann, G. Kuczera, J.E. Ball, Streamflow data preparation for regional flood frequency analysis: lessons from south-east Australia. *Aust. J. Water Resour.* **14**(1), 17–32 (2010)
50. K. Haddad, A. Rahman, F. Ling, Regional flood frequency analysis method for Tasmania, Australia: A case study on the comparison of fixed region and region-of-influence approaches. *Hydrol. Sci. J.* (2014) doi: [10.1080/02626667.2014.950583](https://doi.org/10.1080/02626667.2014.950583)