

Received 25.10.2012  
Reviewed 17.12.2012  
Accepted 20.12.2012

A – study design  
B – data collection  
C – statistical analysis  
D – data interpretation  
E – manuscript preparation  
F – literature search

## Comparison of machine learning methods for runoff forecasting in mountainous watersheds with limited data

Jan ADAMOWSKI<sup>1) ABCDEF</sup>, Shiv O. PRASHER<sup>2) ABCDEF</sup>

<sup>1)</sup> Department of Bioresource Engineering, McGill University, 21 111 Lakeshore Road, Ste. Anne de Bellevue, QC, Canada, H9X 3V9 Tel: 514 398 7786, e-mail: [jan.adamowski@mcgill.ca](mailto:jan.adamowski@mcgill.ca) Fax: 514-398-83873

<sup>2)</sup> Department of Bioresource Engineering, McGill University, 21 111 Lakeshore Road, Ste. Anne de Bellevue, QC, Canada, H9X 3V9; e-mail: [shiv.prasher@mcgill.ca](mailto:shiv.prasher@mcgill.ca) Fax: 514-398-8387

**For citation:** Adamowski J., Prasher S.O. 2012. Comparison of machine learning methods for runoff forecasting in mountainous watersheds with limited data. *Journal of Water and Land Development*. No. 17 p. 89–97

### Abstract

Runoff forecasting in mountainous regions with processed based models is often difficult and inaccurate due to the complexity of the rainfall-runoff relationships and difficulties involved in obtaining the required data. Machine learning models offer an alternative for runoff forecasting in these regions. This paper explores and compares two machine learning methods, support vector regression (SVR) and wavelet networks (WN) for daily runoff forecasting in the mountainous Sianji watershed located in the Himalayan region of India. The models were based on runoff, antecedent precipitation index, rainfall, and day of the year data collected over the three year period from July 1, 2001 and June 30, 2004. It was found that both the methods provided accurate results, with the best WN model slightly outperforming the best SVR model in accuracy. Both the WN and SVR methods should be tested in other mountainous watershed with limited data to further assess their suitability in forecasting.

**Key words:** *Himalayas, machine learning, runoff forecasting, support vector regression, wavelet networks*

### INTRODUCTION

Accurate runoff forecasting allows for effective planning and use of water resources. However, in mountainous regions the complexity of rainfall-runoff relationships and difficulty in obtaining accurate data for many of the required parameters of process or physically based runoff forecasting models limits the use and applicability of such models. Data driven hydrological models may be an alternative to process based models for runoff in such regions due to their minimal information requirements and relatively fast development and processing times. Previously, data driven models have been statistically based with autoregressive moving average (ARMA) and multiple linear regression (MLR) models being two of the

most popular techniques. More recently, machine learning type models such as artificial neural networks (ANN) and support vector regression (SVR) have been applied to hydrological forecasting.

Comparisons of ANN models with conceptual models have been undertaken for runoff forecasting by a number of researchers, with promising results. KANG *et al.* [1993] used both ANN and ARMA models to predict flows on a daily and hourly basis, reporting that ANNs could be successfully applied for flow forecasting. ZEALAND *et al.* [1998] compared ANNs with conventional methods for a portion of the Winnipeg River system, finding the ANN model outperformed the conventional models. More recently, CASTELLANO-MENDEZ *et al.* [2004] compared ANN and Box-Jenkins models for flow forecasting in

Spain, and found that the ANN models produced superior results. NILSSON *et al.* [2005] compared ANN with conceptual runoff modeling, reporting superior results for the ANN models. Superior forecasting by ANN models over conceptual models was also reported by PANDA *et al.* [2009] in a comparison between ANN and MIKE 11 models for simulation of river levels. Supporting these results, SHAMSELDIN [2010] developed ANN models based on different input parameters and concluded that ANN models show good potential and promise to be used for flow forecasting in developing countries. Also apparent from the review of this literature is that the multi-layer perceptron model optimized with back propagation has become the most popular ANN model used in flow forecasting.

Coupled wavelet networks (WN) are a hybrid version of the ANN model in which the input data is decomposed via wavelet analysis with the outputs of the analysis forming the inputs to the ANN model. WN for runoff forecasting has been explored in several studies. CANNAS *et al.* [2006] investigated WN models with continuous and discrete wavelet transforms for a Mediterranean watershed with a large data-set, reporting good predictive results from the WN models over those without wavelet pre-processing. PARTAL and KISI [2008] and KISI [2009] explored WN for flow forecasting in Turkey, finding the wavelet transform model performing significantly better than other models. WU *et al.* [2009] developed WN models for predicting flow one, two and three days ahead, finding the WN performance superior to the ANN models. SANG *et al.* [2009] compared ANNs with WN models for hydrological forecasting and found that WN models performed better for complex hydrologic series and long forecast periods. ADAMOWSKI and SUN [2010] developed WN models for short-term flow forecasting in Cyprus, finding improved forecasting abilities of the WN models over ANN models. And finally, KRISHNA *et al.* [2011] applied WN models to the Malaprabha River basin in India, and found that they provided better forecasting for daily river flow series, and in particular for peak flows in comparison to the ANN models.

Support vector regression (SVR) is an alternate machine learning method to ANNs that have recently been applied to flow forecasting situations. ASEFA *et al.* [2005] investigated the application of SVR models for stream flow at the seasonal and hourly timescales, reporting good results for these applications. BEHZAD *et al.* [2008] investigated the application of SVR models in predicting stream flow one day ahead and compared the results to ANN and ANN-GA (ANN genetic algorithm) models. They reported that the SVR models produced results at least as good as those predicted by the ANN and ANN-GA models. WANG *et al.* [2009] compared the ability of SVR models in predicting long term monthly river flow discharges

with three other AI models, reporting strong predictive results from the SVR models.

The purpose of this study is to investigate the relative forecasting abilities of SVR and WN methods in total runoff forecasting in a mountainous watershed with limited data. To the best knowledge of the authors, a comparison of the SVR and WN methods for runoff forecasting in mountainous watersheds with limited data has not been reported in the literature. SVR and WN models were applied for daily total runoff forecasting in the Sainji watershed in the Himalayan region of India, based on daily rainfall and runoff data collected by one of the authors from the watershed over the three year period from July 1, 2001 and June 30, 2004. The data for the three watersheds was randomized for training and testing.

## METHODS

### SVR

The SVR model used in this analysis was built on the basis of Vapnik's statistical learning theory. A brief synopsis of the SVR theory used in developing the SVR model is presented here, however, for a more complete derivation of the SVR model the reader is referred to VAPNIK [1995], BURGESS [1998], and CRISTIANINI and SHAWE-TAYLOR [2000].

SVR is based upon non-linear mapping of the original values to a linear function in a higher-dimensional feature space, and optimizing the model through minimization of an error function. The linear function has the form of:

$$y = f(x) = \langle \mathbf{w} \bullet \mathbf{x} \rangle + b \quad (1)$$

where:  $w$  and  $b$  are regression parameters, and  $\langle \mathbf{w} \bullet \mathbf{x} \rangle$  represents the dot product of  $w$  and  $x$ . If  $x_m$  and  $y_m$  are the  $m^{\text{th}}$  input vector and target output,  $m = 1, \dots, M$ , the error function is given by:

$$J = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{m=1}^M |y_m - f(x_m)|_{\epsilon} \quad (2)$$

The optimal regression function can thus be obtained by minimizing this error function [SMOLA, SCHOLKOPF 1998]. The second part of this error function is known as the  $\epsilon$  - insensitive loss function. The  $\epsilon$ -insensitive loss function was proposed by VAPNIK [1995] as a robust loss function to reduce sensitivity to the outliers by focusing on optimizing a bound around the regression function. A SVR regression model based on this function calculates the difference between the predicted and the actual values, and if the differences are less than  $\epsilon$  the regression function is considered to be acceptable [SMOLA, SCHOLKOPF 1998].

Using Lagrangian multipliers, the solution to the optimization problem of equation (2) can be transformed into the so-called dual problem [GUNN 1998; VAPNIK 1995]:

Maximize

$$\sum_{i=1}^M \alpha_i^* (y_i - \varepsilon) - \alpha_i (y_i + \varepsilon) - \frac{1}{2} \sum_{i,j=1}^M (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(\mathbf{x}_i, \mathbf{x}_j) \quad (3)$$

Subject to:

$$\sum_{i=1}^M (\alpha_i - \alpha_i^*) = 0 \text{ and } 0 \leq \alpha_i, \alpha_i^* \leq C$$

where  $\alpha_i$  and  $\alpha_i^*$  are the Lagrange multipliers, and  $K$  a kernel. The introduction of the kernel allows the processing of non-linear functions. The commonly-used kernels are the radial basis function (RBF) kernels, sigmoid kernels, and polynomial kernels [CHANG, LIN 2001; GUNN 1998]. The RBF kernel, most commonly used in SVR approaches, is defined as follows:

$$K(x, y) = e^{-\gamma(x-y)^2} \quad (4)$$

where  $\gamma$  is a kernel parameter. The solution of this problem will yield  $\alpha_i$  and  $\alpha_i^*$  for all  $i = 1, 2, \dots, M$ . For all the training points within the  $\varepsilon$  - sensitive zone will yield  $\alpha_i$  and  $\alpha_i^*$  equal to zero.

To build the SVR model, it is first trained using a portion of the data set (e.g. 80%). During training the model is run with different sets of  $\gamma$ ,  $\varepsilon$ , and  $C$  values set by the user, with the optimum values determined by optimizing the error during five-fold cross validation. Once the model has been trained the remainder of the data is used to test the predictive performance of the model.

#### COUPLED WAVELET NETWORKS (WN)

Coupled wavelet network models utilize wavelet analysis to decompose the input data with the wavelet outputs used as inputs into an ANN model. The discussion is therefore split into these components. The WN models used in this study were developed following the procedures outlined by CANNAS *et al.* [2006], PARTAL [2009], and ADAMOWSKI and CHAN [2011].

#### Wavelet analysis

Wavelets provide time-scale representation to analyze time series datasets that contain non-stationarities. Wavelets decompose an original time series into sub-series of different resolution levels [TIWARI,

CHATTERJEE 2010]. This decomposition can highlight the different effects of variables contained within the original data that may not be otherwise apparent. Various wavelet outputs may then be chosen to be combined and used as the ANN input to generate a more accurate process than what may be generated using only raw input data.

The mother wavelet of a signal  $x(t)$  is the transforming function in the wavelet process and may be defined as:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (5)$$

Wavelet components can be generated by breaking down the original data set using either the Continuous Wavelet Transform (CWT) or Discrete Wavelet Transform (DWT). The CWT of a signal  $x(t)$  is defined as follows:

$$CWT_x^\psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} x(t) \psi^* \left( \frac{t-\tau}{s} \right) dt \quad (6)$$

where  $s$  is the scale parameter,  $\tau$  is the translation parameter and ‘\*’ denotes the complex conjugate [CANNAS *et al.* 2006]. However, CWT calculations require a large amount of computation time and resources and as such DWT calculations which require a smaller amount of computational time and resources are the preferred method. DWTs are achieved by modifying the wavelet representation to:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{|s|}} \psi \left( \frac{t - k\tau_0 s_0^j}{s_0^j} \right) \quad (7)$$

where  $j$  and  $k$  are integers and  $s_0 > 1$  is a fixed dilation step [CANNAS *et al.* 2006]. The effect of discretizing the wavelet is that the time-space scale is sampled at discrete intervals rather than in a continuous function. The DWT scales and positions are usually based on powers of two (dyadic scales and positions), as these combinations are usually the simplest and most efficient cases.

In undertaking the DWT, the original time series is passed through high-pass and low-pass filters. This process separates the data into low frequency components showing trends and high frequency components showing detail. Due to this separation, wavelets allow for the analysis of characteristics that may not otherwise be detected.

#### Neural networks

Neural networks are a machine learning method capable of solving complex relationships between input and outputs. They differ from process based

models as they develop relationships by performing input-output mapping without the necessity of understanding the nature of the phenomena being modeled. To do this, neural networks are composed of simple processing elements (nodes) linked by weighted synaptic connections [MULLER, REINHARDT 1991]. They can construct complex relationships through combining multiple simple relationships between many nodes, thus mapping out an overall more complex scenario.

Multi-layer perceptrons (MLP) are the most widely used type of neural network model and the

model chosen for this research. These models consist of neurons and nodes. The neurons are organized in layers, with each neuron only connected with neurons in the adjacent layers. Each node  $j$  receives a weighted input being the output from every node  $i$  in the previous layer. The effective incoming signal propagates forward through a nonlinear activation function, towards the nodes in the next layer, thus building a nonlinear relationship connecting the inputs and outputs. Figure 1 provides a visual representation of this process.

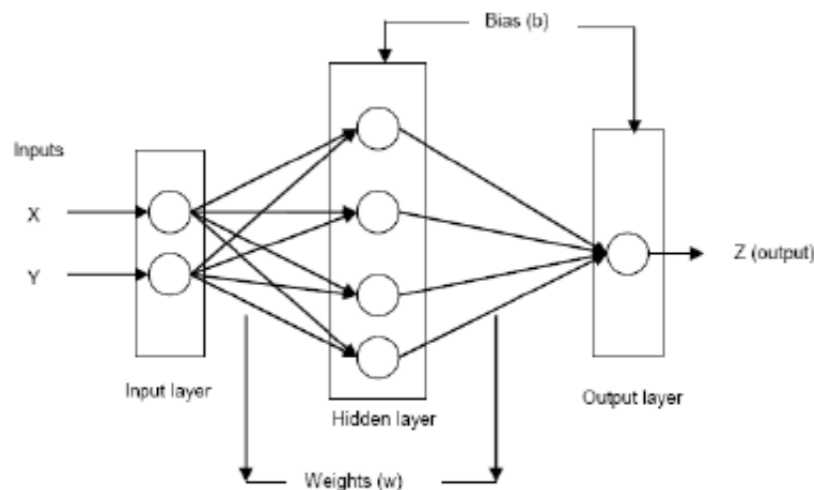


Fig. 1. ANN architecture with one hidden layer [PRAMANIK, PANDA 2009]

The mathematical form of a three-layer feed forward ANN is given as:

$$O_k = g_2 \left[ \sum_j w_{kj} g_1 \left( \sum_i w_{ji} I_i + w_{jo} \right) + w_{ko} \right] \quad (8)$$

where  $I_i$  is the input value to node  $i$  of the input layer and  $O_k$  is the output at node  $k$  of the output layer [TIWARI, CHATTERJEE 2010]. The strength of the connection between the input node  $i$  and the hidden node  $j$  is controlled by  $w_{ji}$ , and  $w_{kj}$  controls the strength of the connection between the hidden node  $j$  and the output node  $k$ .  $g_1$  and  $g_2$  are activation functions for the hidden layer and the output layer, respectively.

The Levenberg-Marquart (LM) algorithm was utilized to train the ANN models (in MATLAB) since it has been shown to be fast, accurate, and reliable [ADAMOWSKI, KARAPATAKI 2010; PRAMANIK, PANDA 2009].

#### MODEL PERFORMANCE COMPARISON

The performance of the models were evaluated by analyzing the simulated results against the observed data and using three commonly used statistical

tools, the coefficient of determination ( $R^2$ ), the root mean squared error (RMSE), and modeling efficiency (EF).  $R^2$  indicates the degree of correlation between the simulated and observed values, and comparisons of  $R^2$  values may indicate the relative ability of different models to accurately simulate observed data.  $R^2$  values range from zero to one, with one indicating a perfect fit and zero indicating no correlation.

The root mean squared error represents the mean error between the simulated values and the observed values. A small RMSE indicates less mean error between the simulated and observed results than a large RMSE. However, RMSE can be very sensitive to extreme values with a single large error result capable of resulting in a high RMSE. If the data is strongly biased, the modeling efficiency (EF) gives a more acceptable analysis than  $R^2$ . A negative value for EF indicates very poor model performance.

#### STUDY AREAS AND DATA

The study watershed is located between  $78^{\circ}0'24''E$  and  $78^{\circ}1'21''E$  longitudes and  $30^{\circ}28'58''N$  to  $30^{\circ}30'42''N$  latitudes in the Uttaranchal State, India [SHARDA *et al.* 2006]. This watershed is known as the Sainji watershed due to the situation of the village of

Sainji within it. The main watershed (WS1) includes two sub-watersheds (WS2 and WS3) having third order and second order streams, respectively, and areas of 255, 52 and 163 ha respectively (Fig. 2). All three watersheds have slopes of between 62 and 66%. WS1 is predominantly a mixed forest and scrub watershed, WS2 predominantly agriculture and scrub forest, and WS3 predominantly dense and mixed forest (Fig. 2).

Runoff, antecedent precipitation index, rainfall and day of the year data were obtained on a daily basis for a three year period between July 1, 2001 and June 30, 2004. The data from the watershed was obtained specifically for this study by one of the authors while working for the Central Soil and Water Conservation Research and Training Institute in Dehradun, Uttarakhand.

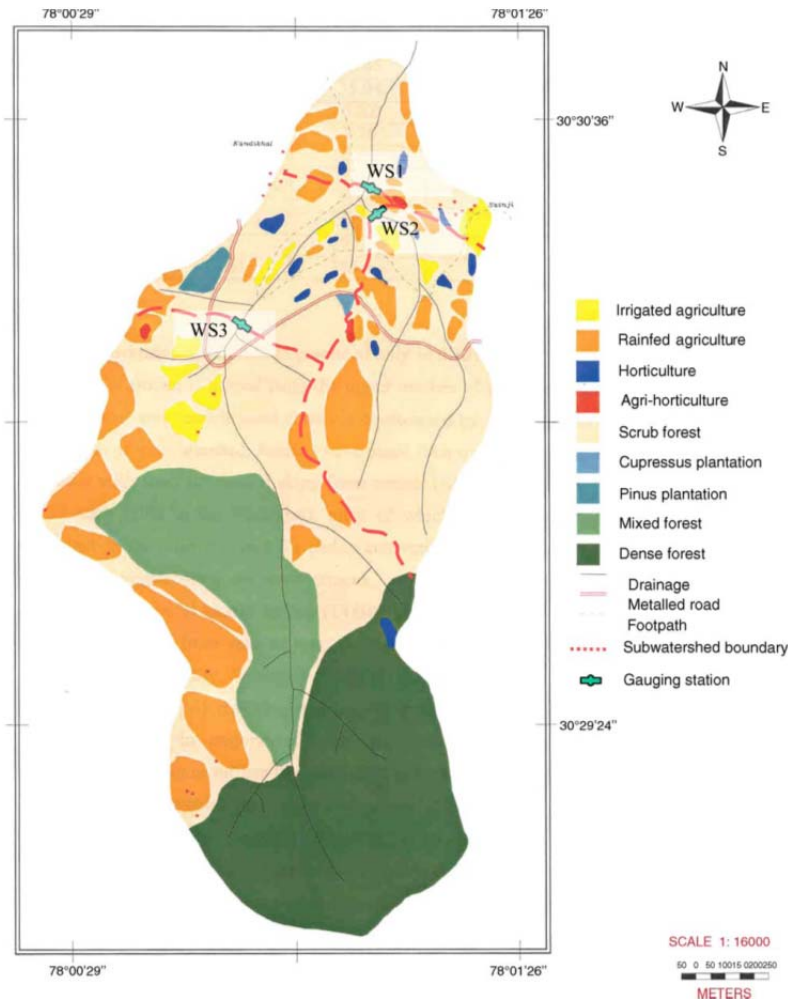


Fig. 2. Land use map of Sainji watershed and location of gauging stations in three sub-watersheds [SHARDA *et al.* 2006]

## MODEL DEVELOPMENT

### SVR MODELS

The data set used for training the SVR models consisted of many input variables and one output. The input variables consisted of: day of the year, rainfall, Antecedent Precipitation Index (API<sub>5</sub>), and runoff. The output (dependent) variable was total runoff. Input variables from the current day to five days before were explored for use in the models.

A five-fold cross validation procedure was applied to check the generalization ability of the model. In this procedure, the data collected over the three

watersheds was randomized and divided into five equal parts, generating five datasets that did not represent a single watershed but rather the characteristics of all three. The models were trained using four parts of the data (80%), and tested with the remaining “unseen” fifth part (20%) to evaluate model performance. The procedure was repeated for all five possible combinations. This procedure facilitated the development of generalizing models, that if proven accurate could be used in predicting flow for ungauged watersheds of similar geographical and climatic characteristics where past rainfall/runoff are not available. This generalizing ability was based upon the work by SHARDA *et al.* [2006] in determining the most important water-

shed features that affect the rainfall-runoff relationships (curve number, rainfall, antecedent moisture condition and day of the year).

**COUPLED WAVELET NETWORK MODELS**

The original data (day of the year, rainfall, Antecedent Precipitation Index, and runoff) was decomposed using a modified version of the 'a trous' wavelet algorithm (to ensure that future data is not used). The original time series for each variable was first decomposed into an approximation and accompanying detail signal. The decomposition process was then iterated with successive approximation signals being decomposed in turn, and in this way the original time series was broken down into lower resolution components. For this study four wavelet levels were selected for each variable. A new series was obtained for each variable by adding the details and approximate series for a specific variable, and these series were then used as inputs to the ANN models. Input variables from the current day to five days before were explored for use in the models.

The ANN networks that were developed consisted of an input layer, a hidden layer and one output layer. Each model was tested to determine the optimum number of neurons in the hidden layer following a trial and error procedure. Following the wavelet decomposition and development of the ANN model, the WN models were trained and tested using a 5 fold cross-validation consistent with the method used in developing the SVR models. In this process the same data in each fold that was used for the SVR model was again used to train and test the WN model, with the same 80/20 split for training and testing data.

**RESULTS AND DISCUSSION**

**SVR MODELS**

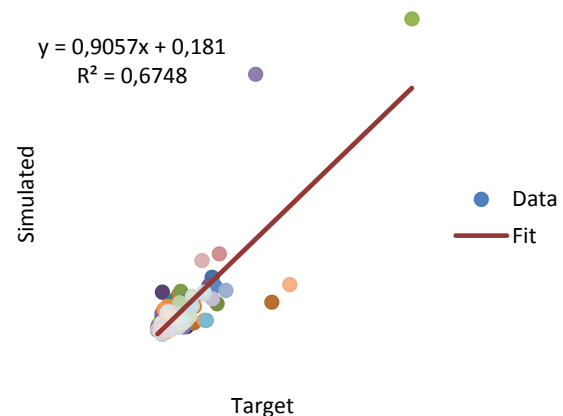
From Table 1, it can be seen that the best SVR model's coefficient of determination ( $R^2$ ) for testing

datasets was 0.88, with a *RMSE* value returned of 0.89 and *EF* value of 0.88 (fold 1).

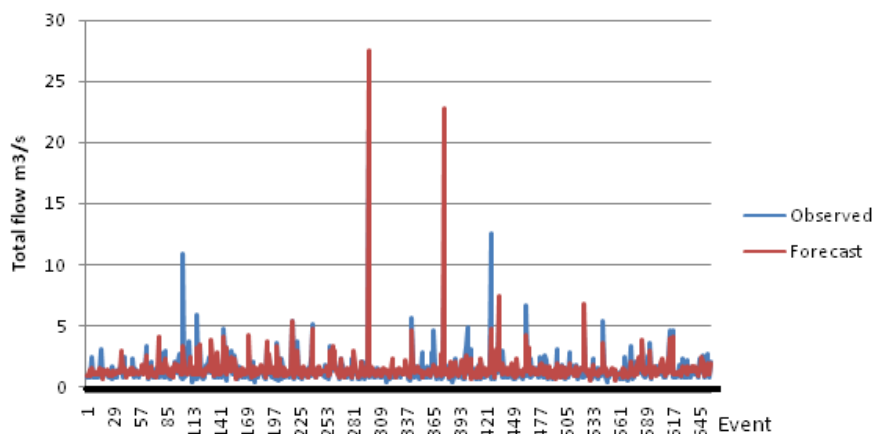
**Table 1.** SVR model results over all 5 folds

Fold	1	2	3	4	5
<b>Training</b>					
$R^2$	0.86	0.86	0.93	0.86	0.84
<i>RSME</i>	0.69	0.78	0.54	0.63	0.84
<i>EF</i>	0.84	0.85	0.93	0.86	0.83
<b>Testing</b>					
$R^2$	0.88	0.70	0.68	0.85	0.77
<i>RSME</i>	0.89	0.74	0.86	1.24	0.87
<i>EF</i>	0.88	0.69	0.59	0.78	0.55

These results indicate that the model was able to learn the relationship between the inputs and the total runoff, and also able to apply it successfully to unseen datasets. The large *RSME* value may be explained by the significant underestimation of a number of the large flow events in the data – significantly increasing the *RSME* value. Figures 3 and 4 show the measured and predicted values, with the significant overestimation of a 9 m<sup>3</sup>·s<sup>-1</sup> flow event apparent.



**Fig. 3.** Scatterplots comparing observed and forecasted flows using the best SVR model (testing)



**Fig. 4.** Best SVR model: Comparison of observed and forecasted flow for each event (testing)

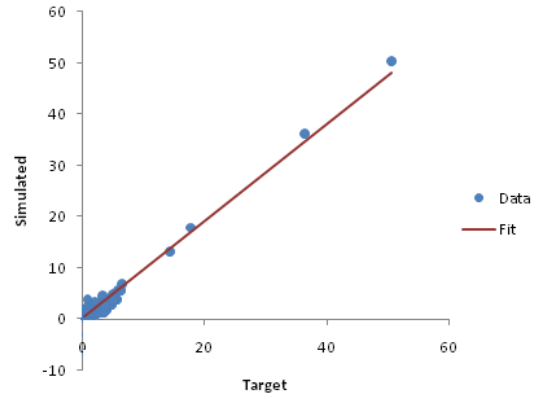
**WN MODELS**

The best WN model is a function of rainfall from the current day, and the two days preceding the current day, the runoff number from the previous day, and the antecedent precipitation from the previous day ( $R_t, R_{t-1}, R_{t-2}, Run_{t-1}, A_{t-1}$ ). This WN model had six neurons in the hidden layer. This WN model had a testing coefficient of determination ( $R^2$ ) of 0.93 (Tab. 2), a *RMSE* value for the testing period of 0.69 and an *EF* of 0.93 (fold 3).

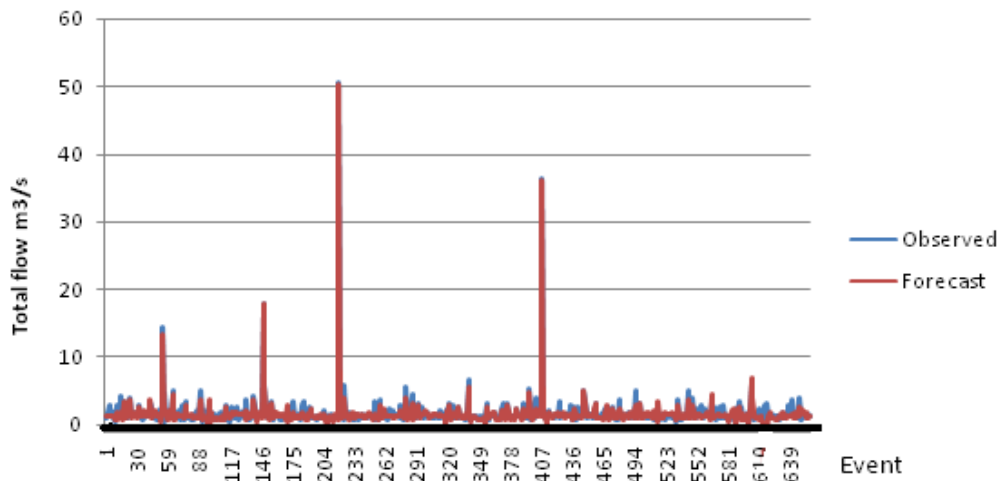
**Table 2.** WN model results over all 5 folds

Fold	1	2	3	4	5
<b>Training</b>					
$R^2$	0.81	0.88	0.85	0.82	0.81
<i>RSME</i>	0.76	0.72	0.79	0.73	0.90
<i>EF</i>	0.81	0.87	0.85	0.82	0.81
<b>Testing</b>					
$R^2$	0.91	0.70	0.81	0.93	0.83
<i>RSME</i>	0.75	0.73	0.58	0.69	0.53
<i>EF</i>	0.91	0.69	0.81	0.93	0.83

Graphical representations of the measured and predicted values are shown in Figures 5 and 6. These results indicate that while the model was able to learn and predict most of the situations accurately, it was not able to accurately predict a number of the large flow events. This can be seen in Figure 6, which shows the simulated and observed total runoff in the Sainji watershed. A number of underestimations by the model can be seen for the large flow events.



**Fig. 5.** Scatterplots comparing observed and forecasted flows using the best WN model (testing)



**Fig. 6.** Best WN model: Comparison of observed and forecasted flow for each event (testing)

**SVR VERSUS WN MODELS**

The best WN model had a testing  $R^2$  of 0.93 compared to the best SVR model that had a testing  $R^2$  of 0.88 (Tab. 3). It can be seen that both models returned promising  $R^2$  values, with the WN model outperforming the SVR model.

**Table 3.** Comparison of the best SVR and best WN testing results

Model	Best SVR statistical results (fold 1)	Best WN statistical results (fold 4)
$R^2$	0.88	0.93
<i>RMSE</i> (mm)	0.89	0.69
<i>EF</i>	0.88	0.93

For *RSME* results, the best WN model had a testing *RSME* of 0.69, compared to the best SVR model which had a testing *RMSE* of 0.89. A lower *RMSE* value indicates the best WN model had smaller differences between total runoff forecasted and the total runoff actually observed from the Sainji watershed, particularly when forecasting large flow events. Similarly, the best WN model outperformed the best SVR model in the *EF* index, recording a value of 0.93 versus 0.88.

Interestingly, the best models did not correspond to being produced over the same data. The best performing WN model was developed with the fold 4 data, while the best performing SVR model was with the fold 1 data. However the WN model pro-

duced with fold 1 data still outperformed the fold 1 SVR model in terms of  $R^2$ ,  $RMSE$  and  $EF$  (Tab. 4).

**Table 4.** Comparison of the fold 1 (best) SVR and fold 1 WN testing results

Model	Best SVR statistical results (fold 1)	Corresponding WN statistical results
$R^2$	0.88	0.91
$RMSE$ (mm)	0.89	0.75
$EF$	0.88	0.91

## CONCLUSIONS

WN and SVR are both promising new methods for runoff forecasting in mountainous regions where limited data is available. In this research the two methods were compared against each other to assess the relative predictive capabilities of each model for prediction of daily total runoff in a Himalayan watershed. The results of the testing indicated that both models can accurately predict daily total runoff in a mountainous watershed. The comparison of the statistical performance results determined that the best WN model outperformed the best SVR model, recording superior statistical results over the  $R^2$ ,  $RMSE$  and  $EF$  results indices. In addition, as the WN model developed with the dataset corresponding to the best SVR model (fold 1) also outperformed the SVR model, it can be established that the WN model was the superior model.

The work undertaken in this research study was based upon a specific watershed and so future research needs to be undertaken in comparing these two machine learning methods for watersheds with different topography and characteristics in other mountainous regions of the world with limited data. Other recommendations for future work include: exploring multistep ahead prediction; exploring different input variables; exploring the use of the SVRs with localized multiple kernel learning; and exploring the use of the bootstrap method coupled with SVR and WN models.

## ACKNOWLEDGEMENTS

This research was partially supported by an NSERC Discovery Grant held by Jan Adamowski.

## NOTATION

The following symbols are used in this paper:

- $CWT$  – continuous wavelet transform  
 $N$  – number of data points used  
 $s$  – scale parameter  
 $x(t)$  – signal  
 $\bar{y}_i$  – mean value taken over  $N$   
 $y_i$  – observed peak weekly water demand  
 $\hat{y}_i$  – forecasted peak weekly water demand  
 $\tau$  – translation parameter

- \* – complex conjugate  
 $\psi(t)$  – mother wavelet

## REFERENCES

- ADAMOWSKI J., CHAN H.F. 2011. A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology*. No 407 p. 28–40.
- ADAMOWSKI J., CHAN H.F., PRASHER S.O., SHARDA V.N. 2011. Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for rainfall-runoff forecasting in Himalayan micro watersheds with limited data. *Journal of Hydroinformatics*. No 3 p. 731–744.
- ADAMOWSKI J., KARAPATAKI C. 2010. Comparison of Multivariate Regression and Artificial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. *Journal of Hydrology*. No 15 p. 729–743.
- 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*. No 390 p. 85–91.
- ASEFA T., KEMBLAWSKI M., MCKEE M., KHALIL A. 2005. Multi-time scale stream flow predictions: The support vector machines approach. *Journal of Hydrology*. No 318 p. 7–16.
- BEHZAD M., ASGHARI K., EAZI M., PALHANG M. 2009. Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Systems with Applications*. No 36 p. 7624–7629.
- CANNAS B., FANNI A., SIAS G., SEE L. 2006. Data preprocessing for river flow forecasting using neural networks: Wavelet transforms and data partitioning. *Physics and Chemistry of the Earth*. No 31 p. 1164–1171.
- CASTELLANO-MENDEZ M., GONZALEZ-MANTEIGA W., FEBRERO-BANDE M., PRADA-SANCHEZ J., LOZANO-CALDERON R. 2004. Modelling of the monthly and daily behavior of the runoff of the Xallas River using Box-Jenkins and neural networks methods. *Journal of Hydrology*. No 296 p. 38–58.
- CHANG C., LIN C. 2001. LIBSVM: a Library for Support Vector Machines [online]. [Access 20.10.2012]. Available at: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- CRISTIANINI N., SHAW-TAYLOR J. 2000. An introduction to support vector machines and other kernel-based learning methods. New York. Cambridge University Press. ISBN 0-521-78019-5 pp. 204.
- GUNN S. 1998. Support vector machines for classification and regression. Technical report. Southampton. University of Southampton. Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science pp.54.
- KANG K.W., KIM J.H., PARK C.Y., HAM K.J. 1993. Evaluation of hydrological forecasting system based on neural network model. In: *Proceedings of the 25th Congress of the International Association for Hydraulic Research*. Delft, The Netherlands p. 257–264.
- KISI O. 2009. Neural networks and wavelet conjunction model for intermittent streamflow forecasting. *Journal of Hydrologic Engineering*. No 14 p. 773–782.
- KRISHNA B., RAO Y., NAYAK P. 2011. Time series modeling of river flow using wavelet neural networks. *Journal of Water Resource and Protection*. No 1 p. 50–59.
- MULLER B., REINHARDT J. 1991. Neural networks: An introduction. Ser. *Physics of Neural Networks*. Vol. 2. New York. Springer-Verlag. ISBN 0387523804 pp. 266.



- NILSSON P., UVO C., BERNTDSSON R. 2005. Monthly runoff simulation: Comparing and combining conceptual and neural and network models. *Journal of Hydrology*. No 321 p. 344–363.
- PANDA R.K., PRAMANIK N., BALA B. 2009. Simulation of river stage using artificial neural network and MIKE 11 hydrodynamic model. *Computers and Geosciences*. No 36 p. 735–745.
- PARTAL T. 2009. River flow forecasting using different artificial neural network algorithms and wavelet transform. *Canadian Journal of Civil Engineering*. No 36 p. 26–38.
- PARTAL T., KISI O. 2008. Wavelet and neuro-fuzzy conjunction model for precipitation forecasting. *Journal of Hydrology*. No 342 p. 199–212.
- PRAMANIK N., PANDA R.K. 2009. Application of neural network and adaptive neuro fuzzy inference systems for stream flow prediction. *Hydrological Science Journal*. No 54 p. 247–260.
- SANG Y., WANG D., WU J. 2009. Comparative study of some improved ANN models for hydrologic time series forecast. *Intelligent Systems, GCIS '09, WRI Global Congress*. May 19–21, 2009. Xiamen, China. Los Alamitos. IEEE Computer Society. Vol. 4 p. 63–67.
- SHAMSELDIN A.Y. 2010. Artificial neural network model for river flow forecasting in a developing country. *Journal of Hydroinformatics*. No 12 p. 22–35.
- SHARDA V., PATEL R.M., PRASHER S.O., OJASVI P.R., PRAKASH C. 2006. Modeling runoff from middle Himalayan watersheds employing artificial intelligence techniques. *Agriculture Water Management*. No 83 p. 233–242.
- SMOLA A., SCHOLKOPF A. 1998. A tutorial on support vector regression. *NeuroCOLT2 Technical Report NC2-TR-1998-030* pp. 71.
- TIWARI M.K., CHATTERJEE C. 2010. A new wavelet-bootstrap-ANN hybrid model for daily discharge forecasting. *Journal of Hydroinformatics*. No 13 p. 500–519.
- VAPNIK V. 1995. *The nature of statistical learning theory*. Ser. Information Science and Statistics. New York. Springer Verl. ISBN 978-0-387-98780-4 pp. 314.
- WANG W., CHAU K.C., CHENG L.Q. 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*. No 374 p. 294–306.
- WU C., CHAU K., LI Y. 2009. Methods to improve neural network performance in daily flows prediction. *Journal of Hydrology*. No 372 p. 80–93.
- ZEALAND C., BURN D.H., SIMONOVIC S.P. 1998. Short term streamflow forecasting using artificial neural networks. *Journal of Hydrology*. No 214 p. 32–48.

---

**Jan ADAMOWSKI, Shiv O. PRASHER**

### **Porównanie metod uczenia maszynowego do prognozowania splywu w zlewniach górskich na podstawie ograniczonych danych**

#### **STRESZCZENIE**

**Słowa kluczowe:** *Himalaje, prognozowanie splywu, regresja wektora wsparcia, sieci falkowe, uczenie maszynowe*

Prognozowanie splywu z obszarów górskich z użyciem programowanych modeli jest często trudne i niedokładne z powodu złożonych zależności między opadem a splywem i problemów związanych z pozyskaniem niezbędnych danych. Modele uczenia maszynowego stwarzają alternatywę dla prognozowania splywu z takich regionów. W pracy analizowano i porównano dwie metody uczenia maszynowego – metodę regresji wektorów nośnych (SVR) i sieci falkowych (WN) do dobowego prognozowania splywu w górskiej zlewni Sianji, usytuowanej w indyjskiej części Himalajów. Modele opracowano na podstawie danych o splywie, wskaźniku poprzednich opadów, opadzie i kolejnym dniu roku za trzyletni okres od 1 lipca 2001 r. do 30 czerwca 2004 r. Stwierdzono, że obie metody zapewniają dokładne wyniki, przy czym najlepszy model WN nieco przewyższa najlepszy model SVR pod względem dokładności. Obie metody powinny być testowane w innych zlewniach górskich o ograniczonej liczbie danych, aby lepiej ocenić ich przydatność do prognozowania.