CSCE Annual Conference
Growing with youth – Croître avec les jeunes

Laval (Greater Montreal)
June 12 - 15, 2019

COMPARISON OF FOUR INPUT VARIABLE SELECTION METHODS FOR ARTIFICIAL NEURAL NETWORK BASED FLOOD FORECASTING MODELS

Snieder, E.[1], Shakir, R.[1], Khan, U.T.[1,2],
[1]Department of Civil Engineering, York University, Toronto ON, Canada
[2]usman.khan@lassonde.yorku.ca

Abstract: Artificial neural networks (ANNs) are increasingly used for flood forecasting. The performance of these models relies on the selection of relevant and non-redundant input variables. However, input variable selection (IVS) is typically performed using expert knowledge or simple linear methods. This study evaluates four IVS methods based on the following criteria: partial correlation (PC), partial mutual information (PMI), input omission (IO), and combined neural pathway strength (CNPS). Each method is to inform models for the Bow River for daily water level forecasts of 1, 2, and 3 timesteps. The results of this study indicate that CNPS produces the best performing ANNs. Additionally, this study demonstrates that termination criteria do not reliably identify the optimum number of inputs for the ANNs. Instead, it is recommended that the number of inputs be determined by a model-based optimization, where the IVS method is only used to rank input usefulness.

1 INTRODUCTION

The risk of urban flooding is expected to increase due to climate change and rapid loss of permeable areas in cities (ASCE 2000a, Wilby 2006). Damage caused by flooding can be mitigated using early flood warning systems, which can provide advance warning of flood risk to local authorities and floodplain occupants, reducing the risk of damage to property and loss of life (Shrubsole et al. 1993, Yin et al. 2004). Typically, flood warning systems rely on rainfall-runoff models that estimate future stage or discharge levels.

Conventionally, physically based rainfall-runoff models have been used for hydrological forecasting. These models typically rely on simplifications of complex, nonlinear hydrological processes, which also exhibit high temporal and spatial variability (ASCE 2000b, Wijesekara et al. 2012, Khan and Valeo 2016). Throughout the past two decades, data-driven models (DDMs) have emerged as capable alternatives to physically based models for describing rainfall-runoff systems (ASCE 2000b, Dawson and Wilby 2001, Shrestha and Nestmann 2009). While physically based models rely on principles of physics to describe a system, DDMs use mathematical relationships and patterns in the data to characterize the systems (Solomatine and Ostfeld 2008). Among DDM methods, artificial neural networks (ANNs) are the most widely used for flood forecasting applications (Solomatine and Ostfeld 2008).

1.1 Overview of input variable selection

While ANN models have repeatedly demonstrated their aptness for predicting flow, the selection of input variables (a process commonly referred to as Input Variable Selection, IVS) is not widely used in model development and is often named as a gap in research (Maier and Dandy 2000, Maier et al. 2010, May et al. 2011, Abrahart et al. 2012). IVS methods are typically used to identify the most useful features from a larger set of candidate input variables. Useful inputs are defined as having maximum relevancy to the output while minimizing redundancy with other candidate inputs (May et al. 2011). Minimizing the number of model inputs is important for lowering computational demand, reducing output variability caused by local minima on the error surface, and can inform on behaviour of underlying hydrological processes (Šindelář and Babuška 2004, Bowden et al. 2005a, May et al. 2008a).

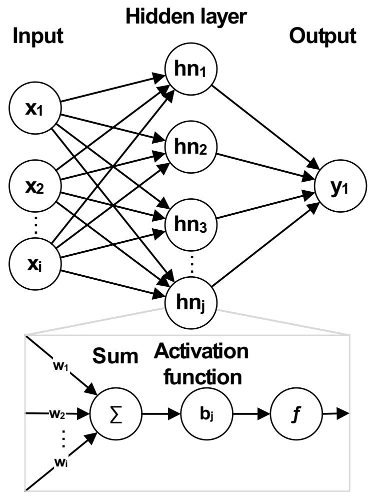

Figure 1: A schematic showing the components of a typical feed-forward multi-layer perceptron ANN model including the inputs (x_i), nodes in the hidden layer (hn_j), the output (y), the weights (w_i), the biases (b_j), and the activation function (f).

Many applications of ANNs for flow forecast predictions do not describe a systematic or rigorous IVS process; or rely on methods such as *a priori* knowledge of the system, trial-and-error, or linear cross-correlation for input selection (Maier and Dandy 2000, Bowden et al. 2005a, Abrahart et al. 2012). Each of these approaches has severe limitations. Expert knowledge of hydrological systems is not dependably available for many systems and may be unreliable or incomplete. Trial-and-error approaches, where a brute-force method is used to determine the best inputs, are computationally intensive, especially for ANN systems with a large number of candidate inputs (May et al. 2011). Finally, linear cross-correlation, which is the most commonly used data-driven IVS approach (ASCE 2000b, Abrahart et al. 2008, Nanda et al. 2016), is limited to identifying discrete, linear relationships between the output and individual candidate inputs. While linear cross-correlation may be useful for providing modellers with the approximate lag times between monitoring points within a system, it is not capable of capturing the interdependencies, redundancies, and nonlinearities typical of hydrological systems and therefore unsuitable for identifying optimum inputs for ANNs (Abrahart et al. 2012). Consequently, there is a clear need for more robust IVS methods that do not rely on *a priori* knowledge or assumptions about the system, are computationally inexpensive, and can characterise nonlinear and interdependent relationships between input candidates. IVS has been the focus of multiple review papers that provide a comprehensive overview of its taxonomy and methods in environmental systems modelling (Bowden et al. 2005a, Maier et al. 2010, May et al. 2011). IVS methods can be grouped into two broad classes: model-free methods and model-based methods, which are described in the sections below, followed by the overall objective of the research is presented, and finally a detailed description of four IVS method used in this study.

1.1.1 Model-Free Methods

Model-free IVS methods are methods that do not rely on an established model (Bowden et al. 2005a, May et al. 2011). The majority of IVS methods fall into this classification, and it includes common methods such as *priori* knowledge and linear correlation. Two model-free IVS methods are applied in this study: Partial Correlation (PC) and Partial Mutual Information (PMI).

1.1.2 Model-Based Methods

Model-based IVS methods where the selection is dependant on an existing model. Input usefulness is determined based on their behaviour within the model and once useful inputs are identified, models are recalibrated using only the selected inputs (May et al. 2011). The main limitations of this method are its high computational demand if the candidate input set is very large, and that it requires assumptions about the model configuration. For example, ANN models with varying architectures may be reliant on different inputs. This is addressed to some degree by using ensembles of neural networks, discussed further in section 0. Two model-based methods are used in this study: Input Omission (IO) and Combined Neural Pathway Strength (CNPS).

The IVS methods selected for this study, whether model-based or model-free, are conceptually distinct in their approach to selecting the most relevant inputs from a larger candidate dataset. An important factor for each method is the termination or stopping criterion used to determine which inputs are considered relevant and which are excluded. The termination criterion that is used has a major impact in the final configuration, the effort needed to train, and hence, the performance of an ANN model. Therefore, each IVS method is applied with and without a termination criterion. When a termination criterion is not used, the each IVS method selects a fixed number of inputs.

1.2 Objectives

The ultimate objective of this research is to further develop, refine and compare IVS methods described above: PC, PMI, IO, and CNPS and to provide modellers with improved tools for ANN model development. First, we propose two novel advancements of the IO and CNPS methods: the new IO method builds on previous IO methods and is adapted to quantitatively identify non-useful inputs; the CNPS builds on work by Duncan (2014) and is improved by eliminating the requirement of an arbitrary significance threshold. We compare these two model-based methods to two commonly used model-free methods (PC and PMI). Such a comparison is necessary to demonstrate the advantages and limitations of each approach, and a comprehensive comparison of performance efficacy has not been published in literature before. These methods were chosen based on their distinctive characteristics, which allow for a comparative analysis. Other than comparing model-free and model-based IVS approaches, the selection of the aforementioned IVS methods also allows for the comparison of linear (PC) and non-linear (PMI, CNPS, IO) methods, and low computational effort (PC, CNPS, IO) versus high computational effort (PMI) methods.

Secondly, the impact of the standard, predetermined termination criterion used for each IVS method on ANN model performance is quantified. The model performance with termination criteria-based selected inputs is compared to model performance using a pre-defined number of inputs using each IVS methods. This analysis will help answer questions related to optimizing model complexity (measured as the number of inputs used) and model performance. Lastly, the ANN model performance using IVS is compared to performance using a full set of candidate inputs in an effort to quantify the improvement in model performance by elimination those input variables that are not relevant in predicting the model output.

This study is focused on the Bow River watershed, with predictions made for Calgary, Alberta. The following section provides details of the study regions, the mathematical development of each IVS method, the structure of the ANN used, the candidate input variables, and model performance metrics used for the comparison.

2 METHODS

The following section outlines the methods used for this study. All computations were performed using MATLAB 2018b.

2.1 Site description and data source

This study used hydro-meteorological data from the Bow (Upper and Central) watershed, located in western Canada. The target gauge for forecasts is situated in Calgary, and other inputs include temperature and precipitation gauges in Calgary, and a flow gauge located over 100km upstream. Data in November to April is removed due to snow conditions. The Upper and Central Bow River watersheds have headwaters fed by the Rocky Mountains and have predominantly natural and agricultural land use. The data used in this study was collected and distributed by Environment Canada, the Canadian Water Survey, or the City of Calgary.

2.2 ANN Model configuration

A generalized process flow diagram describing the modelling process is provided in A feedforward multi-layer perceptron ANN is used for this research, which is the most widely used type of ANN flood forecasting (Maier et al. 2010, Abrahart et al. 2012). Choosing a widely used model type facilitates comparison with existing studies. Moreover, both model-based IVS methods are designed specifically for feedforward ANNs.

Upstream water level and local meteorological stations are used as candidate input variables at several lag times to predict the downstream (target) stage at several lead times. The objective of this structure is to be able to predict the stage in advance to highlight the risk of high peak flow or flooding conditions using available upstream data. The number of time lags applied to upstream stations for selecting candidate inputs was determined based on a linear cross-correlation. The cross-correlation based candidate input selection was very lenient, as it only serves to provide an approximate idea of which time lags to include in the candidate set. In total there are 30 candidate inputs, which are summarized in Table 1. The lag times in the table are denoted such that for an input X_{t-L} is corresponds to a forecast Y_{t+F}; where L and F correspond to the lag and lead shifts (times), respectively.

Table 1: Summary of all candidate input variables for the Bow River system

	Station ID	**Data type**	**Data source**	**Lag times**	**Total inputs**	**Lead times**
	BB001, BH004	Max, min, mean daily water level	WSC	0:1:2 d	18	
Bow River	3031093	Cumulative daily precipitation	City of Calgary	0:1:2 d	3	1, 2 & 3 d
	3031093	Max, min, mean daily temperature	City of Calgary	0:1:2 d	9	

All model configurations use a single hidden layer with 25 nodes, which is sufficiently sized such that it does not restrict the performance of models with a large input set. While there exists various ‘rule of thumb’ methods for selecting the number of hidden neurons, there is high variability between these methods (Maier and Dandy 2000). The hidden layer could be easily be optimized based on a systematic trial-and-error approach, it would be computationally demanding to repeat this process for each unique IVS method (Khan et al. 2018). Moreover, the model-based IVS methods would require a coupled optimization as their selection is influenced by the hidden layer size. For the sake of simplicity and comparison between methods, the hidden layer size is fixed for all model configurations. It is likely that optimizing the hidden layer size on a case-by-case basis would improve model performance, as an oversized hidden layer size may produce poor convergence during training. The second-order Levenberg-Marquardt algorithm is used to train the neural network. While the first-order backpropagation algorithm is the most widely used algorithm for flood forecasting ANNs, the Levenberg-Marquardt is commonly used and is considered more efficient than the simpler backpropagation algorithm (Maier et al. 2010, Abrahart et al. 2012).

Typically, for ANNs, datasets are partitioned into three subsets: training, validation, and testing. The training subset is used to calibrate the ANN’s weights and biases, the validation subset is used to terminate training (and prevent overfitting), and the testing subset is used to evaluate the model performance. In this study, the dataset is partitioned into training-validation-testing blocks of 60%-20%-20%, respectively. K-Fold Cross-Validation (KFCV) is used to train ensembles of ANNs instead of single ANNs. In this study, KFVC has four unique training-validation configurations, and the testing block is constant in each instant. Next, since ANNs are initialized with random values for weights and biases, each unique fold is trained 250 times, (referred to as multi-start). Collectively, these techniques capture some of the uncertainty associated with the model configuration during calibration. Since model ensembles are being used, results are typically displayed with confidence limits or as distributions.

Best practices in hydrological modelling suggest that model performance be assessed the based on multiple performance measures, as different measures capture different model characteristics (Maier et al. 2010, Ewen 2011). There exists a wide variety of performance measures; this study utilises four common measures, including Root Mean Squared Error (RMSE), Nash-Sutcliffe Efficiency (NSE), Mean Absolute Error (MAE) and the Persistence Index (PI). The PI is similar to the NSE, but instead of the mean observed value, it uses the observed value lagged by the model's lead time (Kitanidis and Bras 1980, De Vos and Rientjes 2005, Abrahart et al. 2008). These error measures are chosen from three distinct error measure taxonomies used for neural network assessment, being based on square residuals (RMSE and NSE), absolute residuals (MAE), and timing (PI) (Maier et al. 2010). Also, while NSE and RMSE are very similar, they are both included in this study due to their widespread use (Ewen 2011).

2.3 Input variable selection

The four IVS methods are described in detail in this section. All the IVS methods are applied using the calibration (testing and validation) partition and the ANN performance is assessed based on the test partition, which is completely independent from the calibration procedure. Firstly, each IVS method is used to identify useful inputs until a termination criterion is met, which is typical for IVS algorithms. However, using a non-optimum termination criterion may cause an over- or under-selection of input variables and subsequently, poor model performance. The poor performance may be attributable to the severity of the termination criterion and not the capability of the IVS method itself in identifying useful inputs.

In order to address the risk of non-optimum termination criteria, each method is used to select the most relevant inputs (which we have selected as 3 and 6 inputs) without regard for any termination criteria. This will showcase the capability of each IVS method to identify the most useful inputs and allows for a direct comparison between input reduced models that have the same number of input parameters.

2.3.1 Partial Correlation

The first IVS method examined in this research is Partial Correlation (PC), which is based on linear correlation between inputs. Unlike most linear correlation based methods, the PC criteria reduces the likelihood of linear redundancy between input variable methods (May et al. 2011). PC is typically used in a forward-selection algorithm, such that the calculation of partial correlation at any step depends on the previously selected inputs (Sharma 2000, May et al. 2008b, He et al. 2011). This approach is criticized by May et al. (2008), who explains that linear selection methods are unsuitable for non-linear models such as ANNs.

This method uses forward selection algorithm where inputs are systematically moved from a set of candidates to a set of selected inputs. At each step of the algorithm, the correlation between each remaining candidate input and the output is calculated, controlling for the effects of the set of inputs that have already been selected. This is known as the partial correlation and is calculated as the correlation between the residuals of the least squares estimates for the output and candidate inputs, based on the selected inputs. The partial correlation values are squared such that both positive and negative correlations are equally indicators of usefulness.

2.3.2 Partial Mutual Information

The second IVS method examined in this study is Partial Mutual Information (PMI). The PC algorithm is conceptually similar to PC, but uses partial mutual information instead of partial correlation as the selection criterion for the forward selection algorithm. This IVS method has been widely used and refined in many studies for identifying useful inputs for environmental models (Bowden et al. 2005b, 2005a, He et al. 2011). Notably, May evaluated several different termination criteria for PMI's forward selection algorithm, and identified an AIC-based termination criterion as a simple and effective stopping criterion, which is the termination criterion used in this study (May et al. 2008b, 2008a).

This selection algorithm is conceptually very similar to the PC-based method described above but is distinguished by its capability to consider non-linear relationships between variables. Instead of least squares regression, non-parametric kernel regression is used to calculate residuals. The partial mutual information is then calculated based on the probability density estimates of the residuals.

2.3.3 Input Omission

Input Omission (IO) identifies useful and non-useful inputs based on the significance (or insignificance) of the error caused by the omission of the input being considered (Setiono and Liu 1997). Setiono and Liu (1997) use input omission (referred to as feature selector in their paper) with retraining at each step to inform an IVS forward selection algorithm, which measured the change in classification accuracy due to input omission as a selection criterion. Next, Abrahart et al. (2001) use input omission (referred to as saliency analysis in their paper) to explain ANN behaviour and identify important inputs; timeseries plots are generated from IO to infer the effect of each input on the model output, for example how omitting precipitation impacts the rising limb of the modelled hydrographs. This analysis provided a defense against the 'black box' criticism against ANNs by demonstrating that omitting highly relevant input parameters creates a physically explainable impact on the model output. It is likely that the concept of IO has been used on a wider scale for IVS, however due to inconsistencies with the method name, it is difficult to find such cases.

The method proposed in this research for IO is adapted from Abrahart et al.(2001); however, in which IO is applied without retraining. Similar to Setiono and Liu (1997), input usefulness is determined based on the performance of the calibration set for the modified ANN. AIC is used as the performance function for thi method, because it permits for a small increase in error to be negated due to a slightly lower model complexity.

2.3.4 Combined Neural Pathway Strength

Finally, Combined Neural Pathway Strength (CNPS) is a model-based approach that relies on an approximation of the combined neural pathway strength of each input to estimate its usefulness. Generally speaking, the strength is defined as the absolute magnitude of the weights associated with each input: the higher the magnitude the stronger or more relevant a particular input is in predicting the desired output. Nash et al. (1997) provide an early example of this method, where the neural pathway strength was calculated as the relative, absolute, matrix multiplication between the first and second sets of weights of a 3-layer ANN. The CNPS of each input was expressed as a percentage of the overall strength of all inputs, giving the relative strength of each input compared to the others.

More recently, Duncan (2014) calculated the CNPS simply as the matrix product of the weight layers, and used the sign of the combined pathway strength values to indicate whether the input has an excitatory (positive correlation with output) or inhibitory (negatively correlated with output) effect on the output. An ensemble of ANNs is then used to estimate the consistency of the CNPS values throughout the ensemble. Using an ensemble of models allow for methods such multi-start and KFCV, which is described in section 0, to be used to estimate uncertainty associated with the ANN configuration. Duncan (2014) demonstrated the CNPS approach by estimating input usefulness as a function of its consistency (whether excitatory or inhibitory) and the variance of the combined neural pathway magnitudes amongst ensemble members. While this approach has limitations, such as the approximate neural strength ignoring the effects of the activation functions and biases within the ANN model, it has demonstrated strong performance for identifying useful inputs (Duncan 2014, Khan et al. 2018, Laureano-Rosario et al. 2018). The CNPS method used in this research is a modified version of the method proposed by Duncan (2014), in which greater emphasis is placed on consistent behaviour or inputs (excitatory or inhibitory) compared to the variance of the combined neural strength values. The CNPS values are calculated as the matrix multiplication of the input-hidden and hidden-output weight matrices.

3 RESULTS AND DISCUSSION

Firstly, ANN models were trained using all inputs, and IVS-reduced inputs (both using the termination criteria and a fixed number of inputs) to predict the stage at the target station, to serve as a comparison point for the input reduced models. A comparison of model performance using the four performance metrics:

RMSE, NSE, MAE and PI, is summarised in Figure 3 for the Bow River 1-day lead time. While overfitting is not a concern, performance measures were calculated based on the test dataset, which is completely independent from the data used for calibration. In addition, the results are for all ensemble models (including the cross-validation and multi-start scenarios) to quantify the uncertainty of the predictions. Thus, the figures show the 25th and 75th percentile values (blue boxes), the median value (red line), and outliers (red crosses) of the predictions rather than deterministic results. These figures show the performance of models ensembles with different input combinations, as determined by the four IVS methods. On the horizontal axis of each subplot, the left most values are from the base model (which includes all the candidate inputs), followed by each of the four IVS methods, first using the termination criteria, followed by the fixed number of inputs.

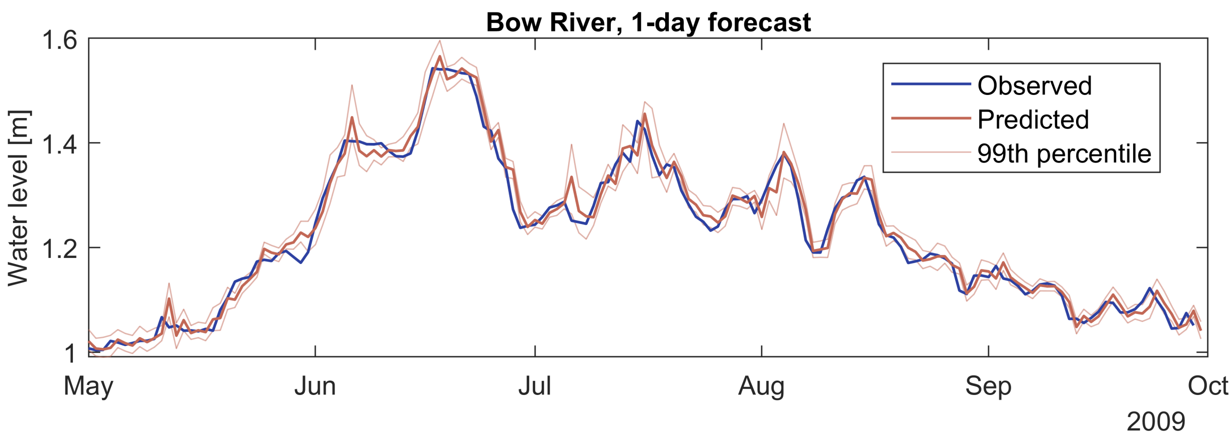

Figure 2: Timeseries plots of observed and predicted (using all candidate inputs) water levels for the Bow River, for a section of the test dataset; note that the predicted values include the model ensemble median and 99th percentile range

Overall, there results demonstrate that it is possible to achieve high performance of flow forecasting models using an ANN approach. However, an analysis of the impact of selecting inputs (based on different IVS methods) can help further refine these models, which are detailed in the following sections.

3.1 Model performance with all candidate inputs

The predicted water level for a 1-day lead time is shown in Figure 2, along with the 99% confidence bands owed to the multi-start and KFCV. Figure 3 highlights the effectiveness of using an ANN approach for flood forecasting: the error metrics indicate high performance for each metric. The RMSE and MAE are low (roughly 1% of the observed data), where the NSE is high, with median values about 0.9 for the Bow. Similarly, the median PI values are positive indicating that the predicted values are an improvement over the last known flow values. As the lead time increases, the models exhibit poorer performance, as expected, yet is still within an acceptable range. These models are improved using IVS methods and discussed in detail in the following section. As discussed in section 1.1, reducing the number of model inputs can improve convergence during training (as illustrated by the reduced ensemble variance in Figure 3), and lower data requirements and complexity of the models.

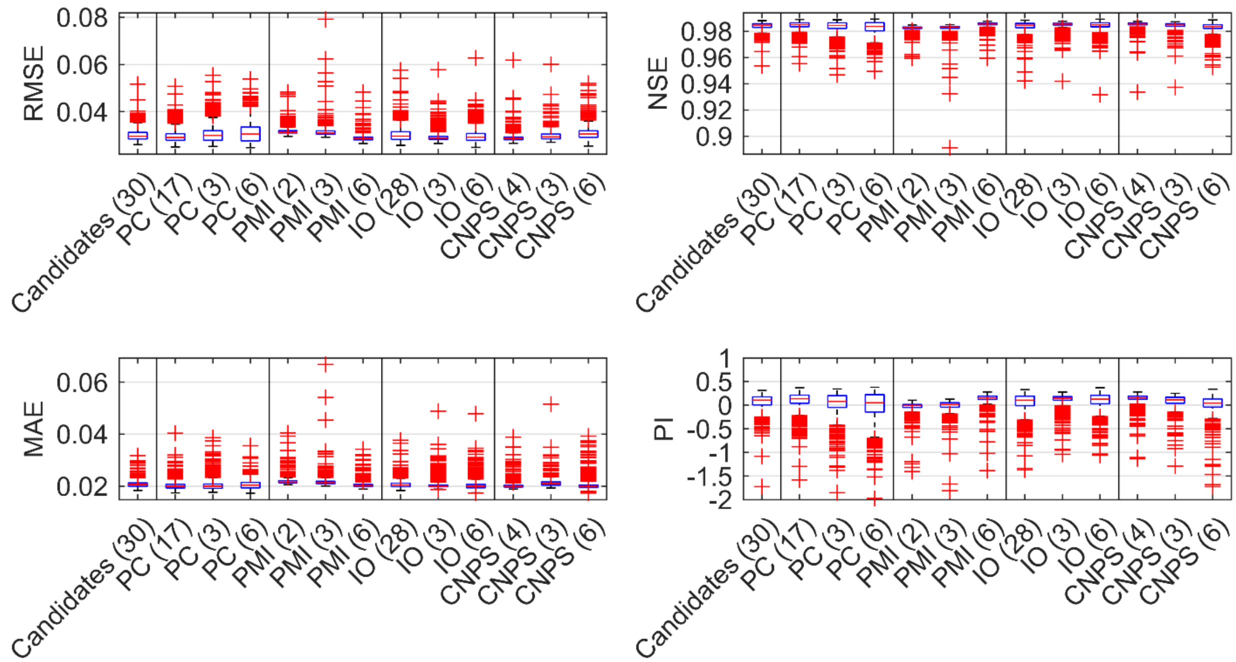

Figure 3: Comparison of ANN model performance (RMSE, NSE, MAE, and PI) for the Bow River for the 1-day lead time for models that use all candidate inputs (30), termination criteria based inputs (variable), 10% of all inputs (3), and 20% of all inputs (6) for each of the 4 IVS methods

3.2 Model performance with IVS

The majority of IVS models exhibit stronger performance compared to the base models. For some input sets, the median performance improves slightly and the variance within the model ensembles decreases, as is expected for models with fewer non-useful inputs. In some instances, the IVS models exhibit a decrease in performance when too few inputs are selected, as is the case for the PMI models where the termination criterion produces a model with only two inputs. For IVS methods where a termination criterion is used, models typically perform reasonably well across the different IVS methods; however, the number of inputs varies drastically between methods. From these models, it is evident that strong performance is possible with as little as 2 input parameters. Likewise, if the termination criterion is too lenient, such is the case for IO, the model performance is not dissimilar to the base model where the full candidate set is used.

Next, each IVS method was applied to select a fixed number (3 and 6) inputs, which allows for comparison of the capability of each IVS method for selecting useful inputs. Furthermore, for IVS methods where the termination criterion is very lenient (PC and IO), constraining the input permits for usefulness of the first 3 or 6 input selections to be assessed. PC exhibits a decrease in performance when the number of selected inputs is constrained, which suggests that the method is poor for ranking useful inputs. Next, PMI exhibits stronger performance, indicating that the method is capable of ranking useful inputs, yet is constrained by a string termination criterion. IO exhibits improved performance when the inputs are constrained, suggesting strong capability for ranking useful inputs and a termination criterion that is much too lenient. Lastly, CNPS exhibits a minor decrease in performance when the number of inputs is fewer or greater than the number determined by termination, indicating that the termination criterion processes the best performing models. Overall, models with 3-6 inputs tend to exhibit the strongest performance.

4 CONCLUSIONS AND RECOMMENDATIONS

This study evaluated four different IVS methods for ANN models for the Bow River. PC demonstrated reasonable performance for both watersheds, however the termination criteria over-selects inputs its selection is outperformed by other IVS methods in most instances. PMI suffered early-stopping, as the termination-based selection was too strict. PMI also favoured autoregressive inputs, which resulted in poor model performance, most notably in the Don River models. It is possible that modifications to PMI such as using a non-Gaussian kernel, or a scaling factor, may yield improvements in input selection. IO demonstrated reasonably strong performance, however the termination criterion used in this study is not recommended, as it was too lenient and inconsistent. IO may be improved by making changes to the method, by exploring topics such as retraining after omission, multiple input omission, or utilizing difference performance criteria. CNPS demonstrated the strongest and most consistent performance amongst the IVS methods evaluated. Next, this study demonstrated that the use of a termination criterion to is not consistent in selecting the optimum number of inputs. Instead, it is recommended that IVS be used to rank inputs, after which the number of inputs is determined on a case by case basis, based on a systematic evaluation of model performance. Future research topics may include the coupled optimization of ANN model inputs and hidden layer size and further refinements to CNPS-based IVS methods.

REFERENCES

Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Shamseldin, A.Y., Solomatine, D.P., Toth, E. and Wilby, R.L. 2012. Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography*, **36**(4): 480–513.

Abrahart, R.J., See, L. and Kneale, P.E. 2001. Investigating the role of saliency analysis with a neural network rainfall-runoff model. *Computers and Geosciences*, **27**(8): 921–928.

Abrahart, R.J., See, L.M. and Solomatine, D.P. 2008. *Practical Hydroinformatics*. Available from www.springer.com/series/6689.

ASCE. 2000a. Artificial Neural Networks in Hydrology. II: Hydrologic Applications. *Journal of Hydrologic Engineering*, **5**(April): 124–137. doi:10.5121/ijsc.2012.3203.

ASCE. 2000b. Artificial Neural Networks in Hydrology. II: Hydrological Applications. J*ournal of Hydrologic Engineering*. doi:10.5121/ijsc.2012.3203.

Bowden, G.J., Dandy, G.C. and Maier, H.R. 2005a. Input determination for neural network models in water resources applications. Part 1 - Background and methodology. *Journal of Hydrology*, **301**(1–4): 75–92.

Bowden, G.J., Maier, H.R., and Dandy, G.C. 2005b. Input determination for neural network models in water resources applications. Part 2. Case study: Forecasting salinity in a river. *Journal of Hydrology*, **301**(1–4): 93–107.

Dawson, C.W.W., and Wilby, R.L.L. 2001. Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, **25**(1): 80–108.

Duncan, A. 2014. *The Analysis and Application of Artificial Neural Networks for Early Warning Systems in Hydrology and the Environmen*t. University of Exeter,. Available from http://files/78/Duncan_2014_The Analysis and Application of Artificial Neural Networks for Early Warning.pdf.

Ewen, J. 2011. Hydrograph matching method for measuring model performance. *Journal of Hydrology*, **408**(1–2): 178–187.

He, J., Valeo, C., Chu, A. and Neumann, N.F. 2011. Prediction of event-based stormwater runoff quantity and quality by ANNs developed using PMI-based input selection. *Journal of Hydrology*, **400**(1–2): 10–23.

Khan, U.T., He, J. and Valeo, C. 2018. River flood prediction using fuzzy neural networks: an investigation on automated network architecture. Water Science and Technology, **2017**(1): 238–247. doi:10.2166/wst.2018.107.

Khan, U.T. and 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.

Kitanidis, P.K. and Bras, R.L. 1980. Real-time forecasting with a conceptual hydrologic model: 2. Applications and results. *Water Resources Research*, **16**(6): 1034–1044.

Laureano-Rosario, A., Duncan, A., Mendez-Lazaro, P., Garcia-Rejon, J., Gomez-Carro, S., Farfan-Ale, J., Savic, D. and Muller-Karger, F. 2018. Application of artificial neural networks for Dengue Fever outbreak predictions in the Northwest coast of Yucatan, Mexico and San Juan, Puerto Rico. *Tropical Medicine and Infectious Disease*, **3**(1): 5.

Maier, H.R. and Dandy, G.C. 2000. Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environmental Modelling and Software*, **15**(1):

101–124.
Maier, H.R., Jain, A., Dandy, G.C. and Sudheer, K.P. 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling and Software*, **25**(8): 891–909.
May, R., Dandy, G. and Maier, H. 2011. *Review of Input Variable Selection Methods for Artificial Neural Networks. In Artificial Neural Networks - Methodological Advances and Biomedical Applications*. InTech. doi:10.5772/16004.
May, R.J., Dandy, G.C., Maier, H.R. and Nixon, J.B. 2008a. Application of partial mutual information variable selection to ANN forecasting of water quality in water distribution systems. *Environmental Modelling and Software*, **23**(10–11): 1289–1299.
May, R.J., Maier, H.R., Dandy, G.C. and Fernando, T.M.K.G. 2008b. Non-linear variable selection for artificial neural networks using partial mutual information. *Environmental Modelling and Software*, **23**(10–11): 1312–1326.
Nanda, T., Sahoo, B., Beria, H. and Chatterjee, C. 2016. A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products. *Journal of Hydrology*, **539**: 57–73.
Nath, R., Rajagopalan, B. and Ryker, R. 1997. Determining the Saliency of Input Variables in Neural Network Classifiers. *Computers & Operations Research*, **24**(8): 767–773.
Setiono, R. and Liu, H. 1997. Neural Network Feature Selector. *IEEE Transactions on Neural Networks,* **8**(3): 654-662.
Sharma, A. 2000. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1 - A strategy for system predictor identification. *Journal of Hydrology*, **239**(1–4): 232–239.
Shrestha, R.R. and Nestmann, F. 2009. Physically based and data-driven models and propagation of input uncertainties in river flood prediction. *Journal of Hydrologic Engineering*, **14**(12): 1309–1319.
Shrubsole, D., Kreutzwiser, R., Mitchell, B., Dickinson, T. and Joy, D. 1993. The history of flood damages in ontario. *Canadian Water Resources Journal*, **18**(2): 133–143.
Šindelář, R. and Babuška, R. 2004. Input selection for nonlinear regression models. *IEEE Transactions on Fuzzy Systems*, **12**(5): 688–696.
Solomatine, D.P. and Ostfeld, A. 2008. Data-driven modelling: some past experiences and new approaches. *Journal of Hydroinformatics*, **10**(1): 3. doi:10.2166/hydro.2008.015.
De Vos, N.J. and Rientjes, T.H.M. 2005. Constraints of ANNs for rainfall-runoff modelling Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation Constraints of ANNs for rainfall-runoff modelling. *HESSD Earth Syst. Sci. Discuss*, **2**(2): 365–415. doi:https://doi.org/10.5194/hess-9-111-2005.
Wijesekara, G.N., Gupta, A., Valeo, C., Hasbani, J.G., Qiao, Y., Delaney, P. and 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**–**413**: 220–232. doi:10.1016/j.jhydrol.2011.04.018.
Wilby, R. L. 2006. A Review of Climate Change. *Built Environment*, **33**(1): 31–45.
Yin, X., Zhang, J. and Wang, X. 2004. Sequential injection analysis system for the determination of arsenic by hydride generation atomic absorption spectrometry. *In Fenxi Huaxue*. Springer, Berlin. doi:10.1017/CBO9781107415324.004.