Application of Neural Network to flood forecasting, an examination of model sensitivity to rainfall assumptions

Seyyed Adel Alavi Fazel, PhD Candidate
School of Information and Communication Technology, Griffith University, Gold Coast Campus, Australia, seyyedadel.alavifazel@griffithuni.edu.au;

Hamid Mirfenderesk, Associated Professor
Team leader, Waterway and Flood Management, The Gold Coast City Council, Gold Coast, Australia, hmirfenderesk@goldcoast.qld.gov.au;

Michael Blumenstein, Professor
School of Information and Communication Technology, Griffith University, Gold Coast Campus, Australia, m.blumenstein@griffith.edu.au

Rodger Tomlinson, Professor
School of environment, Griffith University Gold Coast Campus, Australia, r.tomlinson@griffith.edu.au

Abstract: This paper describes the development of a back-propagation Neural Network model for predicting flood and its application to a short response catchment on the Gold Coast, Australia. The main body of knowledge on flood forecasting is based on traditional physically based and conceptual methods. These methods, despite being based on robust physical laws, have limitations. Data-driven models are plausible alternatives to physically based methods for certain flood forecasting applications. However, there is still a need for further demonstration of their ability in flood forecasting in order to build enough confidence for their application in practice.

In this study, existing uncertainties during real-world flood event is emulated. Four scenarios are considered for this study, rainfall is assumed known, rainfall is naively predicted, rainfall is treated as hidden variable and rainfall is predicted using axillary ANN. The study shows that the proposed ANN is adequately skilled for short-term flood predictions; however variability in rainfall within span of few hours limits reliability of predictions as time horizon increases. The study suggests that although variability in rainfall is a major component in incorrect predictions, commonly utilized optimisation goal is not fully aligned with purpose of modelling and produces unintentional bias. The study shows that proposed ANN models is sensitive to rainfall assumptions (beyond the time that the forecast is performed) and temporal/spatial resolution of the rainfall data. This paper speculates that by changing the optimization goal, the model can be improved.

Keywords: Rainfall-runoff modelling, neural networks, uncertainty, assumption.

1. INTRODUCTION

During a flood emergency, authorities rely on models that forecast water levels for a particular lead time. These models map observable variables such as rainfall to flood levels. Due to differences between catchments, models inevitably contain catchment-specific parameters. Decision support systems are designed based on the three modelling approaches (Beven, 2012):

Physically inspired models: these models are developed based on governing laws of physics; the parameters of these models can be deterministically evaluated by laboratory observations. These models are ideal for catchments where limited observations are available.

Conceptual models: these models are based on the relationship between important variables in a catchment; these models require calibration based on a subset of observation time series. These models are most commonly operational models due to simplicity in development and low computational costs.
Data-driven models: these models are based on the relationship between variables and require limited assumption about physics of the problem. These models have limited restrictions about nature of relationship between variables and allow inclusion of variables with ease. Regardless of the approach taken, it is essential for the flood forecasters to understand limits of utilized models and sources of uncertainty (Refsgaard, van der Sluijs, Højberg, & Vanrolleghem, 2007). A literature review shows that underlying reason for existence of uncertainty and error is commonly attributed to errors in input recording data, model structure, model parameters and lack of sufficient calibration data (Han, Kwong, & Li, 2007; Zhang, Liang, Yu, & Zong, 2011). Bootstrap and Bayesian methods have been used to address these types of uncertainties (Kohavi, 1995; Tiwari & Chatterjee, 2010). In addition to the abovementioned uncertainties, a challenge in flood forecasting is due to the uncertainty that is induced by assumptions regarding the unseen rainfall from prediction time to prediction horizon. This uncertainty grows as forecasting horizon increases. This paper is focused on investigating this uncertainty and its significance in flood forecasting using Artificial Neural Network (ANN). Commonly physically based and conceptual models assume that rainfall is a known quantity during training and benchmarking without taking into account operational limitation. The issue of implication of making assumptions regarding unseen rainfall is not well addressed. Notably, conceptual and physical based model do not provide ideal test bed for understanding this uncertainty (Refsgaard, Henriksen, Harrar, Scholten, & Kassahun, 2005). The contribution of this research is two folds, first, development of a data-driven rainfall-runoff model for Tallebudgera Creek Road using ANN and second, assessment of impact of making assumptions regarding unseen portion of rainfall.

2. STUDY AREA

Study area is the Tallebudgera catchment in the Gold Coast, Australia. This station measures both water level and rainfall. Two more rainfall station upstream of this station is included in this study. The catchment of the Tallebudgera Road station has an area of approximately 64 square kilometres, is very steep and not affected by tide. The following data set is used in this study:

- Water level measurements at the Tallebudgera Creek road gauging station at 15 minute intervals.
- Measured rainfall at the three gauging stations, i.e. Tallebudgera Road, Tallebudgera Dam and Springbrook. In this study water level less than 1 metre Australian Height Datum (AHD) is regarded as a non-flood event and filtered out.

Available dataset is partitioned into training and testing portions. To create training examples, dataset is divided into ‘flood’ events defined when variation in water levels is more than one meters. Each event consists of observations from the time that rainfall occurs up until water levels stabilize. Events are ranked based on highest water levels reached during the event and the highest water level and every third ranking event were reserved for testing and the remainder were utilized for model calibration. The aerial map of the studied area is presented at figure 2.

3. METHODOLOGY

ANN is one of the most commonly utilized data-driven models for hydrological modelling (Maier, Jain, Dandy, & Sudheer, 2010). Neural networks are biologically inspired data-driven models that allow identification of a mapping between input variables and output variables. Architecture of neural networks is demonstrated at Figure 1. The building block of ANN is information processing units referred to as neurons. Each neuron receives information from inner layer and performs a weighted summation over all information received. In most common applications the neuro-structure do not contain feedback loop and information flow from inner to outer layers only. The output of each neuron is determined by value of a (nonlinear) function of this summation. This procedure continues until the information from sensory input travel to output layer. By giving training examples to the network, synaptic weights that of a network adjust in a way to best mimic the relationship between input and output and minimize a user defined penalty function. Details of implementation of neural networks can be found in work of Haykin (Haykin, 2009). The training process of neuro-structure consists of a method for adjusting the synaptic weights. The most commonly used method is referred to “back-propagation” (BP). BP method consists of the following steps:

1) Initialization of weights of the neuro-structure
2) Steepest descent optimization
3) Performance evaluation

Nueo-structure commonly contain step 2-4 and the procedure is continued until a user-defined stopping criterion is achieved. As demonstrated in steps for ANN model development, limited assumptions are made in regards to physics of the problem. It is known that given sufficient training examples, ANN is able to identify the mapping between input and output space. Assuming that ‘sufficient’ training examples are available, these structures allow modeller to identify arbitrarily complex mappings between variables. Levenberg-Marquardt training algorithm has been found as the most consistent method for training ANN in hydrological modelling and adopted herein (Piotrowski & Napiorkowski, 2011).

![Figure 1. An example of neural networks](image1)

![Figure 2. Arial photo of the study area, Tallebudgera Creek Road, Australia](image2)

Four rainfall prediction methods, as shown in Table I, were tested to examine the effect of various rainfall prediction methods on the performance of the ANN flood forecasts. Two types of inputs were selected in the regression model namely, antecedent water levels and rainfall.

<table>
<thead>
<tr>
<th>Scenario Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No assumption for future rainfall. In this approach the network is trained with rainfall up until the time that forecast is carried out and unseen rainfall is treated as hidden variable</td>
</tr>
<tr>
<td>B</td>
<td>Rainfall from prediction time to the prediction horizon is known</td>
</tr>
<tr>
<td>C</td>
<td>Naïve prediction of rainfall for unseen rainfall from prediction time to prediction horizon. In this scenario the rainfall is assumed to continue at a constant rate (i.e. last observation) in the period between the prediction time and prediction horizon.</td>
</tr>
<tr>
<td>D</td>
<td>The rainfall between the prediction time and prediction horizon is estimated by applying ANN to the available rainfall data until prediction time.</td>
</tr>
</tbody>
</table>

Scenario “B” cannot happen in practice, as in a real situation we do not have access to measured rainfall at the time that prediction is carried out beyond the prediction time. This scenario is tested to provide a performance benchmark for the ANN model and understanding limitation of meaningful forecasting horizon so other rainfall scenarios can be assessed in terms of their ability to provide accurate flood forecasts.

For each of the scenarios, two sub-scenarios are considered, 1) considering the spatial distribution of rainfall in upstream catchments 2) giving rainfall as an average value over the entire upstream catchment

Scenario "A" is developed by the following formulation:

\[ w(t + h) = f(w(t), w(t - 1), ..., r(t), r(t - 1), ...) \]  

(1)
where "w" refers to water levels, "r" to rainfall series, "t" is the current time index, "h" is the time steps in future, and "f" is the mapping function between observations and future water levels. In this formulation, observed variable are rainfall and antecedent water levels. Using trial and error, it was found that 4 antecedent water lags is optimum.

The remaining scenarios were developed, using the following formulation:

\[ w(t + h) = f(w(t), w(t - 1),..., r(t + h), r(t + h - 1),...) \]  

(2)

To realize scenario “B” the rainfall component of \( r(t) \) to \( r(t+k) \) in equation (2) is substituted with observed rainfall. Scenario “C” is based on a naïve prediction of rainfall, more specifically, \( r(t) \) to \( r(t+k) \) in equation (2) is substituted with the latest available rainfall recording at \( r(t) \). In this approach, it is assumed that the rainfall will continue at the same rate and magnitude as occurred in the last time step until the forecast time horizon. Scenario “D” uses a rainfall input that is generated by rainfall prediction algorithm in simple univariate setting. In this algorithm input rainfall is selected from the past three time steps from the current time and the output is the unseen rainfall from prediction time to forecasting horizon.

For Development of ANN, number of layers, transfer function that maps layers, number of neurons in each layer, objective function and stopping criterion have to be chosen by modeller. Despite extensive successful application of ANN, no concrete rules for each of these steps can be found in literature. Most commonly, no more than one hidden layer is employed in application of ANN (Napolitano, Serinaldi, & See, 2011); the most common transfer function between input and hidden layer is tangent hyperbolic and between hidden to output layer is linear; the objective function is Root Mean Square error and number of neurons commonly do not have significant effect (it is set to 10 for this study). The most stable and commonly utilized optimization method for optimization of weight space in “small” neural networks is Levengberg-Marquardt (LM) method (Napolitano et al., 2011; Piotrowski & Napierkowski, 2011), hence adopted herein. LM method is an iterative learning procedure that requires a stopping criterion. For each epochs or iteration of training, the RMSE over training set defines the gradient of weight adjustment until a stopping criterion is achieved. The stopping criterion adopted herein is suggestion by Moradkhani et al (Moradkhani, Hsu, Gupta, & Sorooshian, 2004). According to the proposed stopping criterion, available training examples are randomized and divided into “k” portions, ‘k-1’ are given for training and subsequent to 10 successive failures in improvement of performance of the model when deployed for ‘kth’ portion, training is stopped. To maximize utility of available training examples, ‘k’ such models are developed each trained over ‘k-1’ part of available training examples and final output is obtained by averaging over outputs produced by all models.

4. RESULTS AND DISCUSSIONS

Back propagation is a iterative method, hence one realization of a particular optimization result of ANN does not confirm superiority of a particular input sets. Hence ANNs with 50 different initial conditions for each of the scenario are trained and the performance of each is tested in the test set. To ensure reproducibility of the results, root mean square error of each of the experimental scenarios for different forecasting horizons is demonstrated in boxplots of Figure 3 to 5. The boxplot shows the variation around the average root mean square of error term. Scenarios that have considered an average rainfall over the catchment area identified by indices 1 and scenarios which contain spatial distribution of rainfall are shown with indices 2.

Figure 3. Boxplot of the experiments for each of the investigated scenarios for one hours forecasts
It can be seen that scenario “B” produces the best results, as it expected. This scenario uses the measured rainfall in the period between the prediction time and prediction horizon and is used only as a bench mark to assess the accuracy of realistic scenarios “A”, “C” and “D”. The qualitative results, demonstrated at Figures 6 to 8, show that the performance of Scenario B does not deteriorate similar to the remaining scenarios for all prediction horizons. This indicates that utilized ANN model is structurally sound and given information. Scenarios A, C and D do not perform as well as scenario B, as these scenarios are not informed with the measured rainfall between the prediction time and prediction horizon. This suggests that longer forecasting horizons rely on the presence of accurate rainfall forecasts for the period between the time of prediction and the horizon time. Any attempt to reduce the forecasting error inevitably needs to address this issue first.

Figure 4 shows that for one hour forecast, all three methods perform similarly. It also shows that consideration of spatial distribution of rainfall over the catchment (shown by indices 1) increases the variance of the results substantially for all scenarios. This could be attributed to the increase in the number of model parameters in this alternative. Naïve prediction of rainfall method is the worst performer. Scenario “D” shows slightly better performance both in terms of accuracy and variance. Figures 6 to 8 show the performance of the same above-mentioned scenarios in time domain. The inadequacy of RMSE as a performance indicator can be seen through comparing figures 5 and 8. Figure 5 suggests that scenario D performs better than scenarios A and C, whereas if peak approximation is considered as the goal of modelling, Figure 8 suggests otherwise. The reason for this contraction is that RMSE does not distinguish between errors during unimportant low flow and important high flow times. From a flood forecasting perspective, only high flow situations are important due to imbalance of dataset [15].

Figures 6 to 8 show that scenario C (i.e. naïve rainfall forecast) produces better forecasts than scenarios D and A during peak flood and for long prediction horizons. An explanation for this observation is that RMSE penalizes under and over prediction equally and also do not take into account the distribution of error. Therefore, a model with slight errors spreading out over the whole time series scores the same RMSE as a model with errors concentrated around peak values. The explanation for ‘better’ quantitative results obtained in scenario D can be explained by taking into account the imbalance of dataset towards low rainfall events. For this reason, the univariate rainfall forecasting model consistently generates low rainfalls. This in turn, minimizes the error as the assumption is correct most of the time. However this comes at the cost that inevitably, peak values will be missed and the error will concentrate on peak values rather than being spread out.

Figure 6. Forecasting water levels at Tallebudgera Creek road one hour in advanced, lumped model
5 CONCLUSION AND FUTURE WORK

The paper explains the development of an ANN model for flood forecasting in catchments with a short response time. To test the skill of the proposed ANN Model, the system is run by the measured rainfall (beyond the time that forecast is performed). The predictions are satisfactory, indicating that the model has adequate skill to forecast flood.

The model is then tested with three different assumptions regarding rainfall between prediction time and prediction horizon. The assumptions are: i) no assumption about the rainfall, i.e. unseen rainfall is treated as a hidden variable in the ANN model ii) a naïve prediction of rainfall (i.e. rain will continue with the same rate as the last observation during the period between prediction time and prediction horizon), iii) The rainfall between the prediction time and prediction horizon is estimated by applying ANN to the available rainfall data until prediction time.

The model makes satisfactory decisions for short forecasting horizons in all the above mentioned scenarios, although the variance of the predictions is different in alternate scenarios. This confirms limited usefulness of the model for short term forecasting which is approximately 2 hours for this case study. Two reasons for poor long-term forecasting skill are identified namely:

1) Uncertainty associated with input to the model. This includes temporal and spatial resolution of rainfall prior to prediction time; and characteristics of the unseen rainfall between prediction time and prediction horizon
2) Uncertainty associated with model structure and parameters. This includes uncertainty generated due to a poor choice of optimisation method

The number of measurement points usually limits the degree that spatial resolution of rainfall can be improved. The study shows that improving the rainfall temporal resolution has a modest positive impact on the variance of the errors in flood forecasts to the catchment under study at short forecasting horizons.

In terms of usability as a decision support tool, the proposed ANN model (possibly any decision support model in general) has limitation in predicting long-term forecasts without the knowledge of unseen rainfall. In terms of the model structure and parameters, the proposed model has a robust structure and performs well when it is driven by measured rainfall during the period between prediction time and prediction horizon.

There are two possible ways to improve the model performance, i) develop a methodology that provides better predictions of rainfall between the prediction time and prediction horizon, ii) improve the model structure/parameters so it can use the pattern in the available data and provide a good forecast, in the absence of unseen rainfall input iii) rebalancing the dataset to allocate more learning capacity to peaks rather than to low flows.

A data driven model requires an in-built performance measure utility (i.e. optimisation method) for fitting free parameters of the model to observations in order to make good decisions. A literature review shows that optimisation method (goal) has been a less explored area in the field of ANN-based flood
forecasting. Existing literature has a focus on minimisation of RMS error as the main optimisation goal. As indicated by quantitative analysis of results, this generic optimisation goal is not appropriate for flood forecasting, as a flood forecaster is more interested in the performance of the model during peak flood, timing of flood and rate of rise and descend of flood water. It appears that there is some room for improving model structure through developing optimisation techniques that resemble the goal of the decision support tool. The following suggests a few alternative optimisation goals for future studies:

i) RMS error minimisation, by rebalancing the dataset – In this approach low flow data are filtered out of the optimisation scheme. This indicator aims at improving the magnitude of peak water level prediction.

ii) Timing of the peak. In this approach, the lag time between rainfall and water level can be used as a sole or secondary goal for optimisation. This indicator aims at improving the timing of the peak flood prediction.

iii) Rate of rise and descend of flood water. In this approach the rate at which water level rises and/or reduces can be used as a sole or secondary goal for optimisation. This indicator aims at improving the prediction of catchment response to rainfall.

iv) Risk minimisation – In this approach the model uses the consequence of bad decisions (by the ANN) as a feedback to the optimisation scheme.

ACKNOWLEDGMENT

This study is supported by the ARC industry linkage grant, project ID LP110200832. The authors of the paper would like to thank the Australian Research Council and Gold Coast City Council for funding this study. They also would like to thank Gold Coast City Council for the provision of data that have been used in this study.

REFERENCES


