Leadership in Sustainable Infrastructure
Leadership en Infrastructures Durables

Vancouver, Canada
May 31 – June 3, 2017/ *Mai 31 – Juin 3, 2017*

AUTOMATED FEATURE SELECTION FOR FUZZY NEURAL NETWORKS: AN APPLICATION FOR URBAN FLOOD PREDICTION

Khan, Usman T [1,2]
[1] York University, Canada
[2] usman.khan@lassonde.yorku.ca

Abstract: Urban floods are one of the most devastating natural disasters globally and improved flood prediction is essential for better flood management. Today, high-resolution real-time datasets for flood-related variables are widely available. This data can be used to create data-driven models for improved real-time flood prediction. However, data collected for extreme observations have large uncertainty, which must be accounted for in these models. In addition to this, typically the selection of input features and the architecture for data-driven models has been subjective. Addressing these concerns will improve flood prediction and will provide more accurate flood risk assessments. In this research, a new type of fuzzy neural network is proposed to predict peak flow in an urban river. The network uses fuzzy number to account for the uncertainty in the data and model parameters. An algorithm that uses possibility theory is used to train the network. An adaptation of the Automated Neural Pathway Strength Feature Selection (ANPSFS) method is used to select the input features. A search and optimisation algorithm is used to select the network architecture. A number of different inputs are considered including lagged precipitation and mean daily flow rate. The impact of this approach is that network training does not follow the typical ad hoc approach, but is based on objective criteria. Data for the Bow River in Calgary are used to train and test the network. Model performance using different features is used to compare the effectiveness of using the ANPSFS approach.

1 Introduction

In 2013, Calgary and southern Alberta experienced one of the worst natural disasters in Canada: floods caused approximately $6 billion in damage (Environment Canada, 2013), and were responsible for four deaths, and displaced more than 100,000 residents (Alberta Government, 2014; Khan & Valeo, 2016). The floods contributed to the transport of large amounts of sediment and the destruction of river banks, channels and aquatic ecosystems (Environment Canada, 2013). Better flood protection, and timely flood mitigation strategies may have prevented some of the destruction and casualties. Flood prediction models play an integral role in flood protection, and improved models are critically needed to help protect against more frequent and intense floods expected in the future.

The mechanisms behind extreme events in southern Alberta are generally understood and documented (Valeo et al., 2007) however, predicting floods remains a challenge because of the uncertainty in the numerical models. Flood prediction using physically-based and deterministic hydrologic models (such as a rainfall-runoff routing models) rely on simplified, conceptual representation of highly complex, correlated and spatially distributed processes that occur in a watershed (Wijiesekera et al., 2012; Vrugt et al., 2005). Uncertainty in the data used to calibrate model parameters and errors in model structure, compounds the complexity of flood models (Vrugt et al., 2005).

An alternative to physically-based models are data-driven models, which use generalized relationships between input and output datasets (Solomantine & Ostfeld, 2008). The models can characterize a system with limited assumptions and have similar (if not better) performance than physically-based models. A simpler model structure often means that the propagation of uncertainty from different sources is easier to quantify. These types of models (e.g. artificial neural networks, ANNs) have been widely used in hydrology, including for predicting flow rate (Adamowski & Sun, 2010; Duncan et al., 2011; Li et al., 2015). An advantage of using a data-driven approach is that data collected from real-time flow rate monitoring stations can be used to calibrate the model. This type of data is routinely collected by Environment Canada, which means that site specific surveys that are required for many physically-based methods are not necessary.

Data-driven models have intrinsic uncertainties associated with it that are not random or probabilistic in nature, thus, making it well suited for the use of fuzzy number theory (Ozbek & Pinder, 2006). Fuzzy numbers use fuzzy sets and possibility theory to describe uncertain or imprecise quantities, measurements or observations (Khan and Valeo, 2015). They are more suitable when data is missing, incomplete or vague, combined from multiple sites, and to represent uncertainty that is not random (Khan and Valeo, 2016). Using fuzzy numbers instead of probability-based methods means that some of the strict assumptions in many probability models can be relaxed (Kahraman et al., 2006). The literature demonstrates the utility and advantage of using fuzzy numbers in conjunction with data-drive methods. For example, Alvisi & Franchini (2011) proposed a method to create a fuzzy neural network (FNN), where the model coefficients and output of the ANN are fuzzy numbers rather than deterministic. These fuzzy numbers quantify the total uncertainty of the calibrated coefficients. This application is unique compared to the more typical use of fuzzy set based applications that involve ANN. Typical methods use fuzzy logic based methods where automated IF-THEN rules are used to create deterministic outputs (Abrahart et al., 2012; Alvisi & Franchini, 2011). These types of methods do not quantify the uncertainty of the predictions.

A major issue in implementing ANNs is that the choice of network architecture. This includes the selection of and the number inputs used, the number of hidden-layers, the number of neurons in each hidden-layer; the selection of the transfer functions, the amount of data in each subset, the training algorithm used; and the performance metric selected for optimisation. All of these important components of the network affect the values of the model coefficients, and the overall model performance. This means that each ANN model has uncertainty pertaining to the trained parameters. Recent reviews of ANN applications in hydrology have indicated that the lack of uncertainty quantification is a major reason for the limited appeal of ANN by water resource managers (Abrahart et al., 2012). Most ANN applications have a deterministic structure that generate point predictions without a quantification of the intervals corresponding to these predictions (Kasiviswanathan & Sudheer, 2013.). This means that end-users of these models may have excessive confidence in the forecasted values, and overestimate the applicability of the results (Alvisi & Franchini, 2011). However, the characterisation of uncertainty in a model is essential for both research and operational purposes. Without this, model results have limited value (Kasiviswanathan & Sudheer, 2013).

Thus, in this research a new method to account for the various uncertainties in ANN models is presented. First, an updated Automated Neural Pathway Strength Feature Selection (ANPSFS) method is used to select the optimum input data from a larger dataset. Second, a search algorithm is used to find the optimum network architecture (i.e. number of neurons in the hidden layer and fraction of data used for training). Lastly, a possibility-theory based FNN algorithm is used (based on Khan and Valeo, 2016b) to quantify the uncertainty of the model coefficients and outputs. Data from the Bow River in Calgary is used to train and validate the model, and results from the 2013 flood demonstrate the model.

2 Methods

2.1 Site description and data collection

The Bow River basin is located in southern Alberta and has an area of approximately 25,123 km^2. The basin provides significant economic activity for a major urban centre (the City of Calgary), as well as irrigation water for several irrigation districts in southern Alberta. The River originates in Bow Lake in the Rocky Mountains and flows south-easterly through Calgary (drainage area of 7870 km^2) and into Hudson Bay (Robinson et al., 2009). The River averages a 0.4% slope over its 645 km length. The Bow River is

supplied by precipitation accumulated in the snowpack in the Rocky Mountains, precipitation and discharge from shallow groundwater. Runoff peaks in the spring (June) while low flows are seen in the winter (January). Just upstream of Calgary, the River is regulated by Bearspaw Dam. The Bow River flows through many residential communities and the commercial centre of Calgary. Thus, it is extremely important for accurate and timely peak flow prediction of this river for the safety of the residents, and for protection of assets in downtown Calgary. Therefore, for this research the "Bow River at Calgary" (Environment Canada WSC Station ID: 05BH004) flow measuring station was selected for peak flow rate prediction. This station is located centrally within Calgary.

Eleven years of hourly flow rate data for was obtained from Environment Canada for the period January 1, 2000 to December 31, 2010. The annual median flow rate and the annual peak flow rate varied between 47 – 85 m^3/s and 172 – 787 m^3/s respectively. The highest peak flow rate occurred in 2005 (787 m^3/s). The peak flow rates in 2005 are associated with a flood event in Calgary and southern Alberta, which estimated to have cost approximately $400 million in damage (Valeo et al., 2007). This flood event was the last major flood in southern Alberta prior to the floods in June 2013. The peak flow rate for the 2013 event was approximately 1600 m^3/s, significantly higher than annual peaks in the previous decade (Khan & Valeo, 2016). Data from 2013 was also collected and used for model testing. In addition to the hourly flow rate, the daily minimum and maximum temperature and daily precipitation for the same period was obtained from Environment Canada.

The hourly flow rate collected for the station was filtered by removing dates where shift corrections (usually due to ice conditions) were applied by Environment Canada. Any data points within the selected range that had corrections applied were removed from further analysis. In addition to this, a second filter was applied to remove the low flow (below media ice-free median) rate periods, since the primary objective of this research was to predict peak flow rates. This reduced the data from an original, unfiltered set of 4018 days of hourly flow rate data to 1860 for Calgary, i.e. representing about 40% of available data.

2.2 Fuzzy Neural Network

Alvisi & Franchini (2011) proposed a method to create a FNN where the coefficients and the output of the network are fuzzy rather than deterministic numbers. In this method, a feedforward Multi-Layer Perceptron neural network (see Section 2.2.1 for more details) is modified to predict an interval rather than a single value for the coefficients and outputs. Each interval corresponds to an α-cut interval (an interval corresponding to the lower and upper limits of a fuzzy set at a defined membership level α). This is repeated for several α-cut levels to build a discretized fuzzy number at a number of membership levels. A stepwise, constrained optimisation algorithm is used to define these intervals. The constraints are such that to find the upper and lower limits of each coefficient to minimize the width of the predicted interval. However, Alvisi & Franchini (2011) defined the amount of data to be included in each interval arbitrarily, setting a predetermined amount of data to be captured within each interval. Khan & Valeo (2016b) refined this method by using the relationship between possibility theory and probability theory (known as probability-possibility transformations) to define more objective constraints. Using the transformation principles proposed by Dubois et al (2004), this method selects (1 – α) fraction of data to be captured within each interval, where α corresponds to the α-cut interval of a fuzzy number. Theoretical and mathematical details of this method are available in Khan & Valeo (2016b).

This FNN method was implemented in MATLAB (version 2016a). The built-in MATLAB Neural Network Toolbox was used with a two-step optimisation approach: first, the Shuffled Complex Evolution algorithm (commonly known as SCE-UA, Duan et al., 1992) was used to find an initial solution to the minimisation problem. Then, further refinement of the solution was conducted using the built-in MATLAB minimisation function *fmincon*.

2.2.1 Network architecture

For this research a three layer, feedforward Multi-Layer Perceptron architecture was selected to model peak flow rate (Q_P) using several candidate inputs (see Section 2.2.2). This structure is one of the most common structures for ANNs and thus, was selected to be able to compare directly with previous studies

(He et al., 2011). For this architecture, two transfer functions are required, one between the input and hidden-layer which was selected as the hyperbolic tangent sigmoid function, and one between the hidden-layer and output-layer which was a pure linear function (following Alvisi & Franchini, 2011). Network training was completed using a backpropagation algorithm, the Levenberg-Marquardt method, minimising Mean Squared Error (MSE). The input and output data were pre-processed before training, validation and testing. The data was normalised so that input and output data fell within the interval [-1 1].

There is no consistent method for selecting two important components of ANN architecture: the number of neurons in the hidden-layer (n_H), and the amount of data used for training, validation and testing (known as data-division) (Abrahart et al., 2012). Typically, an ad hoc or trial-and-error method is used to select the number of neurons (Maier et al., 2010; Alvisi & Franchini, 2011; He et al., 2011). The n_H selected must balance the complexity and generalisation of the final model; too many neurons increase the complexity and hence the processing speed, while reducing the transparency of the model. Not enough neurons risk reducing model performance and forgoing the ability of modelling non-linear systems. Similarly, the issue of data-division, which can have significant impacts on final model structure, is also determined in an ad hoc or trial-and-error basis (Maier et al., 2010).

For this research, a coupled method to select the optimum n_H and data-division for the ANN model described is proposed. This method uses the independent test dataset for all inferences regarding the error statistic (MSE) to be minimized. The smallest number of neurons and the least amount of data for training is targeted. The first is to reduce computational effort. The second is to prevent the risk of over-fitting to the training data, and to a larger dataset for testing for robust statistical inference of that dataset.

First, the dataset is randomly split into a 50%-25%-25% ratio for training, validation and testing, respectively. For each subset, data is randomly sampled to group into each subset. Then, the network is trained using 1 to 20 neurons. This process is repeated 100 times to account for the different selection of randomly sampled data in each subset. This is because the random initialisation of the ANN can cause variability in overall model performance (Napolitano et al., 2011). The MSE for each test dataset is then calculated and compared, and the number of epochs for each model is measured. Epochs are the number of times each model coefficient is modified in the optimisation algorithm. The configuration that leads to the lowest MSE for the testing dataset and the lowest number of epochs was collected. This process is then repeated by sequentially increasing the amount of data used for training (and thus, reducing the amount of data equally partitioned for validation and testing) by increments of 0.5% from 50% to 75%, conducting each iteration of this change 100 times. Using this process, a number of different circumstances where the highest performance was calculated can be listed, and the best combination can be objectively selected. In doing so, both the processing time and the complexity of the system is accounted for in the final model architecture.

2.2.2 Automated Neural Pathway Strength Feature Selection

The ANPSFS method (Duncan, 2014) is a method designed to reduce the complexity of models by reducing the number of inputs (i.e., the features) from the total available amount to only those that are relevant to the model at hand. While most data-driven models are typically considered black-box models, input feature selection methods can provide additional insights into the system by identifying the features that are the most relevant. In the ANPSFS method, the strength of a particular pathway from a given input to the output (i.e. QP) is calculated using:

[1] $W_{IO} = W_{IH} \times W_{HO}$

where W_{IH} are the weights (the coefficients) between the input and hidden layer, and W_{HO} are the weights between the hidden layer and output later. The higher the value of the W_{IO}, the more significant the associated input. In the present research, W_{IO} is calculated using all the input data for each year individually. This array of W_{IO} (one for each year) is then used to create the ensemble interquartile range (EQR) for each of the inputs as follows:

[2] EQR = $\min(|Q_1|,|Q_3|)/\max(|Q_1|,|Q_3|) \times \text{sign}(Q_1).\text{sign}(Q_3)$

where $Q_1$and Q_3 are the first and third quartile of the W_{IO}. The EQR are ranked in descending order, and the top three ranked inputs are selected as the most significant features. Thus, the ANPSFS method uses W_{IO} and EQR to select the most important features for Q_P (output) prediction given a set of input data.

2.3 Model performance evaluation

The Nash-Sutcliffe efficiency (R^2) (Nash and Sutcliffe, 1970) and the root mean square error (RMSE) were used to asses the model performance for each year:

[3] $R^2 = 1 - \sum(y_m - y_o)^2/\sum(y_o - y)^2$

[4] $RMSE = \text{sqrt}[\sum(y_o - y_m)^2]$

where y_m is the model output, y_o is the observed value, and y is the mean of the observed data.

3 Results

3.1 Network architecture selection

Sample results of the proposed method to select the number of neurons in the hidden layer and the amount of data used for each component are shown in Figure 1. Figure 1a shows the mean MSE (solid black line) of the test dataset for the initial 50%-25%-25% data-division scenario, with n_H varying between 1 to 20 neurons. This simulation was repeated 100 times and the upper and lower limits of MSE for each of these simulations are shown in grey. This figure demonstrates that the number of neurons did not have a noticeable impact on the MSE for this configuration. The most significant outcome of this process is that the variability (the difference between the upper and lower limits) of the performance seems to decrease after $n_H = 6$ and increases again after $n_H = 12$, with the lowest MSE at $n_H = 10$. This result has two important implications: first, increasing the model complexity results in limited improvement of model performance, suggesting that a simpler model structure may be more suitable to describe the system. Second, the variability in performance indicates that the initial selection of data in each subset can highly influence the performance, especially at the lower (i.e., $n_H < 6$) and higher (i.e., $n_H > 12$) ends of the spectrum of n_H. This suggests that an optimum selection of hidden neurons lies within this range.

Figure 1b shows the change in the mean (solid black line) and the variability (in grey) of the number of epochs needed to train the network for the initial data-division scenario, as n_H increases from 1 to 20. The variability of the number of epochs drastically decreases as n_H increases from 1 to 5. This means that a simpler model structure may require more time to train, and the performance of these simpler architectures (n_H = 1 to 5) is more variable. This is likely because the initial dataset selection has a higher impact on the final model performance for less complex models. The lowest number of mean epochs for this analysis occurred at $n_H = 19$, with 26 epochs.

The impact of changing the amount of data used for training, validation and testing on the model performance (MSE) was generally inconclusive (see Figure 1c) as the amount of data used for training was increased from 50% to 75% at 0.5% intervals. For the sample shown, ($n_H = 10$) increasing the amount of data used for training has minimal impact on model performance, indicating that using the least amount of data for training (and thus having a higher fraction available for validating and testing) would be ideal. The number of epochs needed for training the network at different data-division scenarios was inconclusive (see Figure 1d for the n_H =10 sample case). The significance of this analysis is that the amount of computational effort (or time) does not necessarily decrease as a larger fraction of data is used for training.

Based on these results, $n_H = 10$ with a 50%-25%-25% data-division was selected as the optimum architecture for this research. The fact that the mean and the variability of MSE was the lowest at $n_H = 10$ makes it a preferred option over the $n_H = 19$ case, which as a lower number of mean epochs but had higher variability in MSE. Secondly no significant trend was seen as the amount of data used for training,

validation and testing was altered, however lower MSE values were seen at n_H = 10 compared to other at n_H values. Thus, the option that guarantees the largest amount of data for validation and training is preferred.

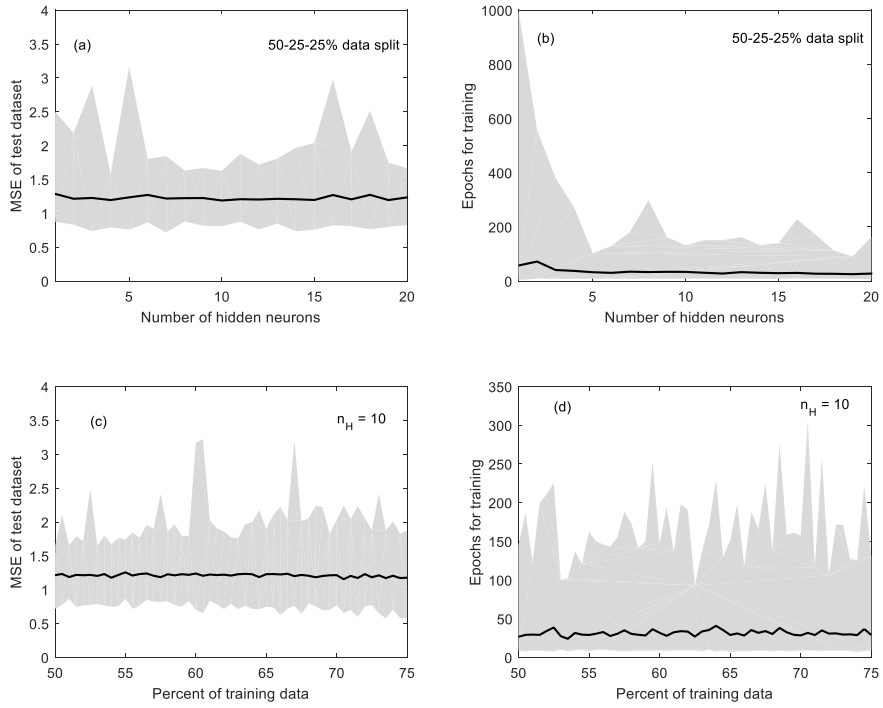

Figure 1: Sample results of the coupled method to determine the optimum number of neurons in the hidden-layer and percentage of data for training:

The overall outcome of this component of this research was that that instead of using the typical trial-and-error based approach to selecting neural network architecture parameters; the proposed method can provide objective results. Systematically exploring different numbers of n_H and division of data can help select a model with the best performance. Once these FNN parameters were identified, subsequent training of the FNN was completed and the results of this are presented in the next section.

3.2 Input feature selection using the ANPSFS method

The input data used for Q_P prediction in the Bow River at Calgary included mean daily flow rate (calculated using the collected hourly data), maximum, mean and minimum air temperature, and total daily precipitation. Each of these datasets were lagged by 1, 2 or 3 days, in order to be able to predict Q_P with 1, 2 or 3 days advance notice. Thus, the total amount of candidate input parameters were 18 (lagged versions of the above parameters, and lagged Q_P for each lag).

Results of the ANPSFS method are shown in Figures 2 – 4 below. Figure 2 shows the values of W_{IO} of each year for all 18 inputs. The figure shows that the importance (i.e. strength) of each candidate input parameter changes on an annual basis. The trend shows that there is no clear input parameter that is consistently ranked high, indicating the changing importance of each parameter in prediction Q_P. Figure 3 shows a box-plot of each candidate input parameter further highlighting this variability. When W_{IO} is below zero, this means that input parameter is not significant in prediction Q_P, i.e. it acts as an inhibitor in predicting QP. Several candidate variables have mean values of W_{IO} below 0 highlighting that their potential as an input is low.

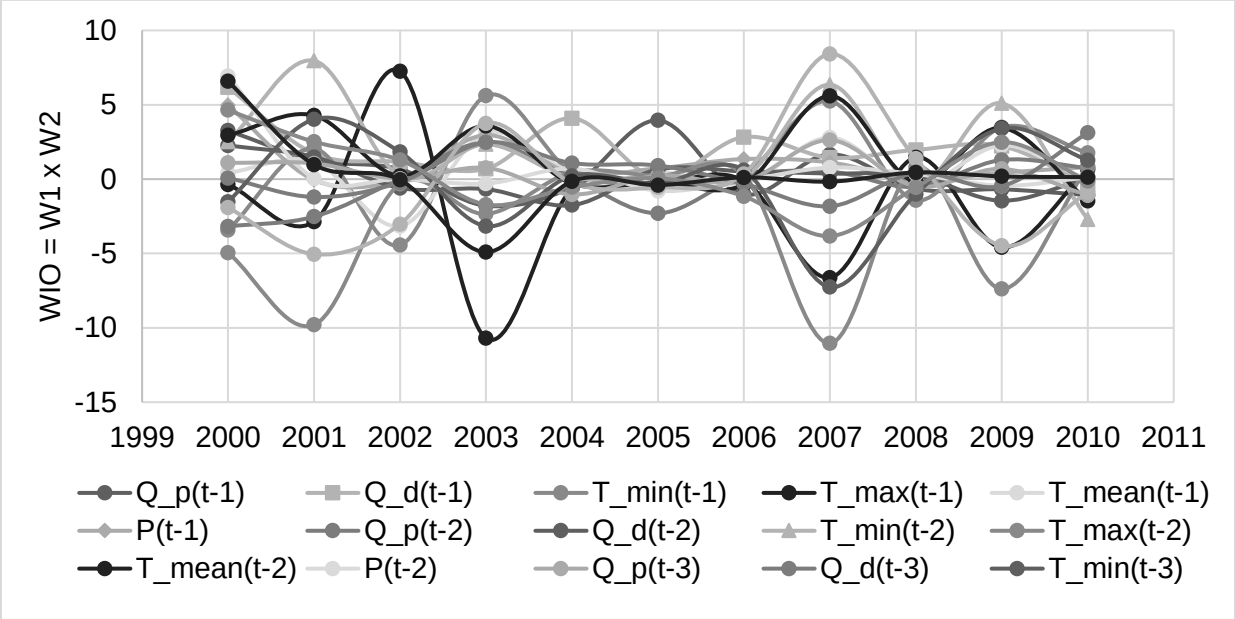

Figure 2: Values of WIO for the entire study period for each of the 18 input variables

Figure 4 shows the EQR values calculated using the W_{IO} for each feature from highest to lower. It shows that only three inputs have a positive (and thus non-inhibiting) effect: 1-day lagged mean flow, 1-day lagged precipitation and 2-day lagged minimum air temperature. Thus, for this research these three inputs are selected as the final inputs to the model to predict Q_P in the Bow River at Calgary. It is worth nothing that as additional data is available and added to the system it can be used to update the ANPSFS to select other input parameters to improve Q_P prediction. The significance of these results is that the method can be used to reduce the complexity of the model by reducing the number of inputs from 18 to 3, removing only those parameters that may not relevant.

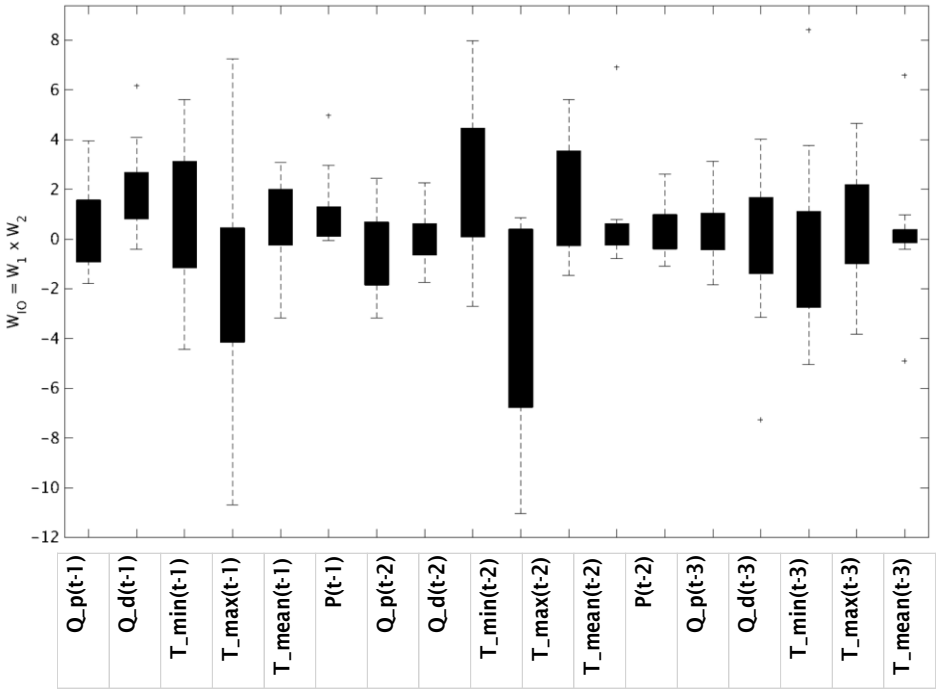

Figure 3: Box whisker plots for WIO for each of the input parameters

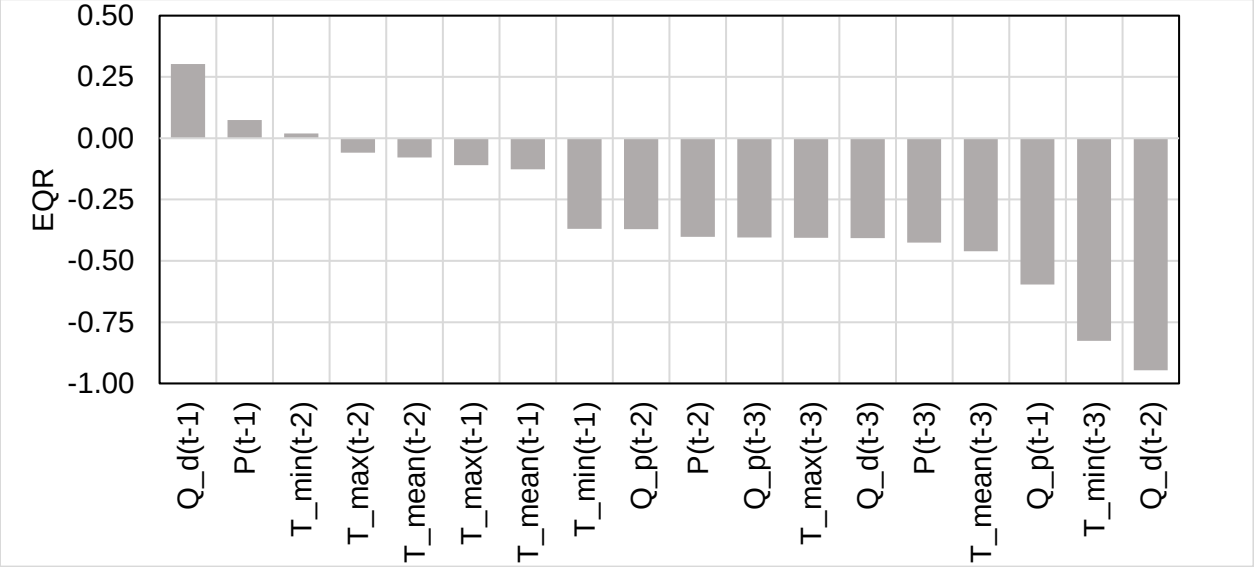

Figure 4: EQR values for each of the 18 candidate features to predict peak flow rate

3.3 Flood prediction

A FNN model was constructed using the method outlined in Section 2.2 using the three input parameters: 1-day lagged mean flow, 1-day lagged precipitation and 2-day lagged minimum air temperature. Figure 5 shows the R^2 and RMSE of predicting Q_P in the Bow River at Calgary for the eleven period. Network performance with these inputs was high: RMSE ranged between 7 and 47 m^3/s (about 10% of mean flow). The R^2 value was above 0.90 for all years except 2005 (which was also a flood year). This approach shows that using an objective rather than ad hoc method to design the architecture of a neural network (for input selection, n_H and data-division) can provide reliable results, which by design are ones with the highest performance. The trained FNN was used to predict the peak flow in the Bow River for the 2013 flood event, which is summarised in the next section.

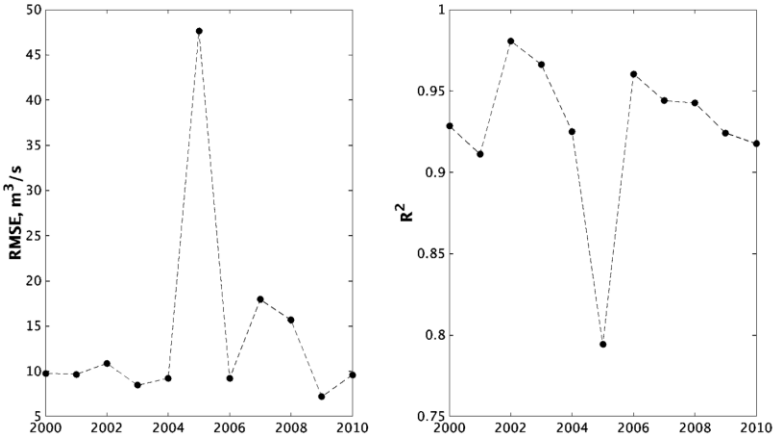

Figure 5: A summary of the RMSE and R2 values of the FNN mode

3.3.1 2013 Calgary floods

Figure 6 shows a comparison of predicted and observed daily peak flow in Calgary for 2013. Note that the predicted values for Q_P are fuzzy numbers rather than discrete numbers. So the Figure shows results for

a membership level of 1 (the central value of the fuzzy number) and for membership level of 0 (the upper and lower bounds of the fuzzy number). The predicted intervals captured all daily peak flow rates, including the highest flows observed that year during the flood. This shows that using an FNN can predict Q_P in the Bow River up to a day in advance with high accuracy. Model outputs can be used for risk assessments to determine the probability of the predicted flow to exceed a given threshold. This data can be used by water resource managers to implement flood defence systems.

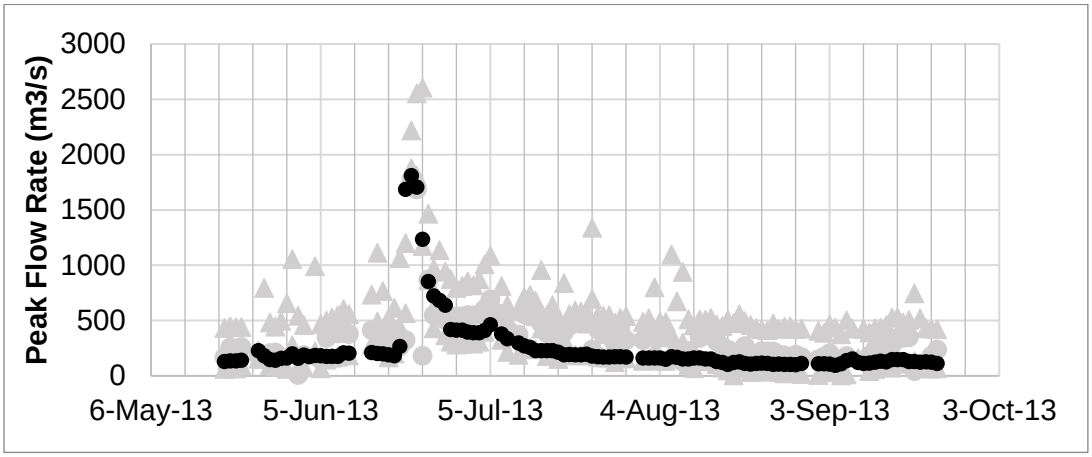

Figure 6: A comparison of observed (black circles) and predicted (grey markers) peak flow rate for Calgary in 2013

4 Conclusion

In this research, three different methods were used to reduce the overall uncertainty in an ANN model used to predict peak flow rate in the Bow River in Calgary. First, an existing fuzzy neural network was adapted using the principles of possibility theory. This allowed uncertainty in the model coefficients and output to be quantified by representing them as fuzzy numbers rather than discrete numbers. Second, an objective method to select the architecture of the FNN was developed. Lastly, the inputs used in the model were selected using a modified Automated Neural Pathway Strength Feature Selection (ANPSFS method). The impact of each of these methods was demonstrated by training a large dataset (11 years of data) for Calgary, and then tested on one extreme event (the 2013 floods in Calgary).

Acknowledgements

The author would like to acknowledge York University for funding, Dr S Alvisi (UNIFE) for the original FNN code, Dr A Duncan (Exeter) for discussions on ANNs; and Environment Canada for providing the data.

References

Alberta Government (2014). Respecting Our Rivers: Alberta's Approach to Flood Mitigation from https://pabappsuat.alberta.ca/albertacode/images/respecting-our-rivers.pdf (accessed July 1, 2014).

Adamowski, J., & Sun, K. (2010). Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. Journal of Hydrology, 390(1), 85-91.

Alvisi, S., & Franchini, M. (2011). Fuzzy neural networks for water level and discharge forecasting with uncertainty. Environmental Modelling & Software, 26(4), 523-537.

Abrahart, R. J., Anctil, F., Coulibaly, P., Dawson, C. W., Mount, N. J., See, L. M., Asaad Y. Shamseldin, A. Y., Solomatine, D. P., Toth, E., & Wilby, R. L.: Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting, Prog. Phys. Geog., 36, 480-513,

Duncan, A., Chen, A. S., Keedwell, E., Djordjevic, S., & Savic, D. (2011). Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks. In Weather Radar and Hydrology IAHS Red Book series no. 351. IAHS. Exeter, UK.

Duncan, A. P. 2014. The Analysis and Application of Artificial Neural Networks for Early Warning Systems in Hydrology and the Environment. PhD Thesis. University of Exeter, Exeter, UK.
Dubois, D., & Prade, H. (1997). Bayesian conditioning in possibility theory. Fuzzy Sets and Systems, 92(2), 223-240.
Dubois, D., Foulloy, L., Mauris, G., & Prade, H. (2004). Probability–possibility transformations, triangular fuzzy sets, and probabilistic inequalities. Reliable Computing, 10(4), 273-297.
Duan, Q., Sorooshian, S., & Gupta, V.: Effective and efficient global optimisation for conceptual rainfall-runoff models, Water Resour. Res, 28(4), 1015-1031,
Environment Canada (2013). Alberta's Flood of Floods on Canada's Top Ten Weather Stories for 2013, http://ec.gc.ca/meteo-weather/default.asp?lang=En&n=5BA5EAFC-1&offset=2&toc=show (accessed July 1, 2014).
He, J., Chu, A., Ryan, M. C., Valeo, C. & Zaitlin, B.: Abiotic influences on dissolved oxygen in a riverine environment. Ecological Engineering, 37, 1804–1814
Khan, U. T., & Valeo, C. (2016). Short-term peak flow rate prediction and flood risk assessment using fuzzy linear regression. Journal of Environmental Informatics. 28(2), 71-89.
Khan, U. T., & Valeo, C. (2014b). A new fuzzy linear regression approach for dissolved oxygen prediction. Hydrological Sciences Journal.
Kahraman, C., Beşkese, A. & Bozbura, F. T. (2006). Fuzzy regression approaches and applications. In C. Kahraman, (Ed.), Fuzzy Applications in Industrial Engineering, pp. 589-615. Springer Berlin Heidelberg. Berlin, Germany
Kasiviswanathan, K. S., & Sudheer, K. P.: Quantification of the predictive uncertainty of artificial neural network based river flow forecast models, Stoch. Env. Res. Risk A., 27, 137-146
Khan, U. T. and Valeo (2016b). Dissolved oxygen prediction using a possibility theory based fuzzy neural network. Hydrology and Earth System Sciences, 20(6), 2267
Li, Z., Huang, G., Han, J., Wang, X., Fan, Y., Cheng, G., Zhang, H., & Huang, W. (2015). Development of a Stepwise-Clustered Hydrological Inference Model. Journal of Hydrologic Engineering, (ahead-of-print).
Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P.: Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions, Environ. Modell. Softw., 25, 891-909,
Napolitano, G., Serinaldi, F., & See, L.: Impact of EMD decomposition and random initialisation of weights in ANN hindcasting of daily stream flow series: an empirical examination, J. Hydrol., 406(3-4), 199-214,
Nash, J. E., & Sutcliffe, J. V.: River flow forecasting through conceptual models: Part I. A discussion of principles, J. Hydrol., 10(3), 282-290. 1970.
Ozbek, M. M., & Pinder, G. F. (2006). Non-probabilistic uncertainty in subsurface hydrology and its applications: an overview. Water, Air, & Soil Pollution: Focus, 6(1-2), 35-46.
Robinson, K. L., Valeo, C., Ryan, M. C., Chu, A., & Iwanyshyn, M. (2009). Modelling aquatic vegetation and dissolved oxygen after a flood event in the Bow River, Alberta, Canada. Canadian Journal of Civil Engineering, 36(3), 492-503.
Solomantine, D. P., & Ostfeld, A. (2008). Data-driven modelling: some past experiences and new approaches. Journal of Hydroinformatics, 10(1), 3-22.
Valeo, C., Xiang, Z., Bouchart, F. J.-C., Yeung, P. & Ryan, M. C. (2007). Climate Change Impacts in the Elbow River Watershed. Canadian Water Resources Journal. 32(4), 285-302.
Vrugt, J. A., Diks, C. G., Gupta, H. V., Bouten, W., & Verstraten, J. M. (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. Water Resources Research, 41, W01017.
Wijesekara, G. N., Gupta, A., Valeo, C., Hasbani, J. G., Qiao, Y., Delaney, P., & Marceau, D. J. (2012). Assessing the impact of future land-use changes on hydrological processes in the Elbow River watershed in southern Alberta, Canada. Journal of hydrology, 412, 220-232.