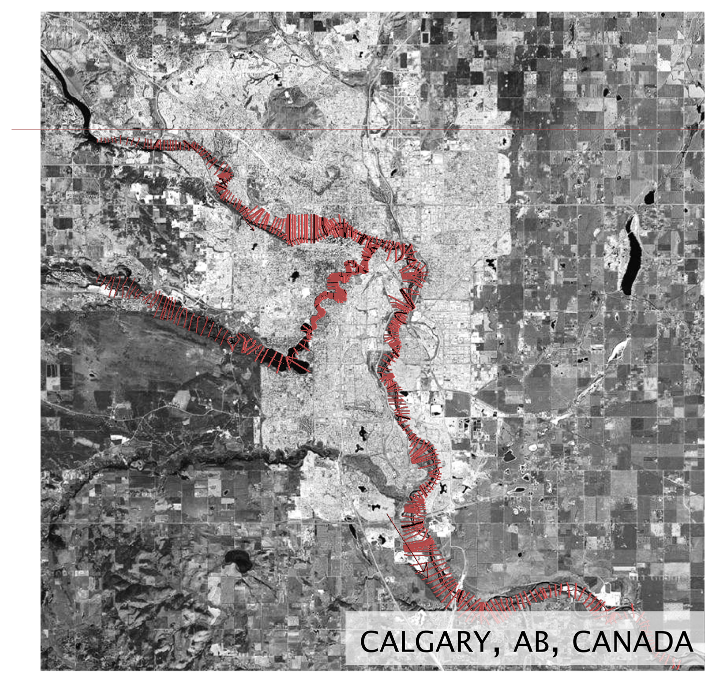
Real-time flood forecasts using a possibility theory based fuzzy neural network

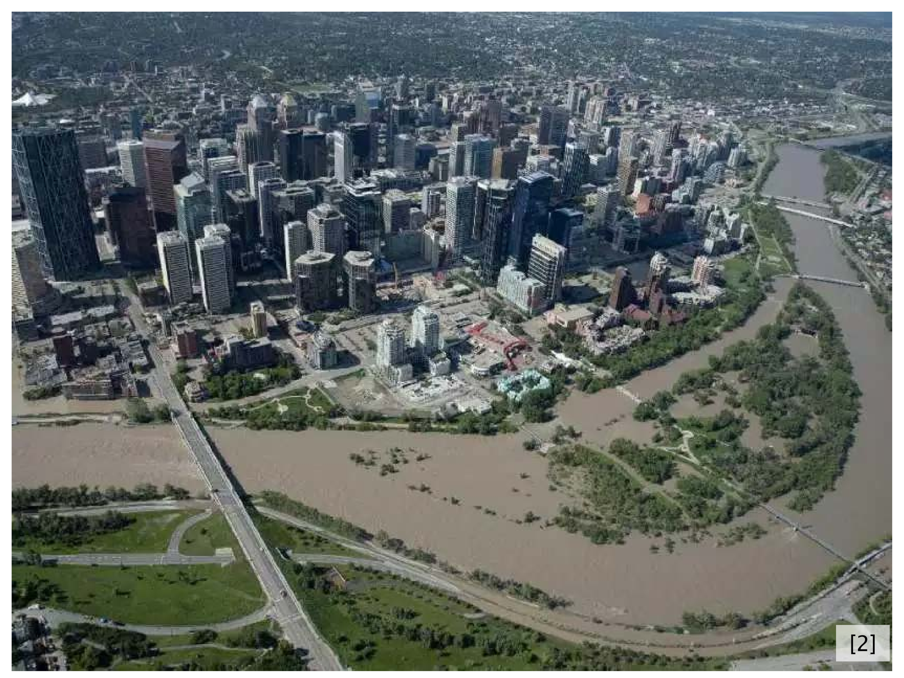
Usman T. Khan | Assistant Professor | Civil Engineering, York University | Toronto, ON
utkhan.info.yorku.ca

Acknowledgements: YorkU for funding this research; Dr Alvisi for the original FNN code; Dr Duncan for in-depth ANN discussions

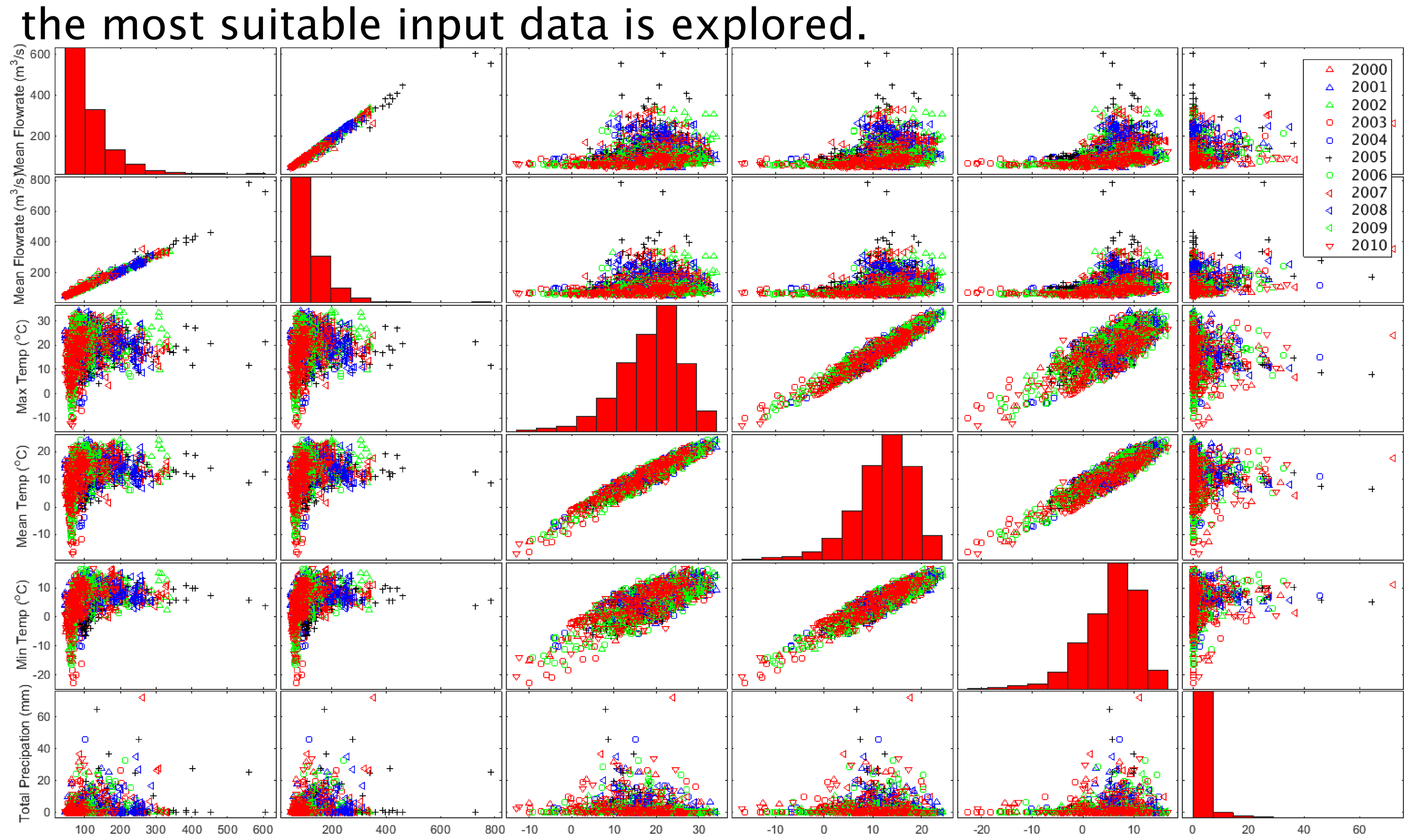
Introduction

The June 2013 floods in Calgary, Alberta were one of the worst natural disasters to occur in Canada. The floods were responsible for four deaths, displaced more than 100,000 Albertans, and caused approximately $6 billion in damage [1]. This event highlighted the importance and necessity of better flood protection, effective and timely flood mitigation strategies, including improved flood prediction.



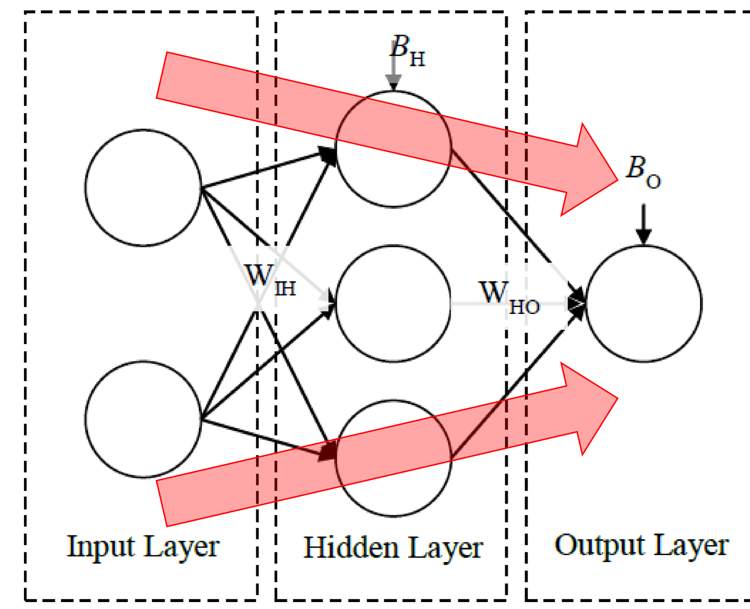


Given the wide availability of high-resolution, real-time meteorological and discharge data in many urban areas, there is great potential to use data-driven methods (such as neural networks) for predicting floods in urban rivers. However, data-driven models have intrinsic uncertainties associated with them that cannot be represented using probability theory exclusively. In this research we present a fuzzy-set based method of quantifying the uncertainty in neural networks. Additionally, a method (the Combined Neural Pathway Strength Analysis, CNSPA, [3]) to select the most suitable input data is explored.



Peak daily and mean daily flowrate, minimum, maximum and mean daily temperature and cumulative daily precipitation for Calgary between 2000 and 2011.

Methods



In the CNSPA method, the strength of a particular pathway from the input and output is calculated using:

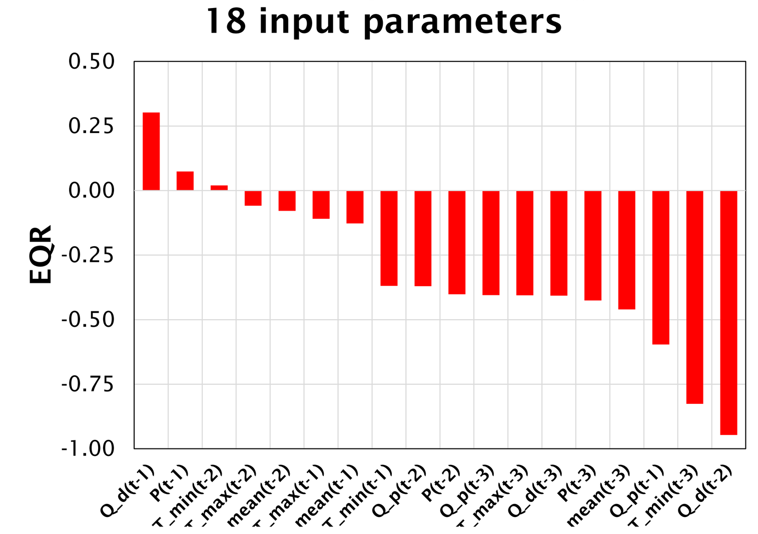
$$W_{IO} = W_{IH} \times W_{HO}$$

where W_{IH} are the weights 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. W_{IO} is then calculated for each years data using lagged (1, 2, and 3 days) input data. The ensemble interquartile range (EQR) is then calculated foe each input as follows:

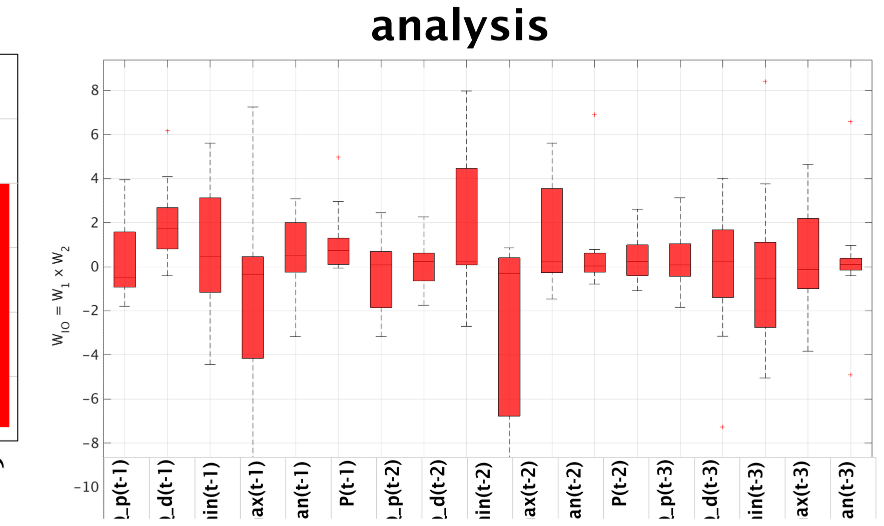
$$EQR = \min(|Q_1|,|Q_3|)/\max(|Q_1|,|Q_3|) \times \mathrm{sign}(Q_1).\mathrm{sign}(Q_3)$$

where Q_1 and Q_3 are the first and third quartile of W_{IO}. The EQR are ranked in descending order, and the top three ranked inputs are selected as the final inputs of the model. The network is designed with one hidden layer, three hidden layer neurons, hyperbolic tangent and linear activation functions between the first and second layers. Half the data was used for training, 25% for validation, and 25% for testing the network. The network was trained using possibility theory based intervals following [4, 5].

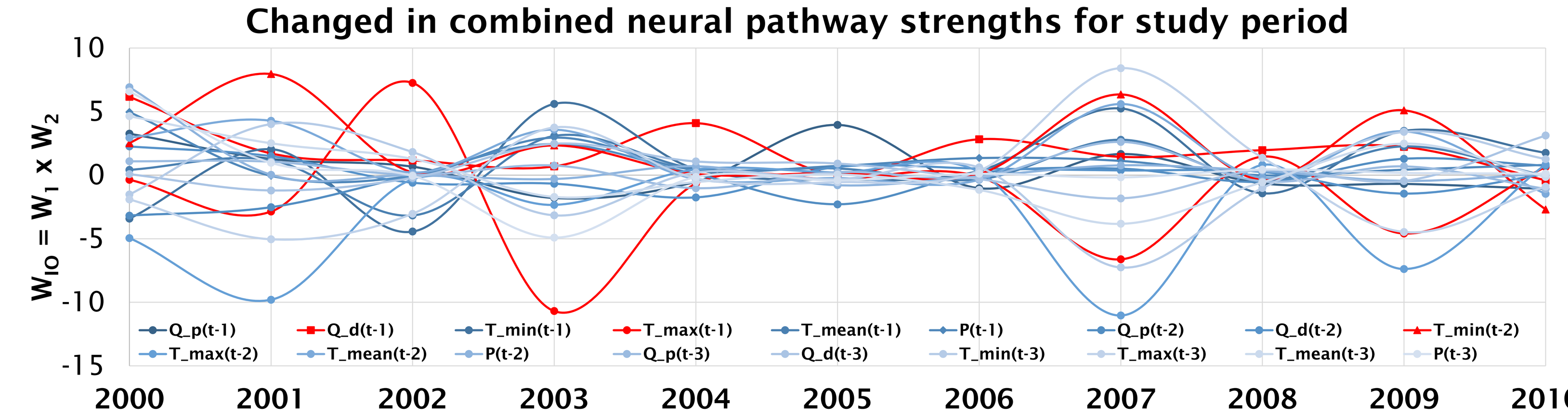
Ensemble Interquartile Range for 18 input parameters



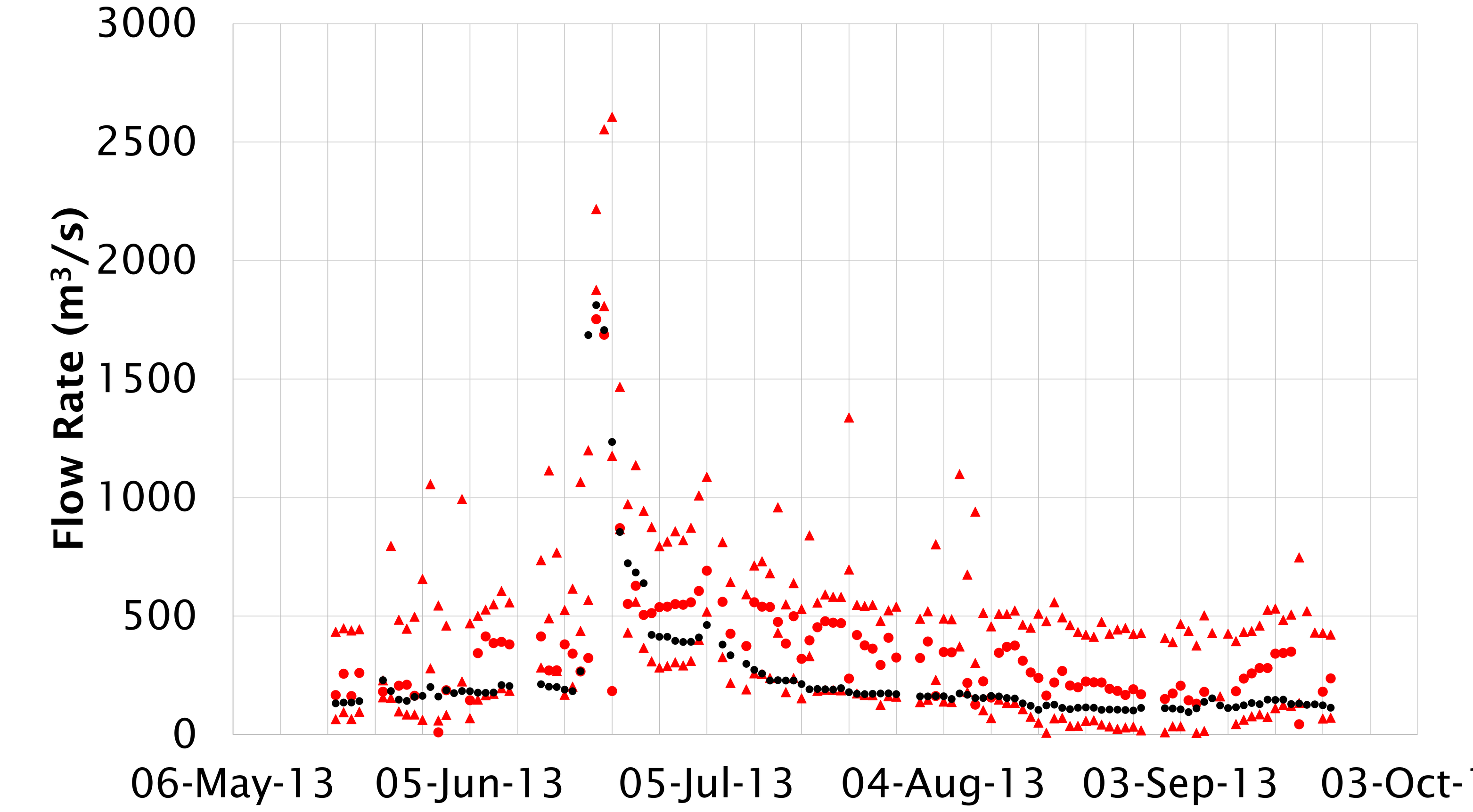
Combined neural pathway strength analysis



Changed in combined neural pathway strengths for study period

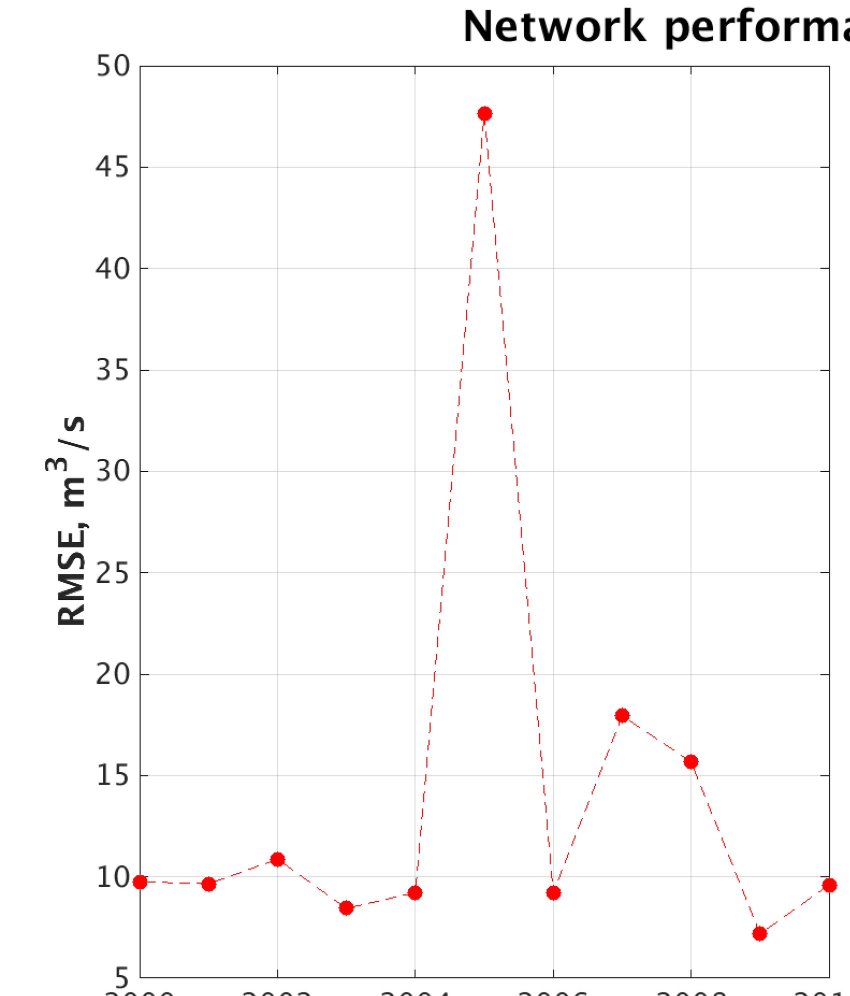


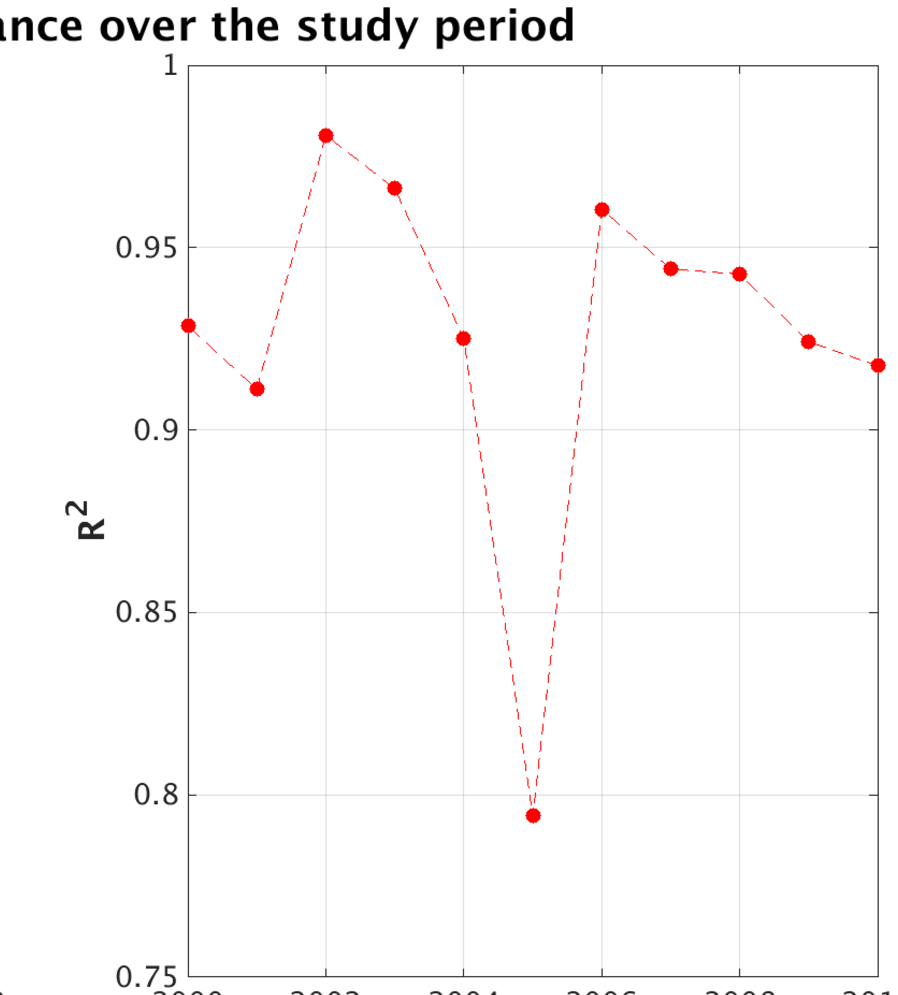
Results



Q_D(t-1), P(t-1) T_{min}(t-2) were selected as the input data using the EQR method. 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. The 2013 flood data set was used to test the fuzzy neural network. The predicted fuzzy intervals captured all daily peak flow rates within its intervals at a membership level of 0. This shows that using an FNN can predict Q_p in a river up to a day in advance with high accuracy. The CNSPA and EQR approach makes selecting network input an objective decision.

Network performance over the study period





[1] Khan, U. T. & Valeo, C., 2016. Journal of Environmental Informatics. [2] Hayward, J., 2013. The Canadian Press. [3] Duncan, A. P., 2014. PhD Thesis. University of Exeter. [4] Alvisi, S. & Franchini, M. , 2011. Environmental Modelling & Software. [5] Khan , U. T. & Valeo, C., 2016, Hydrology & Earth System Sciences.