Comparing univariate ARMA and ARFIMA model for forecasting daily streamflows

WEN WANG[1], XI CHEN[1], JUN MA[2] & SHIQING HUO[2]

1 *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China*
wangwen@hhu.edu.cn

2 *Yellow River Conservancy Commission, Hydrology Bureau, Zhengzhou 450004, China*

Abstract The performances of two types of univariate time series model, i.e. the autoregressive moving average (ARMA) model and the long-memory fractionally integrated autoregressive moving average model (ARFIMA) model, for forecasting daily streamflows, are investigated in the present study. Both models are applied to four daily average discharge series of medium-sized watersheds in cold mountainous regions, and built on the basis of deseasonalized series. The result shows that both the ARMA model and the ARFIMA model work well for forecasting short-term daily average discharges, and the performance of the ARFIMA model is generally slightly better than that of the ARMA model. The use of the ARFIMA model is therefore recommended as an alternative to the conventional ARMA model for modelling univariate daily streamflow processes for the cold mountainous regions from the perspective of forecasting.

Key words streamflow process; ARMA model; long memory; ARFIMA model; time series analysis; hydrological forecast

INTRODUCTION

When modelling streamflow processes in ungauged basins, we may sometimes have to deal with univariate streamflow processes. The most commonly used model for modelling univariate hydrological time series is the autoregressive moving average (ARMA) model (see Hipel & McLeod, 1994), which is a short-memory model, whose autocorrelation function (ACF) decays exponentially. However, since the early work of Hurst (1951), it has been well recognized that many time series, in diverse fields of application, may exhibit the phenomenon of long-memory or long-range dependence. In the hydrology community, many studies have been carried out on the test for long-memory in streamflow processes (e.g. Burlando *et al.*, 1996; Montanari *et al.*, 1997; Rao & Bhattacharya, 1999; Ooms & Franses, 2001; Wang *et al.*, 2007). Burlando *et al.* (1996) and Montanari *et al.* (1997) applied the ARFIMA mode to the modelling of the daily water inflow to Lake Maggiore, and showed that the fit with an ARFIMA model is much better than the one provided by more traditional ARIMA models. But their result is based on the one single streamflow process, and the model performance comparison is based on the model-fitting residuals, without considering the performance for out-of-sample forecasting.

In the present study, the performances of ARMA and ARFIMA model for making 1- to 7-days ahead out-of-sample forecasts are compared for four daily streamflow processes in the cold mountainous regions of British Columbia in Canada.

METHODS OF MODELLING UNIVARIATE TIME SERIES

Autoregressive moving average (ARMA) model

The ARMA model is one of the most important and highly popularized time series models, and has a long history of being applied to streamflow forecasting problems (see Hipel & McLeod, 1994). An ARIMA(p, q) model is given by:

$$(1-\phi_1 B-\cdots-\phi_p B^p)x_t=(1-\theta_1 B-\cdots-\theta_q B^q)\varepsilon_t \qquad (1)$$

where B is the backshift operator, that is, $Bx_t = x_{t-1}$; ε_t is a white noise process with zero mean and variance σ^2. When $q = 0$, then the ARMA(p,q) model reduces to an AR(p) model.

Fractionally integrated autoregressive moving average (ARFIMA) model

The ARFIMA model has not been widely accepted in the hydrology community so far. The general form of an ARFIMA(p,d,q) model is given by:

$$(1-\phi_1 B-\cdots-\phi_p B^p)\nabla^d x_t=(1-\theta_1 B-\cdots-\theta_q B^q)\varepsilon_t,\ |d|<0.5 \tag{2}$$

where $\nabla^d = (1-B)^d$ can be given by means of the binomial expansion, and d denotes the fractional differencing parameter. A larger value of d indicated a stronger intensity of long-memory. When $d = 0$, equation (2) reduces to the classical ARMA(p,q) model. The conventional ARMA process is often referred to as a short memory process because its autocorrelation function (ACF) $\rho(h)$ decays exponentially as $h\to\infty$, whereas the long-memory ARFIMA process is characterized by a hyperbolically decaying autocorrelation function. In fact, it decays so slowly that the autocorrelations can not be summed.

Many methods for estimating the long-memory parameter d are available, which can be categorized into two classes: the heuristic and the parametric methods. The heuristic methods are useful to find a first estimate of d and to test if a long-range dependence in the data exists. The parametric methods obtain consistent estimators of d via maximum likelihood estimation of parametric long-memory models. They give a more accurate estimate of d, but generally require knowledge of the true model. The S-Plus function *arima.fracdiff* fits ARFIMA models to univariate time series data through the approximate Gaussian maximum likelihood algorithm of Haslett & Raftery (1989).

When forecasting from the univariate ARFIMA(p,d,0) model, we can express the process $(1-\phi_1 B-\cdots-\phi_p B^p)\nabla^d x_t=\varepsilon_t$ in the form of an infinite autoregression, where $\nabla^d=\sum_{j=0}^{\infty}\left[\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}B^j\right]$. The infinite autoregression may be truncated at a suitably high lag (say, 200), and the forecasts are then made recursively for multi-steps ahead.

DAILY STREAMFLOW DATA USED

Daily average discharge series of four watersheds in British Columbia, Canada, are used in the present study. All the four watersheds are located in cold mountainous regions where snow-melt water makes a considerable contribution to runoff. The description of the eight gauging stations is given in Table 1. The data are selected data so that the series should be approximately stationary, at least by visual inspection, and the time-span of the data ranges from 23 to 30 years. All the four daily series exhibit strong and regular seasonal variations.

MODEL BUILDING FOR DAILY STREAMFLOW SERIES

ARMA model

The procedure of fitting deseasonalized ARMA models to daily flow series includes two steps: deseasonalization and ARMA model construction. We first log-transform the data series, and

Table 1 Daily streamflow data used.

Gauging stations	Area (km^2)	Latitude	Longitude	Period	Avg. discharge (m^3/s)	ACF(1)
Canoe River below Kimmel Creek	298	52.728	–119.408	1972–1994	14.5	0.9294
Columbia River Near Fairmont Hot Springs	891	50.324	–115.863	1946–1975	11.1	0.9676
Columbia River at Nicholson	6660	51.244	–116.912	1933–1962	107.5	0.9778
Fraser River at Mcbride	6890	53.286	–120.113	1959–1988	197.3	0.9582

Note: The ACF(1) is calculated based on log-transformed and deseasonalized series.

deseasonalize the series by subtracting the seasonal (e.g. daily or monthly) mean values and dividing by the seasonal standard deviations of the log-transformed series. To alleviate the stochastic fluctuations of the daily means and standard deviations, we smooth them using the first eight Fourier harmonics.

In the present study, for simplicity, we considered using the AR(p) model, that is, $q = 0$ in the ARMA(p,q) model, when using the ARMA model. The AR(p) model structure (listed in Table 2) for all the daily flow series are determined according to the Alkaike Information Criterion (AIC). The goodness of fit of the models are examined by inspecting the ACF of the residuals from the models, which indicates that there is generally no significant autocorrelation left in the residuals from all the AR models.

Table 2 AR(p) and ARFIMA(p,d,0) model structure for the streamflow series.

Gauging stations	AR(p)	ARFIMA(p,d,0)	
	p	p	d
Canoe River below Kimmel Creek	10	3	0.31
Columbia River at Nicholson	35	3	0.4392
Columbia River Near Fairmont Hot Springs	6	8	0.4213
Fraser River at Mcbride	9	7	0.1886

ARFIMA model

Before fitting the ARFIMA model to a daily flow series, the existence of long-memory should be confirmed. The first heuristic indication of the presence of long-memory comes from the slow decaying ACF structure of the four deseasonalized daily series, as shown in Fig. 1. In addition, it has been shown (Wang *et al.*, 2007) that all the daily flow series under consideration have significant long-memory.

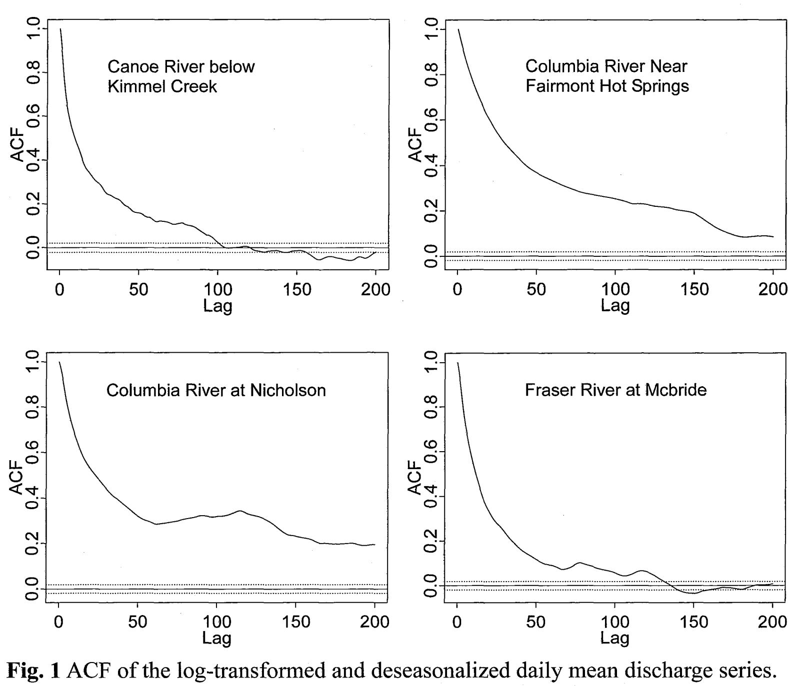

Fig. 1 ACF of the log-transformed and deseasonalized daily mean discharge series.

Similar to constructing the ARMA model to the daily streamflow series, the procedure of fitting ARFIMA models also includes two steps, i.e. firstly deseasonalize the data, and then construct the ARFIMA model. We only consider using the ARFIMA model of the structure (p,d,0). The deseasonalization is done in the same way as that for building the ARMA model. The order of the AR component and estimates of d for the fitted ARFIMA model of all the deseasonalized daily series are also listed in Table 2. The order p of the AR component in the ARFIMA model is determined according to AIC, and the parameters of the AR component and value of d are estimated using the function *arfima.fracdiff* in S-Plus. The examination of the ACF of the residuals from the ARFIMA models shows that no significant autocorrelations are observed in the residual series.

COMPARISON OF MODEL PERFORMANCES

Model performance measures

There is extensive literature on model forecasting evaluation indices. Despite its crudeness and identified weaknesses, the Coefficient of Efficiency (*CE*) introduced by Nash & Sutcliffe (1970) is still one of the most widely used criteria for the assessment of model performance. However, *CE* is a global measure for comparing the predicted value with the overall mean value; it is not efficient enough to evaluate the predictions for those series whose mean values change with seasons, which is almost always the case for hydrological processes. Therefore, a Season-Adjusted Coefficient of efficiency (*SACE*) (Wang *et al.*, 2004) is used here for evaluating the model performance. *SACE* is calculated by:

$$SACE = 1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q}_m)^2} \tag{3}$$

where Q_i is the observed value; $\hat{Q}_i$ is the predicted value; $\bar{Q}_m$ is the mean value of season m; $m = i \bmod S$ (mod is the operator calculating the remainder), ranging from 0 to $S - 1$; and S is the total number of "*season*". Note that, a "*season*" here is not a real season. It may be a month or a day, depending on the timescale of the time series. For daily streamflow series, one season is one day. Furthermore, for daily series, the mean values of 365 days are smoothed using the first eight Fourier harmonics.

Besides the above two performance measures, the root mean squared error (*RMSE*) is also used in the present study.

Comparison of model performances

For each daily flow series, we make 1- to 7-day ahead forecasts for the last 5 years of data to evaluate the performance of the fitted model. Both the AR and ARFIMA models are built on a rolling-forward basis. For instance, we use the daily flow data of the Columbia River near Fairmont Springs in 1946–1970 to fit an AR(6) model, and make forecasts for 1971; then use the data in 1946–1971 to fit another AR(6) model, and make forecasts for 1972; etc. Data are log-transformed and deseasonalized before being used. The evaluation results for 1- to 7-day ahead forecasts are reported in Table 3. According to the values of performance measures, we see that:

(a) The measure CE shows that both models give good forecasts (i.e. the value of CE > 0.8) for 7-day ahead forecasts. But SACE shows that, although for all the cases both models can give forecasts significantly better than seasonal mean values for all lead times (i.e. SACE > 0), satisfying forecasts (i.e. SACE > 0.8) are available only for 1- to 3-days ahead for the Columbia River at Nicholson and Columbia River near Fairmont Hot Springs, for 1-day ahead for the Fraser River at Mcbride, and no satisfying forecasts are available, even for 1-day

Table 3 Multi-step forecasting performances of ARMA and ARFIMA models.

Lead time (days)	Measure	Model	1	2	3	4	5	6	7
Canoe River below Kimmel Creek	RMSE	ARMA	4.119	6.044	6.696	7.151	7.419	7.543	7.607
		ARFIMA	4.122	6.046	***6.692***	***7.143***	***7.406***	***7.525***	***7.583***
	CE	ARMA	0.944	0.879	0.851	0.830	0.817	0.811	0.808
		ARFIMA	0.944	0.879	0.851	0.831	0.818	0.812	0.809
	SACE	ARMA	0.719	0.394	0.256	0.152	0.087	0.056	0.040
		ARFIMA	0.718	0.394	0.257	0.154	0.090	0.061	0.046
Columbia River Near Fairmont Hot Springs	RMSE	ARMA	0.851	1.475	1.984	2.423	2.795	3.113	3.399
		ARFIMA	***0.848***	***1.468***	***1.971***	***2.405***	***2.774***	***3.089***	***3.372***
	CE	ARMA	0.995	0.986	0.975	0.963	0.951	0.939	0.927
		ARFIMA	0.995	0.986	0.975	0.963	0.951	0.940	0.928
	SACE	ARMA	0.974	0.921	0.858	0.788	0.717	0.649	0.582
		ARFIMA	0.974	0.922	0.859	0.791	0.722	0.655	0.588
Columbia River at Nicholson	RMSE	ARMA	6.157	11.962	17.263	21.944	25.997	29.498	32.469
		ARFIMA	***6.102***	***11.844***	***17.084***	***21.729***	***25.723***	***29.220***	***32.220***
	CE	ARMA	0.998	0.991	0.982	0.970	0.958	0.946	0.935
		ARFIMA	0.998	0.991	0.982	0.971	0.959	0.947	0.936
	SACE	ARMA	0.983	0.936	0.868	0.786	0.700	0.613	0.531
		ARFIMA	0.983	0.938	0.870	0.790	0.706	0.620	0.538
Fraser River at Mcbride	RMSE	ARMA	30.893	52.331	66.262	75.012	81.011	85.403	88.747
		ARFIMA	***30.885***	***52.286***	***66.153***	***74.813***	***80.684***	***84.939***	***88.178***
	CE	ARMA	0.979	0.939	0.902	0.875	0.854	0.838	0.825
		ARFIMA	0.979	0.939	0.903	0.875	0.855	0.839	0.827
	SACE	ARMA	0.888	0.677	0.483	0.337	0.227	0.141	0.072
		ARFIMA	0.888	0.678	0.485	0.341	0.233	0.150	0.084

Note: the bold italic values denote better performances of ARFIMA model against ARMA model in terms of *RMSE*.

ahead for the Canoe River. The results illustrate that the measure CE exaggerates the performance of the forecasting models in cases where the streamflow processes have strong seasonality, whereas the SACE is a better alternative.

(b) Among the four cases, the ARFIMA model outperforms the ARMA model in three cases (the Columbia River at Nicholson, the Columbia River near Fairmont Hot Springs and the Fraser River at Mcbride) for all lead times, and outperforms the ARMA model for 3-days and longer lead time forecasts for the Canoe River, indicating its good modelling performance for the four cases under consideration in the present study.

(c) According to the performance measures CE and SACE, we see that the ARFIMA model and ARMA perform much better for the streamflows measured at Nicholson and near Fairmont Hot Springs for the Columbia River than those measured below Kimmel Creek for the Canoe River and at Mcbride for the Fraser River. This is probably because the streamflow processes for the Columbia River not only have significant snow-melt water contributions, but are also regulated by two lakes, i.e. Lake Koocanusa and Columbia Lake, which lead to stronger temporal dependence and stronger intensity of long-memory, as indicated by larger values of ACF(1) (Table 1) and larger values of the long-memory parameter (Table 2) than those for the Fraser River and Canoe River.

The scatter plot of the one-day ahead predicted values *versus* the observed values for the four streamflow series with the ARFIMA models are shown in Fig. 2. From Fig. 2 we see that there is no systematic bias in the forecasts, and the results are generally satisfying, especially for the Columbia River.

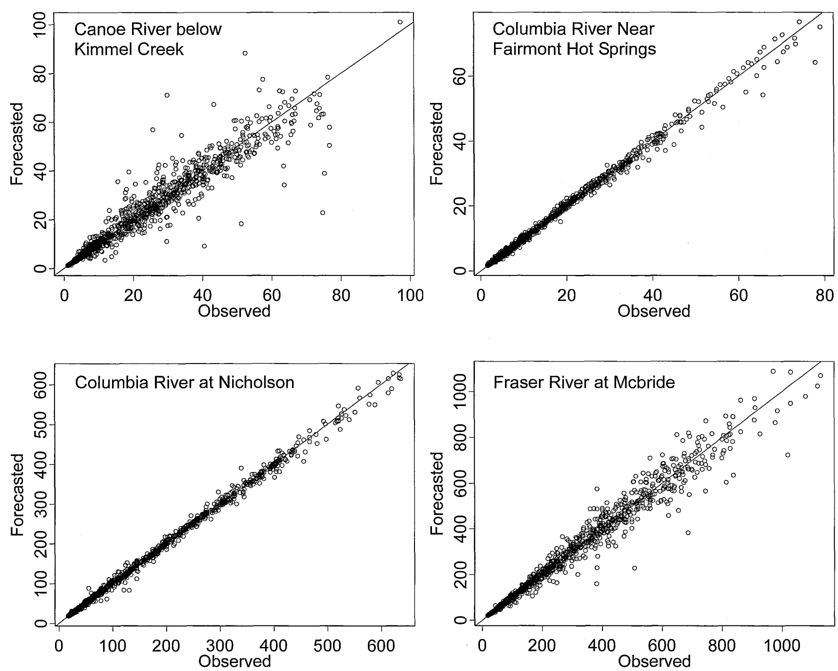

Fig. 2 Observed discharges *versus* 1-day ahead forecasts with ARFIMA models.

CONCLUSIONS

When modelling streamflow processes in ungauged basins, we may sometimes have to deal with univariate streamflow processes. In the present study, two univariate time series models, i.e. the autoregressive moving average (ARMA) model and the fractionally integrated ARMA (ARFIMA) model are applied to four medium-sized watersheds. The result shows that univariate time series models work very well for forecasting daily streamflow processes of the watersheds in cold mountain areas where snow-melt water makes a considerable contribution to runoff. According to the values of coefficient of efficiency (CE), satisfying 7-day ahead forecasts (CE $>$ 0.8) are available for the four cases. According to the season-adjusted coefficient of efficiency (SACE), satisfying forecasts (i.e. SACE $>$ 0.8) can be made for a lead time of three days for two out of four cases (i.e. for the Columbia River at Nicholson and the Columbia River near Fairmont Hot Springs), and one-day ahead satisfying forecasts are available for one case (the Fraser River at Mcbride). And the ARFIMA model built on the basis of deseasonalized daily streamflow series outperforms the ARMA model in all cases and all lead times except for 1- and 2-day ahead forecasts for the Canoe River. The use of the ARFIMA model can therefore be recommended as an alternative to the conventional ARMA model for modelling univariate daily streamflow processes from the perspective of forecasting for the watersheds where snow-melt water contributes considerably to runoff.

Acknowledgements The financial support from the National Science Foundation of China (Project no. 40771039) and the Research Fund for the Doctoral Program of Higher Education (Project no. 20060294003) are gratefully acknowledged.

REFERENCES

Burlando, P., Montanari, A. & Rosso, R. (1996) Modelling hydrological data with and without long memory. *Meccanica* **31**(1), 87–101.

Haslett, J. & Raftery, A. E. (1989) Space-time modelling with long-memory dependence: assessing Ireland's wind power resource (with discussion). *Applied Statistics* **38**, 1–50.

Hipel, K. W. & McLeod, A. I. (1994) *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam, The Netherlands.

Montanari, A., Rosso, R. & Taqqu, M. S. (1997) Fractionally differenced ARIMA models applied to hydrological time series: Identification, estimation, and simulation. *Water Resour. Res.* **33**(5), 1035–1044.

Nash, J. E. & Foley, J. J. (1982) Linear models of rainfall–runoff systems. In: *Proc. Int. Symp. on Rainfall-Runoff Modelling*. (ed. by V. P. Singh), 51–66. Water Resources Publications, Mississippi State University, USA.

Ooms, M. & Franses, P. H. (2001) A seasonal periodic long memory model for monthly river flows. *Environ. Modelling & Software* **16**, 559–569.

Rao, A. R. & Bhattacharya, D. (1999) Hypothesis testing for long-term memory in hydrological series. *J. Hydrol.* **216**(3–4), 183–196.

Wang, W., van Gelder P. H. A. J. M., Vrijling, J. K. & Ma, J. (2004) Predictability of streamflow processes of the Yellow River. In: *Proc. 6th Int. Conf. on Hydroinformatics*, 1261–1268. World Scientific, Singapore.

Wang, W., van Gelder P. H. A. J. M., Vrijling, J. K. & Chen, X. (2007) Detecting long-memory: Monte Carlo simulations and application to daily streamflow processes. *Hydrol. Earth Syst. Sci.* **11**, 851–862.