R. Erinawati and J. D. Fenton (2004) A comparison of computational intelligence systems for river flow forecasting, *Proc. 4th Austral. Stream Management Conf.,* Launceston, Tasmania, 19-22 Oct 2004, Eds. I. D. Rutherfurd, I. Wiszniewski, M. Askey-Doran, & R. Glazik, *pp*217-222, Dept of Primary Industries, Water and Environment, Hobart.

A comparison of computational intelligence systems for river flow forecasting

Ririn Erinawati[1], John D Fenton[2]

1 Department of Civil and Environmental Engineering, The University of Melbourne, Parkville 3052, Victoria, r.erinawati@civenv.unimelb.edu.au
2 Department of Civil and Environmental Engineering, The University of Melbourne, Parkville 3052, Victoria, http://www.civag.unimelb.edu.au/~fenton/, fenton@unimelb.edu.au

Abstract:

Accurate river flow forecasting is important for most problems related to river basin and stream management. Conceptual or physical-based models can lead to a better understanding of hydraulic and hydrological processes, however the majority of these types of models are relatively expensive, require intensive data input, and high levels of user expertise. There are many practical situations where the main interest is to determine accurate stream flow predictions at specific locations. In such situations, system or black-box models are preferred. Typically, these are based on historical measurements of inflows and outflows for system identification, which are then used for system simulation and stream flow prediction.

This study compares artificial neural networks (ANNs), fuzzy rule-based system (FRBS), autoregressive moving average with exogenous (ARMAX), linear transfer function (LTF), and nonlinear transfer function (NLTF) methods. It tests the application of these methods to six study areas in Victoria and one in the USA. The results indicate that in most of the cases ANNs perform very well. ARMAX and FRBS also give satisfactory results. LTF, although providing the best results for the Mitta Mitta River, its performance, to some extent, was inconsistent. However, due to its simplicity and few parameters, LTF can still be recommended as an alternative. Finally, linear and *nonlinear* transfer function methods are compared. These methods have proved to be quite powerful (Tsykin, 1985). The transfer functions are obtained using linear and non-linear time series equations and solving them with optimisation procedures. The problems of selecting equations, objective functions and optimisation methods are also discussed. Several optimisation methods were tested, included linear, non-linear, global, and the genetic algorithm. Results obtained are quite good. It is interesting that the simple linear transfer function gives results which are just as good as the nonlinear transfer function methods.

Key Words

Neural Networks, flow prediction, Fuzzy Logic, ARMAX, linear, non-linear, transfer function, optimisation, river, routing.

Introduction

River flow forecasting is of significant importance for planning and operational purposes in streamflow management, and can be considered a major focus of hydrological research. A large number of models have been proposed to achieve more precise and reliable predictions of streamflow, including many detailed conceptual or physical-based models. While these are believed to provide a better understanding of hydraulic and hydrological processes, most of these approaches are relatively expensive, require intense data input and high levels of user expertise.

In practice, there are many situations where the main interest is to determine accurate stream flow predictions at specific locations. In such cases, system identification models are preferred. Typically, these are based on historical measurements of inflows and outflows for system identification, which are then used for system simulation and stream flow prediction.

In the last decade, there has been significant growth of interest in computational intelligence systems and their application to water related problems, including artificial neural networks, fuzzy set theory, chaos theory, genetic algorithms, and model trees. Studies using these methods to date have reported successful findings. However, most of these studies have only considered one or two samples, which makes it difficult to generalize the capability of the models for flow forecasting. Therefore, this study addresses the application of five system identification techniques: ANNs, FRBS, ARMAX, LTF, and NLTF. ANNs are selected due to their popularity. They are considered one of the most widely used computational intelligence systems in water related problems. There are numerous studies on the application of ANNs in water related problems, such as rainfall-runoff modelling (Minns & Hall, 1998), daily streamflow (Birikundavyi, Labib, Trung, & Rouselle, 2002), rating curves (Bhattacharya & Solomatine, 2000). Fuzzy set theory is also increasing in popularity. This method is commonly used as a control system, such as water level control (Lobbrecht & Solomatine, 1999), and disaster control planning (Esogbue, 1996). ARMAX is an extended form of time series ARMA, and is well accepted in hydrology. The simple linear transfer function model, LTF, is used as a basic comparison for the three aforementioned models. These techniques are applied to predict the stream flow based on information from an upstream river section. A more general method is to add nonlinear terms to the relationship between input and output, and to calculate the corresponding nonlinear transfer function (NLTF), which is also investigated here.

Artificial Neural Networks (ANNs)

ANNs are defined as massive parallel-distributed information-processing systems that resemble biological neural networks of human cognition (ASCE & Govindaraju, 2000). Although McCulloch and Pitts first introduced the idea of artificial neural networks over fifty years ago, significant developments only started in 1982, when Hopfield introduced an iterative procedure for neural networks. There are several types of ANNs. Multi-layer perceptrons (MLP) are a type of 'feed forward' network, with one or more hidden layers. The MLP with three hidden layers is employed in this study, consisting of one input layer, one hidden layer, and one output layer. The MLP uses a common back-propagation algorithm as the learning rule. There are two phases involved in the back-propagation algorithm, a feed forward phase where the information propagates forward to calculate the output signal, and a backward phase where the connection weights are updated to minimize the difference between computed output and the given output.

Fuzzy Rule Based-System (FRBS)

Fuzzy logic was introduced by Lotfi Zadeh in 1965. Fuzzy logic extends from the general form of Boolean logic, using 'true' and 'false' to handle the concept of vagueness. This approach involves taking a value between 1 (full belongingness) and 0 (non belongingness), rather than a crisp value. The degree of belongingness is called membership function. Fuzzy rules are collections of linguistic IF and THEN arguments. The general form of the fuzzy rule can be expressed as IF "X" THEN "Y". X is the provision and Y is the consequence of the rule. Since it is based on verbal arguments, this rule allows for imprecision and uncertainty in variables. There are five steps involved in a fuzzy rule based-system: fuzzify inputs, apply fuzzy operators, apply implification method, aggregate outputs, and defuzzify outputs. In practice, due to the complexity of physical systems, it is often difficult to construct rules. Several methods have been proposed to extract rules directly from numerical data.(Abe, 1997; Abe & Lan, 1995). In this study the Sugeno-type is implemented.

Autoregressive Moving Average (ARMAX)

ARMAX is a variant of time series ARMA, with additional information. Hydrologists have commonly used this type of system identification for 30 years. A general form of ARMAX can be written as:

$$A(q)y(t) = B(q)u(t) + C(q)e(t) \tag{1}$$

where *y(t)* is flow at time t, $u(t)$ is input from upstream at time *t*, *e(t)* is a white noise at time *t*, *A(q)*, *B(q)* and *C(q)* are the autoregressive, exogenous, and moving average parameters. Where,

$$A(q) = 1 + a_1 q^{-1} + ... + a_n q^{-n}, \tag{2}$$

$$B(q) = b_1 + b_2 q^{-1} + ... + b_m q^{-m+1}, \tag{3}$$

$$C(q) = 1 + c_1 q^{-1} + ... + c_p q^{-p}, \tag{4}$$

where q^{-1} is the back shift operator, n, m and p are the order of the autoregressive, exogenous, and moving average respectively.

Linear Transfer Function

The LTF is a general method for solving arbitrary physical systems, where the relation between input and output is linear. It is assumed that the output at any instant depends on a weighted sum of the input values over a given preceding period. In hydrology it is best known as the unit hydrograph method, where for a given input time series of rainfall values and a given output time series of runoff values, usually streamflows, the transfer (weight) function value between the two is calculated for a number of time differences between input and output. Then, it is assumed that the sequence of transfer function values holds for all input series, and is then used to calculate the corresponding output series for any given sequence of inputs.

In this work the idea is extended to the case where, instead of the input being rainfall, it is streamflow at an upstream station. The method has the potential of giving a simple numerical description of the nature of wave propagation in the stream, as well as a means of computing downstream flows in the case of floods. The linear combination of weighted previous input values leading to output value y_t at time t is written

$$y_t = \sum_{p=1}^{P}\left[\sum_{j=0}^{J-1} h_j^p u_{t-j}^p\right], \tag{5}$$

where, $u_{t-j}^{(p)}$ is the input value at a previous time step $t-j$, at which $h_j^{(p)}$ is the transfer ("weight") function, P is the number of input sequences, p is the input sequence index, J is the number of transfer functions, j is the transfer function index. In practice equation (5) cannot be satisfied exactly. The h values that satisfy it were obtained from a training process, using optimisation methods. The model parameters were obtained by minimising or maximising an objective function. The least squares objective function and adaptive cluster covering optimisation were implemented.

Non-linear Transfer Function

In the nonlinear case the NLTF is written as

$$y_t = \sum_{j=0}^{N} h_{1n} u_{t-j} + \sum_{j=0}^{N} h_{2j} u_{t-j}^2 + \sum_{j=0}^{N} h_{3j} u_{t-j}^3 + \dots, \tag{6}$$

which allows for nonlinearities in the system, such as the speed of propagation of flood waves depending on the magnitude of the flood, and so on.

Study Areas and Problem Formulation

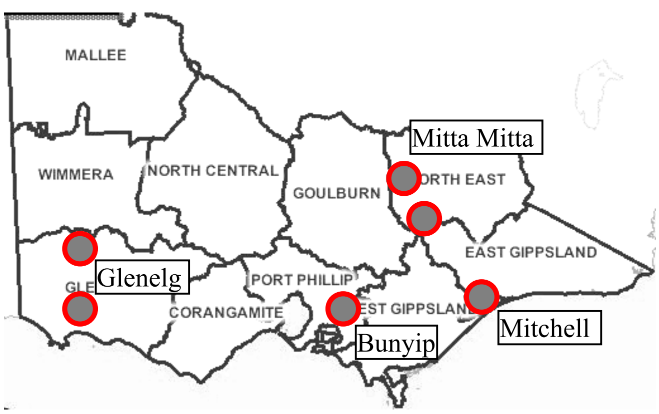

Figure 1. The Study Areas

Six sections of river in Victoria were modelled; the Bunyip river, the Mitchell river, two sections of the Mitta Mitta river and two sections of the Glenelg river (Figure 1). The Mitchell river is located in south-eastern Victoria, in the Gippsland region, draining headwaters of the Wonnangatta river and the high plains. The Mitta Mitta river is situated in north-eastern Victoria. It flows northward from the Great Dividing Range for about 270 km to join the Murray river. There are two regions of interest on the Mitta-Mitta river. The first is on the upper part of Dartmouth dam, roughly a 100km section, and the second lies in the middle part of the river. The Bunyip river is located in southern Victoria, and data sets were obtained from three gauging stations. The Glenelg river is situated in western Victoria. It flows from the southeast corner of the Victoria Ranges. The typical characteristics of this region are flat volcanic plains, similar to other regions in the western district. Two areas of the Glenelg were modelled, one is in the upper Glenelg, and another one is in the lower Glenelg. Except for the Mitchell river, where the daily data was employed, the rest of the Victorian rivers used three hourly data. The South Fork River is located at South Carolina. The river reach runs from Chili Bar to Salmon Falls; hourly data are used.

Four techniques were applied to predict the stream flow based on historical information from the upstream part of the rivers. In this particular study, the downstream flow at time *t*, *y(t)* was assumed to be related to the past input from upstream, *u(t-j)*,

$$y(t) = f\big(u(t-1), u(t-2), \ldots, u(t-j), \ldots,\ u(t-J) + e(t)\big), \tag{7}$$

where f is the unknown mapping function, t is the time index, J is the unknown number of the past input, j is the past input index, and e is the error to be minimised. The performance indicators of the models are given as the root mean square errors (RMSE), the mean absolute errors (MAE), and the coefficient of determination (R^2). The results are shown in Table 1.

Table 1. Comparison of errors of the various methods

	Linear Transfer Function			ARMAX			ANNs			FRBS		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
Bunyip	37.9	19.4	0.34	8.3	3.8	0.78	26.9	16.7	0.53	15.6	11.1	0.6
Glenelg 1	13.4	5.2	0.48	4.3	1.8	0.94	9.4	4	0.8	14.6	7.6	0.42
Glenelg 2	15.2	6.5	0.42	13	5.1	0.69	0.6	0.2	0.99	1.4	0.6	0.92
Mitchell	26	8.8	0.75	9.8	4.6	0.84	2.3	1	0.99	2.5	1	0.99
Mitta Mitta 1	14.7	9.5	0.93	15.7	13.9	0.87	14.6	11.1	0.84	19.2	11.2	0.82
Mitta Mitta 2	4.9	2.2	0.99	14	10.2	0.85	2	1.4	0.998	4.3	2.8	0.99

Examination of the results in the table shows that in most cases, the ANNs performed very well, while ARMAX and FRBS also gave satisfactory results. The simple LTF, although it gave best results for the Mitta Mitta River, its performance, to some extent, was inconsistent. However, due to its simplicity and having fewer parameters, LTF can still be used as an alternative, provided its performance is monitored.

Linear versus nonlinear transfer function methods

Objective Function:

A least-square objective function is used in this study, such that the total square error, the sum of the squares of the errors of the individual equations is minimised. The main reason for the popularity of the least square method is its direct applicability to any model without the need for taking any cognisance of the stochastic property involved (Diskin & Simon, 1977). Kuczera (1983) argues that the use of LS OF can be considered reliable as long as the errors have zero mean and constant variance, the errors are independent of each other, and the errors are normally distributed.

Optimisation Method:

There are many optimisation techniques available. According to Jayasuriya (1991) there are five types of optimisation techniques mostly used for hydrology problems; simplex (non linear), pattern-search method, steepest-descent, Newton, Gauss and Gauss-Marquardt. She tested four optimisation methods for rainfall-runoff models: steepest descent, simplex, pattern search, and Gauss-Marquardt method. The results do not show the superiority of one optimisation method over another. Based on the goodness of fit statistic, she concluded that the pattern-search method performs the best as compare to three other methods. Her comparison did not include the efficiency of the algorithms.

Solomatine (1998) compared the use of global optimisation algorithms in calibration problems. He compared nine different algorithms, CRS2, CRS4, GA, multis, ACCO, ACCOL, ACD and ACDL. His comparison was based on three performance indicator: effectiveness, efficiency (running time), and reliability (robustness). Among nine different algorithms, he concluded that ACCOL and CRS4 show the highest effectiveness, efficiency and robustness. ACCO can be the first choice to obtain a reasonable optimiser assessment. ACDL is proved to be efficient and effective, however, in some cases the reliability needs to be improved. He further stated that GA, CRS2, and Multis might provide reasonable results as well. However, they require more function evaluations. GA also experiences some premature convergence before it reaches the global minimum. M-Simplex can be considered performing well for low dimension problems, however it often converges prematurely in high dimensions.

In this study seven methods are tested, namely Multis, M-Simplex, Genetic Algorithm (GA), Control Random Search (CRS), Improved Control Random Search (CRS4), Adaptive Cluster Covering with Local search (ACCOL), and Adaptive Cluster Descent with Local Search(ACDL).

The Mitchell River was selected for testing, using daily data. The length of the transfer function was 10. The optimisation tool "GLOBE" was used and combined with the LTF program written using Delphi 7. Since the optimisation method is sensitive to the initial value and the range of variables, different initial values and initial variables were tested. Two performance indicators are presented here: the root mean square error (RMSE) and the mean absolute error (MAE).

The results vary considerably for each method; all were found to be very sensitive to the selection of the upper and lower limit of the variables Although the initial values have influence, the results are not as sensitive to them as the range of variables. Therefore a normalization procedure is considered quite important. When the data are scaled, the range of variables can be set relatively constant.

To perform the optimisations three different scaling methods are considered. First, data are normalised by dividing by a constant number (type 1). However, when the range of data between input and output is large, this type of normalisation will also produce a large range. This problem can be solved by using a different denominator for each set of data; the input data are divided by the maximum input and output data are divided by the maximum output (type 2). However, the second type of normalisation also has a drawback. If the minimum value of data is not zero, this normalisation cannot imitate the distribution of the actual data. Therefore, the normalisation of type 3 is applied. In this normalisation, the normalised data are distributed as original data. The type 1 gave unsatisfactory results, while type 2 and 3 performed better. Both type of normalisation produced approximately the same results. As long as the range of variables is small, the optimisation can give reasonable results.

The end results are summarized in Table 2. Both methods perform relatively well, except in the case of the Glenelg River. Even in one case (Mitta Mitta section 1), LTF gives slightly better performance.

Table 2. Comparison between linear and nonlinear transfer function methods

	Linear Transfer Function			Nonlinear Transfer Function		
	RMSE	MAE	R^2	RMSE	MAE	R^2
Glenelg	15.1652	6.5172	0.5230	15.1709	6.5412	0.5226
Mitchell	2244.3106	761.6434	0.7507	2107.2205	724.5321	0.7802
Mitta Mitta 1	1266.5498	824.8829	0.9255	1276.2055	821.8015	0.9244
Mitta Mitta 2	426.0657	192.4022	0.9945	363.9829	138.2525	0.9960
Southfork	0.2697	0.1996	0.9957	0.2665	0.1893	0.9958

Conclusion

Flow forecasting provides crucial information for many problems related to stream management. In practice, there are many cases where the main interest is to determine accurate flow predictions at specific locations. In such situations, system identification models are preferred. In this study, ANNs, FRBS, ARMAX, and LTF were applied to identify stream flow based on upstream information. In general, ANNs showed better results than the others. ARMAX and FRBS also gave satisfactory results. However, ANNs, FRBS, and ARMAX require fundamental knowledge of system identification to set their parameters. LTF, on the other hand, although it may not always perform well, is relatively easy to apply and involves only a few parameters. This paper also discussed both linear and nonlinear models and the selection of the optimisation method. Although the initial values of optimised parameters have influence on the optimisation, they are found to be not as important as the range of variables considered . Scaled data are highly recommended. It is found that both LTF and NLTF models give reasonable results, and the more general nonlinear model is not superior to the linear one.

Acknowledgments:

This study is part of research toward a PhD in the University of Melbourne funded by AusAID .

References:

ASCE, & Govindaraju, R. S. (2000). Artificial Neural Networks in Hydrology. I: Preliminary Concepts. *Journal of Hydrologic Engineering* **5**(2), 115-123.

Abe, S. (1997). *Neural networks and fuzzy systems : theory and applications*. Boston : Kluwer Academic.

Abe, S., & Lan, M.-S. (1995). Fuzzy rules extraction directly from numerical data for function approximation. *IEEE Transactions on Systems, Man and Cybernetics* **25**(1), 119-129.

Bhattacharya, B., & Solomatine, D. (2000). Application of artificial neural networks in stage-discharge relationship. Paper presented at the *4th International Conference on Hydroinformatics*, Iowa, USA.

Birikundavyi, S., Labib, R., Trung, H. T., & Rouselle, J. (2002). Performance of neural networks in daily streamflow forecasting. *Journal of Hydrologic Engineering* **7**(5), 392-398.

Diskin, M. H., & Simon, E. (1977). A procedure for the selection of objective functions for hydrologic simulation models, *Journal of Hydrology* **34**, 129-149.

Esogbue, A. O. (1996). Fuzzy sets modeling and optimization for disaster control systems planning. *Fuzzy Sets and Systems* **81**, 169-183.

Jayasuriya, L. N. N. (1991). *A Methodology for Estimating Yield from Small Ungauged Rural Catchments*. Unpublished PhD Thesis, The University of Melbourne, Melbourne.

Lobbrecht, A., & Solomatine, D. (1999). Control of water levels in polder areas using neural networks and fuzzy adaptive systems. *Journal of Water Industry*.

Minns, A. W., & Hall, M. J. (1998). Rainfall-runoff modelling as a problem in artificial intelligence: experience with a neural network. Paper presented at the *Proceeding 4th National Hydrology Symposium*, Cardiff.

Rao, S. S. (1996). *Engineering optimization : theory and practice*. (3rd ed). New York : Wiley.

Solomatine, D. (1998). Genetic and other global optimization algorithms-comparison and use in calibration problems. Paper presented at the *International Hydroinformatics Conference*, Copenhagen.

Tsykin, E. N. (1985). Multiple nonlinear statistical models for runoff simulation. and prediction, *Journal of Hydrology* **77**, 209-226.