River Flow 2004 – Greco, Carravetta & Della Morte (eds.)
©2004 Taylor & Francis Group, London, ISBN 90 5809 658 0

River flow forecast by means of selected black box models

A. Piotrowski, P.M. Rowiński & J.J. Napiórkowski
Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

ABSTRACT: Performance of two different types of approaches were investigated with respect to the daily river flow predictions. The first approach is Multi-Layer Perceptron Artificial Neural Network, the second one, based on the deterministic chaos concept, is so called Phase-Space Reconstruction (PSR) model. Both models were applied to daily river flow data collected from several gauges, located in river reaches in western Canada. Each data set consists of more than 10000 consecutive daily measurements. The method based on the phase-space reconstruction theory has been applied despite the fact that the authors doubt in the existence of deterministic chaos in the river system. Classical phase-space method may be applied only to single time series data. In the present study an extension was proposed, in which time series from additional gauge stations added to the system. Comparison of models' performance was made for this extended and classical approaches for both phase-space reconstruction and Artificial Neural Networks models.

1 INTRODUCTION

An accurate forecasting of river flow is a fundamental problem in hydrological sciences. In recent years application of black box type models to runoff prediction has been gaining more popularity due to their easiness of employment and relatively low number of variables involved. Multi-Layer Perceptron Artificial Neural Networks (ANN) are probably most broadly used for forecasting (Hsu et al. 1995, Campolo et al. 1999, Campolo et al. 2003, Dolling & Varas 2002, Chibanga et al. 2003 and many others). They are the simplest among the non-linear neural network models, nevertheless they are capable of making very accurate forecast, when proper set of input variables is taken into account. On top of that they are easy to be trained, contrary to, for example, radial basis functions networks. ANN approach is sometimes called global approximation approach as it deals with all set of historical data values to optimise network weights, which means that all known information about the system is used to make a relevant forecast.

A methodologically different approach, though still quite popular, pertains to phase-space reconstruction (Jayawardena & Lai 1994, Porporato & Ridolfi 1997, Sivakumar at al. 2002). It is a local approximation approach and it has been developed in the framework of the deterministic chaos theory. The presence of low-dimensional chaotic behaviour in the rainfall-runoff process is a matter of an ongoing debate (see e.g. Schertzer et al. 2002) and it will not be discussed in detail herein. Let us only note that the main idea of this approach corresponds to the possibility of the reconstruction of the phase-space from a discrete set of values for a given observable scalar such as water stages at a given gauge station. The authors of the present study will benefit from the phase-space reconstruction method which in fact is a way of finding the most similar situations in historical data and applying only these selected parts of data set for forecasting. It seems that this method may stand alone as a reliable tool without linking it to the existence or nonexistence of a hypothetical deterministic dynamical system leading to disordered solutions. According to Sivakumar et al. (2002), the phase-space reconstruction method is better than ANN approach when dealing with autoregressive forecast problem. Further in the paper it will be shown that such conclusion is premature and in case of the selected Canadian river reaches both ANN and phase-space reconstruction methods boast similar performance of forecasting daily runoffs. Additionally an extension, further called as quasi-phase-space reconstruction method, will be proposed and this method will allow for the use of the data from more than one gauge station. As expected this method will improve the results considerably.

The comparison of the results of two black box models will be presented according to the following scheme:

- version A: as inputs only historical runoff data from the same gauge are applied for both ANN and phase-space reconstruction approaches;
- version B: forecast is made for the same gauge as in version A, but input data set consists of data from 2 or 3 gauges. Additional gauge for version B is selected at the same river where the forecasting gauge or its main tributary is placed.

For each case 3 consecutive measurements from particular gauge were treated as input variables which means that we have 3 input variables in version A, and 6 or 9 in version B.

2 FLOW DATA

Long enough data sets collected in 5 western Canadian rivers (Figure 1) are applied in the analyses (Environment Canada 2003). Table 1 presents the duration of daily data sets, location of the main gauge (used for the forecasts, denoted by number one) and the additional gauges (applied in version B only, denoted by number two or three). Each data set contains more than 10000 daily measurements, and 5 river reaches are considered in the study. It seems to provide sufficient information to compare the performance of all the considered models. In each case the daily river flow forecast was made for last 5000 measurements. Those 5000 measurements were not taken into account during ANN training process, but they were only used to compare performance of both models at each river reach. The ANNs were trained with the use of the earlier records (larger than 5000 elements in each case).

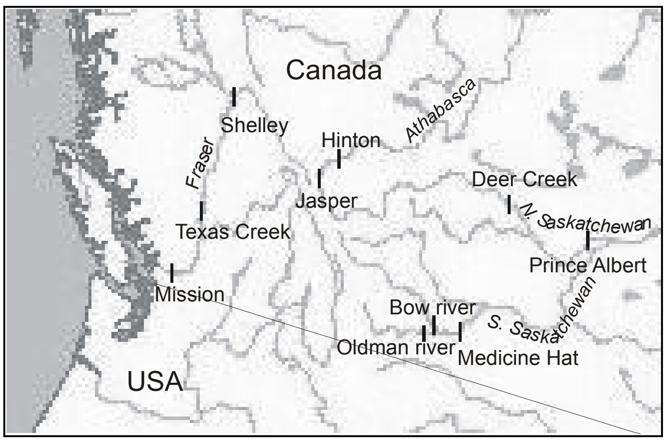


Figure 1. Location of gauge stations.

Table 1. Description of data sets

River	Gauge sites	Length
Athabasca*	1. Hinton 2. Jasper	71.01.01- 02.12.31
Fraser* (upper part)	1. Above Texas Creek 2. Shelley	55.01.01- 02.12.31
Fraser* (lower part)	1. Mission 2 Above Texas Creek	65.05.01- 92.12.31
South Saskatchewan**	1. Medicine Hat 2. Bow river 3. Oldman river	65.01.01- 02.12.31
North Saskatchewan**	1. Prince Albert 2. near Deer Creek	70.01.01- 01.12.31

*mountainous terrain, **steppe

3 MODELS APPLIED

In this section attention is focused on the implementation of some anticipatory nonlinear methods based on Deterministic Chaos and Artificial Neural Networks for prediction of the inflows.

3.1 *Artificial Neural Network approach*

Multi-Layer Perceptron Artificial Neural Networks have become widespread in recent years and the researchers often claim that they provide a useful tool for the predictions of river flow. Three layer networks with sufficient number of hidden nodes are usually applied due to the continuity of the relevant function. Every network contains an appropriate number of input and output nodes which is equal to the number of input and output variables, and the assumed number of hidden nodes. There is no effective rule for the estimate of the number of hidden nodes. In this study it usually turns out to be close to the number of input nodes, but in each case it is experimentally verified.

The ANN nodes in neighbouring layers are linked via weighted connections. The values of those weights can be adaptively modified during the process of training the network. In this study, due to the relatively simple architecture of all the networks, Levenberg–Marquardt non-linear optimisation algorithm was adopted (Press et al. 1989).

Shortly the Multi-Layer Perceptron network (see Figure 2) operates in the following way: signals s_i ($i = 1,..N$) from the input nodes (e.g. values of input variables normalized to 0-1 interval) are multiplied by proper weights w_{ji} ($j = 1,\ldots K$), connecting the neuron from which signal has been dispatched and a suitable neuron in the second layer.

In the second layer the weighted sum of all the inputs are computed and then transformed by logistic function

$$f(x) = \frac{1}{1+e^{-ax}} \tag{1}$$

giving the output value of a neuron in the second layer:

$$z_j = f\left(\sum_{i=1}^{N} s_i w_{ji}\right) \tag{2}$$

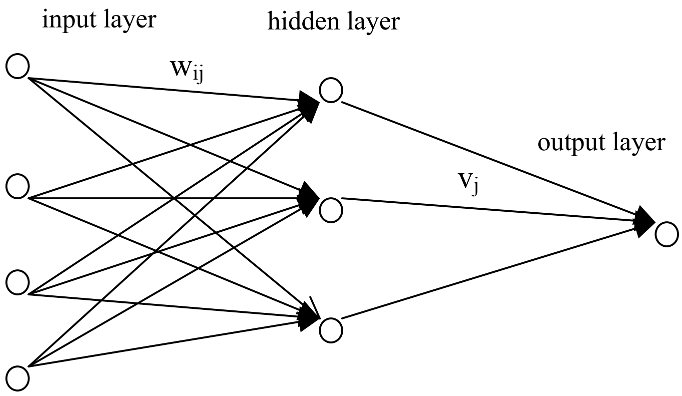


Figure 2. Multi-Layer Perceptron Artificial Neural Network scheme.

Afterwards the weighted signals z_j (multiplied by proper weights v_j), are transferred to the neuron of the third layer. In the neuron of the third layer the new weighted sum is computed

$$y = \sum_{j=1}^{K} v_j f\left(\sum_{i=1}^{N} s_i w_{ji}\right) \tag{3}$$

and after de-normalization of the output, the sought (forecasted) value may be determined. This is a feed-forward network, which means that there is only one direction of the flow of information, from the input to the output layer.

3.2 *"Phase-space reconstruction" approach*

Another method used in runoff forecasting studies is motivated and based on the theory of dynamical systems (Kantz & Schreiber 1997), i.e. the time evolution is defined in some phase-space. For a purely deterministic system, once its present state is fixed, the states at all future times are determined as well. So it is essential to establish a phase-space for the system such that specifying a point in this space specifies the state of the system and vice versa. Then one can study the dynamics of the system by studying the dynamics of the corresponding phase-space points.

Characteristic for chaotic systems is that in many cases the corresponding attractors turn out to be strange attractors of the fractal structure with a non-integer dimension (Kudrewicz 1993, Ott 1993).

Since it is natural to describe a deterministic dynamical system as an object in phase-space, it is also most natural to use a phase-space description for approximation of the dynamic of the system. Such approximate dynamics can be useful for prediction.

The reconstruction of a vector space which is equivalent to the original state space of a system from a scalar time series is the basis of almost all nonlinear methods. Scalar measurements is a projection of unobserved internal variables of a system onto an interval on real axis. It is obvious that even with a precise knowledge of the measurements it may be impossible to reconstruct the state space of the original system from the data. Fortunately, we are rarely keen of obtaining a whole, precise description of the process. It is usually enough to determine its good approximation. Hence a reconstruction of the original space is not really necessary for the data analysis. It is sufficient to construct a new space such that the attractor in this space is "equivalent" to the original one (Kantz & Schreiber 1997).

The classical situation, as treated in hydrological sciences, though being a matter of serious debate (e.g. Schertzer et al. 2002) is the following. A sequence $\{x_i\}$, $x_i \in R^1$, i = 1, ..., N, of measured values is given. In order to find an approximation of the deterministic system one considers the function F defined by the following relationship

$$y_m^{i+\Delta} = F(y_m^i) = F(x_i, x_{i-1}, ..., x_{i-(m-1)}) \tag{4}$$

with a properly adjusted number m, called *embedding dimension* and a given time delay Δ. Hence, the so-called quasi phase-space composed of m-element subsequences y_m^i is considered

$$y_m^i = (x_i, x_{i-1}, ..., x_{i-(m-1)}) \tag{5}$$

The function F in Equation (4) is a dynamical process in the space R^m, which according to the embedding approach (Takens 1981) forms an attractor in R^m, if the original process is a deterministic chaos. To determine proper embedding dimension from a finite sample of points one determines, for example, correlation integral C_N(m, r) for several embedding dimensions (Grassberger & Procaccia 1983). Then correlation dimension D is determined as the slope of function ln C_N(m, r) with respect to ln(r), in a respective range of sufficiently small r, such that the function behaves as a linear one (one expects C to scale like power law, $C(r) \propto r^D$). Now using Takens theorem, one can put m=2D+1 as the searched embedding dimension.

One can proceed then to the stage of determining the prediction model for the relationship F in Equation 4. It is possible for deterministic chaos case: as the process is really deterministic and due to existence of an attractor. The considered prediction model has the form of a function such that it approximates the function F, or even less – a „component" of F, being prediction of a future value of state.

$$x_{i+\mathbf{T}} = f_{\mathbf{T}}(x_i, x_{i-1}, ..., x_{i-(m-1)}) \tag{6}$$

where T is a prediction horizon. Such a function depends on the time instant i of making prediction, and on the horizon T of this prediction. Thus, one can seek for a function

$$f_{\mathrm{T}}^{i}(y_{m}^{i}) = f_{\mathrm{T}}^{i}(x_{i}, x_{i-1}, \ldots, x_{i-(m-1)}) \tag{7}$$

that would determine a good approximation of the value x_{i+T} of the given sequence $\{x_i\}$.

Our computations do not confirm the straightforwardness of the application of the described approach and we become skeptical about its correctness. Correlation integral for data described in section 2 revealed no obvious scaling region as the indicator of self-similar geometry. The saturation value of D (which is just the lower bound of box dimension) with increasing m was not clear at all. We are also aware of the occurrence of noise in the applied data, so the question arises whether those data can be analyzed within the deterministic system.

As a consequence of the above we applied a part of the method derived within the deterministic chaos theory, but as mentioned in Introduction there is no need to decide whether the chaos exists in the system under consideration. We keep the name of the method as the "phase-space reconstruction" approach (Jayawardena & Lai 1994, Sivakumar et al. 2000) to show its origin but we do not claim we have applied the embedding theorem. A method of delays as a realization of the "phase-space reconstruction" approach has been selected for the purpose of the present study. In order to make a forecast one constructs m-dimensional data vectors from measurements spaced equidistant in time (temporal sequence of measured values at the selected gauge) which creates an analogy to the phase-space.

Further the principle of the applied method lies in the search of K-points from the d-dimensional point set that are placed at the smallest distance (according to some assumed measure) from the points representing the current situation. In other words we are interested in finding K vectors of the length m from the past, most of all resembling the current situation. We do assume one day time delay which allows us to use the consecutive recordings from the gauge station. Thus we obtain an autoregressive forecast for each day (version A). In the present study it was assumed that m=3 which proved to produce reasonable results.

3.3 *Quasi "phase-space reconstruction" approach*

We proceed similarly as in Section 3.2 but this time we test whether the similar forecast is possible using data collected from more than one gauge. Intuitively more information should lead to a better forecast. Being remote from the rigorous treatment of the deterministic chaos we may "extrapolate" the previous considerations from scalar observables (water stages) to vectors. This time we build 2m-dimensional quasi "phase-space" where the first d-vector is made of the consecutive measurements from the first gauge and the second d-vector is made of the consecutive measurements in the second gauge station (time delay is again assumed as one day).

One should note that the measurements from different gauges may have different standard deviations, so the scale related to the first three components and the last three components may be heterogeneous, that can impact the search for N closest neighbours in this space. Fortunately in our case the performance of forecast made using data sets divided by standard deviation is almost the same as for raw data and therefore no normalization was necessary. The above describes the technique for daily runoff forecast in version B.

4 COMPARISON OF APPLIED METHODS

The comparison will be made for both versions A and B by means of correlation coefficient applied to measured and predicted daily or (in one case) 4-daily runoff volume increments (see Table 2). Such approach shows the performance of the model much better than the correlation applied to just actual runoff volumes where almost all results are close to unity independently of the quality of results.

Table 2. Correlation coefficients of forecasted and measured (daily, except noted case) runoff changes.

	Ver. A		Ver. B		
river	ANN	PSR	ANN	Quasi PSR	No.of gauges
Athabasca	0.674	0.667	0.904	0.899	2
Fraser (upper part)	0.738	0.735	0.928	0.922	2
Fraser (lower part)	0.478	0.512	0.660	0.507	2
Sth Saskatchewan	0.518	0.507	0.979	0.796	3
Nth Saskatchewan (1 day forecast)	0.601	0.603	0.736	0.760	2
Nth Saskatchewan (4 day forecast)	0.364	0.317	0.856	0.847	2

As it was expected inclusion of additional gauge station (Version B) improves the results quite significantly for both "phase-space reconstruction" as well as ANN (Figures 3-8). The only exception is the lower Fraser river. The addition of data from the gauge station at Texas Creek did not help the "phase-space reconstruction" technique. The exemplification of this fact is seen in Figure 4. Note that all the Figures (3-8) are prepared for a selected period of about one month only, out of 13 years time series. For clarity the relevant period was chosen to reveal especially large changes in the daily river runoff.

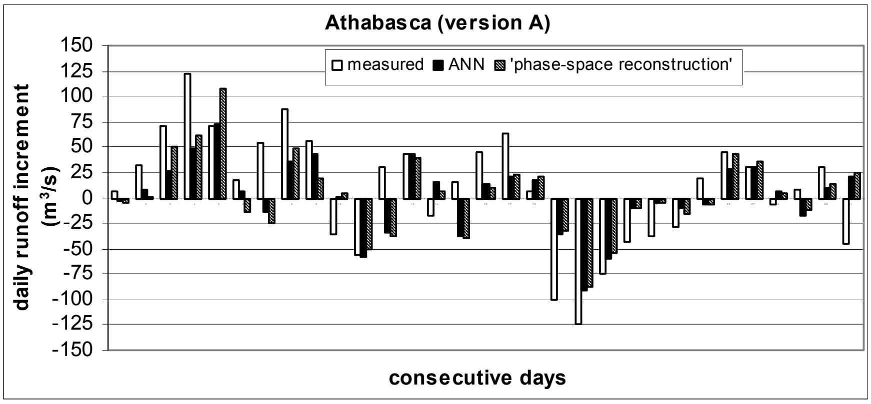
This is worth mentioning that in version A all ANN and phase-space models show similar performance. This result is different from the one of Sivakumar et al. 2002 who considered daily river flow data set from only one gauge station in Thailand. A very short time series was analysed in that study, and the authors removed half of the data set because large differences were observed in those two parts of time series. In that study the phase-space model performance was much better than ANN.

Note that in version B in case of South Saskatchewan river prediction made by ANN is very good indeed and better than that by quasi "phase-space reconstruction" (see Figure 6). In case of quasi "phase-space reconstruction" the difficulty lies in the fact that it is quite hard to find enough vectors (3 delayed data from three gauges) in 9-dimensional space similar to the current data allowing for the evaluation of relevant linear regression coefficients.

An astonishing fact is that in version B in respect to North Saskatchewan River, 4-daily runoff forecast (Figure 8) proves to be better than the daily predictions (Figure 7, Table 2). It is the result of the selection of additional gauge at the distance corresponding to flow routing during the time period of 4 days.

5 CONCLUSIONS

The present study that was based on long enough time series from five river reaches shows that the method analogical to the phase-space reconstruction derived in the framework of deterministic chaos theory is useful for runoff predictions even if we have no evidence about the existence or non-existence of chaos in the considered river systems. It has also been shown that contrary to what is suggested in some hydrological articles, Artificial Neural Networks may provide accurate enough forecasts, in many cases even better than those based on "phase-space reconstruction". On top of that it has been shown that the use of additional information from other gauges improves significantly the forecasts based on both ANNs and "phase-space reconstruction". Such proposal has not been considered in respect to "phase-space reconstruction" in literature so far.



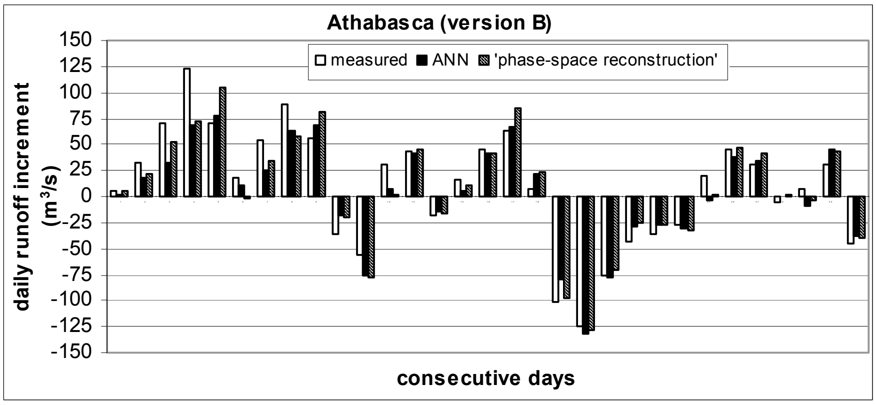
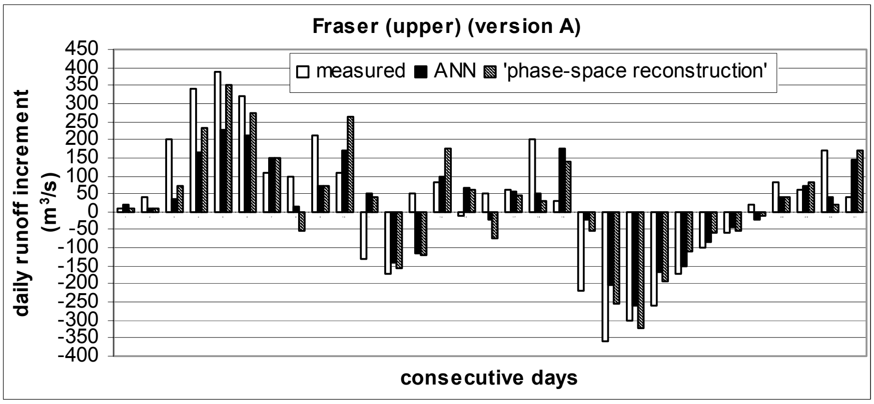


Figure 3. Comparison of one lead day runoff forecast for Hinton gauge station, on Athabasca river, obtained from ANN and quasi "phase-space reconstruction" models in versions A and B



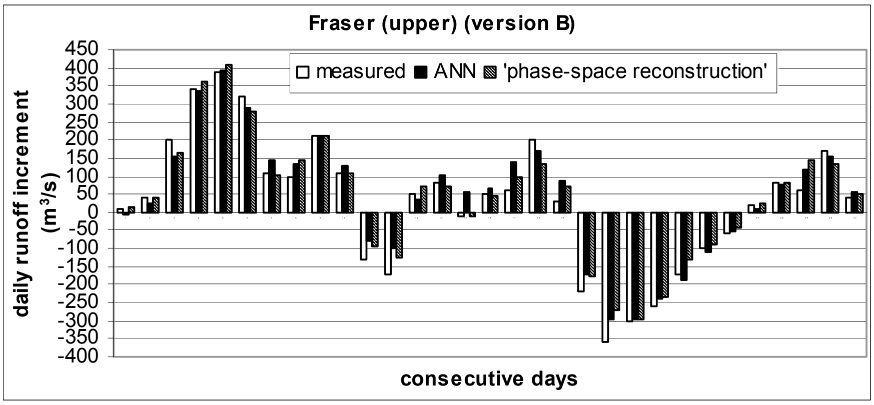
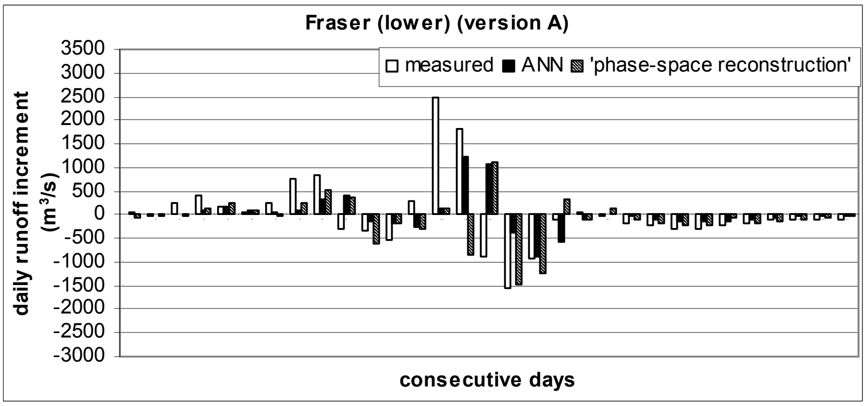


Figure 4. Comparison of one lead day runoff forecast for gauge station near Texas Creek (BC), on upper part of Fraser river, obtained from ANN and quasi "phase-space reconstruction" models in versions A and B



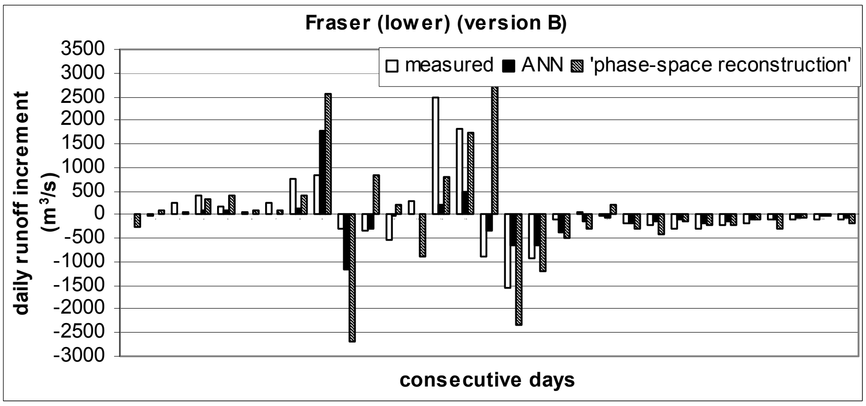
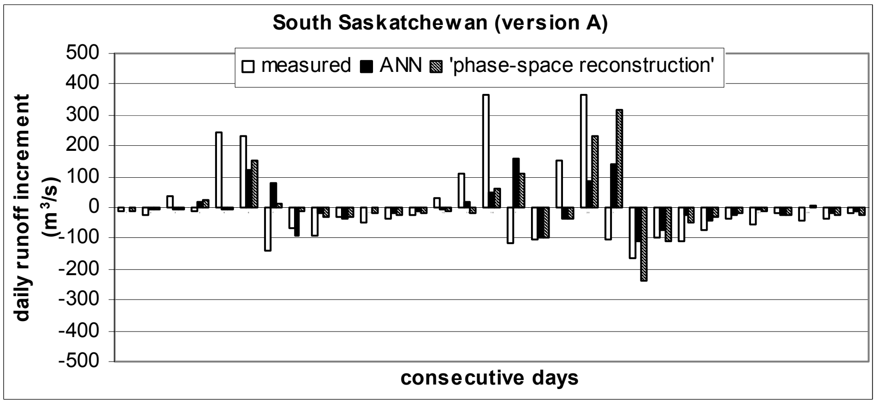


Figure 5. Comparison of one lead day runoff forecast for gauge station near Mission (BC), on lower part of Fraser river, obtained from ANN and quasi "phase-space reconstruction" models in versions A and B



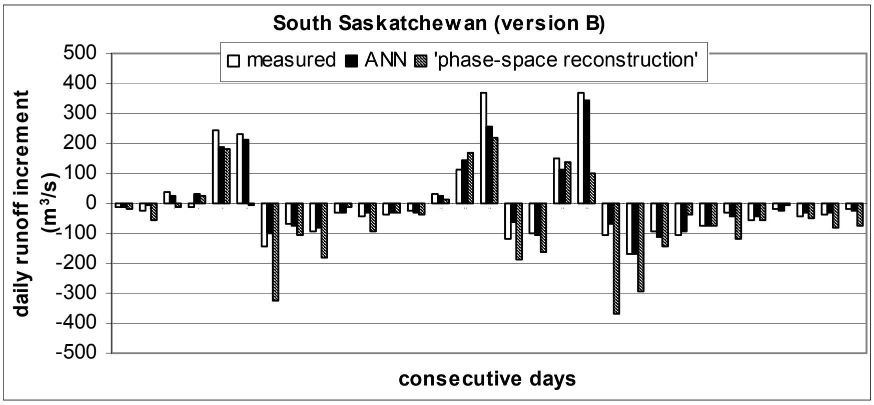
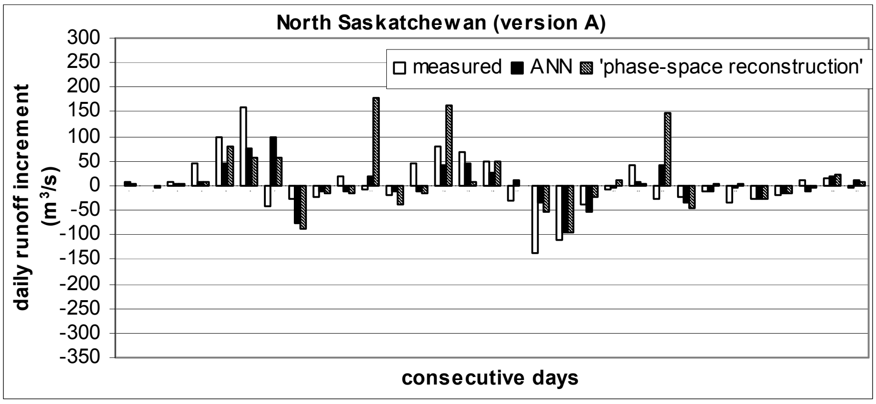


Figure 6. Comparison of one lead day runoff forecast for Medicine Hat gauge station, on South Saskatchewan river, obtained from ANN and quasi “phase-space reconstruction” models in versions A and B



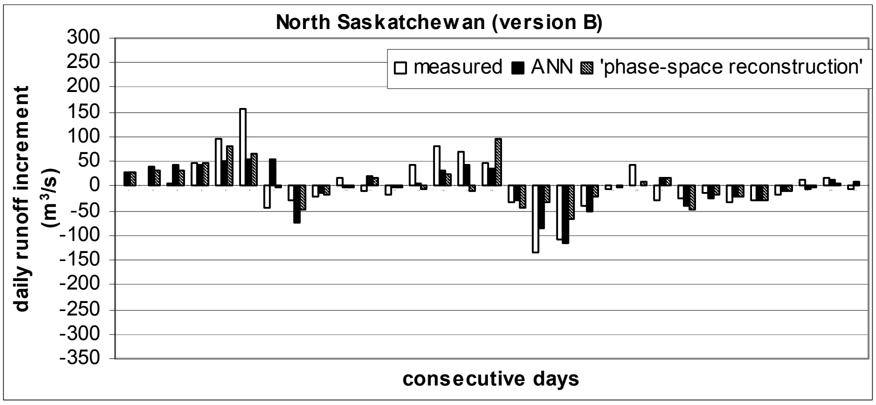
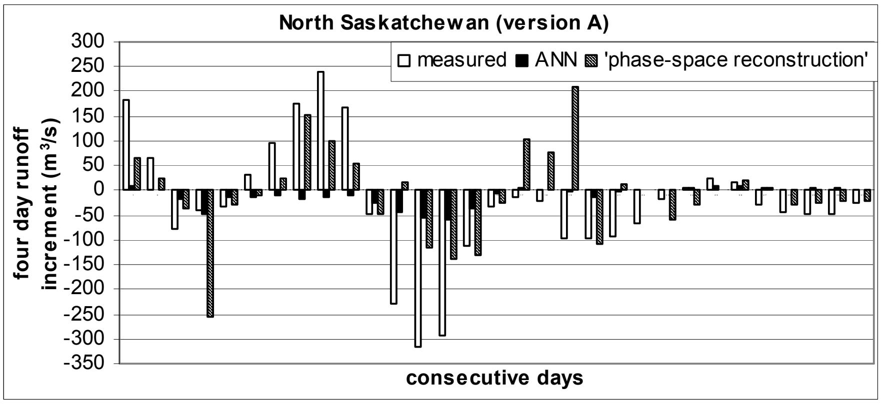


Figure 7. Comparison of one lead day runoff forecast for Prince Albert gauge station, on North Saskatchewan river, obtained from ANN and quasi “phase-space reconstruction” models in versions A and B



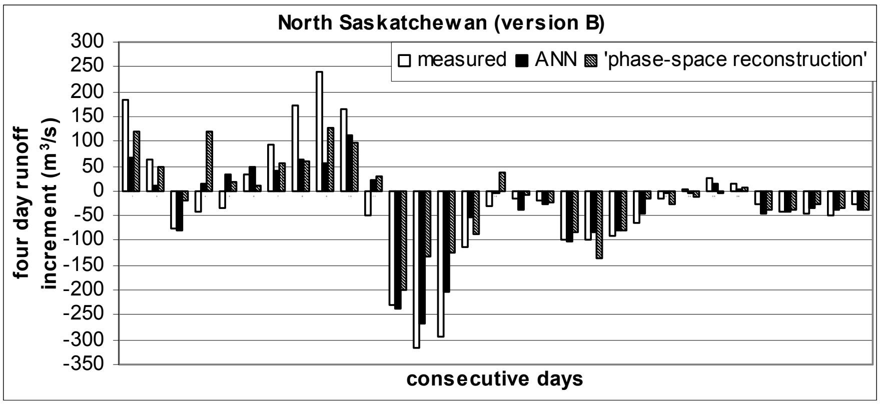


Figure 8: Comparison of four lead day runoff forecast for Prince Albert gauge station, on North Saskatchewan river, obtained from ANN and quasi "phase-space reconstruction" models in versions A and B

6 REFERENCES

Campolo, M., Andreussi, P. & Soldati, A. 1999. River flood forecasting with a neural network model. *Water Resources Research*, 35(4): 1191-1197.

Campolo, M., Soldati, A. & Andreussi, P. 2003. Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences Journal,* 48(3): 381-398.

Chibanga, R., Berlamont, J.& Vandewalle, J. 2003. Modelling and forecasting of hydrological variables using artificial neural networks: the Kafue River sub-basin. *Hydrological Sciences Journa*, 48(3): 363-379.

Dolling, O.R., Varas, E.A. 2002. Artificial neural networks for streamflow prediction. *Journal of Hydraulic Research*, 40(5): 547-554.

Environment Canada web page: http://citech.pyr.ec.gc.ca/aterweb/

Grassberger, P. & Procaccia, I. 1983. Measuring the strangeness of strange attractors. Physica 9(D): 189-208.

Hsu, K-L., Gupta, H.V. & Sorooshian, S. 1995. Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research*, 31(10): 2517-2530.

Jayawardena, A.W. & Lai, F. 1994. Analysis and prediction of chaos in rainfall and stream flow time series. *Journal of Hydrology*, 153, pp 23-52

Kantz, H. & Schreiber, T. 1997. Nonlinear Time Series Analysis. *Cambridge University Press*, Cambridge.

Kudrewicz, J. 1993. Fractals and Chaos, *WNT*, Warsaw.

Ott, E. 1993. Chaos in Dynamical Systems. Cambridge University Press

Porporato, A. & Ridolfi, L. 1997. Nonlinear analysis of river flow time sequences. *Water Resources Research*, 33, 1353-1367.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T. 1988. Numerical recipes: The art of scientific computing (Fortran Version). Cambridge University Press

Schertzer, D., Tchiguirinskaia, I., Lovejoy, S., Hubert, P., Bendjoudi, H. & Larcheveque, M. 2002. Discussion of "Evidence of chaos in the rainfall-runoff process". *Hydrological Sciences Journal*, 47(1): 139-148.

Sivakumar, B., 2000. Chaos in hydrology: important issues and interpretations. Journal of Hydrology, 227: 1-20.

Sivakumar, B., Jayawardena, A.W. & Fernando, T., M., K., G. 2002. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. *Journal of Hydrology*, 265: 225-245.

Takens, F. 1981. Detecting strange attractors in turbulence. In: Rand, D., A. & Young, L.S. (Eds.), *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, 898: 366-381. Springer, Berlin.