ELSEVIER

PII: S0266-9838(96)00050-5

Environmental Modelling & Software, Vol. 12, No. 1, pp. 93–104, 1997

1364-8152/97 $17.00 + 0.00

A combined deterministic and self-adaptive stochastic algorithm for streamflow forecasting with application to catchments of the Upper Murray Basin, Australia

S. Yu. Schreider,[a] A. J. Jakeman,[a*] B. G. Dyer,[b] R. I. Francis[b]

[a]Centre for Resource and Environmental Studies, Australian National University, Canberra 0200, ACT, Australia
[b]Murray–Darling Basin Commission, Canberra 2601, ACT, Australia

Abstract

This paper describes the results of runoff modelling for nine catchments of the Upper Murray Basin (Basin 401) of the Murray–Darling Drainage Division (MDDD), Australia. The work aimed firstly to provide adequate models for long-term streamflow prediction in nine catchments of this Basin feeding the Hume and Dartmouth reservoirs. The development and testing of flow forecasting algorithms for operational management by the Murray–Darling Basin Commission was another purpose of the work reported here.

The conceptual lumped parameter rainfall-runoff model IHACRES (Jakeman *et al.*, 1990, 1993; Jakeman and Hornberger, 1993) was selected as the modelling tool for streamflow prediction in the catchments.

The conceptual rainfall-runoff model IHACRES (with a snow melt/formation module in snow-affected catchments) and a self-adaptive linear filtering approach for the IHACRES residuals were combined and applied for forecasting daily streamflow in the Upper Murray Basin catchments. Different orders of AutoRegressive Integrated Moving Average (ARIMA) models for the residuals were considered in order to select the most appropriate forecasting algorithm. Linear filtering of the conceptual model residuals provides considerable improvement in forecasting for both low and high values of streamflow for developing the operational streamflow forecast system. © 1997 Elsevier Science Ltd.

Keywords: Streamflow modelling; operational streamflow forecasting; ARIMA algorithm

Software availability

Program title:	**pc-ihacres**
Contact address:	Dr Tony Jakeman, CRES, Australian National University, Canberra 0200 ACT, Australia
Available:	August, 1996
Hardware:	IBM compatible 386 or higher (8 MB RAM)
Software required:	Windows 3.1
Cost:	$US 500.

1. Introduction

1.1. Background

Overton and Meadows (1976) define three basic categories of streamflow forecasting methods based upon: (1) regression analysis, (2) time series analysis and (3) flow frequency analysis. The regression-type analysis uses an optimisation procedure where a causal model is structured as a linear or slightly non-linear approximation. Least squares (or modified least squares) regression serves as a tool for this approximation of modelled values against empirical data. Although not strictly a regression model, the IHACRES

*Corresponding author.

model can be classified as belonging to this type of method.

Time series analysis, applied in the present work for operational streamflow forecasting, "analyses a continuous time series of runoff and draws an inference as to the underlying generating mechanism" (Overton and Meadows, 1976). A comprehensive review of the literature on applications of time series analysis techniques in hydrology is out of the scope of the present work but several publications must be mentioned here. Two classical monographs are specially devoted to the problem of time series analysis (Box and Jenkins, 1976; Bras and Rodriguez-Iturbe, 1985) where a detailed description of the ARIMA algorithm used in this work can be found.

Frequency analysis is entirely probabilistic, and is usually used for catchments where streamflow data are continuously recorded but no rainfall records are available. Primarily it is applied in order to evaluate the probabilities of extreme flow values, both high and low.

The self-adapting model applied to flash flood forecasting is described in a range of publications (Wood, 1989; Georgakakos, 1987). An approach combining a deterministic seven-parameter model SM2 with ARIMA was applied by Jamieson *et al.* (1972). The residual variance of the composite model was significantly less than that attainable by using the deterministic model only. A similar methodology to that here was suggested by Brath and Rosso (1993). The main difference between their approach and the method applied here is that in our approach, calibration of the conceptual model is made once in a two-year period, then its parameters are considered unchanging for the period of observation. The linear filtering algorithm is then applied to the model residuals. In Brath and Rosso (1993), the stochastic algorithm was applied solely.

Whereas ARIMA modelling considers the streamflow time series solely in order to provide the optimal forecast, some linear filtering algorithms analyse the time series of the precipitation and streamflow together. The filtering of input precipitation data using an ARMAX modification of the ARIMA method is described, for example, in Karlsson and Yakowitz (1987). The ARMAX (X means the use of *eXogenous* variables, precipitation, for instance) algorithm is a compromise between prediction techniques based on a deterministic relation of streamflow with rainfall input (affected by errors of measurement), and ARIMA models applied solely, where no information on precipitation is used.

More sophisticated techniques for streamflow forecasting are based on Kalman filtering (Kalman, 1960). An example of the application of this technique can be found in Sen (1991), where orthogonal Walsh series, used for describing the periodic component, were combined with the Kalman filter. This method was applied to the prediction of monthly flow for two catchments in Turkey and the United States and for monthly rainfall prediction in Saudi Arabia. The Kalman filter technique is used in the European Flood Forecasting Operational Real-Time System (EFFORTS), widely applied to water resource management in Europe and worldwide (Todini, 1996). This method is based on two linear, interactive Kalman filters, one in the space of the state vector and another in the space of the parameters.

Another new technique widely used for streamflow forecasting is the Nearest Neighbouring Method (NNM). The NNM, closely related to techniques of non-linear dynamics, has been developing quickly in the last decade (Olason and Watt, 1986; Mack and Rosenblatt, 1979; Yakowitz, 1987; Yakowitz and Karlsson, 1987; Galeati, 1990). This method is based on the assumption that the streamflow time series is an output of a deterministic dynamic system with stochastic noise. Kember and Flower (1993) reported that the NNM provides improvement in forecasting compared with the ARIMA model.

In conclusion, it should be explicitly stated that the ARIMA algorithm was chosen among other filtering methods because, for a modest investment, it provides an effective combination with deterministic predictions of streamflow (using the IHACRES model here).

1.2. The Upper Murray Basin description

The Upper Murray Basin is located in the south-eastern part of the MDDD and covers 15,300 km^2 of territory in the states of Victoria (the Mitta-Mitta River catchment and the left bank of the Murray River with a total area of 10,000 km^2) and New South Wales (Fig. 1). River flows are regulated by Hume and Dartmouth reservoirs operated by the Murray–Darling Basin Commission (MDBC). The right bank of the Murray catchment belongs to the New South Wales part of the Basin. The largest water contributors in this area are the right side of the Upper Murray River at Biggara (1165 km^2) and the Tooma River catchment at Pine Grove (1819 km^2). These two catchments drain the western slopes of the Snowy Mountains which is the highest region in Australia (Mt Kosciusko, 2228 m AHD). The outlet of an inter-basin water transfer, via the Snowy Mountains Hydroelectric Scheme, with 580,000 ML mean annual transfer, is located in this part of the Basin. The Jingellic Creek catchment is located in the north of the New South Wales part of the Basin.

The climatology of this Basin is also very heterogeneous and topography strongly influences the spatial variations in the climate. Average annual precipitation reduces with decreasing elevation from more than 2400 mm per year in Mt Bogong on the border with the Kiewa Basin to about 700 mm in the northern part of the Basin along the Murray River. Snow provides a considerable amount of water to the Basin. The large

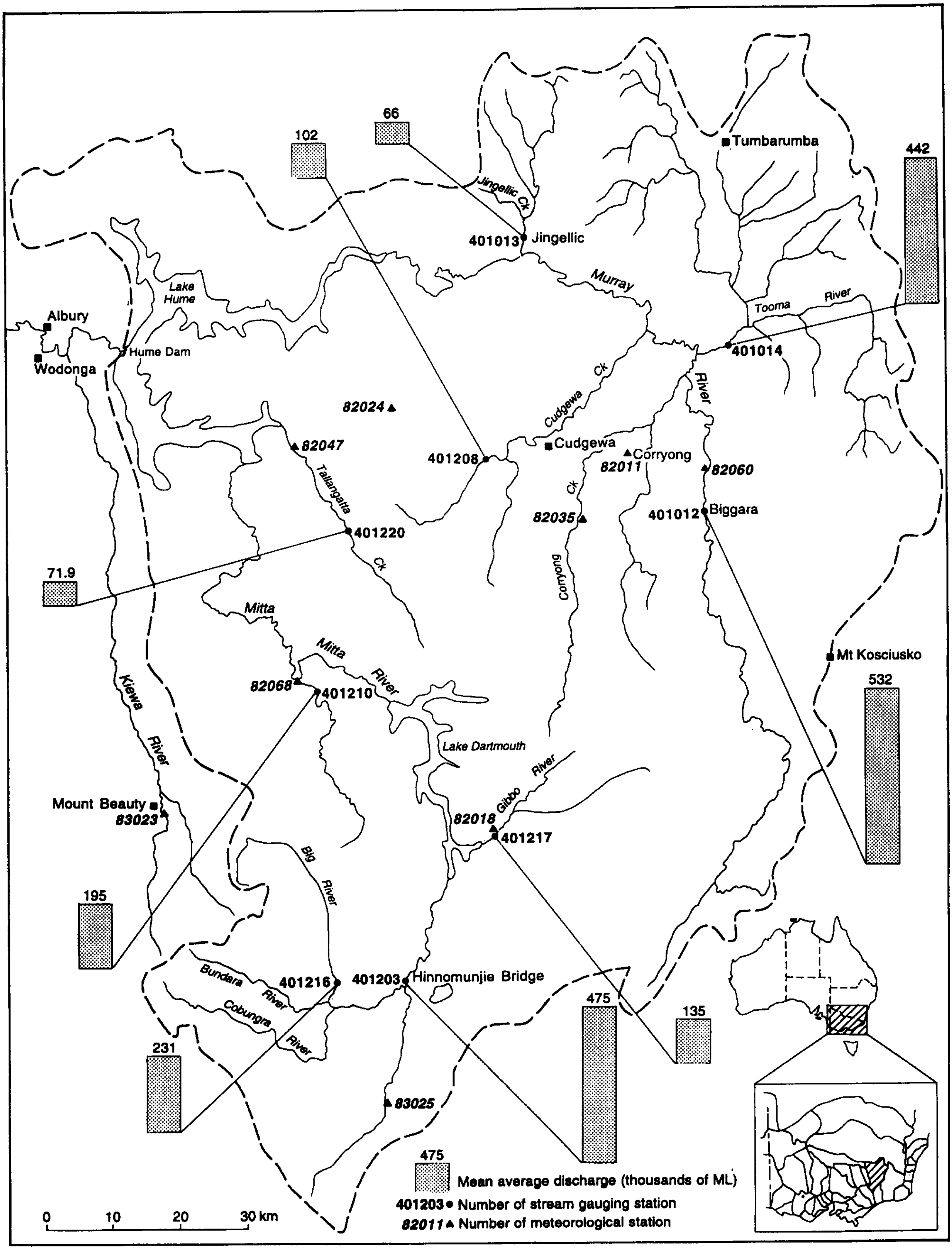

Fig. 1. River network, meteorological and discharge stations for the catchments under consideration in the Upper Murray Basin.

flows in October are attributed to melting snow in its high area.

About 80% of the area of the Upper Murray Basin is forested, although all the major valleys which lie in its north have been cleared for agriculture (Water Victoria, 1989). Water use in this basin is small (4830 ML per year on average). However, it is one of the major contributors of water resources in the MDDD.

The streamflow information for the nine catchments selected for analysis here are used by the MDBC for calculating inlet flows to the Hume and Dartmouth reservoirs. The mean annual discharge and areas of these catchments are presented in Table 1.

1.3. The model

A general description of the IHACRES model can be found in Jakeman *et al.* (1990) and Jakeman and Hornberger (1993). The description of the particular model structure applied here was provided in Schreider *et al.* (1996a). The IHACRES rainfall-runoff model is a dynamic lumped parameter model consisting of two modules: a non-linear loss module which transforms measured rainfall to excess rainfall; and a linear module defined as a recursive relation at time step k (daily here) for modelled streamflow y_k, calculated as a linear combination of its antecedent values and excess rainfall. A more detailed description of the linear module of the IHACRES model is given in Section 4.2.

The non-linear loss module allows one to take into account the effect of antecedent weather conditions on the current status of catchment storage wetness index s_k and vegetation conditions, and evapotranspiration effects, in order to calculate the excess rainfall from the measured precipitation taking into account information on current temperature values.

The snow melt/accumulation module described in Schreider *et al.* (1996c) for modelling daily processes is based on modification of the empirical degree-day approach developed by Whetton *et al.* (1996) for modelling snow melt/accumulation on a monthly timestep. The main advantage of the suggested approach is that it does not require any additional input data except daily temperature and precipitation, which is especially useful for modelling of snow processes in regions with a lack of regular snow observations.

The major model-fit criterion used in the present work is the model efficiency statistic E, or proportion of observed streamflow explained by the model, and is defined by the formula:

$$E = 1 - \Sigma(y_i - y_i')^2/\Sigma(y_i - y_{\text{mean}})^2,$$

where y_i is the daily observed streamflow, y_i' is the daily modelled streamflow and y_{mean} is the long-term mean value of observed streamflow. The closer the efficiency value is to one, the better the fit provided by the model (Nash and Sutcliffe, 1970).

2. Results of runoff modelling for the snow-free catchments

The Tallangatta and Cudgewa Creek catchments were considered as snow free. Tallangatta Creek was modelled using the streamflow data from McCallums gauging station (401220) for the period 1976–90 and the Bullioh station (401218) for the period 1954–75; precipitation data were from the meteorological station at Tallangatta (82047). The streamflow data for Cudgewa Creek were taken from the station at Berringama (401208) and precipitation from the station at Corryong (82011). Temperature data for both catchments were taken from the Corryong station. These two catchments were calibrated on 10 (practically) non-overlapping calibration periods (CPs), each of two years. Successful calibrations, with $E > 0.700$, were obtained for five CPs for Tallangatta Creek and three

Table 1
Characteristics of modelled catchments in the Upper Murray Basin and meteorological stations used (Fig. 1)

Streamgauge station number	River and station location	Mean annual discharge (ML)	Area (km^2)	Meteorological stations (precipitation/temperature)
401203	Mitta-Mitta River at Hinnomunjie	475,000	1533	83025/83023
401220	Tallangatta Creek at McCallums	71,900	464	82047/82011
401208	Cudgewa Creek at Berringama	102,000	350	82011/82011
401012	Murray River at Biggara	532,000	1165	82035/82011
401217	Gibbo River at Gibbo	135,000	389	82018/82011
401210	Snowy Creek at Granite Flat	195,000	407	82068/82011
401013	Jingellic Creek at Jingellic	66,000	328	82024/82011
401014	Tooma River at Pine Grove	442,000	1819	82060/82011
401216	Big River U/S of Joker Creek	231,000	356	83023/83023

Table 2
Model efficiency values E for calibration of the IHACRES model for the Tallangatta and Cudgewa Creeks catchments

Catchment and station number CP	Tallangatta Creek 401220	Cudgewa Creek 401208
1 25/02/72–24/02/74	0.789	–
2 24/02/74–23/02/76	0.837	–
3 4/03/76–3/03/78	–	0.768
5 10/12/79–9/12/81	0.903	0.872
6 9/12/81–8/12/83	0.858	–
8 28/12/85–17/12/87	0.809	0.840

– denotes model diverged ($E < 0.700$).

CPs for Cudgewa Creek (see Table 2). The model efficiency statistics for the best calibration periods were 0.903 for Tallangatta Creek and 0.872 for Cudgewa Creek. A simulation (or validation) test was performed. That is, the values of the parameters τ_w, f, c, in the non-linear module and the coefficients in the linear module of the IHACRES model, optimised during the calibration runs, were used for modelling the streamflow using the rainfall and temperature series for the whole period where continuous records of temperature, precipitation and streamflow were available. The efficiency statistics obtained for these two catchments for the simulation test were 0.611 and 0.638, respectively.

The Jingellic Creek catchment was considered as snow free to a first approximation. It was modelled using, as an input, precipitation records from the single meteorological station at Koetong (82024). Streamflow data were taken from station 401013 at Jingellic. Temperature data were also taken from the Corryong station. This catchment was calibrated on eight practically non-overlapping calibration periods, each of two years. Calibrations, with efficiency coefficients $E > 0.700$, were obtained for three CPs (see Table 3). The

Table 3
Model efficiency values E for calibration of the IHACRES model for the Jingellic Creek catchment

CP	Model efficiency E
5 20/12/79–19/12/81	0.709
7 19/12/83–18/12/85	0.703
8 28/12/85–26/12/87	0.797

calibration results obtained on CP 8 are shown in Fig. 2. The divergence of the model on the other CPs is related to defects in streamflow records of gauging station N 401013. Jingellic Creek is an ephemeral river. This is an explanation for the long-term simulation test providing poor results ($E = 0.533$). The selected structure of the IHACRES model cannot approximate streamflow when the flow in this creek reaches zero. However, such periods are relatively short for this catchment and its calibration, using a version of the model extended to ephemeral rivers (Ye et al., 1995), does not seem justified.

3. Results of runoff modelling for the snow-affected catchments

Model calibration for the snow affected Mitta-Mitta catchment was described in Schreider et al. (1996c). The base meteorological stations 83025 at Omeo and 83023 at Mt Beauty were selected for precipitation and temperature, respectively. Successful calibrations were obtained and the calibration results, in terms of model efficiency, are summarised in Table 4. The simulation results are summarised in Table 5. The results described above show that the IHACRES model, applied to the snow-affected catchments and combined with the snow melt/accumulation module, provides a fit of the model to the observed data of about the same quality as this model (without this module) applied in snow-free basins; cf. Schreider et al. (1996a,b) where the results of the IHACRES application to the practically snow-free Goulburn and Ovens Basins are described.

Model calibration for the snow-affected catchments of the Upper Murray at Biggara, Gibbo River and Snowy Creek was performed with daily time series of:

(1) the equivalent precipitation, estimated by the snow melt/accumulation module (Nariel Creek meteoro-

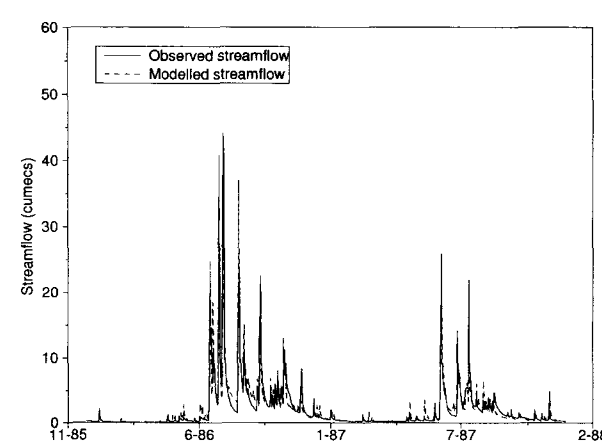

Fig. 2. Observed (solid line), modelled (dashed line) streamflow (cumecs) for CP 8 (1986–1987) for the Jingellic Creek catchment.

Table 4
Model efficiency values E for calibration of the IHACRES snow-runoff model for the Mitta-Mitta catchment

Catchment and station number CP	The Mitta-Mitta River 401203
1 4/02/66–3/02/68	0.827
2 5/01/68–4/01/70	–
3 23/02/70–22/02/72	0.718
4 24/03/72–23/03/74	0.848
5 14/03/74–13/03/76	–
6 2/02/76–1/02/78	0.727
7 2/02/78–1/02/80	–
8 11/02/80–10/02/82	0.859
9 10/02/82–9/02/84	0.910

– denotes model diverged for this calibration period ($E < 0.700$).

Table 5
Simulation statistics with nine two-year calibrated models over the whole period of observation (1965–1985) for the Mitta-Mitta catchment: E and bias (mean daily error in cumecs). Bold values denote the best selected model

Catchment and station number Model number	The Mitta-Mitta River 401203 *E*/bias
1	0.627/0.13
2	–
3	0.649/–0.76
4	**0.669/0.50**
5	–
6	–
7	–
8	0.605/4.13
9	0.681/0.87

– denotes model diverged for this calibration period (cf. Table 4).

logical station 82035 was selected as the basis for the Upper Murray catchment, the Gibbo River Park station 82018 for the Gibbo River and the Mitta-Mitta Forestry station 82068 for Snowy Creek);
(2) temperature, interpolated for each grid cell of the catchment and then integrated over this catchment (Corryong meteorological station 82011 was selected as the basis for the whole region under consideration); and
(3) streamflow during the period 1973–1987.

The period 1973–1987 was divided into seven CPs, each with a duration of two years. The selected CPs do not overlap substantially, with the one exception of CP 7 for the Gibbo River because of lack of streamflow data after 30 June 1986. Successful calibrations were obtained and the calibration results, in terms of model efficiency, are summarised in Table 6.

In order to check the consistency of the results obtained, simulation runs were performed over the 10-year period 1977–1987 with one of the calibrated models for each catchment. The efficiency coefficients E and mean daily bias for these simulation tests are shown in Table 7.

Model calibration for the snow-affected catchments of the Tooma River at Pine Grove and the Big River upstream of Joker Creek was performed with daily time series of:

(1) the equivalent precipitation, estimated by the snow melt/accumulation module (Towong Upper meteorological station N 82060 was selected as the basis for the Tooma River and Mt Beauty station 83023 for the Big River);
(2) temperature, interpolated for each grid cell of the catchment and then integrated over this catchment (Corryong meteorological station 82011 was selected as the basis for the Tooma River catchment and the Mt Beauty station 83023 for the Big River); and
(3) streamflow during the period 1973–1987 for Tooma River and 1965–1985 for Big River.

For the Tooma River, the period 1973–1987 was divided into seven CPs, each with a duration of two years. The selected CPs do not overlap substantially. Successful calibrations were obtained for five of seven CPs and the calibration results, in terms of model efficiency, are summarised in Table 8.

The Big River is a tributary of the Mitta-Mitta catchment. Its watershed is located in the north-western part of the Mitta-Mitta catchment. The period 1965–1985, when meteorological data were available, was subdivided into nine non-overlapping calibration periods each of two years, the same as for the Mitta-Mitta River. Successful calibrations were obtained for six of the nine CPs and the calibration results are summarised in Table 9. Figure 3 shows the calibration results for this catchment obtained for the CP 1 (4/02/1966–3/02/1968).

In order to check the consistency of the results obtained, simulation runs were performed. For the Tooma River it was accomplished over the 14-year period 1973–1987 for each of five models with calibration efficiencies greater then 0.700. For two models, calibrated on CP numbers 2 and 4, the simulation results provided efficiencies higher than 0.650. The efficiency coefficients E and mean daily bias for these simulation tests are shown in Table 10. A simulation test for the Big River was performed over the 20-year period, 1965–1984, for every model obtained. The

Table 6
Model efficiency values E for calibration of the IHACRES snow-runoff model for the Upper Murray, Gibbo River and Snowy Creek catchments

Catchment and station number CP	The Upper Murray River 401012	The Gibbo River 401217	Snowy Creek 401210
1 01/01/73–31/12/74	–	0.799	0.741
2 01/01/75–31/12/76	–	0.710	–
3 06/01/77–06/01/79	–	–	–
4 420/11/78–19/11/80	0.810	0.850	0.792
5 01/01/81–22/12/82	0.852	0.892	0.792
6 13/01/83–12/01/85	–	0.801	–
7* 13/01/85–13/01/87	0.792	0.773	0.856

– denotes model diverged ($E < 0.700$).
*19/10/84–30/06/86 for the Gibbo River.

Table 7
Simulation results for the Upper Murray, Gibbo and Snowy catchments over the 10-year period from 01/01/1977

Catchment	Number of CP where the model was calibrated (Table 6)	Efficiency E	Bias (mean daily absolute error) cumecs/day
Murray River at Biggara	7	0.649	0.37
Gibbo River at Gibbo	4	0.692	0.63
Snowy Creek at Granite Flat	5	0.729	0.06

Table 8
Model efficiency values E for calibration of the IHACRES snow-runoff model for the Tooma River

Catchment and station number CP	Model efficiency E
1 01/01/73–31/12/74	0.703
2 01/01/75–31/12/76	0.782
3 06/01/77–06/01/79	–
4 20/11/78–19/11/80	0.702
5 01/01/81–22/12/82	0.803
6 13/01/83–12/01/85	–
7 13/01/85–13/01/87	0.773

– denotes model diverged ($E < 0.700$).

efficiency coefficients and bias for that catchment are shown in Table 11. The simulation results of the two-year period, 1966–1967, for the model calibrated on CP 5 (14/03/1974–13/03/1976), are presented in Fig. 4.

The results of a long-term historical simulation for the Tallangatta and Cudgewa Creeks catchments are shown in Fig. 5. This figure illustrates that the simulation results are reasonably good for almost all periods except during large flows in 1956 and 1974. The huge underestimations which occurred in these years might be explained by instrumental errors related to periods of very high streamflow (for instance, the poor calibration of the upper part of river profile sections for relevant stream gauging stations). Such results can be improved by filtering of the residuals.

4. Forecasting algorithms

4.1. Structure of the forecasting algorithm

The conceptual rainfall-runoff model IHACRES and a self-adaptive linear filtering approach were combined

Table 9
Model efficiency values E for calibration of the IHACRES snow-runoff model for the Big River catchment

CP	Model efficiency E
1 4/02/66–3/02/68	0.862
2 5/01/68–4/01/70	–
3 23/02/70–22/02/72	–
4 24/03/72–23/03/74	–
5 14/03/74–13/03/76	0.770
6 2/02/76–1/02/78	0.836
7 2/02/78–1/02/80	0.734
8 11/02/80–10/02/82	0.792
9 10/02/82–9/02/84	0.894

– denotes model diverged ($E < 0.700$).

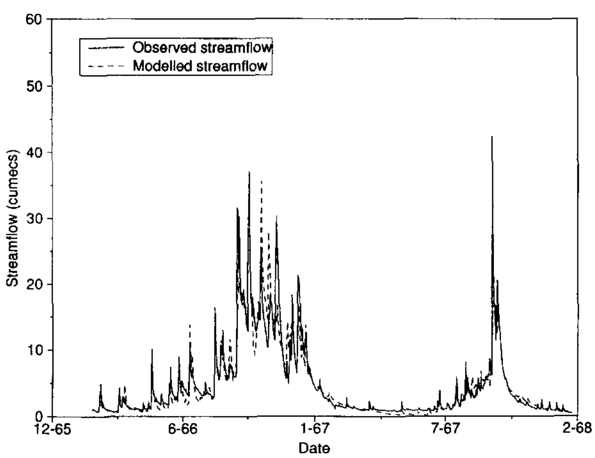

Fig. 3. Observed (solid line), modelled (dashed line) streamflow (cumecs) for CP 1 (1966–1967) for the Big River catchment.

Table 10
Simulation results over the 14-year period from 01/01/1973 for the Tooma River catchment. Bold values denote selected model

Number of calibration period (Table 8)	Efficiency E	Bias (mean daily absolute error) cumecs/day
2	**0.664**	**0.98**
4	0.663	3.25

Table 11
Simulation results over the 20-year period from 01/01/1965 for the Big River catchment. Bold values denote selected model

Number of calibration period (Table 9)	Efficiency E	Bias (mean daily absolute error) cumecs/day
1	0.693	0.76
5	**0.709**	**–0.18**
6	0.648	–1.10
7	0.646	0.47
8	0.657	1.81
9	0.656	0.27

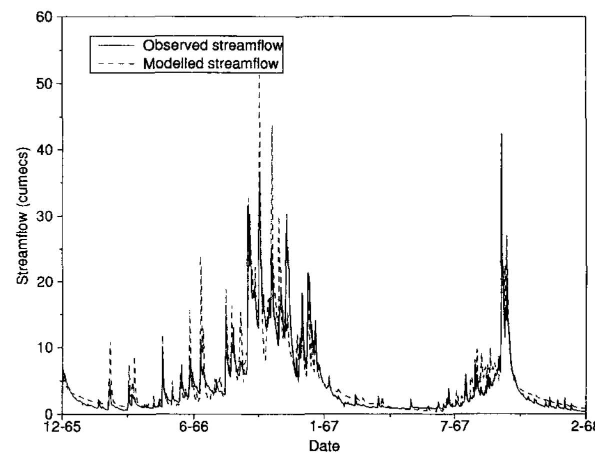

Fig. 4. Observed (solid line), modelled (dashed line) streamflow (cumecs) for the simulation test over the period 1966–1967 for the Big River catchment.

and applied for forecasting daily streamflow for the nine catchments in the Upper Murray Basin (Table 1). Different types of models were considered in order to select the most appropriate forecasting algorithm.

A disadvantage of rainfall-runoff models applied solely is that the residuals of the model are not white noise. The mean value of residuals of such models may be zero but the variance tends to change through time (e.g. seasonally) and residuals are strongly autocorrelated.

Linear filtering can be applied in order to decompose the residuals into a systematic component and white noise. The AutoRegressive Integrated Moving Average (ARIMA) model was selected here as an instrument for filtering the residuals.

Schematically, the model applied may be represented as a combination of two steps:

(1) The deterministic conceptual model IHACRES, providing for each time step (daily here) k the modelled value of streamflow y_k, which can be expressed through its measured value x_k as:

$$x_k = y_k + \xi_k, \text{ where } \xi_k \text{ are the residuals.}$$

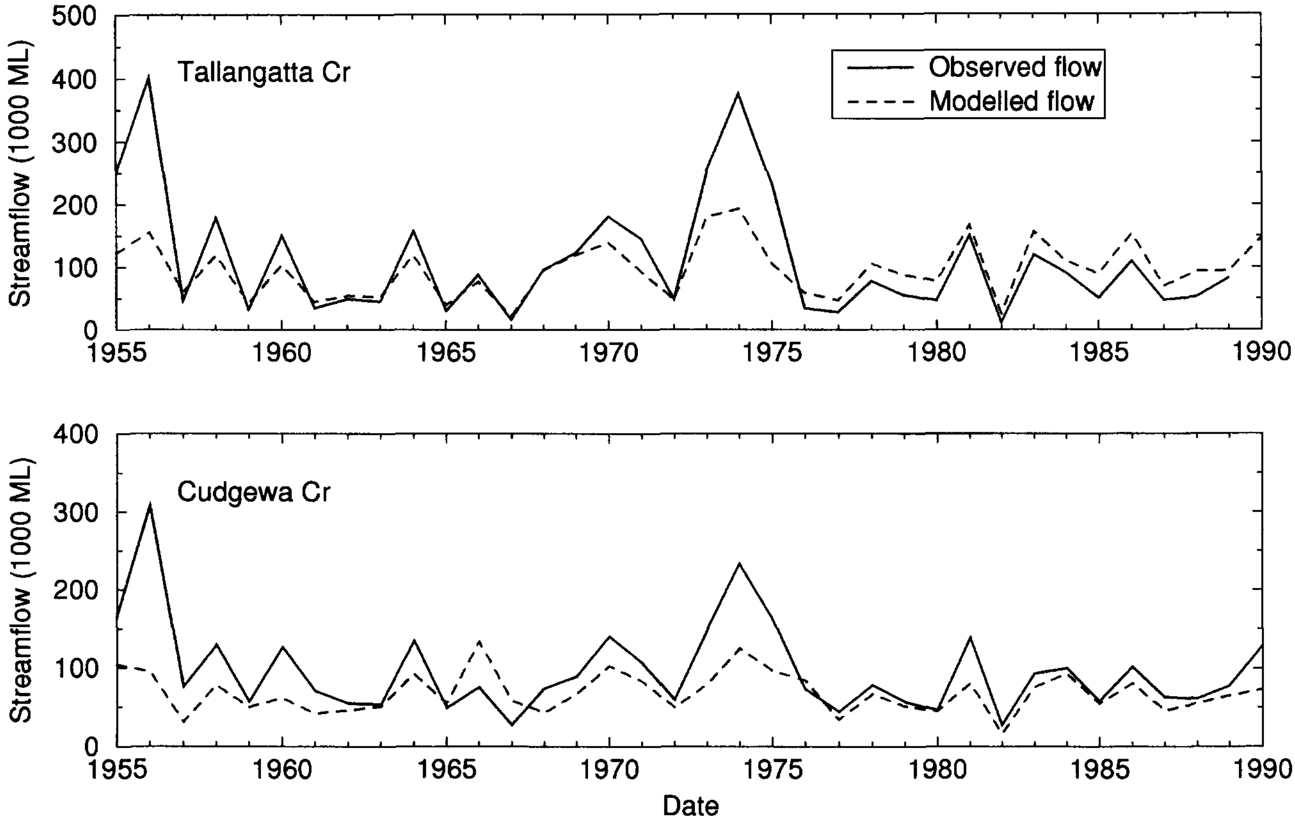

Fig. 5. Historical simulation of annual streamflow for the Tallangatta and Cudgewa Creeks using the IHACRES model.

The discrete random function ξ_k is strongly autocorrelated and its variance has seasonal fluctuations. Therefore it is logical to filter these residuals, or decompose them into a combination of systematic and white noise components.

(2) The residuals ξ_k of the IHACRES model, linearly filtered using the ARIMA model.

4.2. Deterministic part (IHACRES)

The conceptual dynamic lumped parameter model IHACRES has two modules: a non-linear loss module which transforms measured rainfall to effective rainfall using the temperature data, and a linear module defined as a recursive relation at time step k for modelled streamflow y_k, computed as a linear combination of its previous values and excess rainfall. The loss module is used to account for the effect of antecedent weather conditions on the current status (s_k) of soil moisture and vegetation conditions, and evapotranspiration effects. The effective rainfall u_k is calculated from the measured precipitation r_k and temperature t_k by formulae given in Schreider *et al.* (1996a).

The version of this model based on the Simple Refined Instrumental Variable (SRIV) technique uses the previous values of the modelled flow for recurrent estimation of its current value. An approach based on a least squares technique uses, for such recurrent relationships, the previously measured values of streamflow. This method is not applicable for periods where the measured flow is unknown but can provide more accurate forecasts for the periods where the observed data can be updated. The particular form of the linear module used in this work, which is based on a two parallel storages approximation (superposition of quick and slow flow recessions), is

$$y_k = -a_1 y_{k-1} - a_2 y_{k-2} + b_0 u_k + b_1 u_{k-1} \quad (1)$$

or the SRIV algorithm, where y_i is modelled streamflow, and

$$y_k = -a_1 x_{k-1} - a_2 x_{k-2} + b_0 u_k + b_1 u_{k-1} \quad (2)$$

for the so-called least squares algorithm, where x_{k-1} and x_{k-2} are previous measured values of flow. The total number of parameters in this version of IHACRES, including its linear and non-linear modules, is six.

4.3. Stochastic part (ARIMA)

The residuals of the model $\xi_1, \xi_2, \ldots, \xi_n$ were considered as the stochastic time series to be filtered. The ARIMA (p,d,q) model is defined by the following relations (Box and Jenkins, 1976; Bras and Rodriguez-Iturbe, 1985):

$$\phi(B)\,((1-B)^d\,\xi_k - D) = \Theta(B)\,a_k,$$

where B is a backward shift operator such that $B\xi_k = \xi_{k-1}$, $(1-B)^d$ is the difference operator, D is the mean value of the differenced series and $\phi(B)$, $\Theta(B)$ are the polynomial expressions for p autoregressive and q moving average values:

$$\phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \phi_p B^p$$

$$\Theta(B) = 1 - \Theta_1 B - \Theta_2 B^2 - \ldots - \Theta_q B^q$$

Modelling of seasonal periodicity was excluded from consideration here. The sum $p + q + d$ defines the number of parameters to be optimised for the stochastic module.

4.4. Application of the ARIMA model

Linear filtering of the residuals of the conceptual model IHACRES was performed using moving windows, each with a duration of 40 days, and a 1 day time step. The ARIMA models were calibrated separately on each window and the forecasts of the residuals for L days forward were calculated. The statistical significance of the approximation was controlled for each step using the Box–Ljung portmanteau statistic (McLeod, 1978).

In order to test the forecast procedure, it was assumed that the rainfall for L days forward must be known; see Eqs (1) and (2).

The residuals of the SRIV and least squares versions of the IHACRES conceptual model were used as inputs for the linear filtering algorithm. To assess how informative these prediction algorithms are, an additional test was suggested: the residuals of the 'naive' forecasting algorithm, where the predicted value of flow is equal to the streamflow at the previous time step ($y_k = x_{k-1}$), were also considered to provide forecasts of streamflow.

A range of different ARIMA structures was considered for each deterministic model (SRIV, least squares and 'naive') to select the best forecasting algorithm. The values for the number of autoregressive parameters p, moving average parameters q and differencing d were limited to a maximum of 2, 2 and 1, respectively, in order to avoid overparametrisation. Values (0,0,0) correspond to the conceptual model itself, when considered as a forecasting algorithm.

4.5. Results and analysis

Table 12 summarises the results of testing the different (p,d,q) parameters on the two-year calibration period (1980–1981) for Tallangatta Creek. The values of mean absolute and mean square errors for one day ahead forecasts calculated over this two-year period were selected as a measure of the quality of the model identification. The model was considered to fail if it diverged for more than 5% (35 of approximately a 700-day period) of the windows. The best results were provided for ARIMA (1,0,0) and (1,1,0). Figures 6 and 7 show the results of SRIV IHACRES performance on the periods of low and high flow, respectively, and their improvements after filtering by an ARIMA process.

Table 12
The mean absolute and mean square forecast errors obtained for a range of ARIMA parameters and three different conceptual models

(p,d,q)	SRIV	Least squares	'Naive' model
(0,0,0)	1.02; 3.58	0.82; 7.57	0.73; 5.12
(1,0,0)	0.65; 2.34	0.85; 7.24	0.75; 5.71
(0,0,1)	0.89; 3.13	*	0.79; 6.00
(2,0,0)	0.69; 2.59	0.82; 6.80	0.79; 6.00
(1,0,1)	0.74; 2.81	*	*
(1,1,0)	0.65; 2.60	1.22; 15.50	1.00; 9.80
(0,1,1)	0.70; 2.60	*	*
(1,1,1)	*	*	*
(1,2,0)	0.90; 5.23	2.05; 38.42	1.46; 20.15
(2,1,1)	0.67; 2.55	1.16; 11.88	0.99; 9.07
(2,2,0)	0.89; 4.54	1.72; 24.49	1.32; 15.64

*means model diverges.

Table 13 illustrates how the quality of forecast depends on how many days forward (L) it is provided. The results are illustrated for two catchments, for Tallangatta and Cudgewa Creeks, respectively. The results show that the quality of forecast obtained by linear filtering of the residuals is better than the forecast obtained by using the conceptual model solely, for values of L up to 3 (the Cudgewa Creek) and 5 (the Tallangatta Creek). The results of the forecasting algorithm application to all the catchments considered

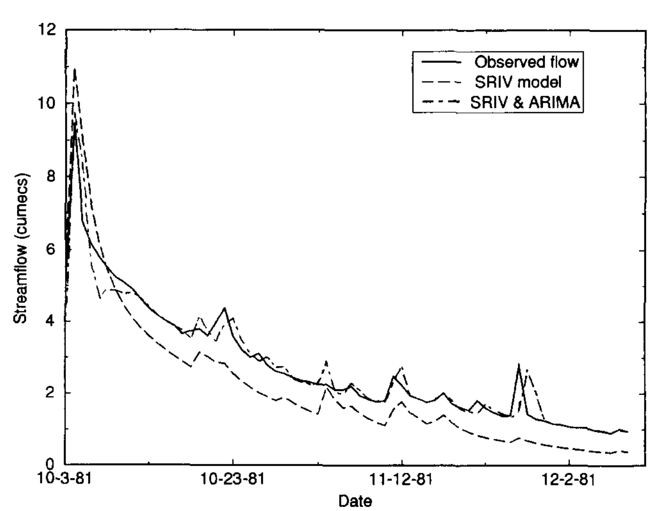

Fig. 6. 1-day ahead forecast for a low-flow period in the Tallangatta Creek catchment.

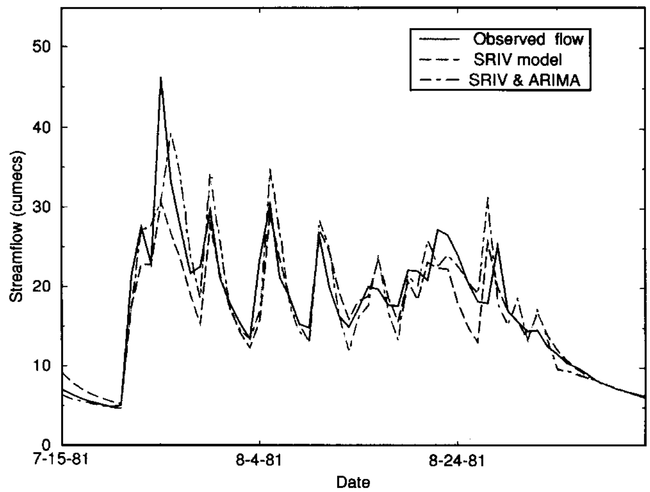

Fig. 7. 1-day ahead forecast for a high-flow period in the Tallangatta Creek catchment.

Table 13
The quality of forecast obtained for the linearly filtered ARIMA (1,0,0) residuals of the conceptual model SRIV IHACRES

L	1	2	3	4	5
The Tallangatta Creek	0.65 2.34	0.81 3.27	0.92 3.35	0.96 3.47	1.03 3.58
The Cudgewa Creek	0.63 2.50	0.74 2.80	0.75 2.80	0.76 2.86	0.73 2.64

Mean absolute and mean square errors are given.

are summarised in Table 14. The efficiency statistics and bias for the IHACRES model, applied solely, are presented for the periods specified in Sections 2 and 3. The selection of different periods for the simulation test in different catchments is explained by the availability and quality of rainfall data at these sites. The combined forecasting algorithm was applied for the whole period of availability of streamflow records. For the periods when precipitation is not recorded, it provides forecasting values using information about streamflow solely. Even for such a test, comparison of their values illustrates the considerable improvement obtained after application of the linear filtering procedure to the residuals of the IHACRES model, even in the case of the Jingellic Creek catchments, where a problem related to its ephemeral nature exists.

5. Discussion and conclusions

Successful calibration of the IHACRES model was performed for nine catchments of the Upper Murray Basin. The results described above show that the IHACRES model provides consistently good estimates of daily streamflow for the snow-free as well as for the snow-affected catchments in the Upper Murray Basin.

A method for combining a conceptual rainfall-runoff model and a self-adaptive linear filtering approach was developed and applied to forecast streamflow for nine catchments in the Upper Murray Basin. Considerable improvement was achieved compared with prediction based on the use of the conceptual model only: the errors of the forecast 3–5 days forward for the combined method are comparable with the errors of a 1-day forward prediction provided by the conceptual model only. The (1,0,0) structure of the ARIMA algorithm was selected as the most appropriate for forecasting in the region under consideration. This allows the use of a combined deterministic and self-adaptive stochastic approach for better approximation of river discharge and for operational forecasting, in particular for reservoir management.

Table 14
Efficiency statistics (*E*) and bias for long-term simulations in all nine catchments considered. Results obtained from the IHACRES simulation, and IHACRES combined with an ARIMA algorithm are presented

Station number	River and station location	IHACRES model applied solely *E*	bias	IHACRES model combined with ARIMA (1,0,0) *E*	bias
401203	Mitta-Mitta River at Hinnomunjie	0.669	0.50	0.835	−0.03
401220	Tallangatta Creek at McCallums	0.611	0.35	0.807	0.04
401208	Cudgewa Creek at Berringama	0.638	0.62	0.744	−0.01
401012	Murray River at Biggara	0.649	0.37	0.835	−0.25
401217	Gibbo River at Gibbo	0.692	0.63	0.805	0.22
401210	Snowy Creek at Granite Flat	0.729	0.06	0.809	0.07
401013	Jingellic Creek at Jingellic	0.533	0.17	0.602	−0.01
401014	Tooma River at Pine Grove	0.664	0.98	0.841	0.06
401216	Big River U/S of Joker Creek	0.709	−0.18	0.845	0.04

References

Box, G. E. P. and Jenkins, G. M. (1976) *Time Series Analysis: Forecasting and Control*. Holden Day, San Francisco.

Bras, R. L. and Rodriguez-Iturbe, I. (1985) *Random Functions and Hydrology*. Addison-Wesley Publishing Company, Reading, Massachusetts.

Brath, A. and Rosso, R. (1993) Adaptive calibration of a conceptual model for flash flood forecasting. *Water Resources Research* **29**(8), 2561–2572.

Galeati, G. (1979) A comparison of parametric and non-parametric methods for runoff forecasting. *Hydrological Science Journal* **35**, 79–94.

Georgakakos, A. P. (1987) Real time flash-flood prediction. *Journal of Geophysical Research* **92**(D8), 9615–9629.

Jakeman, A. J., Chen, T. H., Post, D. A., Hornberger, G. M., Littlewood, I. G. and Whitehead, P. G. (1993) Assessing uncertainties in hydrological response to climate at large scale. In *Macroscale Modelling of the Hydrosphere (Proceedings of the Yokohama Symposium*, July 1993), IAHS Publ. No. 214, pp. 37–47.

Jakeman, A. J. and Hornberger, G. M. (1993) How much complexity is warranted in a rainfall-runoff model? *Water Resources Research* **29**(8), 2637–2649.

Jakeman, A. J., Littlewood, I. G. and Whitehead, P. G. (1990) Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology* **117**, 275–300.

Jamieson, D. G., Wilkinson, J. C. and Ibbit, R. P. (1972) Hydrologic forecasting with sequential deterministic and stochastic stage. *Proceedings of the International Symposium on Uncertainties in Hydrologic and Water Resource Systems*, University of Arizona, pp. 177–187.

Kalman, R. E. (1960) A new approach to linear filtering and prediction problem. *Journal of Basic Engineering*, **82D**, 35–45.

Karlsson, M. and Yakowitz, S. (1987) Rainfall-runoff forecasting methods, old and new. *Stochastic. Hydrol. Hydraul.* **1**, 303–318.

Kember, G. and Flower, A. C. (1993) Forecasting river flow using non-linear dynamics. *Stochastic. Hydrol. Hydraul.* **7**, 205–212.

Mack, Y. P. and Rosenblatt, M. (1979) Multivariate K-nearest neighbour density estimates. *Multivariate Analysis* **9**, 1–15.

McLeod, A. I. (1978) On the distribution of the residual autocorrelations in Box–Jenkins models. *J. Roy. Statistics Soc. Ser. B.* **40**, 296–302.

Nash, J. E. and Sutcliffe, J. V. (1970) River flow forecasting trough conceptual models, Part I—A discussion of principles. *Journal of Hydrology* **10**, 282–290.

Olason, T. and Watt, W. E. (1986) Multivariate transfer function–noise model of river flow for hydropower operation. *Nordic Hydrology* **17**, 185–202.

Overton, D. E. and Meadows, M. E. (1976) *Stormwater Modelling*. Academic Press, New York, 358 pp.

Schreider, S. Yu., Jakeman, A. J. and Pittock, A. B. (1996a) Modelling rainfall-runoff from large catchments to basin scale: the Goulburn Valley, Victoria. *Hydrological Processes* **10**, 863–876.

Schreider, S. Yu., Jakeman, A. J., Pittock, A. B. and Whetton, P. H. (1996b) Estimation of possible climate change impacts on water availability, extreme flow events and soil moisture in the Goulburn and Ovens Basins, Victoria. *Climatic Change* (in press).

Schreider, S. Yu., Whetton, P. H., Jakeman, A. J., Pittock, A. B. and Li, J. (1996c) Runoff modelling for snow-affected catchments in the Australian Alpine Region, Eastern Victoria. *Journal of Hydrology* (in press), also in CRES Working Paper, ANU; Canberra, 1996/5, 26 pp.

Sen, Z. (1991) Adaptive prediction of hydrologic series by Walsh–Kalman model. *Journal of Hydrology* **122**, 221–234.

Todini, E. (1996) The ARNO rainfall-runoff model. *Journal of Hydrology* **175**, 339–382.

Water Victoria (1989) *A Resource Handbook*. Department of Water Resources, Melbourne, Australia, 311 pp.

Whetton, P. H., Haylock, M. R. and Galloway, R. (1996) Climate change and snow-cover duration in the Australian Alps. *Climatic Change* **32**, 447–479.

Wood, E. F. (ed.) (1989) *Workshop on Real Time Forecasting/Control of Water Resource Systems*. Pergamon Press, New York.

Yakowitz S. (1987) Nearest neighbor for time series analysis. *Journal of Time Series Analysis* **8**, 235–247.

Yakowitz, S. and Karlsson, M. (1987) Nearest neighbor methods for time series with application to rainfall-runoff prediction. In *Stochastic Hydrology*, eds. I. B. MacNeil and G. J. Umphries, pp. 149–160. Reidel, New York.

Ye, W., Jakeman, A. J. and Barnes, C. (1995) A parametrically efficient model for prediction of streamflow in an Australian benchmark catchment with complex storage dynamics. *Environmental International* **21**(5), 539–544.