ADAPTIVE LINEAR MODELS FOR REAL-TIME FLOOD FORECASTING

G. E. Amirthanathan

Report 96/3
September 1996

COOPERATIVE RESEARCH CENTRE FOR
CATCHMENT HYDROLOGY

ADAPTIVE LINEAR MODELS FOR REAL-TIME FLOOD FORECASTING

G. E. Amirthanathan

Report 96/3

September 1996

COOPERATIVE RESEARCH CENTRE FOR
CATCHMENT HYDROLOGY

Amirthanathan, G. E., 1952 -

Adaptive linear models for real-time flood forecasting.

Bibliography.
ISBN 0 7326 1111 3.

1. Hydrologic models. 2. Rain and rainfall - Mathematical models. 3. Runoff - Mathematical models. 4. Flood forecasting - Australia - Mathematical models. 5. Real-time data processing. I. Cooperative Research Centre for Catchment Hydrology. II. Title (Series: Report (Cooperative Research Centre for Catchment Hydrology); 96/3)

551.489015118

Keywords

Modelling (Hydrological)
Flood Forecasting
Flood Warnings
Modelling (-Specific Names-II)
Rainfall/Runoff Relationship
Catchment Areas

ISSN 1039-7361

SUMMARY

Adaptive linear models have been used in the past for on-line forecasting of floods. However, there have been few investigations devoted to determining the suitability of these models for real-time flood warning systems for catchments in Australia. This report presents two linear models, namely, the Unit Hydrograph (UH) and ARMA-type models for real-time flood forecasting. As the parameters of the models can vary from storm to storm and also within a storm, a linear sequential optimal estimation technique (Kalman filter) is used to update the parameters of these models. A simple loss model consisting of initial and continuing loss components suitable for real-time operational purposes is used. The models are used to forecast runoff from rainfall for three Australian catchments. Linear river routing models are also developed and applied and a multiple input-single output (MISO) model is used to analyse a river reach with several tributary inflows. Proper *a priori* estimates of noise statistics are important for optimal performance of the filter. A procedure more suitable for short series of data is proposed and applied successfully. The real-time forecasting performance of these linear models is compared with that of non-adaptive methods and some of the reasons for the difference in performance are discussed.

PREFACE

An important component of any flood management strategy is the provision of warning prior to a flood event. If people are given adequate notice of the timing and peak level of impending flood inundation, they can take actions which will save lives, livestock and property. The goal of CRCCH Project D4 "Development of an improved real-time flood forecasting model" is to increase both the accuracy of the prediction of flood level, and increase the warning time of the event.

This report describes the work performed by Dr Amirthanathan during the early stages of the D4 project. He has applied adaptive models (i.e. models which continuously adjust their parameters based on their past forecasting accuracy) to Australian data, and has shown their considerable potential for improved forecasts in many locations.

Other models will be examined during the course of this project. The final outcome will be a knowledge base of model performance from which firm recommendations will be made. Dr Amirthanathan's work is an important contribution to that knowledge base.

Russell Mein
Program Leader, Flood Hydrology
Cooperative Research Centre for Catchment Hydrology

ACKNOWLEDGEMENTS

This research study was conducted at the University of Melbourne and at the Hydrology Division (FWPO), Bureau of Meteorology, Melbourne, from May to December 1992. I am grateful to Professor T.A. McMahon of the University of Melbourne, Assoc. Professor R.G. Mein of Monash University, and Mr Jim Elliott of the Hydrology Division, Bureau of Meteorology, for their valuable advice and help in numerous ways during this project. The staff of the Department of Civil & Environmental Engineering, University of Melbourne, and the staff at FWPO, Bureau of Meteorology, helped me in many ways during the research study and I am thankful to all of them. I am grateful to the University of Melbourne, the Centre for Environmental Applied Hydrology and the Co-operative Research Centre for Catchment Hydrology for funding this project.

TABLE OF CONTENTS

4.2 /...

1. INTRODUCTION

A large number of rainfall-runoff models have been developed and are being used for on-line forecasting purposes with the increasing use of telemetry in the control of water resources systems. Models commonly used include linear unit-hydrograph and non-linear network models, as well as conceptual soil moisture accounting models such as the Stanford Watershed Model. The reliability of streamflow forecasts will govern the reliability of decisions pertaining to flood warning, flood control and river regulation, and so influence the associated benefits which accrue from them. The extent to which a model accurately represents the response of a catchment to rainfall will have a major influence on the reliability of forecasts.

Two of the main problems of real-time flood forecasting are:

(a) the estimation of antecedent conditions for individual storms and the determination of the amount and location of rainfall excess (in real-time) for input to the catchment model;

(b) the need to continuously correct forecasts using the observed errors in earlier forecasts.

Real-time flood forecasting models should be developed using criteria that are important in forecasting, and long-term performance of the model should play a lesser role than short-term performance when model structures are being specified. As real-time forecasting is concerned with predicting an event during its occurrence, it should utilize both the measurements of input/output and the state of the system as they become available. The forecasting procedure should also objectively evaluate the errors in forecasts. This requires that the model be amenable to updating, based upon past performance in predicting the event and upon new measurements as these become available. The feedback information is most valuable in improving the real-time forecasting performance of a model, both by adapting the model structure as well as the parameters, taking into account the uncertainty in the input/output measurements.

1.1 Sources of Error

It is useful to make note of some of the sources of error in the modelling process. In modelling various components of the rainfall-runoff process, several simplifications have to be made that introduce an error which is reflected in the predictions. The input data required for the operation of a rainfall-runoff model include quantitative precipitation measurements and, in some cases, precipitation forecasts. Errors in precipitation input include the measurements themselves as well as averaging errors, while there is a considerable error in forecast rainfall.

There is a further source of error in the model forecasts due to the selected set of model parameters. Some of these parameters are chosen through calibration of the model, while others are found from the physical characteristics of the basin. Errors are also introduced due to imperfect knowledge of the initial state of the hydrologic system, although since the memory of the system is of finite length, this error could be minimised in flood forecasting by beginning the operation of the model well ahead of the start of flood events. In general, water level measurements are the only measurements used in estimating runoff in a flood forecasting system. This introduces a further error, since the relationship between water level and runoff (namely, the rating curve) is often unstable and it is sometimes difficult to establish this relationship, especially in the range of high flood flow levels.

1.2 Objectives of the Study

The objectives of this study are to:

(a) develop adaptive forecasting procedures that are suitable for real-time operational purposes;

(b) use these procedures to address some of the problems of real-time forecasting described earlier, including a suitable loss model for real-time operations; and

(c) apply these adaptive procedures using field data from Australian catchments and to compare their performance with that of other non-adaptive methods.

2. OVERVIEW OF ADAPTIVE FLOW FORECASTING METHODS

Rainfall-runoff models developed so far can be classified into three types, namely, distributed physically based models, lumped conceptual models, and input-output or 'black box' models. As the physically based models have not yet been fully developed, the choice of a rainfall-runoff model for many real-time forecasting applications is between a lumped conceptual model and a 'black box' model such as the unit hydrograph model or an ARMA type model. The more complex lumped conceptual models do not readily lend themselves to having their forecasts updated in a computationally efficient manner. On the other hand, the 'black box' models tend to have minimal computational requirements and, because they can be explicitly specified in analytical form, recursive estimation methods can be readily applied to update their forecasts in real-time. Simple adaptive models such as the ARMAX model or transfer function models have been widely used [2,3]. Kitanidis and Bras [5] developed a stochastic conceptual model based on the soil moisture accounting and the channel routing part of the National Weather Service River Forecast System watershed model, and on a non-linear lumped parameter conceptual hydrological model, and showed its usefulness for real-time forecasting of river flows. A considerable amount of research effort has been spent in exploring the use of sophisticated estimation procedures with relatively little effort devoted to the improvement of the underlying models or the quality of the input/output data. This has resulted in insignificant improvement in forecasts using these sophisticated procedures [7,8].

2.1 Real-Time Forecasting Techniques

In the context of real-time use, an adaptive method is defined as one that utilizes the current and previous input and output in computing the future output, by using the deviation or discrepancy between the latest observed and model outputs as feedback. In the case of non-adaptive methods, the model, once calibrated, is assumed to be fixed and ignores the forecasting error or deviation.

Generally, adaptive linear models have been used in three different ways in real-time forecasting:

(a) the parameters of the model are formulated as the state variable in the filter and updated using the prediction error as feedback;

(b) the hydrological variables (runoff) are considered as the state variables and are corrected at each step using the Kalman filter, as this technique optimally estimates the state variable;

(c) two Kalman filters are used, one updating the parameters and the other correcting the hydrological variable. This method is termed the Mutually Interactive State-Parameter method and was proposed by Todini [3].

Throughout this study, the adaptive method used is as described in (a) above, where the parameters are sequentially updated and used for prediction of flows. This method is described in more detail in the following section.

2.2 Adaptive Prediction Algorithm

The best least squares unbiased estimate of the state variables X (see below) can be obtained by applying the linear Kalman filter. The basis of the prediction is the previously calculated prediction error which is fed back into the model as a variable. After receiving the latest observation, the parameter vector and a weighting matrix are updated recursively for every discrete time instant. Then the predictions can be made. The covariance matrix of the parameter estimation error, containing all previous information on the system, is used as the weighting matrix.

The recursive algorithm of the Kalman filter is given below.

- *State model*

$$X_k = X_{k-1} + w(k) \qquad \text{with } w(k) \cong N(0,Q) \ldots\qquad (2.1)$$

- *Measurement model*

$$q_k = H^T{}_k \cdot X_k + v(k) . \qquad \text{with } v(k) \cong N(0,R) \ldots\qquad (2.2)$$

- *State prediction*

$$X_{k/k-1} = X_{k-1/k-1} \qquad (2.3)$$

- *Output prediction*

$$q_{k/k-1} = H^T{}_k \cdot X_{k/k-1} \qquad (2.4)$$

- *Predicted error covariance matrix*

$$P_{k/k-1} = P_{k-1/k-1} + Q \qquad (2.5)$$

- *Predictor gain*

$$G_k = P_{k/k-1} H_k [H^T{}_k \cdot P_{k/k-1} H_k + R]^{-1} \qquad (2.6)$$

- *State estimation using new measurements*

$$X_{k/k} = X_{k/k-1} + G_k v_k \qquad (2.7)$$

where innovation $v_k = q_k - q_{k/k-1}$ (one step ahead)

- *Error covariance matrix after new measurement*

$$P_{k/k} = [1 - G_k H^T{}_k] P_{k/k-1} \qquad (2.8)$$

An important step in applying the Kalman filter for on-line forecasting is the estimation of error covariance matrices R and Q. An optimal filter will produce an innovation sequence that is time-wise uncorrelated, having the following property of white noise:

$$\mathrm{Cov}\left[v_k \; v_{k-\tau}\right] = \begin{cases} \sigma_k^2 & \text{for } \tau=0 \\ 0 & \text{for } \tau\neq 0 \end{cases} \qquad (2.9)$$

There are several methods to estimate R and Q using this property in an on-line mode [15]; however, these are generally suited for reasonably long series of data. Mehra [16] presented an algorithm for the case of constant R and Q and shows that the estimates are asymptotically unbiased and consistent. In the case of rather short flood events, prior estimates of R and Q could be obtained by applying the filter to the data used for calibration for various values of error statistics until the filter performance is satisfactory. This procedure has been successfully applied in the past [17,18] and is used in this study.

The initial values of P_0, R and Q are chosen using the following procedure. The Kalman gain G_k (= $P_{k/k-1}\ H_k\ [H^T_k\ .\ P_{k/k-1}H_k + R]^{-1}$) is a vector which is multiplied by the innovation v_k for correcting each element in the parameter vector X_k. If P_k is chosen such that $H^T_k\ .\ P_{k/k-1}H_k >> R$, then the parameters will be heavily corrected, resulting in $q_k - H^T_k\ .\ X_{k/k} = 0$. This mode is useful for estimating the unknown parameters (self-tuning), but not suitable for optimal filtering, as the parameters are heavily corrected without any filtering taking place. If P_k is chosen such that $H^T_k\ .\ P_{k/k-1}H_k << R$, then the elements in gain G_k will be so small that hardly any correction would be done on parameter vector X_k. For real-time operation, appropriate values of P_0 should be chosen for optimal performance of the filter without falling into either of these extremes. In the case of short flood events, prior estimation of R, Q and P_0, obtained from the data used for calibration, is found to be adequate [9].

A simple procedure is adopted for estimating the statistics of the noise, where the filter is simulated on the data used for calibration for various values of noise statistics until it performs satisfactorily. This procedure was found to be satisfactory for short series of data ranging from 10 to 30 values. The filtering techniques enable information to be extracted from the difference between model forecast and actual behaviour to ensure the best possible estimation of model parameters at all times. Whilst the models used in this study are relatively simple, ideally the models should be as close as possible to the deterministic part of hydrological processes involved, otherwise the forecasting scheme relies too heavily upon feedback rather than its capability to simulate. In critically looking at the usefulness of optimal filtering techniques, one should remember that it is only a helpful tool to complement and not to replace hydrological models of different kinds.

3. REVIEW OF REAL-TIME ADAPTIVE RIVER FORECASTING TECHNIQUES

3.1 Rainfall-Runoff Models

3.1.1 *Unit hydrograph models (UH)*

The WMO [11] identifies the unit hydrograph (UH) as a widely used rainfall based model for flood forecasting. The operational UH is usually developed using data from a large number of observed flood events and an average or typical UH is adopted for operational use. The use of these average parameters is one of the important causes of errors in flood prediction. As observed in practice, the UH parameters (i.e. the ordinates of the UH) can vary from storm to storm and even within a storm. Hino [12] demonstrated the usefulness of the Kalman filter in sequentially updating the parameters of a UH model of the form:

$$q(k) = \sum_{j=1}^{m} h(j) \cdot p(k - j + 1) + v(k) \tag{3.1}$$

where
- q = measure of runoff at the outlet of the catchment at any time k
- h = ordinates of UH (assumed to be time variant)
- p = lumped rainfall over the catchment area
- m = memory of the hydrologic system
- v = measurement error.

Much of the non-linear behaviour of the catchment is due to the feedback system between the excess rainfall and soil moisture. This time varying feedback system which influences losses and excess rainfall (p_e) must be accounted for in an adequate representation of the rainfall-runoff process. Thus the UH model shown in Fig. 3.1 could be written as follows:

$$q(k) = \sum_{j=1}^{m} h(j) \cdot p_e(k - lg - j + 1) + q_b(k) + v(k) \tag{3.2}$$

where
- q_b = baseflow at the outlet of the catchment
- lg = time lag between rainfall excess impulse and its significant response at the catchment outlet
- p_e = rainfall excess.

In order to incorporate the variation in the baseflow during the flood, the baseflow at any time k could be written as $\alpha_{(k)} \cdot q_b^*$, where $\alpha_0 \cdot q_b^*$ is the baseflow at the beginning of a storm. Now the model could be written as:

$$q_k = \sum_{j=1}^{m} h(j) \cdot P_e(k - lg - j + 1) + \alpha_k \cdot q_b^* + v(k)$$

or

$$q_k = H^T{}_k \cdot X_k + v(k)$$

where
- $H^T{}_k$ = $[p_e(k\text{-}lg), p_e(k\text{-}lg\text{-}1), \ldots p_e(k\text{-}lg\text{-}m\text{+}1), q_b^*]$
- X_k = $[h(1), h(2), \ldots h(m), \alpha_k]^T$
- $v(k)$ = measurement error [$\cong N(0, R)$]
- R = variance of measurement error.

The time varying parameters X_k are sequentially updated as new measurements become available. If the contribution of baseflow is insignificant, it may be omitted, resulting in a reduction of time varying parameters.

Proper estimation of rainfall excess is crucial if the UH model is to be used for real-time flood forecasting. In reality, the initial and continuing loss models normally used in practice give a wide range of parameter values for different storm events. Initial loss, which depends on the initial state of the catchment, could be assumed to be equal to the cumulative rainfall up to the lag 'lg' ahead of any significant initial rise in runoff at the outlet of the catchment, as shown in Fig. 3.1(b).

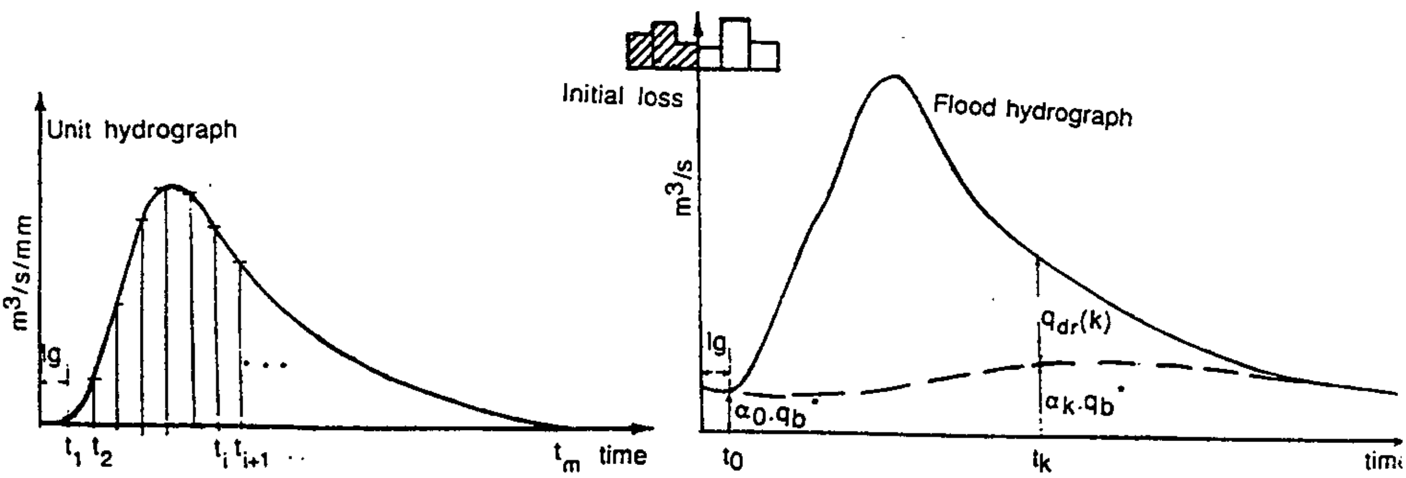


Figure 3.1: (a) Unit hydrograph; (b) Flood hydrograph

The continuing loss is usually assumed to be either a constant rate or a constant fraction of the rainfall. If we assume it to be a fraction of actual rainfall, this unknown fraction could be implicitly incorporated into the time varying parameter X_k and thereby become part of the parameter estimation process. Since the estimation of rainfall excess in real-time flood forecasting is difficult, this feature of adaptive methods is attractive. Thus the modified UH model is given by:

$$q_k = \sum_{j=1}^{m} h^*(j) \,.\, Rf(k - lg - j + 1) + \alpha_k \,.\, q_b^* + v(k)$$

or

$$q_k = H^T{}_k \,.\, X_k + v(k) \tag{3.3}$$

where H^T_k = [Rf(k-lg), Rf(k-lg-1), ...Rf(k-lg-m+1), q_b^*]
X_k = [h^*(1), h^*2), ...h^*m), α_k]^T
v(k) = measurement error [≅ N(0, R)]
Rf = average (lumped) rainfall
h^*(j) = operational UH ordinates.

As the parameters X_k can vary from storm to storm and even within a storm, it is modelled by a multivariate random walk equation given by:

$$X_{k+1} = X_k + w(k) \tag{3.4}$$

where w(k) = parameter error, a vector of zero mean Gaussian noise sequences with covariance matrix Q. The diagonal elements of this matrix are the variances of the corresponding parameters. For the case of time invariant parameters, these elements would be zero and then the procedure is identical to the recursive least squares method.

The adaptive method for a UH model starts with the initial values of the parameters obtained as the average set from the flood events used in the calibration of the model. For the first time step at the beginning of a flood, the forecast is made using these average values. When the first outflow measurement of the forecast flow becomes available, only h(1) and α will be updated, using the Kalman filter, since all the terms in H_k are zero except the first and last terms; the rest of the parameters in X_1 remain unchanged. The new parameters are used to make an adjusted prediction of outflow. When the correction of parameters is done after receiving the next measurement of outflow, h(2) will also be updated in X_2, along with h(1) and α. This procedure continues throughout the event, and after only 'm' steps all the parameters in X_k will be updated simultaneously.

3.1.2 *The ARMA type models*

Autoregressive moving average (ARMAX) type models have been used to model the short-term behaviour of hydrologic systems (13,14,15). It has been shown that an ARMA-like structure of the form of Eqn 3.5 gives an adequate representation of the dynamics of the rainfall-runoff process:

$$q_k + \delta_1 q_{k-1} .... + \delta_r q_{k-r} = \omega_1 p_{e(k-lg)} + \omega_2 p_{e(k-1-lg)} .... + \omega_s p_{e(k-s+1-lg)} + v_{(k)} \tag{3.5}$$

where q = runoff
p_e = rainfall excess
δ = autoregressive (AR) coefficients
ω = moving average (MA) coefficients.

'δ' and 'ω' are parameters analogous to the ARMA parameters of the ARMA process of order (r,s). The parameter 'lg' is the time lag between the first significant response of the system to a given impulse and the impulse itself, v(k) is a zero-mean Gaussian noise process disturbance to account for model error. The weighting parameters of δ and ω in Eqn (3.5) could be interpreted in the following way. As the storm event occurs, the parameters 'ω' will influence the rising limb of the hydrograph and will depend on the basin physiographic characteristics and the state of the basin (soil moisture, infiltration, etc.). Therefore, 'ω' should be modelled as a time variant parameter to capture the proper system response. The autoregressive parameters 'δ' will influence the falling limb or recession of the hydrograph [15]. Since these parameters can vary from storm to storm and also within a storm, it is proposed to model the parmeters by a simple random walk:

$$X_{k+1} = X_k + w(k) \tag{3.6}$$

where w_k is a vector of zero mean Gaussian noise sequences with covariance matrix Q and $X_k = [\delta_1, ...\delta_r, \omega_1, ...\omega_s]^T$.

Eqn (3.5) could be written as:

$$q_k = H^T{}_k . X_k + v(k) \tag{3.7}$$

where
$H^T{}_k = [q_{k-1}, ...q_{k-r}, p_e(k\text{-}lg), ... p_e(k\text{-}lg\text{-}s\text{+}1)]$
$X_k = [\delta_1, ...\delta_r, \omega_1, ...\omega_s]^T$
$v(k)$ = measurement error [$\cong N(0, R)$]
R = variance of measurement error.

3.1.3 *Stochastic rainfall prediction*

A general real-time forecasting model should also allow for input of meteorological forecast rainfalls. When future rainfall values are not considered for real-time forecasting, the result is an under-estimation of discharges due to omission of one or two terms in the model. Thus, when forecasts in advance of one-step ahead are to be made from the above-mentioned models, some assumptions have to be made about the behaviour of rainfall. One approach is to treat rainfall as a stochastic process and an autoregressive model be used to represent rainfall. The model structure can be identified using the rainfall sequences of the historical flood events, where the model order is varied until any further increase in the order does not considerably improve the mean squares of errors. Such an autoregressive model is given below:

$$Rf_{k/k-1} = \gamma(1) . Rf(k-1) + \gamma(2) . Rf(k-2) + ... \gamma(n_r) . Rf(k-n_r) + Vr(k) \tag{3.8}$$

where Vr(k) is a white noise. The Kalman filter is used to update the parameters during the storm. Thus the state equation is given by:

$$Rf(k) = Hr^T_k . Xr_k + Vr(k)$$

where $Hr^T_k = [Rf(k-1), Rf(k-2), Rf(k-n_r)]$

$Xr_k = [\gamma(1), \gamma(2), \gamma(n_r)]^T k$

and the measurement equation is given by:

$$Xr_k = Xr_{k-1} + wr(k)$$

3.2 River Routing Models

Linear flood routing models have been successfully applied in flood forecasting. It can be shown that in more sophisticated models of flood routing, for example, models based on numerical solutions of St Venants equation, the non-linear finite difference equations are linearised at each step to facilitate the solution procedure. When simplified linear models are used for real-time forecasting, similar relinearisation takes place when the new or current data are incorporated by the recursive algorithm. Thus, over the time span required for operational forecasting, a simplified linear model can adequately represent most river dynamics.

Hence a linear model for a river reach can be written in a discretised form given by:

$$Y(k) = \sum_{l=0}^{L} \alpha_l . Y(k-ld-l+1) + \sum_{j=0}^{J} \beta_j . y(k-j-lg) + v(k) \quad (3.9)$$

where
α_l = coefficient of autoregressive (AR) terms
β_j = coefficient of moving average (MA) terms
v(k) = measurement error
ld = lead time
lg = lag time
Y = downstream flow
y = upstream flow.

For the case of the Muskingum method: L=0, J=2, ld=lg=0. In practice, the model order L and J rarely exceed 2 or 3. This model could be written as:

$$Y(k) = H_k . X_k + v(k)$$

where $H_k = [Y(k\text{-}ld), Y(k\text{-}ld\text{-}1), ... y_1(k\text{-}lg_1), ... y_n(k\text{-}lg_n), ...]$
$X^T_k = [\alpha_1, \alpha_2, ..., \alpha_L, \beta_1, ... \beta_J]$
$v(k)$ = measurement error $\cong N(0, R)$.

The parameters can be considered time variant. If there are uncertainties such as an imbalance of flow in the reach due to ungauged inflows, the autoregressive terms adjust to take this into account, thereby improving the forecast.

3.2.1 *Identification of Structure and Parameters of River Routing Models*

The model consists of a structure defined by the orders L, J, the lags lg and the parameters α_l, β_j. The structure of the model could be evaluated using either constrained estimation techniques (CLS) proposed by Natale and Todini in 1976, or by any other recursive algorithm such as the Instrumental Variable method [23] and recursive least squares. The order of the model is increased progressively until the variance of the error reduces significantly. The adopted order is taken as the value at which any further increase in order does not reduce the error variance. The lags for each tributary could be evaluated from knowledge of the hydrology of the watershed, or by using a numerical technique proposed by Wood and Nagy [15], using the CLS method to obtain impulse response ordinates for each tributary. The time interval up to the first significant value of the ordinate or each impulse response function is identified as the lag (lg). This can also be done by the recursive least squares method as well. In the presence of high level coloured noise, the Instrumental Variable method (IVAML) would be very efficient.

3.3 Multiple Input - Single Output Models (MISO)

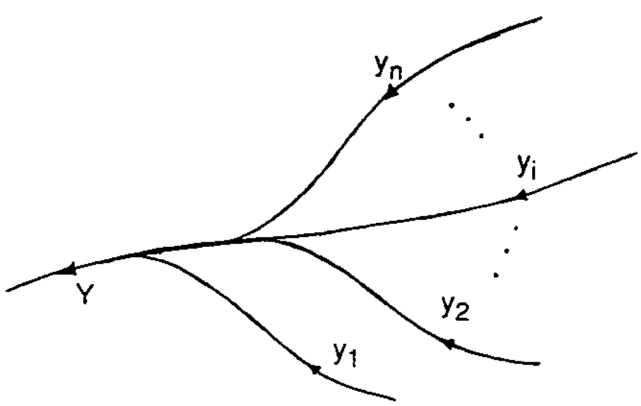


Figure 3.2: River network

A generalised model of a river network is given in Fig. 3.2, which shows a river reach with a number of tributaries joining the main stream between the upstream and downstream gauges. The upstream flow and each of the tributary flows are considered as (multiple) input, while the downstream flow is the output. (For sub-catchments having rainfall data, these average rainfall values could also be considered as one of the inputs and is represented in the same way, as the whole system is considered linear.) Thus the model can be represented by:

$$Y(k) = \sum_{l=0}^{L} \alpha_l . Y(k\text{-}ld\text{-}l) + \sum_{i=1}^{n} \sum_{j=0}^{J_i} \beta_{ji} . y_i(k\text{-}j\text{-}lg_i) + v(k) \qquad (3.10)$$

where α_l = coefficient of autoregressive (AR) terms
$\beta_{i,j}$ = coefficient of moving average (MA) terms
$v(k)$ = measurement error
ld = lead time
lg_i = lag time of i^{th} tributary.

Usually the values of L and J_i rarely exceed 2 or 3. The additional advantage of having autoregressive terms is that it takes into account the other factors such as unmeasured tributary inflows and lateral flows, thus improving the quality of flow forecasts. The model can be written as:

$$Y(k) = H_k . X_k + v(k) \qquad (3.11)$$

where H_k = $[Y(k\text{-}ld), Y(k\text{-}ld\text{-}1) \ldots y_1(k\text{-}lg_1) \ldots y_n(k\text{-}lg_n) \ldots]$
$X^T{}_k$ = $[\alpha_1, \alpha_2 \ldots \beta_{1,1} \ldots \beta_{n,1} \ldots]$

As the parameters of the model vary during the flood event as well as from one flood event to another, these parameters (X) could be assumed to vary according to random walk, given by the following equation:

$$X_k = X_{k-1} + w_k \qquad (3.12)$$

where w_k is a vector of zero mean Gaussian noise sequences with covariance matrix Q.

The identification of the structure and parameters of the model in this case can be done using similar techniques to those discussed in the previous section.

4. APPLICATION OF REAL-TIME ADAPTIVE FLOOD FLOW FORECASTING TECHNIQUES

4.1 Rainfall-Runoff Models

4.1.1 *Tweed River to Murwillumbah*

The Tweed River catchment upstream of Murwillumbah, located in the north-eastern corner of New South Wales, has a basin area of about 630 km^2. The catchment area includes two fairly symmetrical drainage systems, the Tweed and Oxley Rivers. These tributaries originate in the mountainous terrain situated along the boundary of the basin at elevations between 600m and 1100m. Most of the upper catchment is heavily timbered with natural vegetation. Pockets of land in the lower areas have been cleared for banana plantations on the slopes and sugar cane on the flat flood plains. This basin is subject to very intense rainfalls, with 24 hour values up to about 500mm. Variability of rainfall across the basin can be as high as a factor of three. The predictions available from the recorded data appear to reflect the differing rainfall patterns experienced in the catchment.

Ten flood events were considered between 1972 and 1987. Three-hourly rainfall data during these storms was available from rain stations at Kunghur, Chillingham and Murwillumbah. Lumped rainfall using Theissen coefficients was used for the analysis.

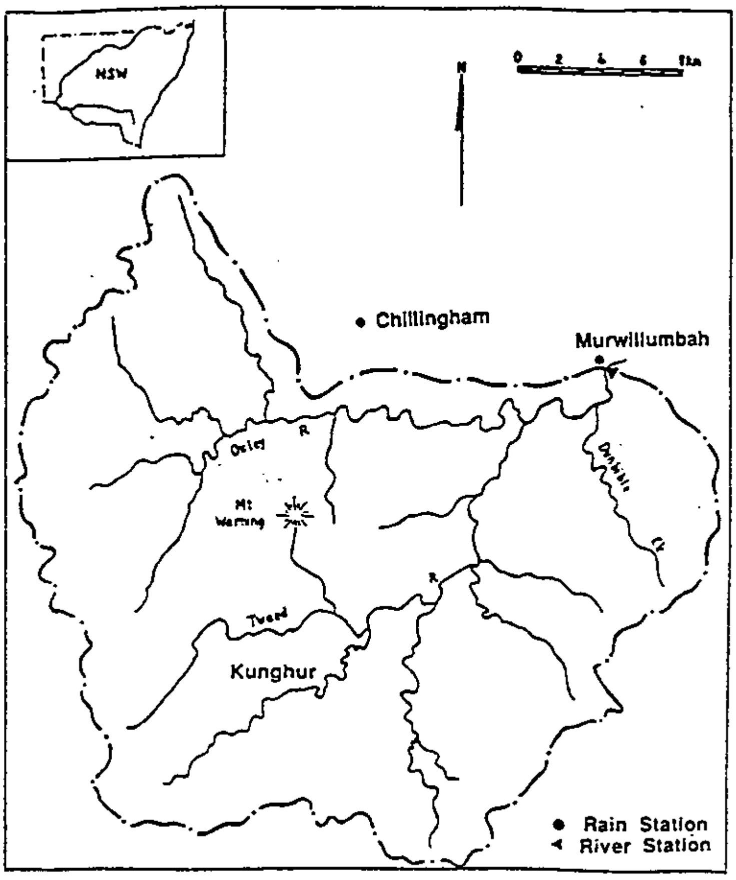


Figure 4.1: Tweed River Basin

4.1.1.1 Unit Hydrograph model

Four events were used for the calibration of the UH model. During the storm, cumulative rainfall up to a lag-time ahead of the initial significant rise in runoff was considered as initial loss. Once the direct runoff (DRO) was obtained from the flood hydrograph, runoff coefficients $\alpha_r(i)$ for rainfall following initial loss were determined for each event; these are given in Table 4.1.

Table 4.1: Runoff Coefficients

Flood Number	Runoff Coefficient (α_r)
1	0.49
2	0.30
3	0.60
4	0.75

Using DRO and rainfall excess Pe (= $\alpha_r(i)$.Rf), the UH ordinates were obtained for each event separately using the recursive least squares method (which is a recursive form of the standard least squares method). Table 4.2 shows the results (see also Fig. 4.1.1).

Table 4.2: Estimation of Unit Hydrograph Ordinates

Flood No	Ordinate Number									
	1	2	3	4	5	6	7	8	9	10
1	0.7	12.3	20.3	9.5	6.5	3.8	1.4	1.2	1.0	1.8
2	0.9	13.2	18.6	10.5	7.3	6.8	5.3	0.8	-1.9	-1.5
3	0.5	9.3	14.2	10.8	7.1	3.3	1.7	2.1	2.5	5.2
4	-0.1	9.8	16.2	9.7	5.5	5.8	3.0	3.0	1.2	0.8
Av*	0.5	11.2	17.3	10.1	6.6	4.9	2.9	1.8	1.4	0.8
(α_{rm})(Av)+	0.3	5.6	8.7	5.1	3.3	2.5	1.5	0.9	0.7	0.4

* Av = average of ordinates

\+ (α_{rm}) = average of runoff coefficients.

Since the $\alpha_r(i)$ varies widely from storm to storm and is difficult to estimate in real-time, the average value of the UH ordinates are obtained and then multiplied by the mean value of α_r (= α_{rm}). This is used as the initial estimate of the operational UH ordinates for all the events considered. Since the operational UH ordinates incorporate the runoff coefficient, the lumped rainfall can be used directly. The UH ordinates

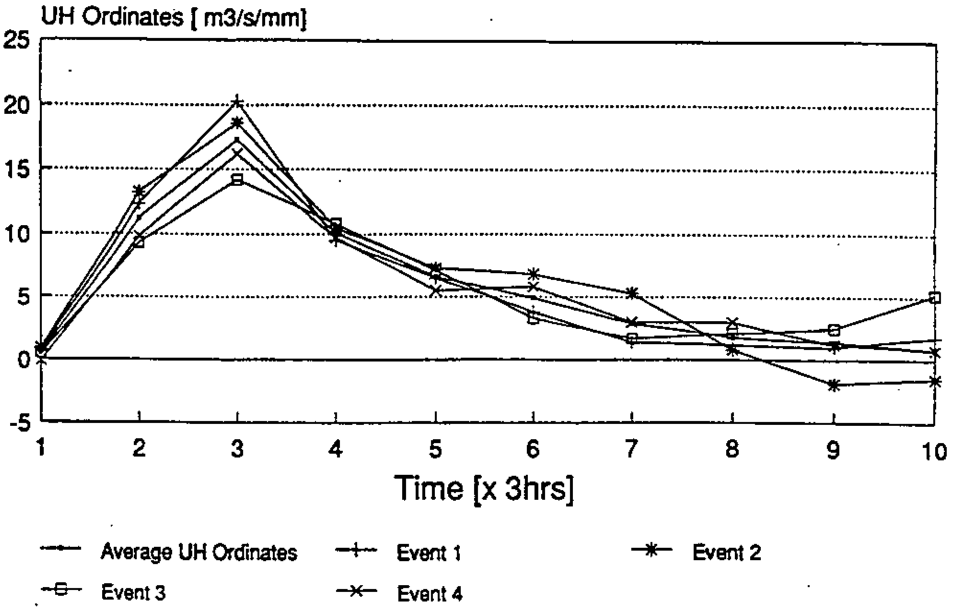


Fig. 4.1.1 Estimation of UH ordinates for Tweed River to Murwillambah

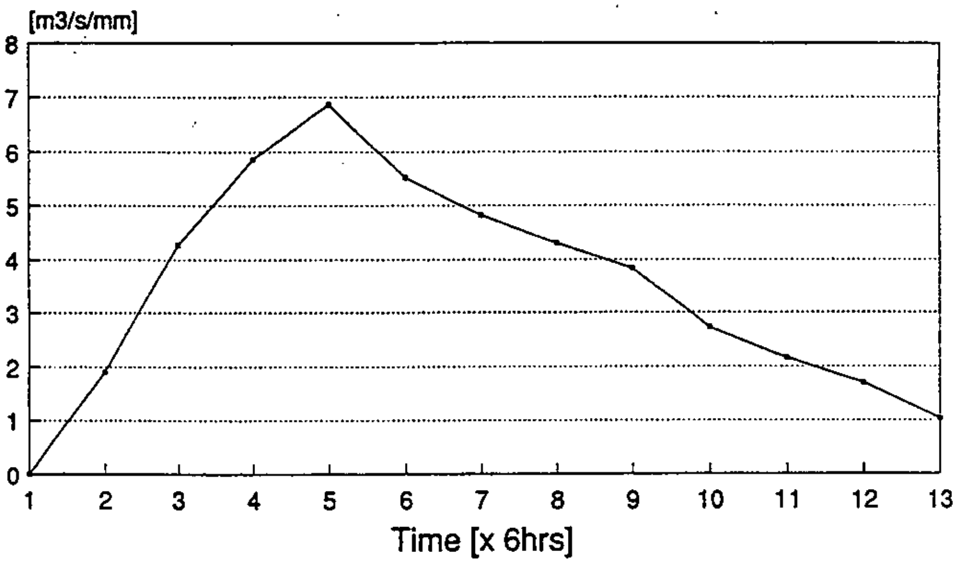


Fig. 4.2.1 Operational UH ordinates for Wilson River to Lismore.

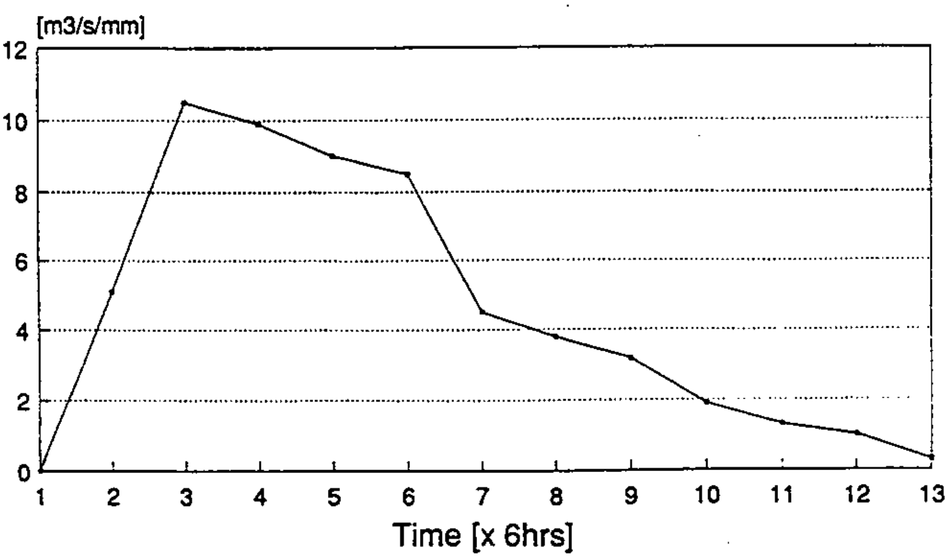


Fig. 4.3.1 Operational UH ordinates for Mitchell River to Glenaladale.

(model parameters) are subsequently updated for each time step during the event. The 3-hour and 6-hour ahead forecasts are estimated using the following equations, rerspectively:

$$q_{k+1/k} = H^T{}_{k+1/k} \cdot X_k \tag{4.1}$$

$$q_{k+2/k} = H^T{}_{k+2/k} \cdot X_k \tag{4.2}$$

where $H^T{}_{k+1/k}$ = [0,Rf(k-lg), Rf(k-lg-1) .... Rf(k-lg-m+1), q_b^*]
$H^T{}_{k+2/k}$ = [0,0,Rf(k-lg) .... Rf(k-lg-m+2), q_b^*]
X_k = $[h^*(1), h^*(2) .... h^*(i), \alpha]^T{}_k$
Rf = lumped rainfall.

4.1.1.2 ARMAX model

The order of the ARMA model and its parameters was estimated by trial and error, using the direct runoff and rainfall excess for the same flood events mentioned earlier. The model order was increased until any further changes did not significantly decrease the mean squares of errors. The final model structure is given below:

$$q_k = \delta_1 q_{k-1} + \delta_2 q_{k-2} + \omega_1 p_{e(k-1)} + \omega_2 p_{e(k-2)} + \omega_3 p_{e(k-3)} + v_{(k)} \tag{4.3}$$

As the estimation of rainfall excess in real-time is difficult, the MA parameters (ω_1, ω_2, ω_3) are multiplied by the mean value of runoff coefficients, giving the initial values of parameters for all the flood events. Thus the model with initial values of parameters is given by:

$$q_k = 1.2q_{k-1} - 0.35q_{k-2} + 4.2Rf_{(k-1)} + 2.5Rf_{(k-2)} - 1.4Rf_{(k-3)} + v_{(k)} \tag{4.4}$$

The corresponding operational UH ordinates of the ARMA model are given in Table 4.3. It can be noted that these values compare closely to the ones used in the UH model (Table 4.2).

Table 4.3: Operational UH Ordinates Corresponding to the ARMA [2,3] Model

	Ordinate value (m^3/s @ 3 hr intervals)											
Interval	1	2	3	4	5	6	7	8	9	10	11	12
Value	0.0	4.2	7.5	6.2	4.8	3.6	2.6	1.5	0.9	0.6	0.5	0.3

The 3-hour and 6-hour ahead forecasts are estimated using the following equations, respectively:

$$q_{k+1/k} = H^T_{k+1/k} \cdot X_k \quad (4.5)$$

$$q_{k+2/k} = H^T_{k+2/k} \cdot X_k \quad (4.6)$$

where $H^T_{k+1/k} = [q_k, q_{k-1}, Rf(k), Rf(k-1), Rf(k-2)]$
$H^T_{k+2/k} = [q_{k+1/k}, q_k, 0, Rf(k-2), Rf(k-3)]$
$X_k = [\delta_1, \delta_2, \omega_1, \omega_2, \omega_3]^T$
Rf = lumped rainfall.

4.1.1.3 Stochastic rainfall prediction for the Tweed catchment

The model structure was identified using the rainfall sequences of the four flood events considered for calibration, where the model order was varied until any further increase in the order did not significantly improve the mean squares of errors. The model based on the mean values of the parameters derived for each of the four events is given below:

$$Rf_{k/k-1} = 1.097Rf(k-1) - 0.252Rf(k-2) - 0.036Rf(k-3) + vr(k) \quad (4.7)$$

where vr(k) is a white noise. These parameter values were considered as initial values for this model and the Kalman filter was used to update the parameters during the storm, as mentioned in Section 3.1.3.

The 3-hours and 6-hours ahead rainfall predictions are given by:

$$Rf_{k+1/k} = Hr^T_{k+1/k} \cdot X\eta_k$$

$$Rf_{k+2/k} = Hr^T_{k+2/k} \cdot X\eta_k$$

where $Hr^T_{k+1/k} = [Rf(k), Rf(k-1), Rf(k-2)]$
$Hr^T_{k+2/k} = [Rf_{k+1/k}, Rf(k), Rf(k-1)]$

These predicted rainfall values are considered only if they are positive, if not, they are taken as zero.

4.1.2 *Wilson River to Lismore*

The Wilson River basin upstream of Lismore is located immediately south of the Tweed River basin in the north-eastern corner of New South Wales and has a catchment area of about 1400 km^2. The river system above Lismore has a fan-shaped drainage pattern with three distinct drainage areas. All three rise in the elevated ranges about 1200m above mean sea level to the north Lismore and radiate southward. Each sub-catchment, individually, can be responsible for flooding at Lismore. The upper reaches are covered by thick rainforest, whilst in the lower reaches much of the vegetation has been removed and is covered by pastoral grasslands. Historically, Lismore is probably the most flood-prone town in Australia, with seventy-one large floods being recorded since 1890. Rainfall over the catchment shows marked temporal and spatial variation, with totals during floods up to 450 mm in the eastern part of the basin and as little as 25 mm in the western part.

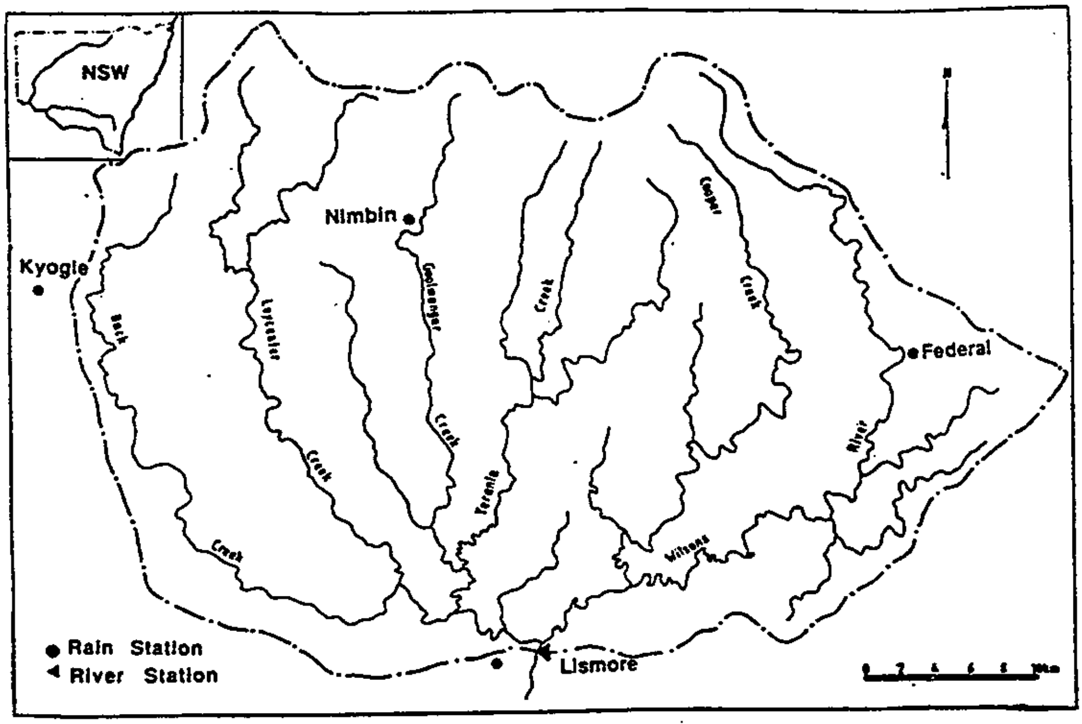


Figure 4.2: Wilson River Basin

4.1.2.1 Unit hydrograph model

Four events were used for the calibration of the UH model, as was the case for the Tweed River. The runoff coefficients $\alpha_r(i)$ varied between 0.6 and 0.7 for these events. The procedure mentioned in Section 4.1.1.1 was adopted to determine the operational 6-hour UH ordinates. The values of the operational 6-hour ordinates are given in Table 4.4 (see also Fig. 4.2.1, p.15).

Table 4.4: Operational UH Ordinates for the Wilson River

	Ordinate value (m^3/s @ 6 hr intervals)												
Interval	1	2	3	4	5	6	7	8	9	10	11	12	13
Value	0.0	1.91	4.27	5.86	6.86	5.52	4.82	4.31	3.84	2.72	2.16	1.70	1.03

The 6-hour and 12-hour ahead forecasts are estimated using the following equations, respectively:

$$q_{k+2/k} = H^T{}_{k+2/k} \cdot X_k \quad (4.8)$$

$$q_{k+4/k} = H^T{}_{k+4/k} \cdot X_k \quad (4.9)$$

where $H^T{}_{k+2/k}$ = [0, Rf(k), Rf(k-2) .... q_b^*]
$H^T{}_{k+4/k}$ = [0,0, Rf(k), Rf(k-2) .... q_b^*]
X_k = $[h^*(1), h^*(2) .... h^*(i), \alpha]^T{}_k$
Rf = lumped rainfall.

4.1.2.2 ARMAX model

The order of the ARMA model and its parameters was estimated by fitting to the operational UH ordinates obtained in Section 4.1.2.1. The model order was increased until any further changes did not significantly decrease the mean squares of errors. An example of this procedure is given in Appendix 1. The model structure is given below:

$$q_k = \delta_1 q_{k-1} + \delta_2 q_{k-2} + \omega_1 Rf_{(k-1)} + \omega_2 Rf_{(k-2)} + \omega_3 Rf_{(k-3)} + \omega_4 Rf_{(k-4)} + v_{(k)} \quad (4.10)$$

The ARMAX model with initial values of parameters is given by:

$$q_k = 0.95q_{k-1} - 0.13q_{k-2} + 2.0Rf_{(k-1)} + 3.3Rf_{(k-2)} - 3.4Rf_{(k-4)} + 3.0Rf_{(k-6)} + v_{(k)} \quad (4.11)$$

The 6-hour and 12-hour ahead forecasts are estimated using the following equations, respectively:

$$q_{k+2/k} = H^T{}_{k+2/k} \cdot X_k \quad (4.12)$$

$$q_{k+4/k} = H^T{}_{k+4/k} \cdot X_k \quad (4.13)$$

where $H^T{}_{k+2/k}$ = [q_k, q_{k-1}, Rf(k), Rf(k-2), Rf(k-4), Rf(k-6)]
$H^T{}_{k+4/k}$ = [$q_{k+2/k}$, q_k, 0, Rf(k-2), Rf(k-4), Rf(k-6)]
X_k = $[\delta_1, \delta_2, \omega_1, \omega_2, \omega_3, \omega_4]^T$
Rf = lumped rainfall (at 6-hour intervals).

4.1.3 *Mitchell River to Glenaladale*

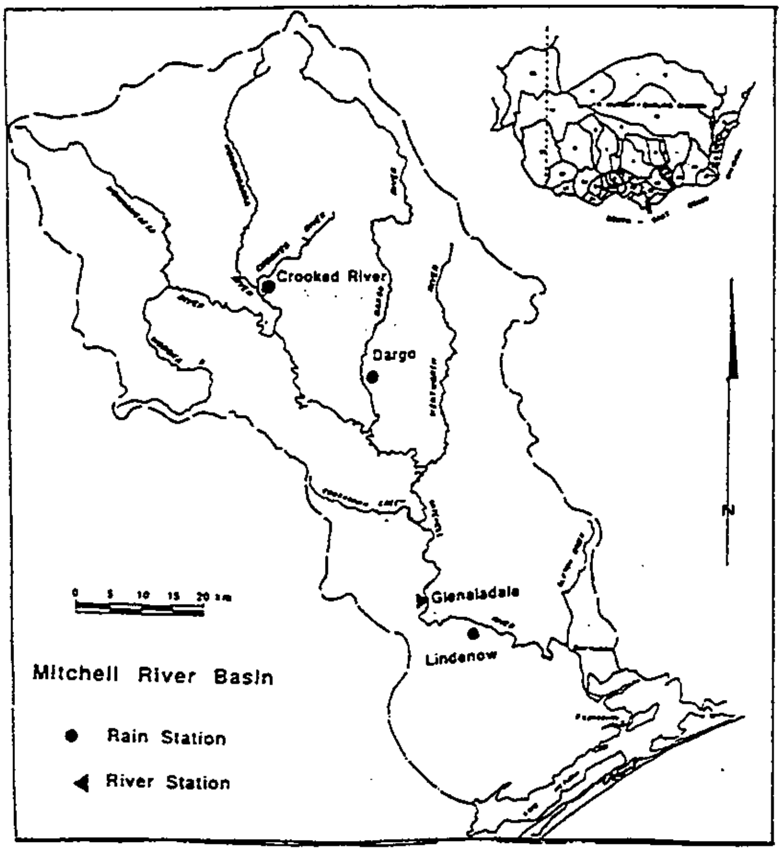


Figure 4.3: Mitchell River Basin

4.1.3.1 Unit hydrograph model

As data for three events only was available, two of the events were used for the calibration of the UH model. The runoff coefficients $\alpha_r(i)$ for the two events were 0.27 and 0.37. The procedure mentioned in Section 4.1.1.1 was adopted to determine the operational 6-hour UH ordinates. The values of the operational 6-hour UH ordinates are given in Table 4.5 (see also Fig. 4.3.1, p.15).

Table 4.5: Operational UH Ordinates for the Mitchell River to Glenaladale

	Ordinate value (m^3/s @ 6 hr intervals)												
Interval	1	2	3	4	5	6	7	8	9	10	11	12	13
Value	0.0	5.1	10.5	9.9	9.0	8.5	4.5	3.8	3.2	1.9	1.3	1.0	0.3

The 6-hour and 12-hour ahead forecasts are estimated using the following equations, respectively:

$$q_{k+2/k} = H^T{}_{k+2/k} \cdot X_k \tag{4.14}$$

$$q_{k+4/k} = H^T{}_{k+4/k} \cdot X_k \tag{4.15}$$

where $H^T_{k+2/k}$ = [0, Rf(k), Rf(k-2) .... q_b^*]
$H^T_{k+4/k}$ = [0,0, Rf(k), Rf(k-2) .... q_b^*]
X_k = [h*(1), h*2) .... h*i), $\alpha]^T_k$
Rf = lumped rainfall.

4.1.3.2 ARMAX model

The order of the ARMA model and its parameters was estimated by fitting to the operational UH ordinates obtained in Section 4.1.3.1. The model order was increased until any further changes did not significantly decrease the mean squares of errors. The model structure is given below:

$$q_k = \delta_1 q_{k-2} + \delta_2 q_{k-4} + \omega_1 Rf_{(k-2)} + \omega_2 Rf_{(k-4)} + v_{(k)} \quad (4.16)$$

The ARMAX model with initial values of parameters is given by:

$$q_k = 0.976q_{k-2} - 0.139q_{k-4} + 5.10Rf_{(k-2)} + 5.52Rf_{(k-4)} + v_{(k)} \quad (4.17)$$

The 6-hour and 12-hour ahead forecasts are estimated using the following equations, respectively:

$$q_{k+2/k} = H^T_{k+2/k} \cdot X_k \quad (4.18)$$

$$q_{k+4/k} = H^T_{k+4/k} \cdot X_k \quad (4.19)$$

where $H^T_{k+2/k}$ = [q_k, q_{k-2}, Rf(k), Rf(k-2)]
$H^T_{k+4/k}$ = [$q_{k+2/k}$, q_k, 0, Rf(k-2)]
X_k = $[\delta_1, \delta_2, \omega_1, \omega_2]^T$
Rf = lumped rainfall (at 6-hour intervals).

4.2 River Routing Model

4.2.1 *Routing upper reach of Mary River (Dagun Pocket to Gympie)*

Data for only two events were available for the study and both were used for calibration of the river routing model. The order L, J and the parameters ld and lg in Eqn (3.9) were identified using the recursive least squares method. The procedure is similar to that mentioned in Section 4.1.1.2. The mean value of the parameters was taken as the initial value for the model. The model is given by:

$$Y_{Gy}(k) = 0.66.Y_{Gy}(k-1) + 0.27.y_{Dg}(k-1) + 0.21.y_{Dg}(k-2) + v(k) \quad (4.20)$$

where Y_{Gy} = runoff at Gympie
y_{Dg} = runoff at Dagun Pocket.

The model can be written as:

$$YGy(k) = H_k \cdot X_k + V(k)$$

where $H_k = [Y_{Gy}(k\text{-}1), y_{Dg}(k\text{-}1), y_{Dg}(k\text{-}2)]$
$X^T_k = [\alpha_1, \beta_1, \beta_2]$
$V(k)$ = measurement error ~ N(0,R).

3-hours ahead prediction is given by:

$$YGy(k+1) = H_{k+1} \cdot X_k$$

where $H_{k+1} = [Y_{Gy}(k), y_{Dg}(k), y_{Dg}(k\text{-}1)]$.

4.3 Multiple Input - Single Output Models (MISO)

4.3.1 *Mary River to Home Park*

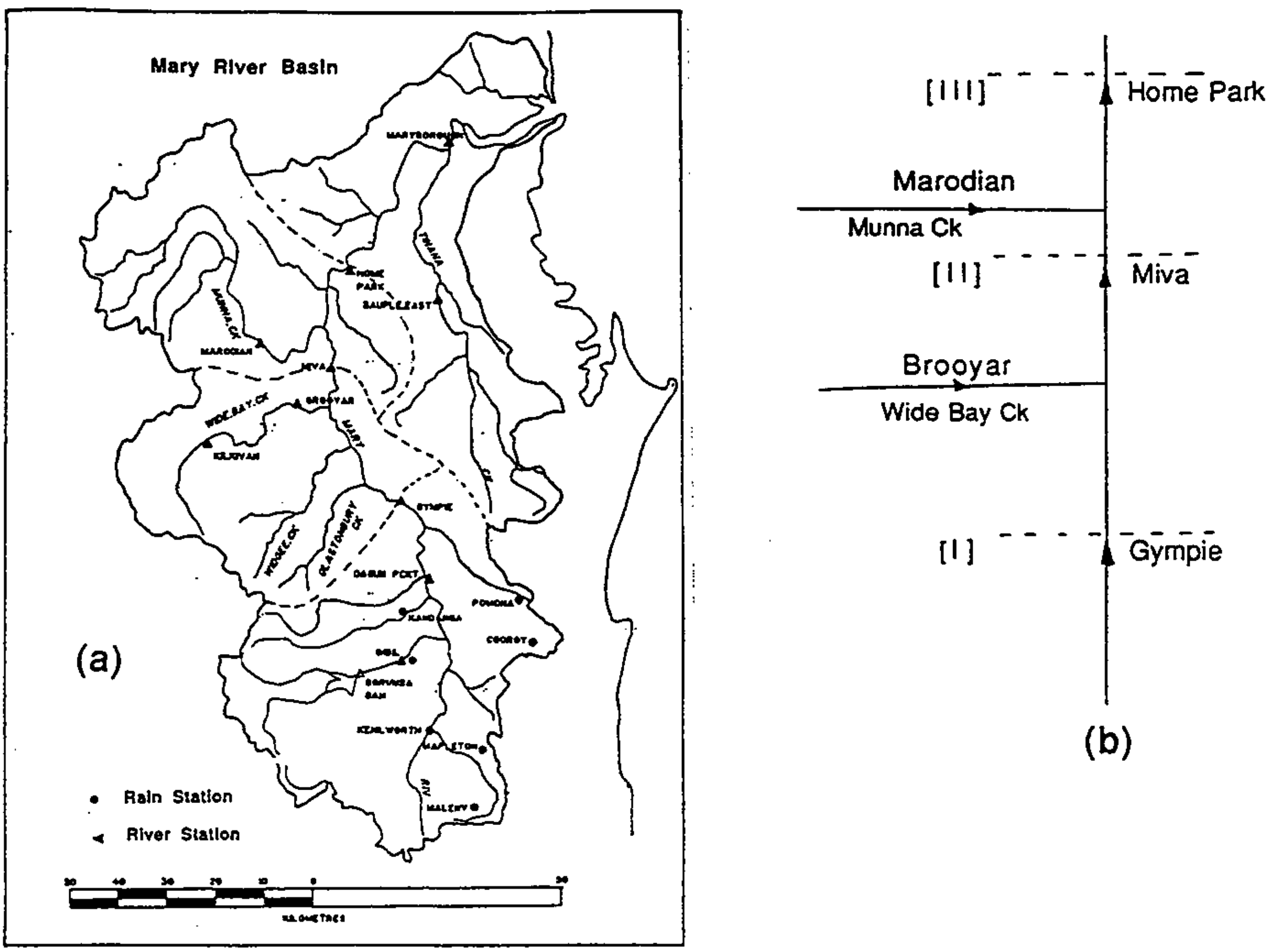


Figure 4.4: Mary River Basin

Fig. 4.4(a) shows the catchment area with stream gauging stations and rainfall stations. Mary River up to Home Park is considered in this study. The schematic diagram of the network is shown in Fig. 4.4(b). The modelling study was approached in three stages. At stage I, a rainfall-runoff model was used to predict the runoff at Gympie. At stage II, the multiple inputs are the flows at Gympie and Brooyar and the output is the flow at Miva. Finally, at stage III, flows at Miva and Marodian are considered as inputs, while the flow at Home Park is considered as an output.

4.3.1.1 Rainfall-runoff model to Gympie

Data for three events were available for the study. Two events were used for the calibration of the model. The procedure given in Section 4.1.1.1 was adopted for obtaining the operational UH. The runoff coefficients $\alpha_r(i)$ varied between 0.69 and 0.77 for these events. The values of the operational 6-hour UH ordinates are given in Table 4.6.

Table 4.6: Operational UH Ordinates for the Mary River to Gympie

	Ordinate value (m^3/s @ 6 hr intervals)												
Interval	1	2	3	4	5	6	7	8	9	10	11	12	13
Value	1.9	5.2	6.0	7.6	15.3	16.5	14.5	9.5	5.3	4.9	4.8	4.5	2.1

The order of the ARMAX model and its parameters was estimated and the model with initial values is given by:

$$q_k = 1.31q_{k-2} - 0.45q_{k-4} + 2.0Rf_{(k-2)} + 3.3Rf_{(k-4)} - 3.4Rf_{(k-6)} + 3.0Rf_{(k-8)} + v_{(k)} \quad (4.21)$$

The 6-hour and 12-hour ahead forecasts were estimated using the following equations respectively:

$$q_{k+2/k} = H^T_{k+2/k} \cdot X_k$$

$$q_{k+4/k} = H^T_{k+4/k} \cdot X_k$$

where $H^T_{k+2/k} = [q_k, q_{k-2}, Rf(k), Rf(k-2), Rf(k-4), Rf(k-6)]$

$H^T_{k+4/k} = [q_{k+2/k}, q_k, 0, Rf(k-2), Rf(k-4), Rf(k-6)]$

$X_k = [\delta_1, \delta_2, \omega_1, \omega_2, \omega_3, \omega_4]^T$

Rf = lumped rainfall (at 6-hour intervals).

4.3.1.2 River routing to Miva

Data for four recent events were available for the study and all of them were used for calibration. The order L, J_1, J_2 and the parameters ld, lg_1 and lg_2 in Eqn (3.10) were identified using the recursive least squares method. The procedure is similar to that given in Section 4.1.1.2. The mean value of the parameters was taken as the initial value of the model. The model is given by:

$$Y_{Mv}(k) = 0.69Y_{Mv}(k-2) + 0.61y_{Gy}(k-2) - 0.31y_{Gy}(k-4) + 0.61y_{Br}(k-2) + v(k) \quad (4.22)$$

where Y_{Mv} = runoff at Miva
y_{Gy} = runoff at Gympie
y_{Br} = runoff at Brooyar.

The model can be written as:

$$Y_{Mv}(k) = H_k . X_k + V(k)$$

where H_k = $[Y_{My}(k-2), y_{Gy}(k-2), y_{Gy}(k-4), y_{Br}(k-2)]$
X^T_k = $[\alpha_1, \beta_{1,1}, \beta_{1,2}, \beta_{2,1}]$
$V(k)$ = measurement error $\cong N(0,R)$.

6-hours ahead prediction is given by:

$$YGy(k+2/k) = H_{k+2} . X_k$$

where H_{k+2} = $[Y_{Gy}(k), y_{Dg}(k), y_{Dg}(k-2), y_{Br}(k)]$.

4.3.1.3 River routing to Home Park

Data for the same four events used in 4.3.1.2 were used for calibration. The order L, J_1, J_2 and the parameters ld, lg_1 and lg_2 in Eqn (3.10) were identified using the procedure given in Section 4.3.1.2. The mean value of the parameters was taken as the initial value of the model. The model is given by:

$$Y_{Hp}(k) = 0.69Y_{Hp}(k-3) + 0.61y_{Mv}(k-3) - 0.31y_{Mv}(k-6) + 0.61y_{Mr}(k-3) + v(k) \quad (4.23)$$

where Y_{Hp} = runoff at Home Park
y_{Mv} = runoff at Miva
y_{Mr} = runoff at Marodian.

The model can be written as:

$$Y_{Hp}(k) = H_k . X_k + V(k)$$

where $H_k = [Y_{Hp}(k-3), y_{Mv}(k-3), y_{Mv}(k-6), y_{Mr}(k-3)]$

$X^T{}_k = [\alpha_1, \beta_{1,1}, \beta_{1,2}, \beta_{2,1}]$

$V(k) =$ measurement error $\sim N(0,R)$.

9-hours ahead prediction is given by:

$$YGy(k+3/k) = \dot{H}_{k+3} . X_k$$

where $H_{k+3} = [Y_{Hp}(k), y_{Mv}(k), y_{Mv}(k-2), y_{Mr}(k)]$.

5. DISCUSSION OF RESULTS

The following least squares criteria were used for comparing the results of different methods.

(i) *Coefficient of efficiency*

A measure of association between predicted and observed flows is given by the sum of the squares of the residuals:

$$S_r = \sum_{j=1}^{n} [q_{obs}(j) - q_{pred}(j)]^2$$

where $q_{obs}(j)$ and $q_{pred}(j)$ are the observed and predicted flows, respectively. A measure of variability of the observed flows is given by:

$$S_q = \sum_{j=1}^{n} [q_{obs}(j) - q_{obs}]^2$$

where q_{obs} is the mean observed flow at the time interval considered (1 to n). Nash and Sutcliffe [1] introduced the efficiency of a model (or the coefficient of efficiency) as the variance of the observed flow accounted for by the model:

$$E = 1 - \frac{S_r}{S_q}$$

(ii) *Coefficient of persistence*

In real-time, the no-model prediction is better represented by the latest measurement of the flow. Thus, a coefficient of persistence P for real-time models can be defined by:

$$P_{(ld)} = 1 - \frac{\sum_{j=1}^{n} \left[q_{obs}(j) - q_{pred}(j)\right]^2}{\sum_{j=1}^{n} \left[q_{obs}(j) - q_{obs}(j - ld)\right]^2}$$

where 'ld' is the predicted lead and P is a function of the forecast lead 'ld'.

5.1 Rainfall-Runoff Models

5.1.1 *Tweed River*

The estimation of the typical UH ordinates clearly indicates that the average time to peak lies around 6-9 hours and this observation is further supported by the ARMA (2,3) model parameters. The model was applied to the four events used in the calibration, and six other events, for testing of the models. In addition, the model was applied with and without forecast of future rainfall.

A. Flood flow forecasting without future rainfall input

Adaptive linear models were applied to all the events, omitting the terms in the models which had future rainfall input. Figs. 5.1.1.1 and 5.1.1.2 show the performance of these two linear models on the events used for calibration. Real-time forecasting performance (3-hours and 6-hours ahead) on the events used for testing is shown in Figs. 5.1.1.3, 5.1.1.4 and 5.1.1.5. The overall performance of the two models was reasonably satisfactory [19]. It was observed that the reason for a minimal warning for event number 8 was that the main rainfall cells associated with this event moved down the catchment towards the outlet of the basin [20]. This results in a short time to peak; the adaptive models tend to perform better for this event to a certain extent, as shown in Fig. 5.1.1.4(b & d). The statistics of forecasting performance of adaptive UH and ARMA models are given in Tables 5.1.1.1 and 5.1.1.2, respectively.

The overall performance of these two linear models was similar, although the ARMAX model required a lesser number of parameters. As a consequence of not considering future rainfall predictions, forecasts were increasingly underestimated as the forecasting lead time increased, due to the omission of some terms in the model.

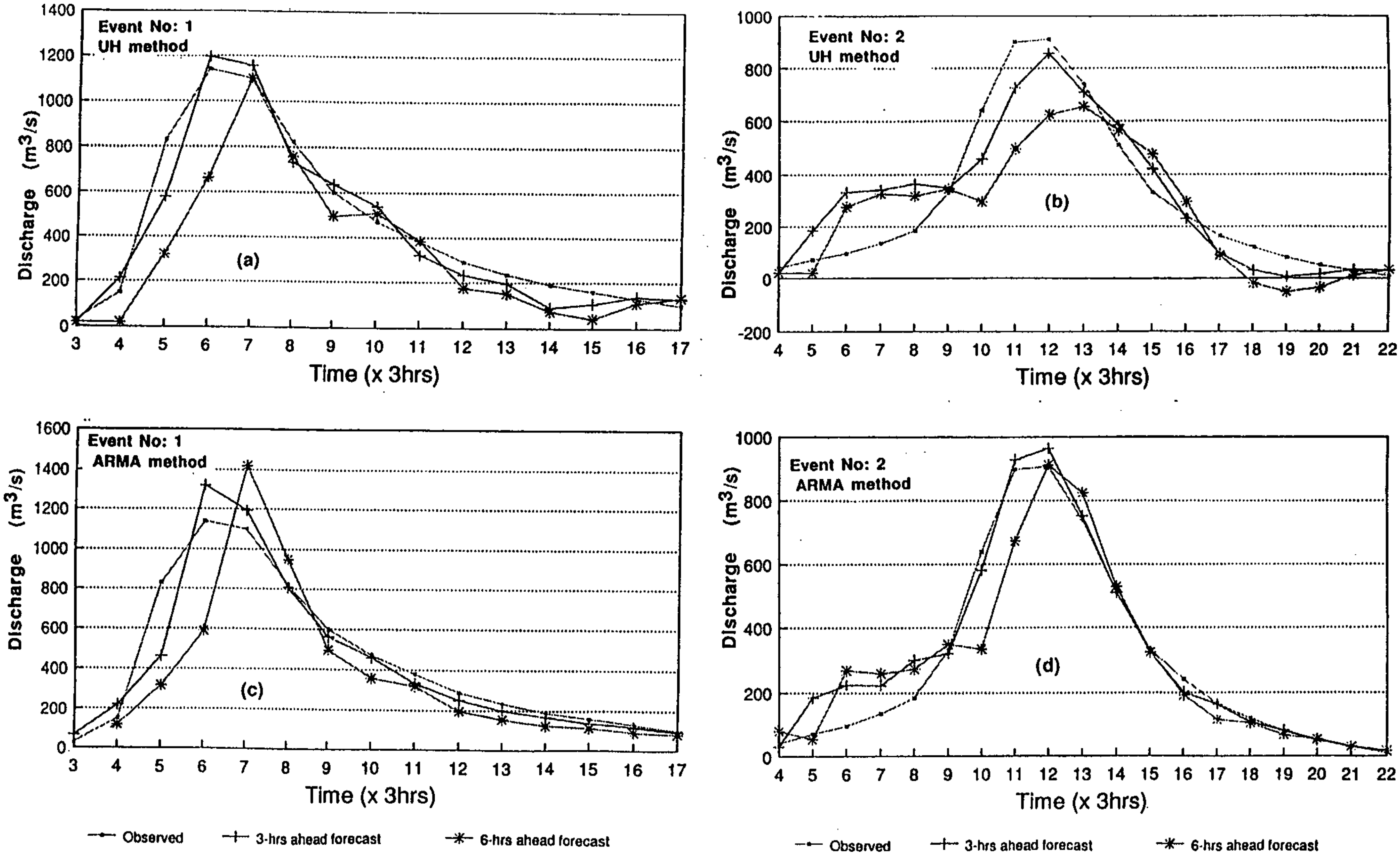


Fig. 5.1.1.1 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (without predicted rainfall)

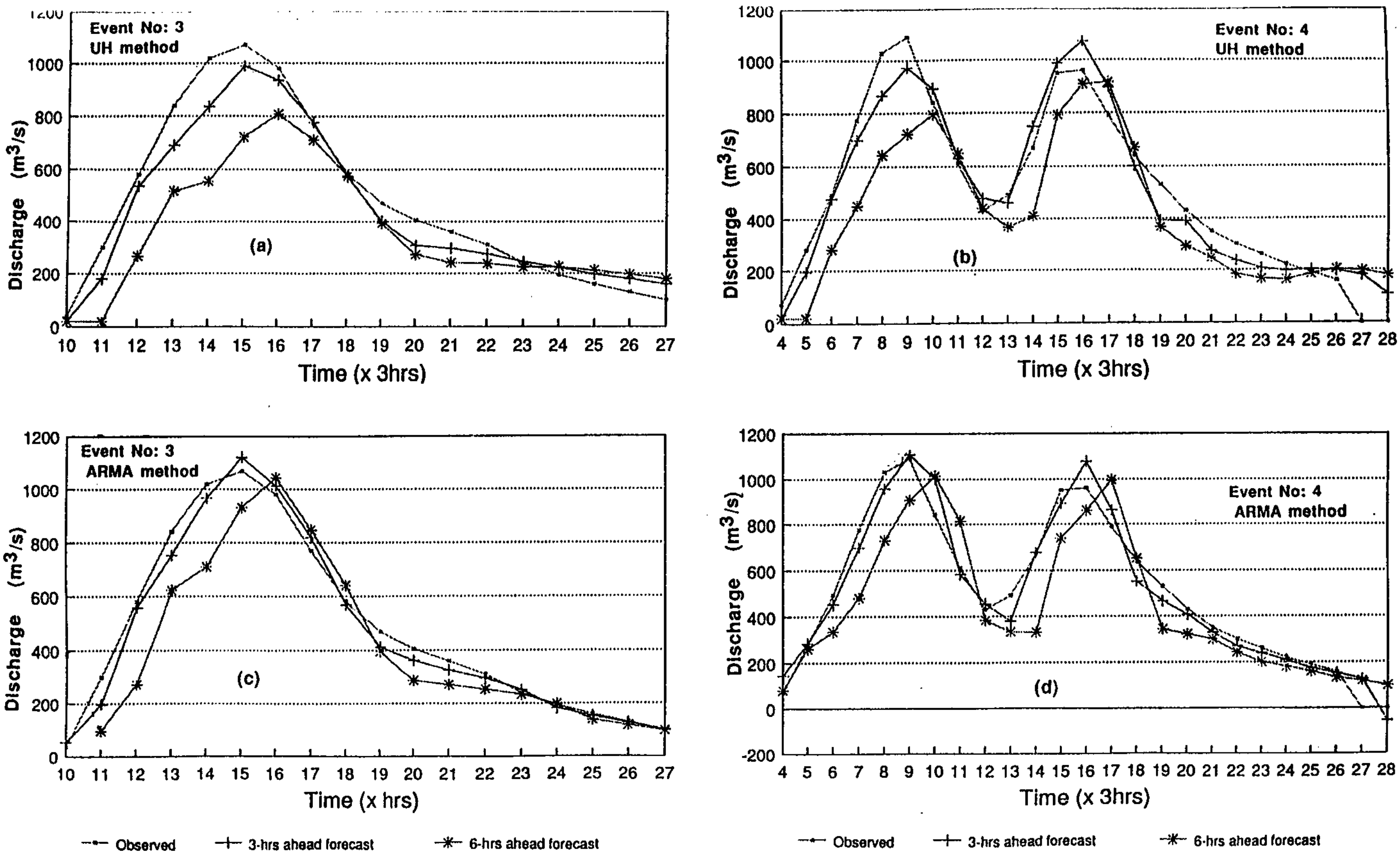


Fig. 5.1.1.2 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (without predicted rainfall)

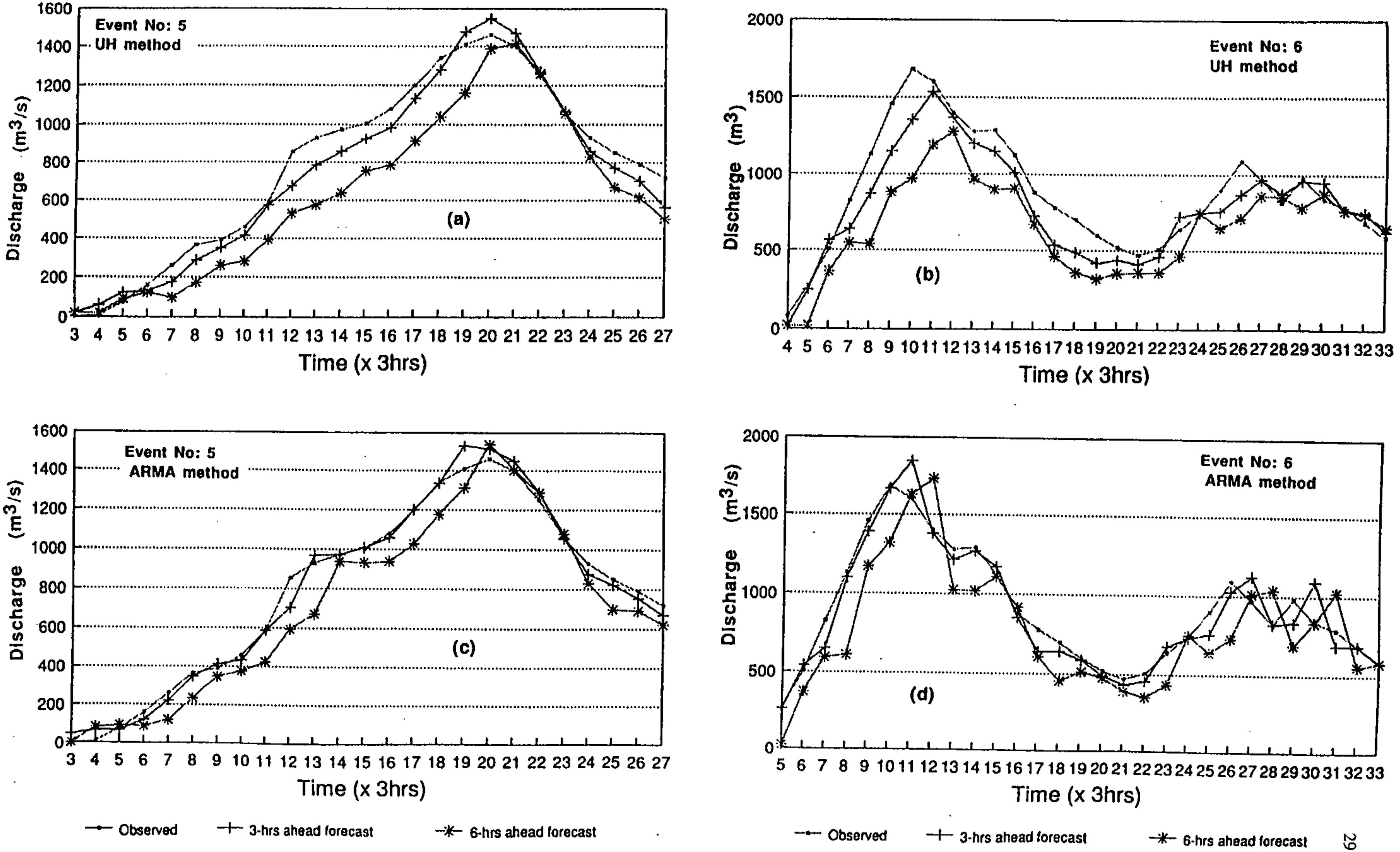


Fig. 5.1.1.3 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (without predicted rainfall)

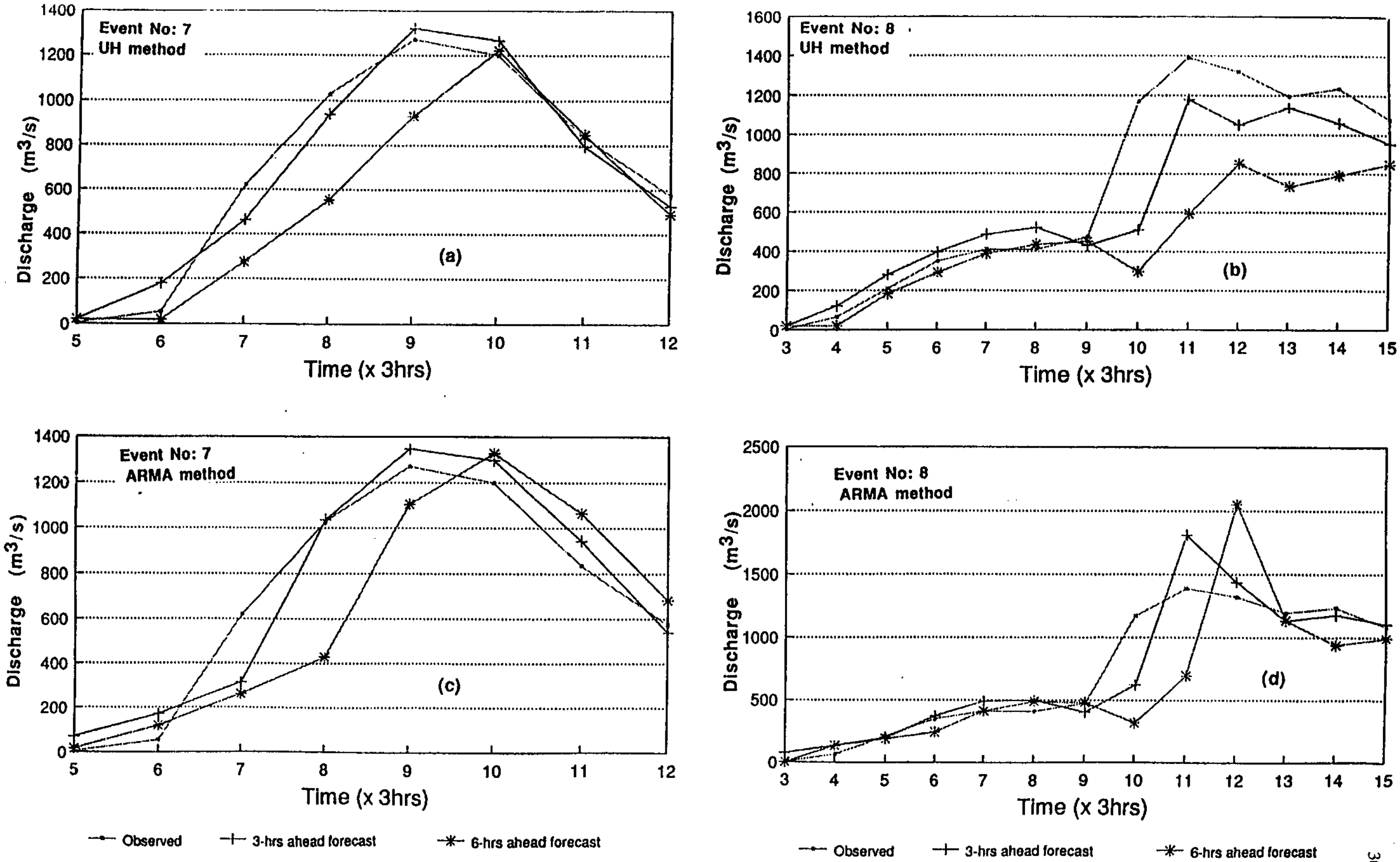


Fig. 5.1.1.4 Observed and predicted flow for Tweed river to Murwillambah

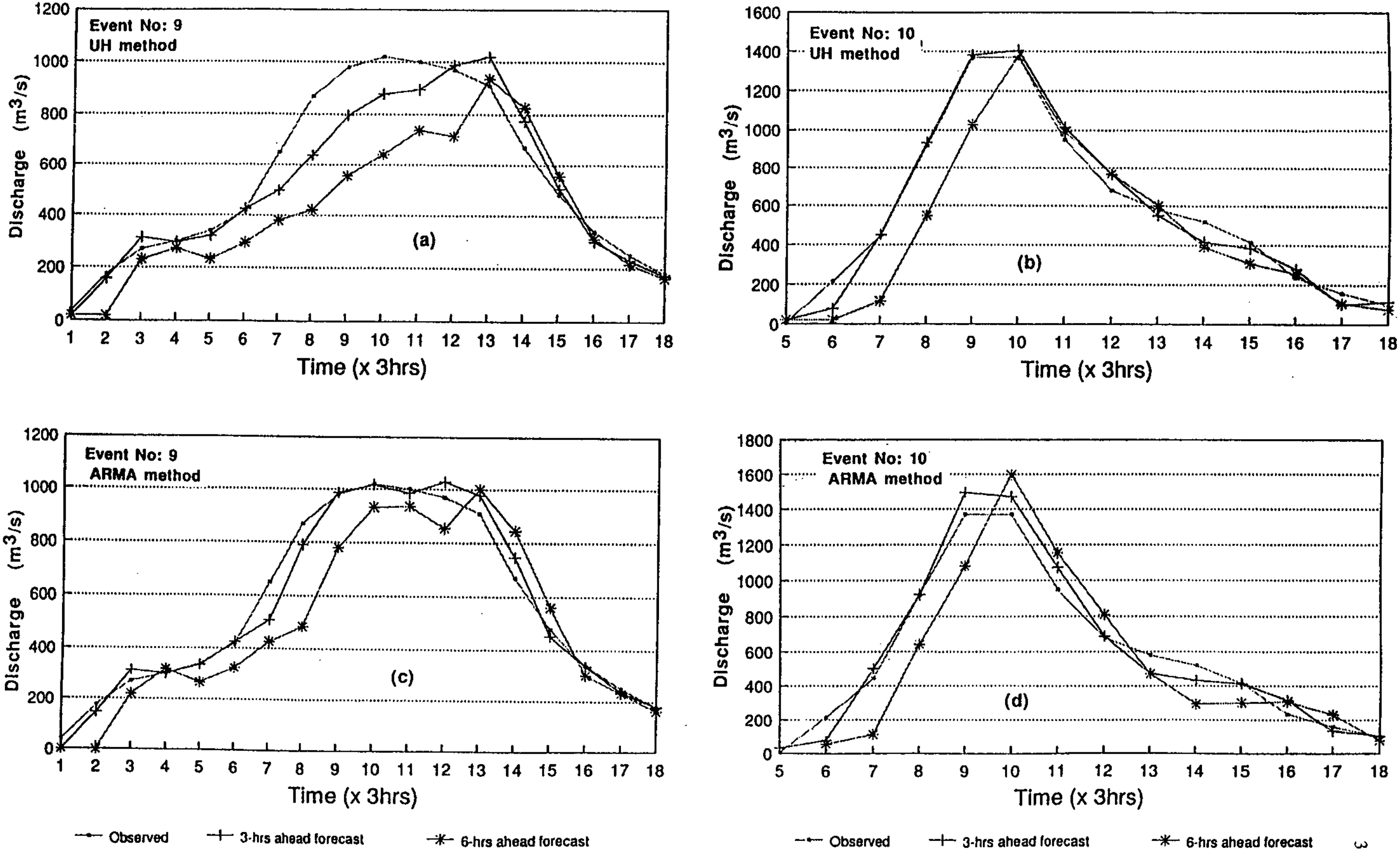


Fig. 5.1.1.5 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (without predicted rainfall)

Table 5.1.1.1: Forecast Statistics for the On-Line UH Model

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		3 hrs	6 hrs	3 hrs	6 hrs	3 hrs	6 hrs
1	484	0.19	0.42	93.7	66.0	85.0	24.0
2	315	0.38	0.55	84.3	60.2	24.9	51.0
3	507	0.16	0.39	93.9	62.1	65.2	13.4
4	526	0.16	0.33	92.8	66.6	70.2	53.3
5	832	0.10	0.25	96.1	73.3	36.7	-1.3
6	910	0.16	0.32	85.5	30.3	34.7	3.3
7	860	0.12	0.30	94.3	13.6	91.9	60.6
8	810	0.30	0.52	74.1	-11.4	-10.5	-39.2
9	590	0.17	0.37	90.0	47.6	37.5	21.6
10	630	0.10	0.28	97.6	79.4	93.2	81.4

Table 5.1.1.2: Forecast Statistics for the On-Line ARMAX Model

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		3 hrs	6 hrs	3 hrs	6 hrs	3 hrs	6 hrs
1	484	0.25	0.48	89.1	54.9	74.0	-0.8
2	315	0.19	0.35	96.1	85.3	81.5	82.0
3	507	0.09	0.27	97.8	81.7	87.5	58.0
4	526	0.13	0.31	95.0	70.5	79.4	58.9
5	832	0.06	0.15	98.6	90.3	77.7	64.3
6	910	0.11	0.24	92.0	57.3	65.3	51.7
7	860	0.17	0.34	88.0	-29.8	82.9	80.3
8	810	0.27	0.49	78.1	-8.2	6.9	-25.2
9	590	0.09	0.24	97.2	78.7	83.4	70.3
10	630	0.14	0.31	96.3	75.9	90.4	83.7

where Qbar = average observed discharge
RMSE = Root Mean Squares of Error.

B. Flood flow forecasts with predicted rainfall

In order to improve on the model performance, the 3-hours and 6-hours ahead predicted rainfall obtained from an adaptive stochastic rainfall model operated in parallel with the adaptive linear model was used.

The performance during the four flood events used in the calibration is given in Figs. 5.1.1.6 and 5.1.1.7. Figs. 5.1.1.8, 5.1.1.9 and 5.1.1.10 show the real-time forecasting performance on the events used for testing. The statistics of forecasting performance of adaptive UH and ARMA models using predicted rainfall are given in Tables 5.1.1.3 and 5.1.1.4, respectively. As might be expected, rainfall is predicted one step beyond the end of the storm, resulting in overestimation and a time delay in the peak flow, although there is a better overall fit in the rising limb of the flood hydrograph. Again, flood number 8 gives anomalous values.

Table 5.1.1.3: Forecast Statistics for the On-Line UH Model (with Predicted Rainfall using the AR(3) Model)

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		3 hrs	6 hrs	3 hrs	6 hrs	3 hrs	6 hrs
1	484	0.20	0.34	93.0	76.8	84.3	52.4
2	315	0.39	0.57	83.4	61.5	24.8	55.6
3	507	0.14	0.23	95.3	86.3	74.7	70.7
4	526	0.13	0.23	93.0	77.0	78.4	77.0
5	832	0.10	0.18	96.1	86.4	39.6	50.6
6	910	0.13	0.24	88.9	56.7	51.4	42.1
7	860	0.13	0.22	93.1	41.6	91.6	83.8
8	810	0.28	0.48	78.1	16.6	14.5	6.2
9	590	0.16	0.29	91.2	66.9	48.6	53.8
10	630	0.13	0.28	96.3	81.2	90.5	84.5

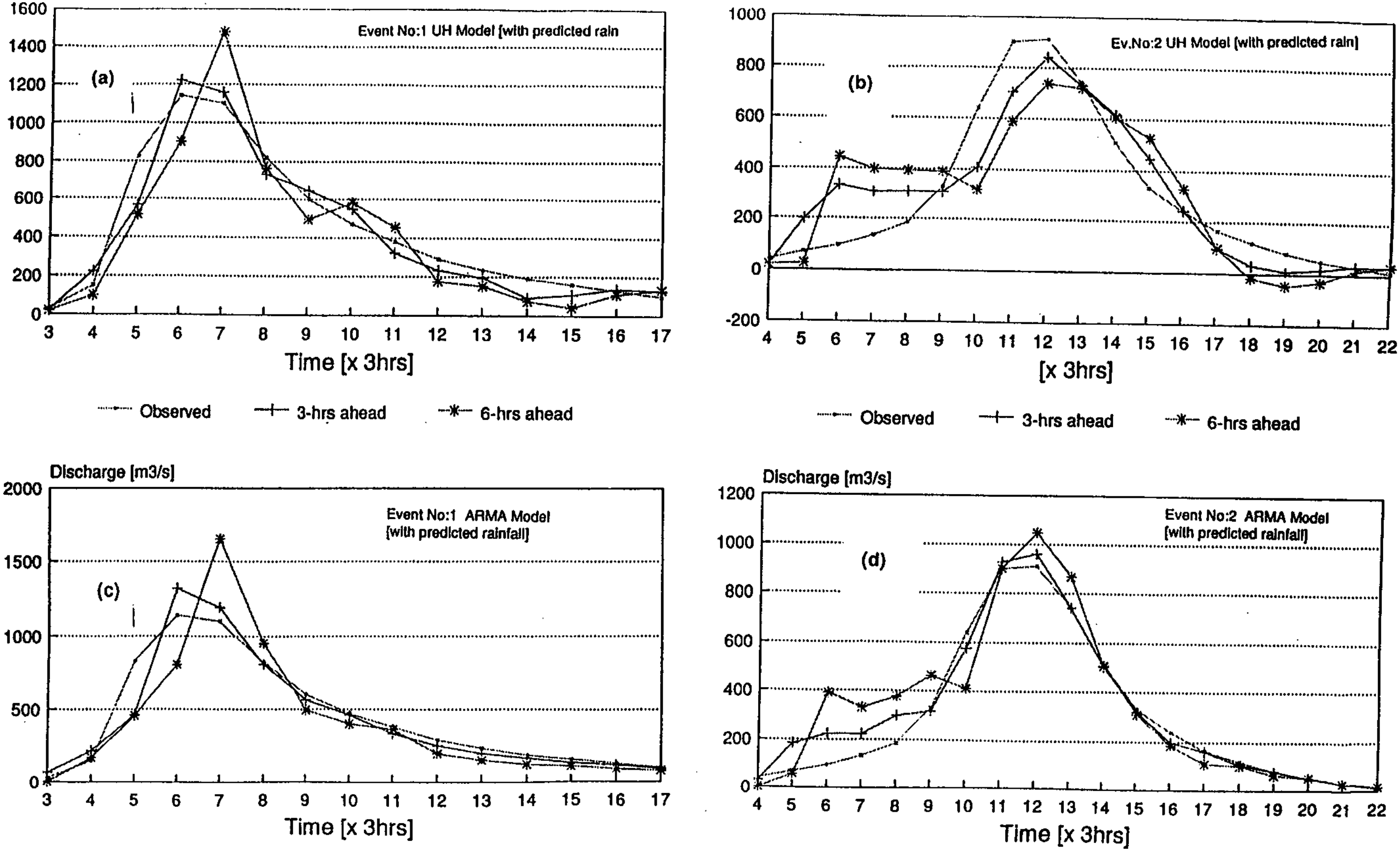


Fig. 5.1.1.6 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (with predicted rainfall)

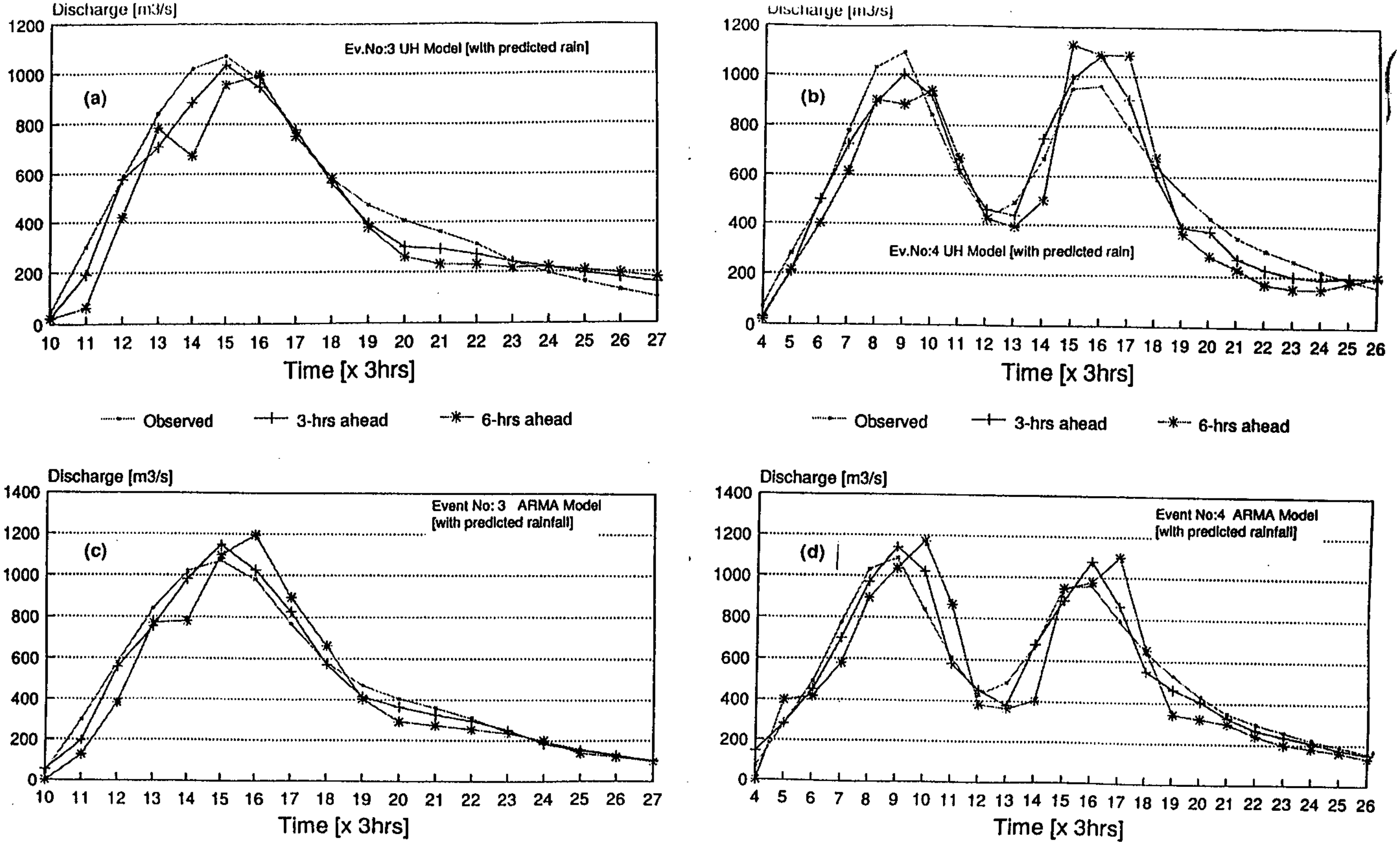


Fig. 5.1.1.7 Observed and predicted flow for Tweed river to Murwillambah
Using adaptive linear models (with predicted rainfall)

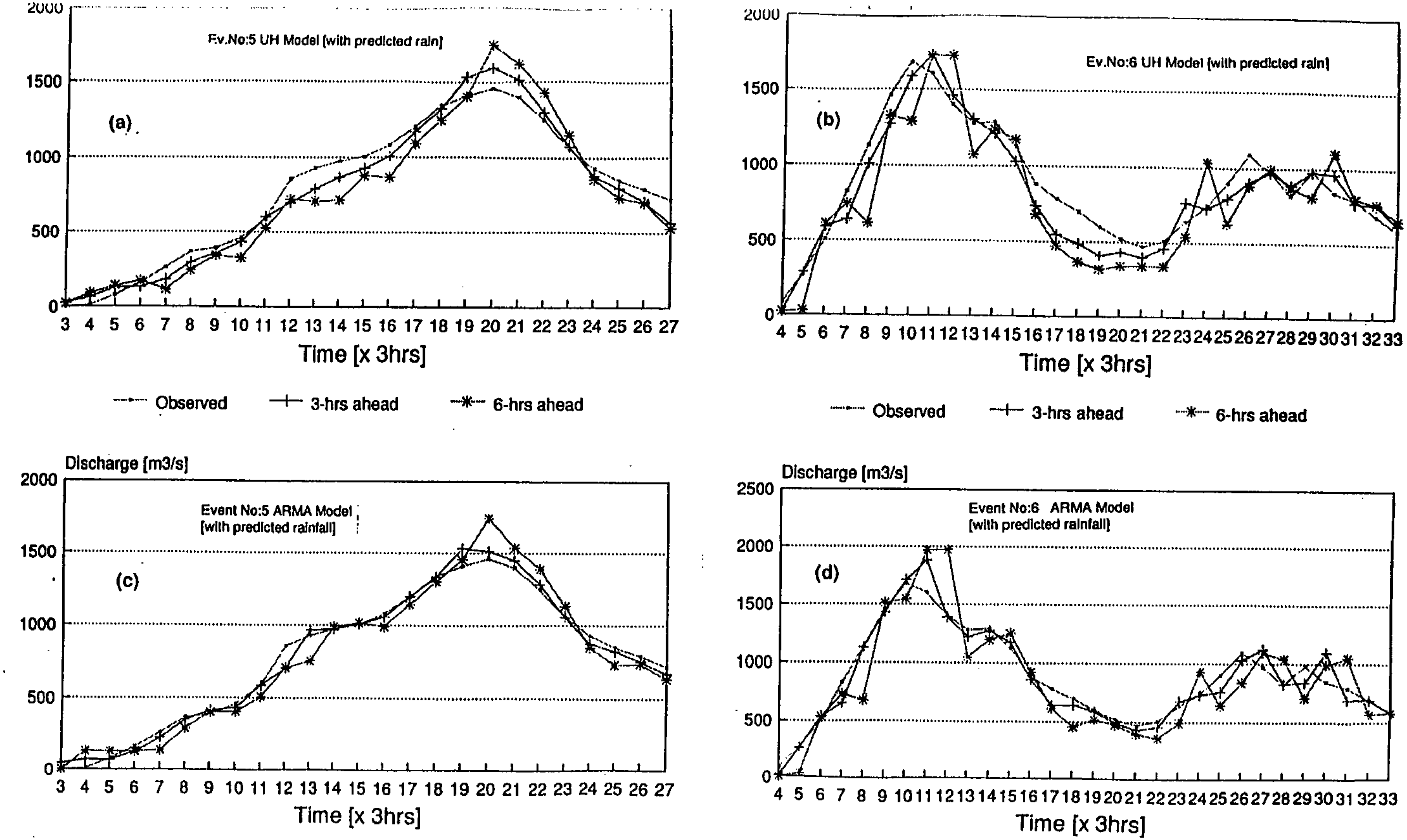


Fig. 5.1.1.8 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (with predicted rainfall)

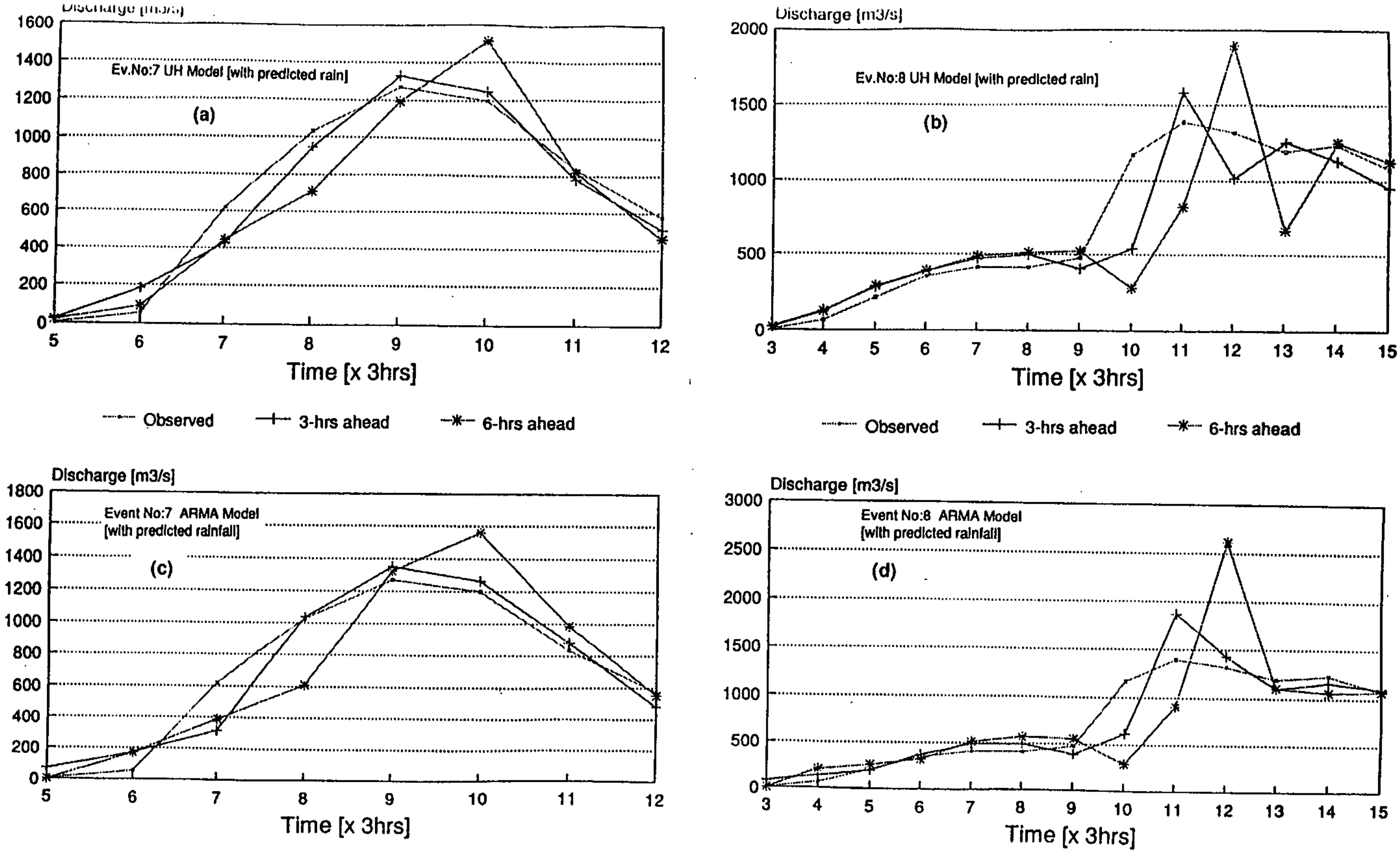


Fig. 5.1.1.9 Observed and predicted flow for Tweed river to Murwillambah using adaptive linear models (with predicted rainfall)

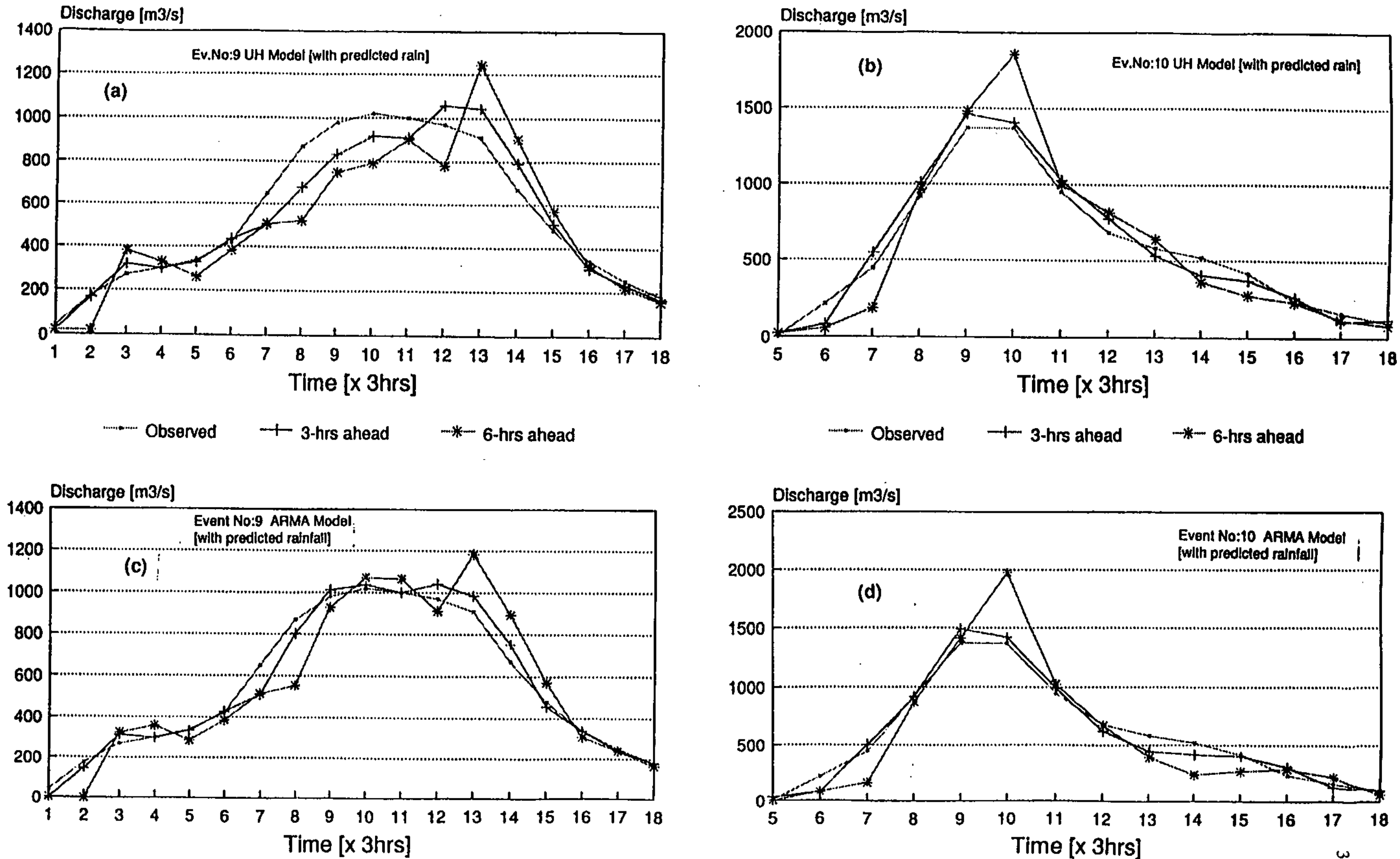


Fig. 5.1.1.10 Observed and predicted flow for Tweed [illegible]

Table 5.1.1.4: Forecast Statistics for the On-Line ARMAX Model (with Predicted Rainfall using the AR(3) Model)

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		3 hrs	6 hrs	3 hrs	6 hrs	3 hrs	6 hrs
1	484	0.25	0.44	89.1	62.1	74.0	22.2
2	315	0.19	0.39	96.1	81.5	81.5	78.7
3	507	0.09	0.22	97.8	88.1	87.5	74.5
4	526	0.13	0.28	95.0	77.8	79.4	70.5
5	832	0.06	0.12	98.6	93.2	77.7	76.0
6	910	0.11	0.24	92.0	57.7	65.3	53.9
7	860	0.17	0.28	88.0	9.9	82.9	89.1
8	810	0.29	0.60	78.1	-30.2	6.9	-35.5
9	590	0.09	0.22	97.2	81.2	83.4	75.5
10	630	0.13	0.35	96.3	70.6	90.4	81.9

where Qbar = average observed discharge
RMSE = Root Mean Squares of Error.

Table 5.1.1.5: Summary of the Results of Peak Runoff Prediction

	Q (%)						T (hrs)					
	UH model				ARMA model		UH model				ARMA model	
	non-adaptive		adaptive		adaptive		non-adaptive		adaptive		adaptive	
	I*	II*	3 hrs ahead	6 hrs ahead	3 hrs ahead	6 hrs ahead	I*	II*	3 hrs ahead	6 hrs ahead	3 hrs ahead	6 hrs ahead
Ave	107	104	97.5	84.5	107.7	107.0	1.7	1.7	1.1	3.5	0.4	3.5
SD	40	15	8.4	13.1	7.9	15.9	2.6	2.6	2.8	2.0	1.9	2.0

I*: using best estimates of IL for each event and median values of CL and basin parameters.

II*: using best estimates of IL and CL for each event and median values basin parameters [19], where

$$Q(\%) = \frac{Q_{(\text{modelled peak})}}{Q_{(\text{observed peak})}}$$

$$T(\text{hrs}) = T_{(\text{modelled peak})} - T_{(\text{observed peak})}.$$

C. Comparison of adaptive models with non-adaptive models

The statistics of the difference between peak discharges and timing of the peaks of the modelled and observed hydrographs for the seven events considered in adaptive and non-adaptive methods are given in Table 5.1.1.5. The performance of the ARMA models is slightly better than that of the UH model.

Considerable improvements brought in by the use of adaptive methods are clearly evident from these tables. Although the average error in peak flow prediction and timing was similar for all the methods, both adaptive models showed a clear advantage in modelling the range of flood events used, as indicated by the lower standard deviation, and so are considered to be more suitable for real-time forecasting.

5.1.2 *Wilson River*

The estimation of the UH ordinates indicate that the average time to peak lies around 21 to 27 hours (Table 4.4). Four out of the nine events were used for calibration of the model. Prediction of future rainfall was not considered here. Figs. 5.1.2.1 and 5.1.2.2 show the performance of these two linear models on the events used for calibration. Real-time forecasting performance (6-hours and 12-hours ahead) on the events used for testing is shown in Figs. 5.1.2.3, 5.1.2.4 and 5.1.2.5.

The overall performance of the two models was satisfactory [19]. For event number 9, rainfall was observed to be only in the eastern part of the catchment, resulting in minimal warning for this flood. This could be due to errors in the estimation of rainfall due to a very localised storm, not properly observed at the rainfall stations. The statistics of forecasting performance of the adaptive UH and ARMAX models are given in Tables 5.1.2.1 and 5.1.2.2, respectively, as well as a comparison for each method with a non-adaptive model (n.adap). The non-adaptive model uses the set of parameters developed at the start of each event and was fixed throughout the event.

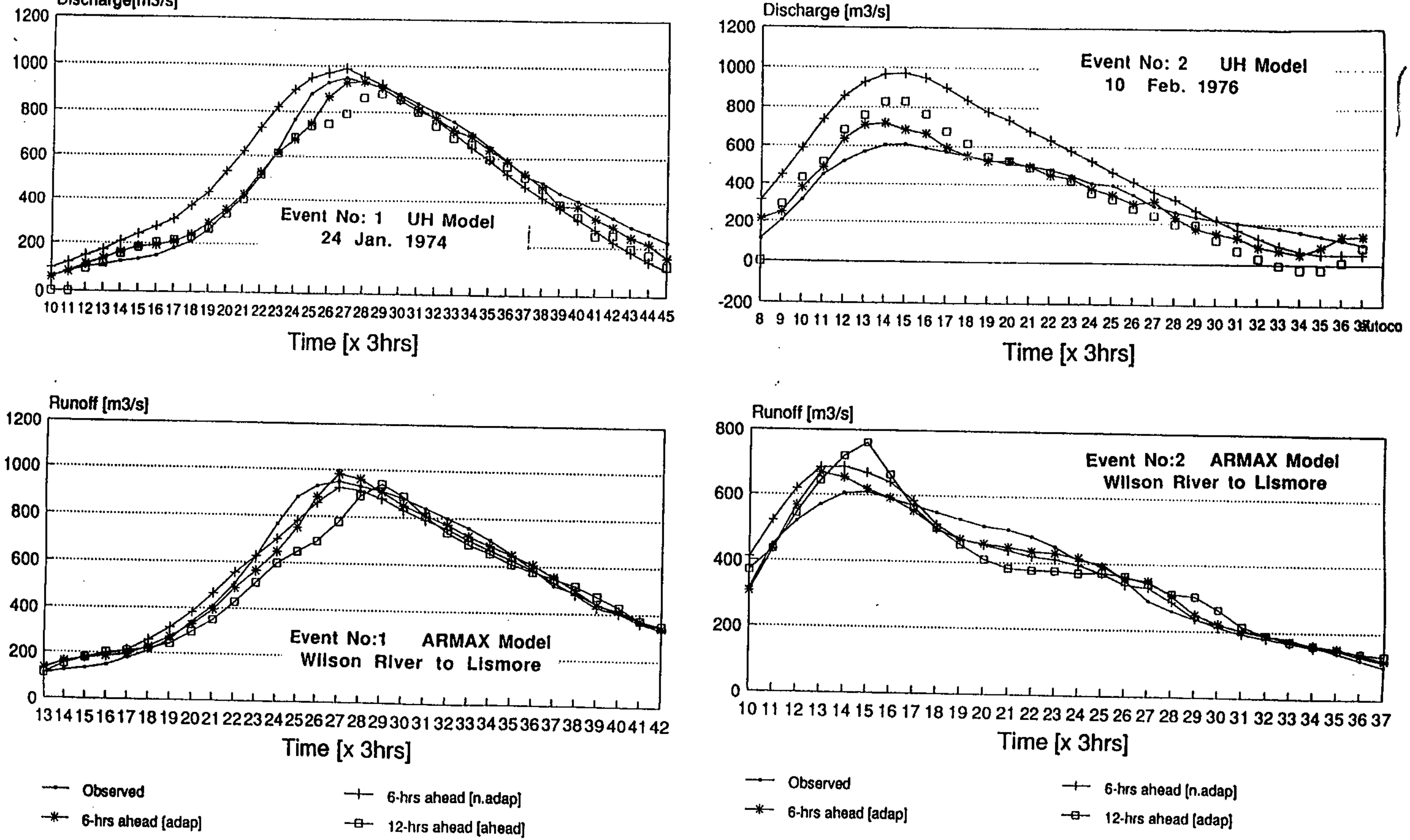


Fig. 5.1.2.1 Observed and predicted flow for Wilson river to Lismore using adaptive linear models

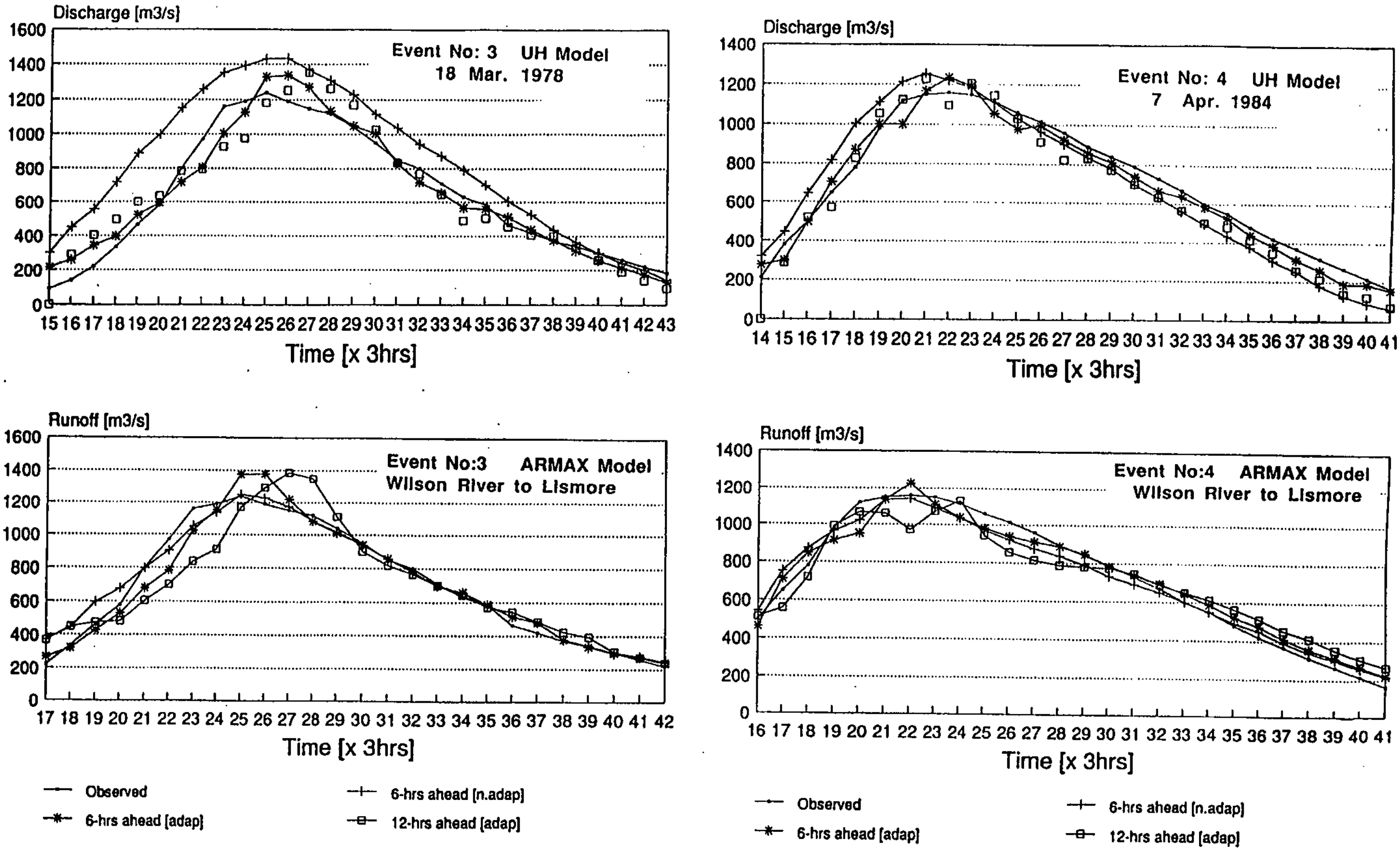


Fig. 5.1.2.2 Observed and predicted flow for Wilson river to Lismore using adaptive linear models

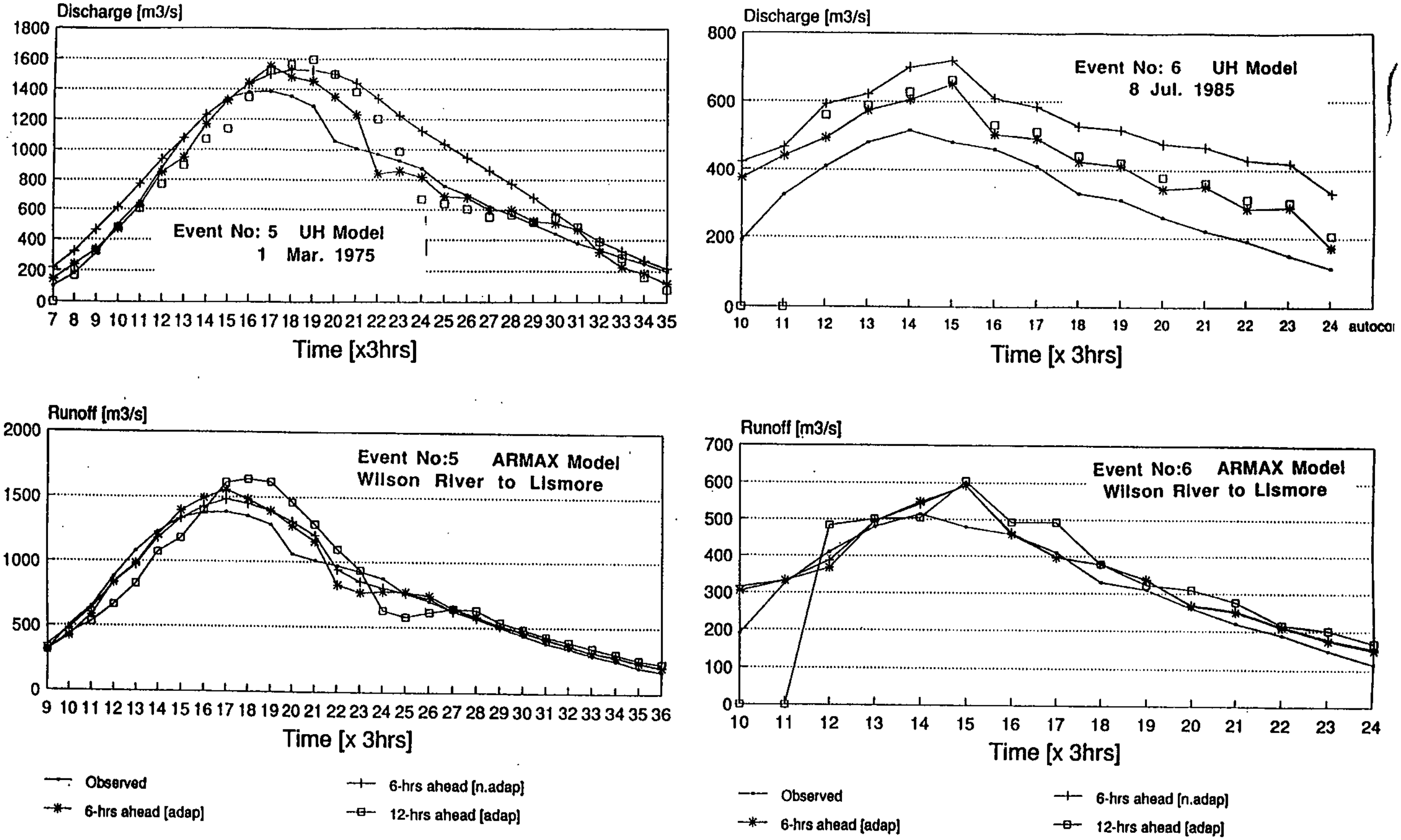


Fig. 5.1.2.3 Observed and predicted flow for Wilson river to Lismore using adaptive linear models

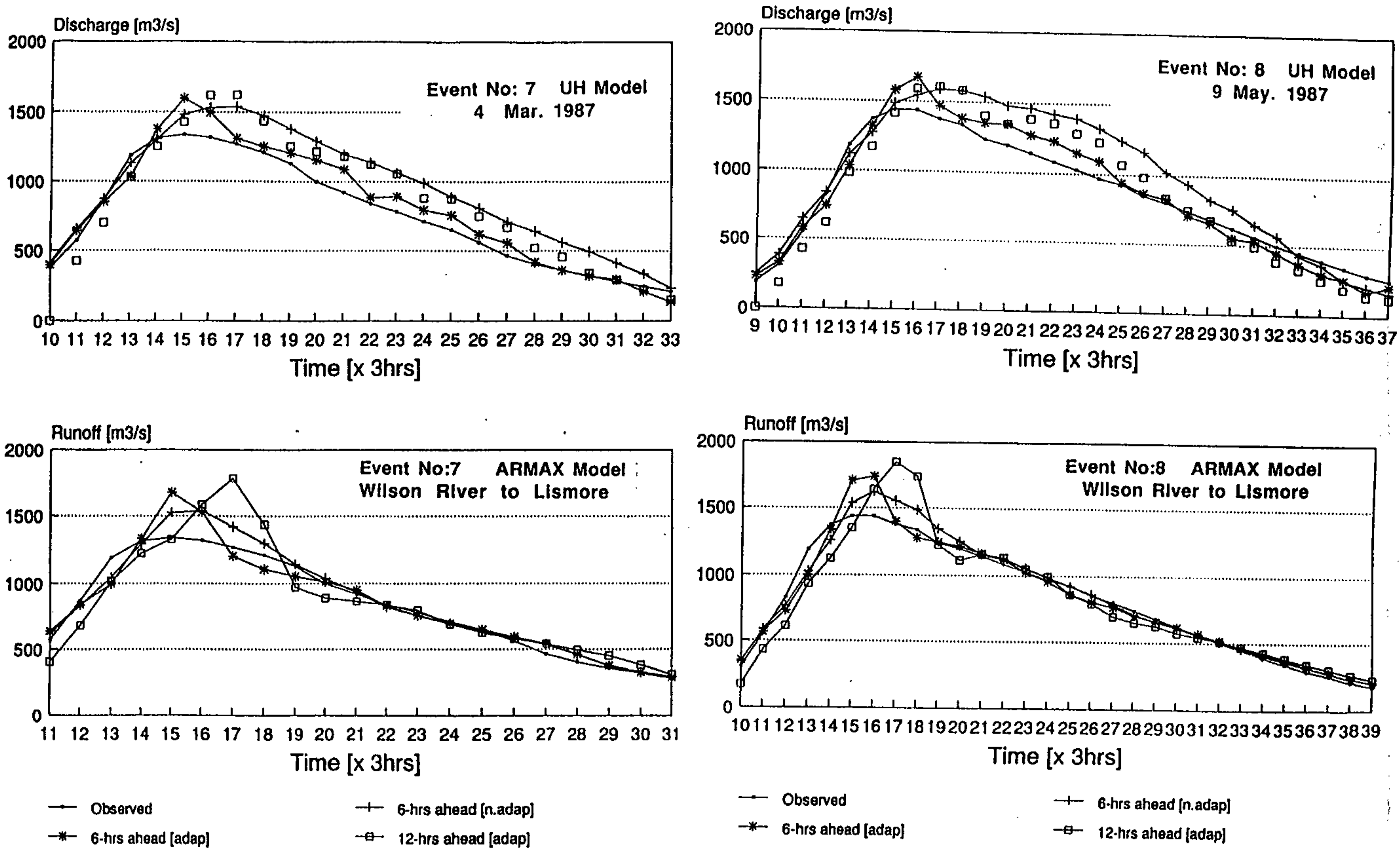


Fig. 5.1.2.4 Observed and predicted flow for Wilson river to Lismore using adaptive linear models

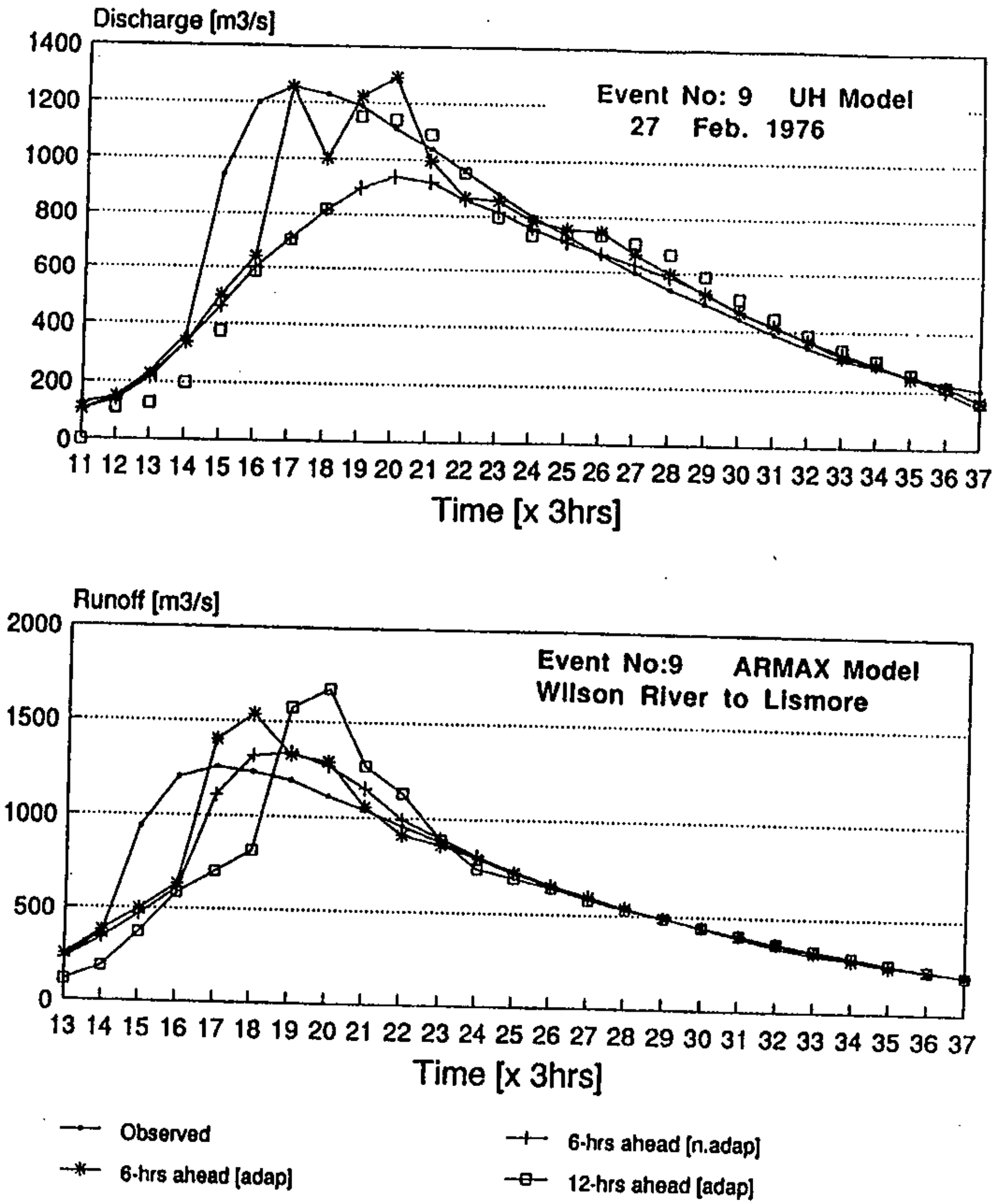


Fig. 5.1.2.5 Observed and predicted flow for Wilson river to Lismore using adaptive linear models

Table 5.1.2.1: Forecast Statics for the On-Line UH Model (for 6- and 12-hour lead times)

Flood No	Qbar (m^3/s)	RMSE Qbar			Coefficient of efficiency (%)			Coefficient of persistence (%)		
		6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap
1	439	.236	.089	.150	87.5	98.2	94.3	2.5	87.1	81.5
2	351	.599	.196	.326	-42.3	84.7	47.7	-420	44.2	-20.6
3	582	.363	.134	.187	70.4	96.0	90.8	-51.4	79.4	80.4
4	659	.162	.084	.117	90.4	97.4	93.9	59.5	89.2	87.0
5	634	.288	.148	.237	82.7	95.4	86.4	9.1	76.0	67.9
6	594	.638	.343	.373	-161	24.4	10.6	-276	-8.7	-15.7
7	323	.282	.142	.259	72.2	94.5	80.0	27.4	81.6	49.2
8	678	.256	.126	.208	83.0	95.9	87.0	27.2	82.5	71.2
9	714	.347	.249	.347	70.4	84.8	64.7	28.9	63.4	61.7

Table 5.1.2.2: Forecast Statics for the On-Line ARMAX Model (for 6- and 12-hour lead times)

Flood No	Qbar (m^3/s)	RMSE Qbar			Coefficient of efficiency (%)			Coefficient of persistence (%)		
		6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap
1	439	.083	.082	.167	98.5	98.5	92.9	87.9	88.1	77.3
2	351	.162	.114	.173	89.6	94.9	85.2	61.9	81.2	65.9
3	582	.115	.125	.210	97.0	96.5	88.4	84.8	82.0	75.3
4	659	.087	.093	.129	97.2	96.9	92.5	88.4	86.8	84.0
5	634	.114	.130	.234	97.3	96.5	86.8	85.9	81.6	68.9
6	594	.154	.150	.175	84.7	85.5	80.4	78.0	79.2	74.6
7	323	.109	.134	.207	96.8	95.1	87.2	89.2	83.5	67.5
8	678	.097	.114	.194	97.6	96.6	88.7	89.6	85.5	75.0
9	714	.254	.259	.407	84.1	83.5	51.5	61.7	60.4	47.3

Comparison of Adaptive Models with Non-Adaptive Methods

The statistics of the difference between peak discharges and timing of the peaks of the modelled and observed hydrographs for the eight events considered in adaptive and non-adaptive methods are given in Table 5.1.2.3. The performance of ARMA models is better than that of the UH model. The improvement brought in by the use of adaptive methods is evident from these tables. The adaptive models, by virtue of the lower standard deviation in both cases, are considered to be more suitable for real-time forecasting.

Table 5.1.2.3: Summary of the Results of Peak Runoff Prediction

	Q (%)						T (hrs)					
	UH model				ARMA model		UH model				ARMA model	
	non-adaptive		adaptive		adaptive		non-adaptive		adaptive		adaptive	
	I*	II*	6 hrs ahead	12 hrs ahead	6 hrs ahead	12 hrs ahead	I*	II*	6 hrs ahead	12 hrs ahead	6 hrs ahead	12 hrs ahead
Ave	107	103	114	115	107	108	-3.0	-3.0	0.8	4.5	0.4	3.5
SD	18	19	8.0	14	7.5	14.5	4.5	4.5	2.0	2.1	1.9	2.0

I*: using best estimates of IL for each event and median values of CL and basin parameters.

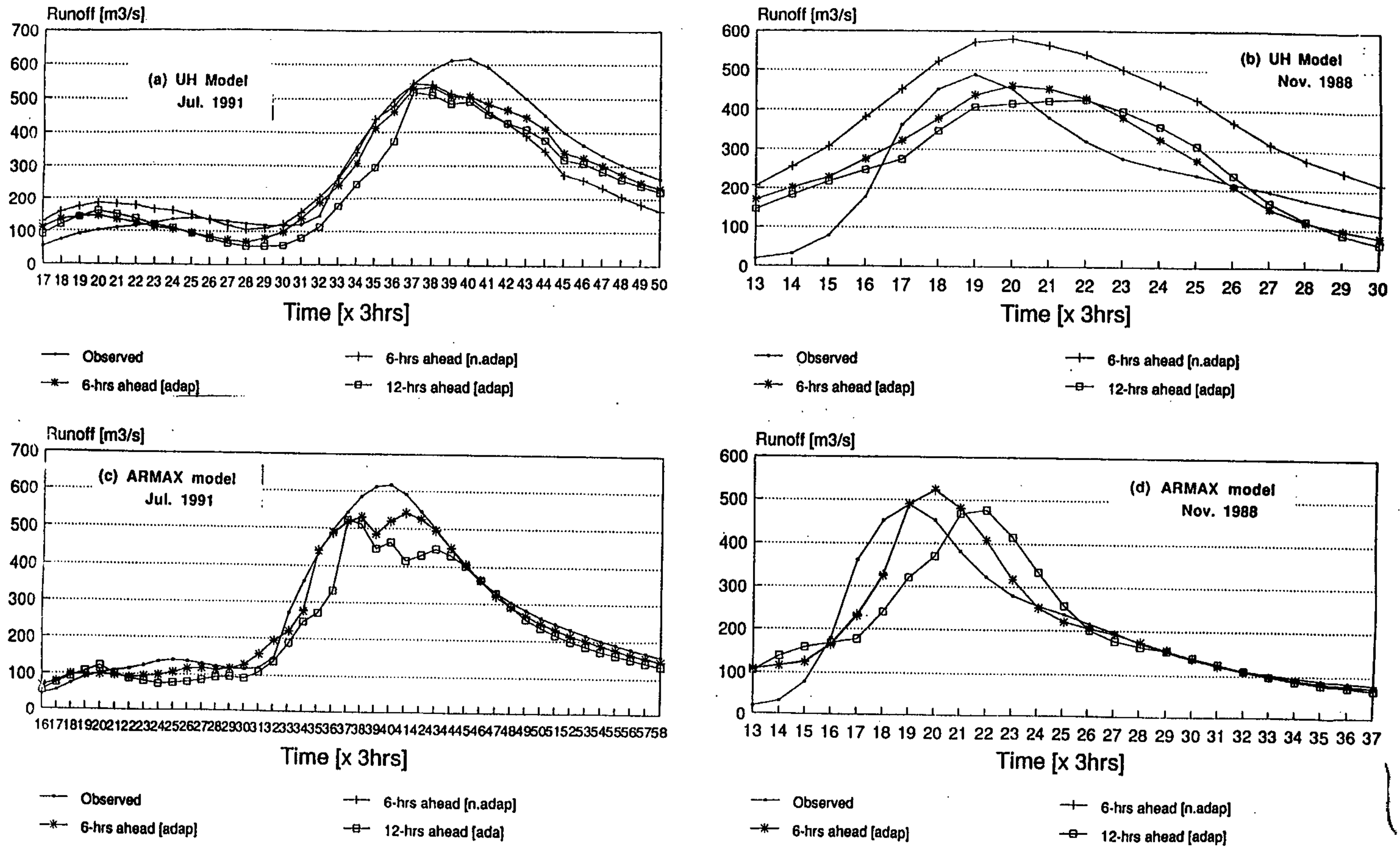
II*: using best estimates of IL and CL for each event and median values basin parameters [19], where

$$Q(\%) = \frac{Q_{(\text{modelled peak})}}{Q_{(\text{observed peak})}}$$

$$T(\text{hrs}) = T_{(\text{modelled peak})} - T_{(\text{observed peak})}.$$

5.1.3 *Mitchell River to Glenaladale*

The estimation of the UH ordinates indicates that the average time to peak lies around 15-21 hours (Table 4.5). Two events were used for calibration of the model and the other two for testing. The uncertainty of the rainfall input is rather high. Fig. 5.1.3.1 shows the performance of the model on the data used for calibration. Fig. 5.1.3.2 shows the real-time forecasting performance 6 hours and 12 hours ahead. The statistics of forecasting performance of the UH and ARMAX models are given in Tables 5.1.3.1 and 5.1.3.2, respectively.



Runoff [m3/s]
(a) UH Model
Jul. 1991
Time [x 3hrs]
Observed
6-hrs ahead [adap]
6-hrs ahead [n.adap]
12-hrs ahead [adap]
Runoff [m3/s]
(b) UH Model
Nov. 1988
Time [x 3hrs]
Observed
6-hrs ahead [adap]
6-hrs ahead [n.adap]
12-hrs ahead [adap]
Runoff [m3/s]
(c) ARMAX model
Jul. 1991
Time [x 3hrs]
Observed
6-hrs ahead [adap]
6-hrs ahead [n.adap]
12-hrs ahead [ada]
Runoff [m3/s]
(d) ARMAX model
Nov. 1988
Time [x 3hrs]
Observed
6-hrs ahead [adap]
6-hrs ahead [n.adap]
12-hrs ahead [adap]

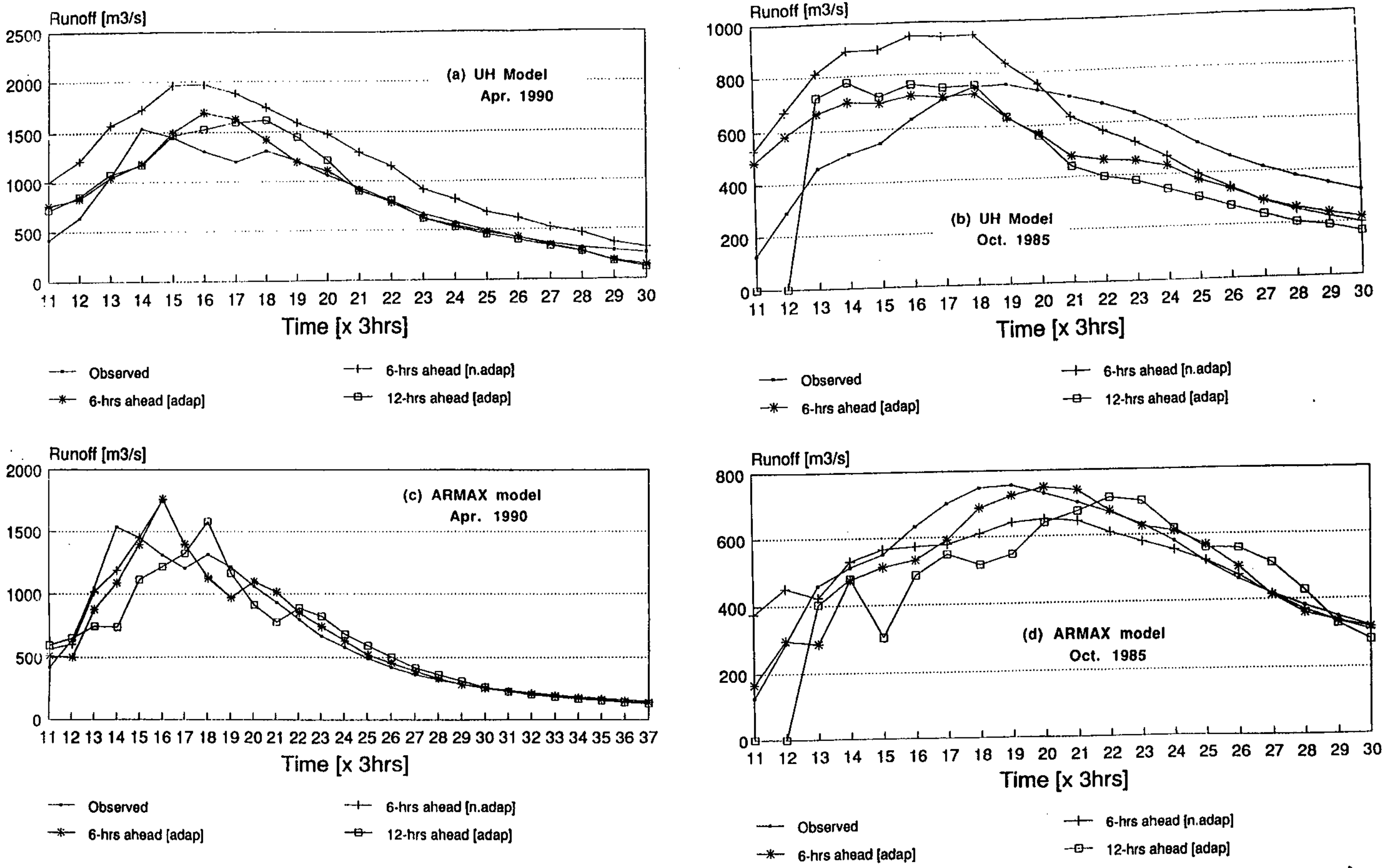


Fig. 5.1.3.2 Observed and predicted flow for Mitchell river to Glenaladale.

Table 5.1.3.1: Forecast Statics for the Mitchell River (UH Model)

Flood No	Qbar (m^3/s)	RMSE Qbar			Coefficient of efficiency (%)			Coefficient of persistence (%)		
		6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap
1	255	0.27	0.19	0.26	84.9	92.7	83.3	6.7	54.9	56.3
2	190	0.74	0.44	0.43	-1.5	64.1	57.8	-93.4	31.6	65.7
3	615	0.56	0.31	0.25	26.5	77.6	81.6	-50.3	54.2	76.9
4	527	0.43	0.33	0.35	-77.0	-2.5	-97.6	-217	-83.0	-105

Table 5.1.3.1: Forecast Statics for the Mitchell River (ARMAX Model)

Flood No	Qbar (m^3/s)	RMSE Qbar			Coefficient of efficiency (%)			Coefficient of persistence (%)		
		6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap
1	255	0.14	0.14	0.26	96.0	96.1	84.1	74.1	74.3	55.6
2	190	0.29	0.29	0.44	84.5	84.0	56.7	70.5	69.6	64.8
3	615	0.26	0.27	0.31	88.1	86.3	80.3	71.0	66.5	68.2
4	527	0.17	0.11	0.21	72.9	88.4	28.5	51.4	79.2	25.5

5.2 River Routing: Dagun Pocket to Gympie

An ARMAX model of three parameters was identified as the best fit for this reach. Figs 5.2.1(a) and (b) show the performance of adaptive and non-adaptive methods for the events used for calibration of the model. The 3-hours ahead real-time forecasting performance for the test event is shown in Fig. 5.2.1(c). Generally, the adaptive method performs better than the non-adaptive (fixed parameter) method for all these events. The statistics of forecasting performance for the adaptive and non-adaptive methods are given in Table 5.2.1.

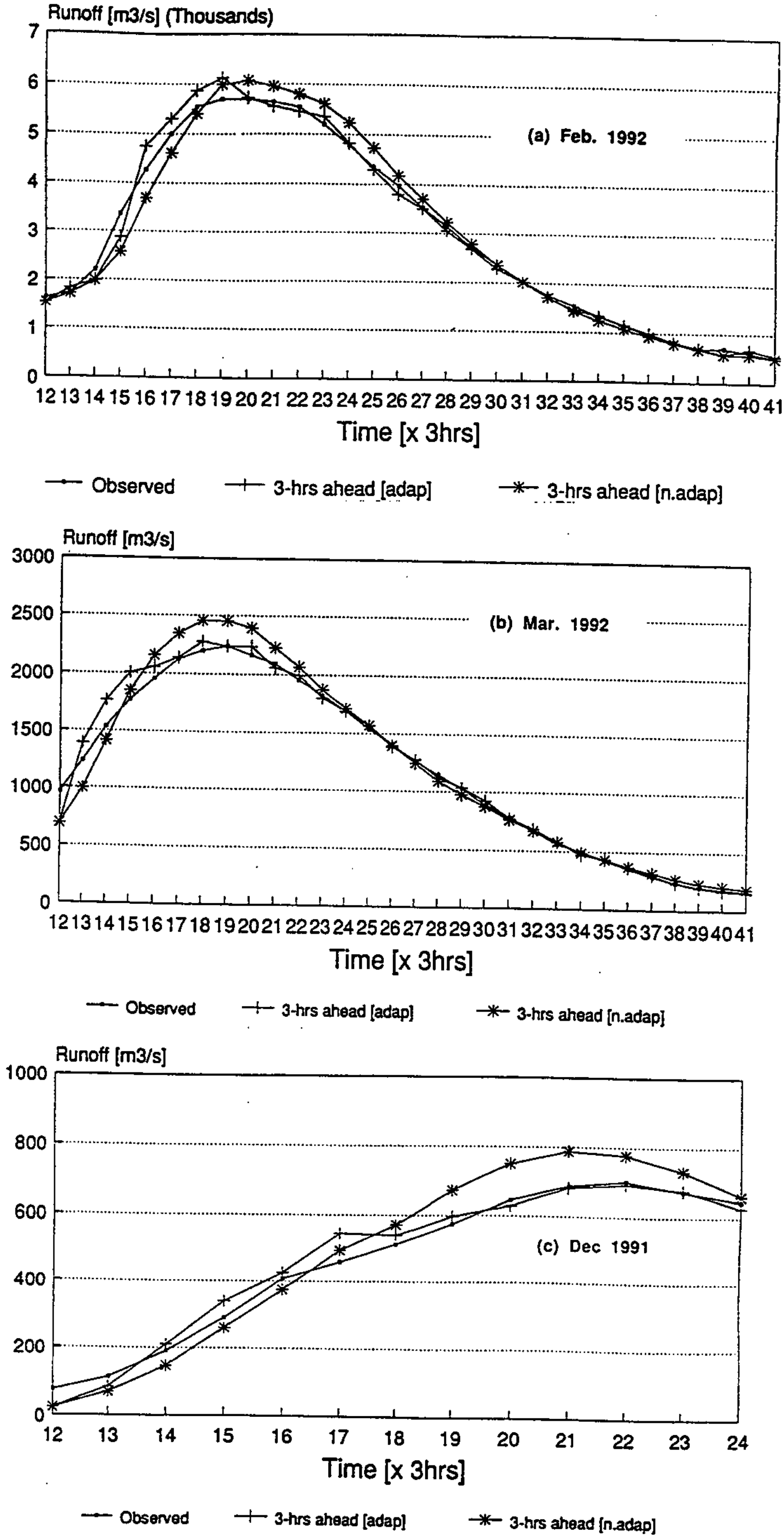


Fig. 5.2.1 Observed and predicted flow for Mary river to Gympie
(River routing model)

Table 5.2.1: Forecast Statistics for River Routing: Dagun Pocket to Gympie

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		3 hrs	6 hrs	3 hrs	6 hrs	3 hrs	6 hrs
1	3041	0.09	0.06	97.7	98.9	55.0	79.4
2	1186	0.09	0.06	97.8	99.1	13.8	63.2
3	525	0.13	0.06	83.2	95.7	-8.0	72.6

where Qbar = average observed discharge
RMSE = Root Mean Squares of Error.

5.3 Multiple Input - Single Output Models (MISO)

As shown in Fig. 4.4(a), the Mary River catchment is divided into three stages, resulting in three models: the first one up to Gympie, the second one up to Miva, and the third one up to Home Park.

5.3.1 *Rainfall-runoff model to Gympie*

From the unit hydrograph, it can be observed that the time to peak lies around 15-21 hours. Figs 5.3.1.1(a) and (b) show the performance of the ARMAX model for the two events used in calibration. For these events, the non-adaptive method seems to perform better than the adaptive method, but the real test of a model is how well it can predict events to which it was not calibrated. The event used for testing is shown in Fig. 5.3.1.1(c), where the adaptive method performs better than the non-adaptive one, although it is difficult to be conclusive without testing the model on other events. The statistics of forecasting performance of the ARMAX model are given in Table 5.3.1.

Table 5.3.1: Forecast Statics for the Mary River to Gympie (ARMAX Model)

Flood No	Qbar (m3/s)	RMSE Qbar			Coefficient of efficiency (%)			Coefficient of persistence (%)		
		6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap	6 hrs n.adap	6 hrs adap	12 hrs adap
1	2671	0.09	0.12	0.27	98.4	96.7	83.0	90.9	81.7	58.1
2	1050	0.08	0.08	0.16	98.4	89.3	92.7	90.5	89.7	79.9
3	459	0.45	0.28	0.62	7.6	65.8	-20.4	-163	2.6	-265

where Qbar = average observed discharge
RMSE = Root Mean Squares of Error.

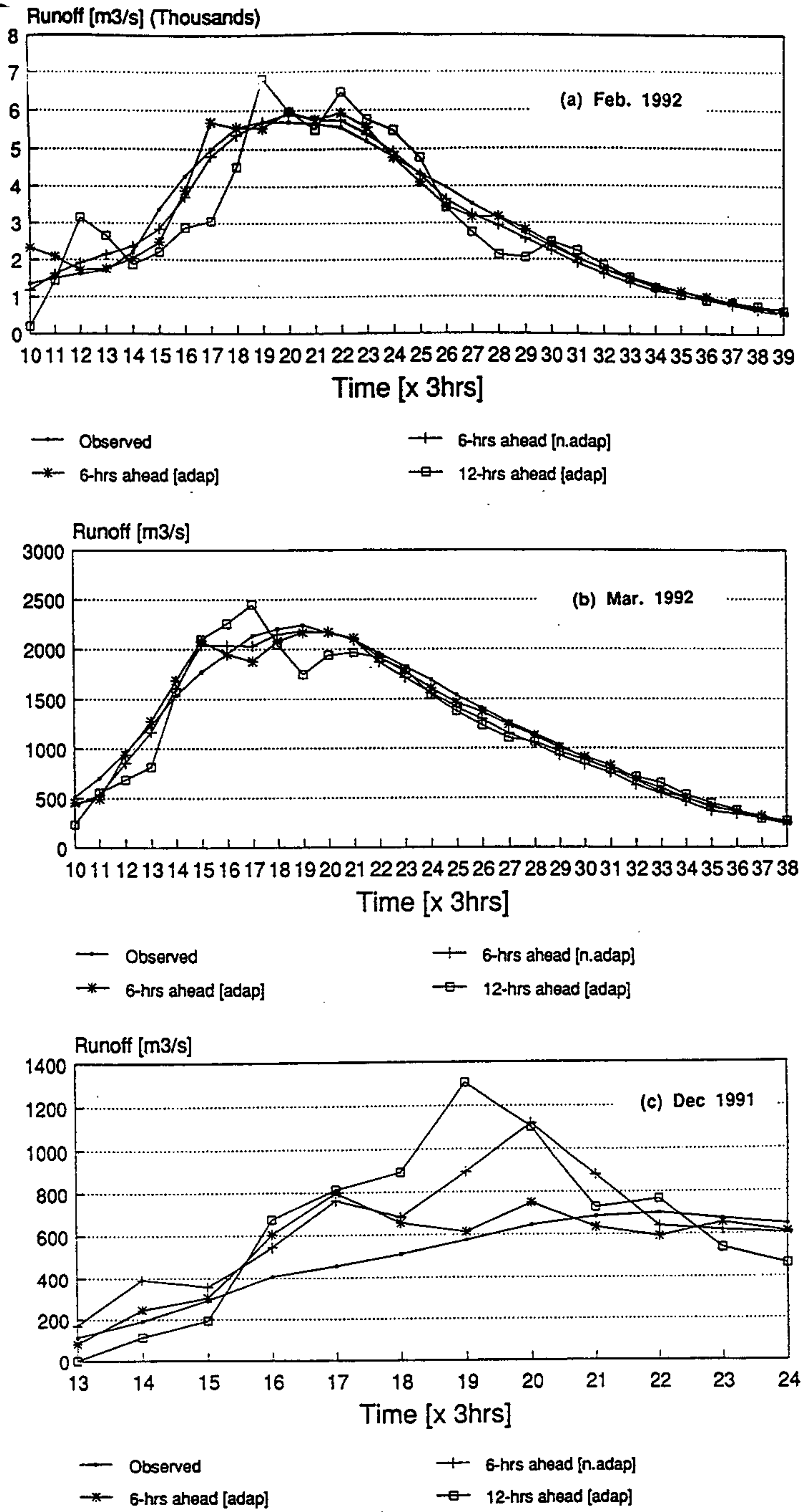


Fig. 5.3.1.1 Observed and predicted flow for Mary river to Gympie, [Rainfall/Runoff model]

5.3.2 *MISO model for Mary River to Miva*

In this case, all four events were used in the calibration. Fig. 5.3.2.1 shows the 6-hours ahead forecast performance for all four of these events. For these events, the non-adaptive method seemed to perform equally as well as the adaptive method, but the real test of a model is how well it can predict events to which it was not calibrated. Therefore, it is difficult to be conclusive without testing the model on other events. The statistics of forecasting performance of the ARMAX model are given in Table 5.3.2.

Table 5.3.2: Forecast Statistics for River Routing: Mary River to Miva

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		6 hrs n.adap	6 hrs adap	6 hrs n.adap	6 hrs adap	6 hrs n.adap	6 hrs adap
1	2948	0.07	0.06	94.2	94.1	84.6	85.9
2	2302	0.09	0.09	91.1	90.5	69.6	71.1
3	1726	0.14	0.13	77.1	72.1	45.5	40.2
4	3125	0.05	0.08	99.1	97.7	94.5	86.6

where Qbar = average observed discharge
RMSE = Root Mean Squares of Error.

5.3.3 *River routing using the MISO model for the Mary River to Home Park*

All four events that were available for the study were used in the calibration of the MISO model. The performance of 9-hours ahead forecasts of the MISO model is shown in Fig. 5.3.3.1. For these events, the adaptive method did not give significant improvements on the forecasts made by the non-adaptive method, but it is difficult to be conclusive without further testing. The statistics of forecasting performance of the ARMAX model are given in Table 5.3.3.

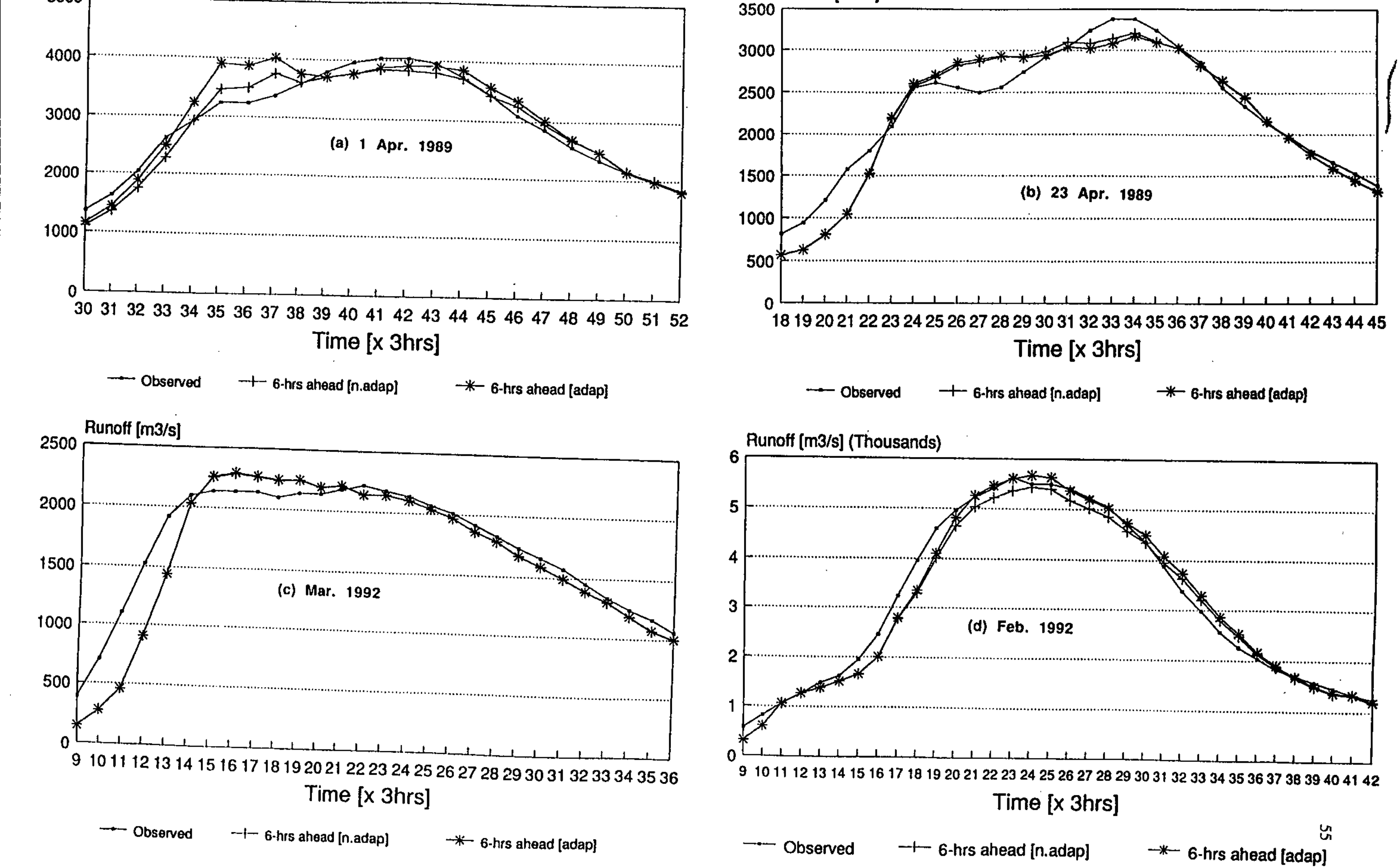
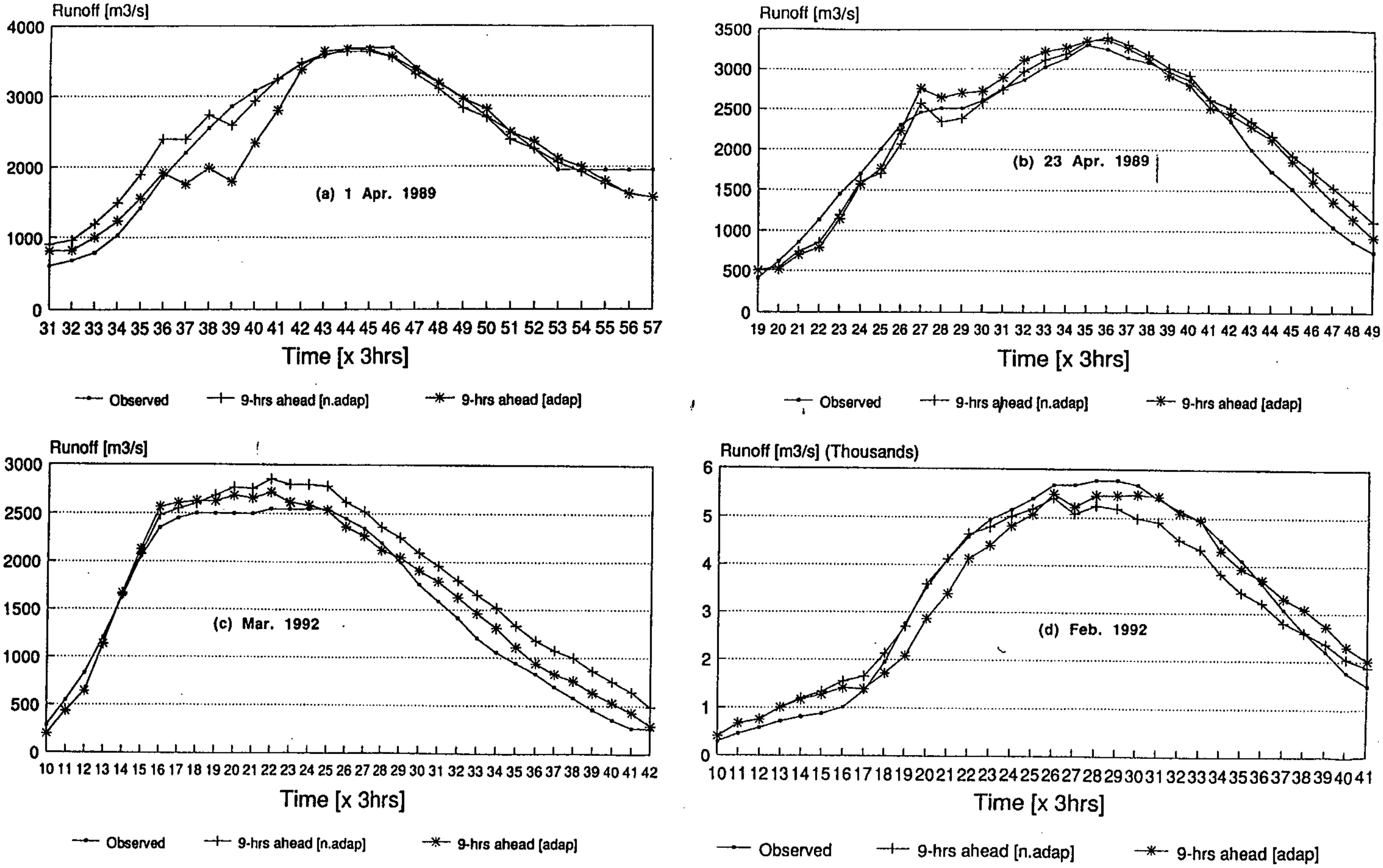


Fig. 5.3.2.1 Observed and predicted flow for Mary river to Miva
(multiple input, single output model)



Flood of 1992.02.20

Fig. 5.3.3.1 Observed and predicted flow for Mary river to Home Park

Table 5.3.3: Forecast Statistics for River Routing (MISO): Mary River to Home Park

Flood no.	Qbar (m^3/s)	RMSE Qbar		Coefficient of efficiency (%)		Coefficient of persistence (%)	
		9 hrs n.adap	9 hrs adap	6 hrs n.adap	6 hrs adap	6 hrs n.adap	6 hrs adap
1	2345	0.10	0.14	93.2	87.7	86.8	76.0
2	2038	0.11	0.10	92.5	94.3	83.5	87.4
3	1584	0.18	0.09	88.9	97.0	72.8	92.7
4	3218	0.12	0.11	95.7	96.2	87.6	88.9

where Qbar = average observed discharge
RMSE = Root Mean Squares of Error.

6. CONCLUSIONS

6.1 Rainfall-Runoff Models

Two adaptive linear models (adaptive unit hydrograph and ARMAX) were applied to the following catchments (observations at 3-hour intervals):

1. Tweed River (630 km^2), which experiences a wide range of runoff coefficient values for different storms.
2. Wilson River (1400 km^2), which has a narrow range of values of runoff coefficient for different storms.
3. Mitchell River (2905 km^2), where the quality of rainfall input is poor, having three rainfall stations located in the lower part of the catchment.

   Given these constraints, the problems faced in real-time forecasting of floods identified earlier were addressed through:

   (a) adopting a simple linear model where the fraction of average rainfall was assumed to represent the continuing loss and was incorporated as a model parameter; and

   (b) using an adaptive method (Kalman filter) to improve the model forecasts using the discrepancy between observed and model output at each time step.

For the Tweed River, both the linear models performed satisfactorily, although the ARMAX model required a lesser number of parameters. Because of the fast response time of the catchment, the use of 3-hours ahead predicted rainfall was needed to improve the performance of the model for the rising limb of the flood hydrograph. The runoff coefficient was found to vary over a wide range for these events, resulting in significant variation in parameters from one event to another. Thus, the adaptive methods performed significantly better than the non-adaptive methods for the catchment.

In the case of the Wilson River, the overall performance of the ARMAX model was slightly better than that of the UH model. The ARMAX model showed no significant improvement with the use of the adaptive method compared to the non-adaptive one. As the runoff coefficient was found to be stable for the events considered, the parameters, too, did not change much from one event to another.

For the Mitchell River to Glenaladale, the ARMAX model performed better than the UH model. This was because the quality of rainfall data affects the performance of the UH model much more than that of the ARMAX model, and the ARMAX model is better suited to situations where uncertainty of rainfall input is high.

6.2 River Routing Method

The same linear models were applied to the river routing problem between the two gauging stations, Dagun Pocket and Gympie on the Mary River (Queensland). The models were considered to perform satisfactorily for all three events. The 3-hours ahead forecasts of the adaptive method was marginally better than that of the non-adaptive method, as the parameters were updated for different events; the adaptive method took into account unmeasured inflows by updating the corresponding parameters. Firm conclusions are not considered possible as only three events were used; however, it would be worthwhile examining the performance of the model for longer lead time forecasts, by either fitting a model with longer lead times or by applying the same model with predicted runoff (3 hours ahead) as input.

6.3 Multiple Input - Single Output (MISO) Models

Mary River to Home Park was considered in three stages: a rainfall-runoff model up to Gympie, followed by MISO models up to Miva and Home Park, respectively. In all four events used for calibration, the adaptive method showed no significant improvement over non-adaptive methods, although without testing the method on events other than those used for calibration, no firm conclusions can be made. By considering the total catchment as the linking together of these three stages, forecasts can be made with a lead time that is equivalent to the summation of lead times at each of the stages, thereby achieving a longer overall lead at the downstream point, although this involves neglecting inflows in the intermediate stages.

6.4 Recommendations

Adaptive methods perform better in situations where uncertainty is high. They are best suited for catchments whose model parameters vary significantly from one event to another.

7. REFERENCES

[1] Nash, J.E. and Sutcliffe, J.V. (1970): River flow forecasting through conceptual models. I: A discussion of principle. *J. Hydrol.*, Amsterdam, **10(3):** 282-290.

[2] Szollosi-Nagy, A. (1976): An adaptive identification and prediction algorithm for the real-time forecasting of hydrologic time series. *Hydrol. Sci. Bull.*, **21:** 163-176.

[3] Todini, E. and Wallis, J.R. (1978): A real-time rainfall-runoff model for an on-line flood warning system. *Proc. AGU Conf. on Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources*, Univ. of Pittsburgh: 355-368.

[4] Ganendra, T. (1979): Real-Time Forecasting and Control in the Operation of Water Resources Systems. PhD thesis, University of London.

[5] Kitanidis, P.K. and Bras, R.L. (1980): Real-Time Forecasting of River Flows. *Report No. 235*, R.M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Cambridge, Mass.

[6] Kitanidis, P.K. and Bras, R.L. (1980): Real-time forecasting with a conceptual hydrologic model. 1: Analysis of uncertainty. 2: Applications and results. *Water Resources Res.*, **16(6):** 1025-1044.

[7] O'Connell, P.E. and Clarke, R.T. (1981): Adaptive hydrological forecasting - a review. *Hydrological Sciences Bull.*, **26:** 179-205.

[8] Bolzern, P., Ferrario, M. and Fronza, G. (1980): Adaptive real-time forecast of river flow-rates from rainfall data. *J. Hydrol.*, **47:** 251-267.

[9] Amirthanathan, G.E. (1989): Optimal filtering techniques in flood forecasting. Proc. Inter. Conf. on FRIENDS in Hydrology, *IAHS Publ. No. 187*: 13-25.

[10] Ambrus, S. (1979): Adaptive forecasting in the Sagyva-Tarna automatic water resources management system. *VITUKI Report No. 7631-1530*, No. 2.

[11] WMO (1975): Hydrological Forecasting Practices. *Operational Hydrology Report No. 6.*

[12] Hino, M. (1973): On-line prediction of hydrologic systems. *Proc. XVth Conf. of IAHR*, Istanbul: 121-129.

[13] Spolia, S.K. and Chander, S. (1974): Modelling of surface runoff by an ARMA model. *J. Hydrol.*, **22(3/4):** 317-332.

[14] Ambrus, S.Z. and Nagy, A.S. (1981): Operational real-time flow forecasting using stochastic prediction algorithm. *Proc. Inter. Symp. on Real-Time Operation of Hydrosystems*, Waterloo, Canada.

[15] Wood, E.F. and Nagy, A.S. (1978): An adaptive algorithm for analyzing short-term structural and parameter changes in hydrologic prediction models. *Water Resources Res.*, **14(4):** 577-581.

[16] Mehra, R.K. (1969): On the identification of variances and adaptive Kalman filtering. *Proc. 1969 Joint Automatic Control Conf.*: 495-505.

[17] Jazwinski, A.H. (1970): *Stochastic Processes and Filtering Theory.* Academic Press, New York.

[18] Fitzgerald, R.J. (1967): Error divergence in optimal filtering problems. *Second IFAC Symp. Automatic Control Space*, Vienna.

[19] Malone, T.A. and Cordery, I. (1989): An assessment of network models in flood forecasting. Proc. Baltimore Symp., New Directions for Surface Water Modeling, *IASH Publ. No. 181*: 115-124.

[20] Avery, D.R. (1989): Application of RORB runoff routing model to a flood warning system in the Tweed Valley, New South Wales. *Hydrology and Water Resources Symp.*, Christchurch, New Zealand: 90-94.

[21] Ibbitt, R.P. (1991): Model calibration and software applications. *Inter. Hydrology and Water Resources Symp.*, Perth, Australia.

[22] Wang Guang-Te et al. (1987): Improved flood routing by ARMA modelling and the Kalman filter technique. *J. Hydrol.*, **93:** 175-190.

[23] Young, P.C. (1974): A recursive approach to time series analysis. *Bull. Inst. Math. and Its Applications (IMA)*, **10(5/6):** 209-224.

[24] Bergman, M.J. and Delleur, J. (1985): Kalman filtering estimation and prediction of daily stream flows. Review, algorithm and simulation experiment. *Water Resources Bull.*, **21(5):** 815-825.

[25] Wood, E.F. and O'Connell, P.E. (1985): Real-Time Forecasting. In: M.G. Anderson and T.P. Burt (Eds), *Hydrological Forecasting*, John Wiley & Sons.

[26] Bras, R.L. and Rodriguez-Iturbe, I. (1985): *Random Functions and Hydrology.* Addison-Wesley, Reading, Mass.

[27] Gelb, A. (1974): *Applied Optimal Estimation.* MIT Press, Cambridge, Mass.

[28] Todini, E. (1978): Mutually interactive state-parsameter (MISP) estimation. In: *Applications of Kalman Filtering to Hydrology, Hydraulics and Water Resources*, Dept Civil Engineering, Univ. of Pittsburgh: 135-152.

APPENDIX 1

REVIEW OF TRANSFER FUNCTION MODELS

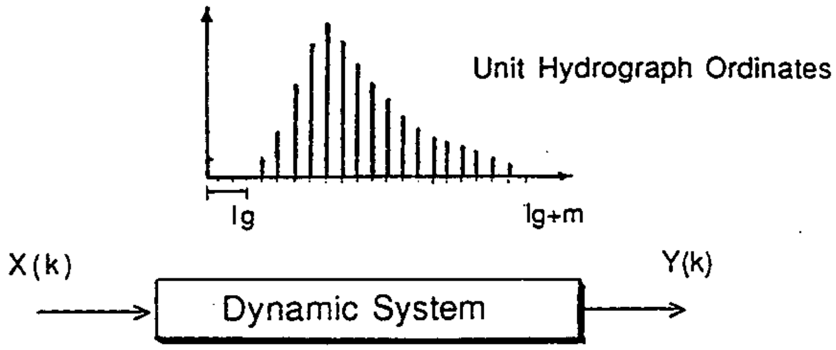
These models can be used to represent the dynamic behaviour of a hydrologic system and are parsimonious in their use of parameters. If Y(k) and X(k) are discrete time measurements of storm runoff and rainfall excess, respectively, a linear relationship between Y(k) and X(k) in current and earlier time periods, X(k-lg), X(k-lg-1), X(k-lg-2) ...., may be written as follows:

$$Y(k) = h_1 . X(k\text{-}lg) + h_2 . X(k\text{-}lg\text{-}1)... + h_m . X(k\text{-}lg\text{-}m)$$

or

$$Y(k) = H[B] . X(k\text{-}lg) \qquad \text{(A1.1)}$$

where $H[B] = [h_1 + h_2.B + h_3.B^2 .... h_{m+1}.B^m]$ is the transfer function of a linear filter, B the backward shift operator, and lg the number of zero ordinates prior to the rising limb. The parameters $h_1, h_2, h_3, ..., h_{m+1}$ are considered as the ordinates of a finite memory impulse response for a catchment. If the series H[B] converges for $|B| < 1$, then the system is stable.



A linear relationship could be found between runoff Y(k) in earlier time periods, Y(k-1), Y(k-2), ... Y(k-r), and rainfall, X(k-lg), X(k-lg-1), ... X(k-lg-s+1), given by:

$$Y(k) = \delta_1 Y(k\text{-}1) + \delta_2 Y(k\text{-}2) .... + \delta_r Y(k\text{-}r) + \omega_1 X(k\text{-}lg) + \omega_2 X(k\text{-}lg\text{-}1) .... + \omega_s X(k\text{-}lg\text{-}s\text{+}1)$$

or

$$\delta[B]\, Y(k) = \Omega[B] . X(k\text{-}lg) \qquad \text{(A1.2)}$$

where
$$\delta[B] = [1 - \delta_1 B - \delta_2 B^2 ... - \delta_r B^r]$$
$$\Omega[B] = [\omega_1 + \omega_2 B ... + \omega_s B^{s-1}]$$

In general, the roots of the characteristic equation $\delta[B]$, with B regarded as variable, should lie outside the unit circle for stability. The transfer function model is $H[B] = \delta^{-1}[B] . \Omega[B]$. If we employ the transfer function given by Eqn(A1.2), then substituting $Y(k) = H[B] . X(k\text{-}lg)$ we obtain:

$$[1 - \delta_1 B - \delta_2 B^2 \ldots - \delta_r B^r] . [h_1 + h_2.B + h_3.B^2 \ldots. h_{m+1}.B^m] . X(k\text{-}lg)$$
$$= \quad [\omega_1 + \omega_2 B \ldots + \omega_s B^{s-1}] . X(k\text{-}lg) \qquad \text{(A1.3)}$$

On equating the coefficients of B, we obtain (for the case of s>r):

$$h_1 = \omega_1$$
$$h_2 = \omega_2 + \delta_1 h_1$$
$$h_3 = \omega_3 + \delta_1 h_2 + \delta_2 h_1$$
$$\bullet$$
$$\bullet$$
$$h_r = \omega_r + \delta_1 h_{r-1} + \delta_2 h_{r-2} \ldots. + \delta_r h_0$$
$$h_{r+1} = \omega_{r+1} + \delta_1 h_r + \delta_2 h_{r-1} \ldots. + \delta_r h_1$$
$$\bullet$$
$$\bullet$$
$$h_s = \omega_s + \delta_1 h_{s-1} + \delta_2 h_{s-2} \ldots. + \delta_r h_{s-r}$$
$$h_{s+1} = \delta_1 h_s + \delta_2 h_{s-1} \ldots. + \delta_r h_{s-r+1}$$
$$\bullet$$
$$\bullet$$
$$h_m = \delta_1 h_{m-1} + \delta_2 h_{m-2} \ldots. + \delta_r h_{m-r}$$

In summary:

$$h_1 = \omega_1 \qquad \text{(A1.4a)}$$

$$h_1 = \omega_1 + \sum_{j=1}^{ir} \delta_j . h_{i-j} \quad \text{for } 2 < i < s$$

where
$$ir = \begin{cases} i & \text{if } i<r \\ r & \text{if } i>r \end{cases}$$

$$h_i = \sum_{j=1}^{r} \delta_j . h_{i-j} \quad \text{for } s+1 < i < m \qquad \text{(A1.4b)}$$

In practice, values of r and s rarely need to exceed 2 or 3.

Estimation of the ARMAX Model Order and Parameters Given the UH Parameters

The order of the model r,s is initially arbitrarily chosen. The autoregressive parameters 'δ' are determined using Eqn(4b) where there are (m-s) equations to determine r number of parameters. The parameters can be estimated using the least squares method. Once the 'δ' parameters are estimated, the 'ω' parameters can be determined by direct substitution in Eqn(4b). The order of the model r,s is determined by the trial and error approach, where the order is increased till the improvement in mean squares of error (MSE) is insignificant.

Example:

The operational UH ordinates of a catchment are given below. Determine the most suitable order (r,s) and parameters for an ARMAX model.

i (6 hr)	1	2	3	4	5	6	7	8	9	10	11	12	13
h_{i-1}	0.0	1.91	4.27	5.86	6.86	5.52	4.82	4.31	3.84	2.72	2.16	1.70	1.03

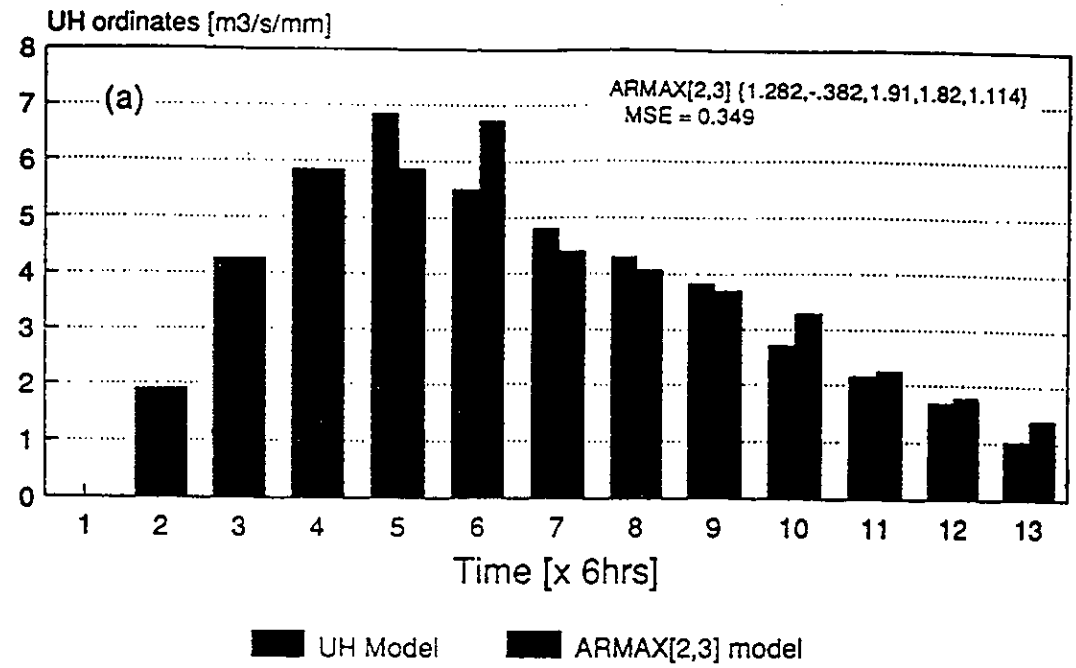
1st Trial: Choose r=2, s=3.

$$h_1 = \omega_1$$
$$h_2 = \omega_2 + \delta_1 h_1 \qquad \text{(A1.5a)}$$
$$h_3 = \omega_3 + \delta_1 h_2 + \delta_2 h_1$$
$$h_4 = \delta_1 h_3 + \delta_2 h_2$$
$$h_5 = \delta_1 h_4 + \delta_2 h_3$$
•
(A1.5b)
•
$$h_{12} = \delta_1 h_{11} + \delta_2 h_{10}$$

δ_1 and δ_2 are estimated from nine equations in Eqn(A1.5b) using the recursive least squares method. The values are estimated to be $\delta_1 = 1.282$ and $\delta_2 = -0.382$. Now these values are substituted in Eqn(A1.5a) to determine ωs. The estimated values are $\omega_1 = 1.91$, $\omega_2 = 1.82$, $\omega_3 = 1.11$, and MSE = 0.349. Fig. A1a indicates that the UH ordinates corresponding to ARMAX [2,3] do not fit well in the peak region. So the model order is increased.

2nd Trial: Choose r=2, s=4.

The same procedure is repeated and the values of 'δ' and 'ω' are estimated to be $\delta_1 = 0.749$, $\delta_2 = 0.074$, and $\omega_1 = 1.91$, $\omega_2 = 2.84$, $\omega_3 = 2.52$, $\omega_4 = 2.16$, and MSE = 0.080. Fig. A1b indicates that the UH ordinates corresponding to ARMAX [2,4] tally reasonably well with the operational UH ordinates. A further increase in the order showed no improvement in the fit and this model was accepted as a good fit.



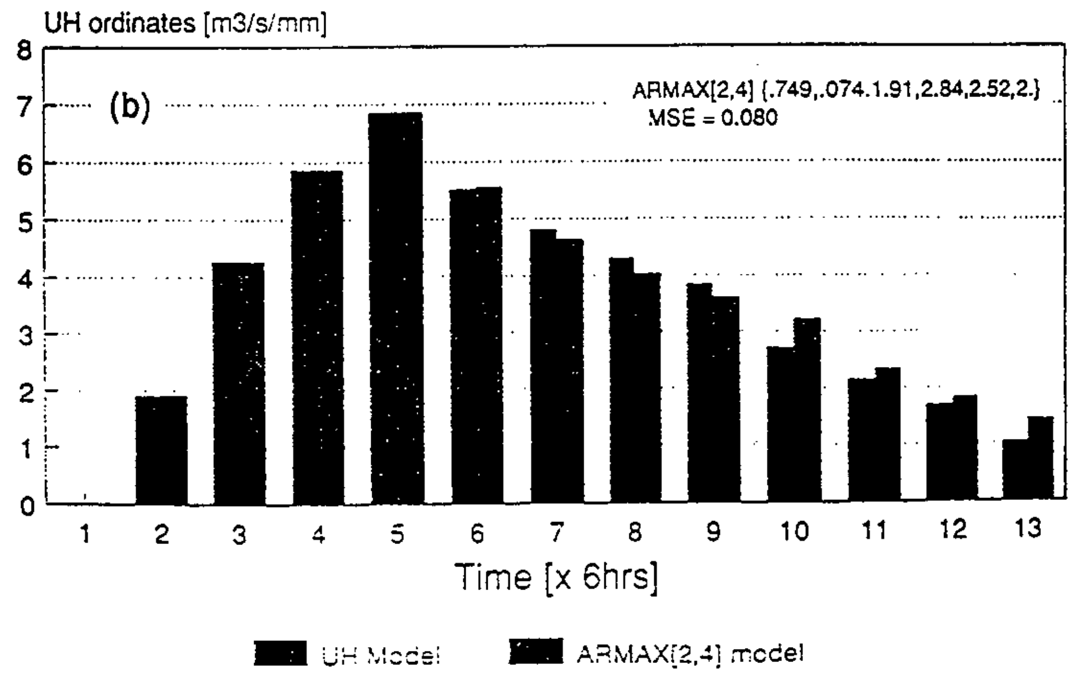


Figure A1: Comparison of actual UH ordinates and the UH ordinates derived from the ARMAX model.

APPENDIX 2

REVIEW OF RECURSIVE LINEAR ESTIMATION METHODS

We begin with a purely deterministic form of the recursive least squares for a linear system with time invariant parameters. We then move on to the dynamic least squares, where the parameters are considered to be variable with time. Finally, the discrete linear optimal estimator, namely, the Discrete Kalman Filter, is considered.

1. Recursive Linear Regression Analysis

Let us consider the problem of linear regression, in which the variable y_o is linearly dependent on n other linearly independent variables, given by:

$$y_o = a_1 . x_1 + a_2 . x_2 .... + a_n . x_n \qquad \text{(A2.1.1)}$$

where a_j, j = 1, 2, .. n constant, but unknown parameters.

Let us suppose that the variables x_j are known exactly, but y_o is observed in the presence of noise ε and there are k such noisy observations y_i, i = 1,2,.. k and the noise sequence is ε_i, i = 1,2,.. k. The observations in general could be written as:

$$y_i = a_1 . x_{i1} + a_2 . x_{i2} .... + a_n . x_{in} + \varepsilon_i \; ; \quad i = 1,2,.. k \qquad \text{(A2.1.2)}$$

or

$$Y = \chi^T . \underline{a} + E \qquad \text{(A2.1.3)}$$

where $Y = [y_1, y_2, ... y_k]^T$

$$\chi^T = \begin{bmatrix} x_{11}x_{12}\cdots x_{1n} \\ x_{21}x_{22}\cdots x_{2n} \\ . \\ . \\ . \\ x_{k1}x_{k2}\cdots x_{kn} \end{bmatrix}$$

$E = (\varepsilon_1, \varepsilon_2, ... \varepsilon_n)^T$

$\underline{a} = (a_1, a_2, ... a_n)^T$

or

$$y_1 = X_i^T . \underline{a} + \varepsilon_i \qquad \text{(A2.1.4)}$$

where $X_i^T = [x_{i1}, x_{i2}, ... x_{in}]$

Let us assume that the sequence of noise ε_i, i = 1,2,.. k has the following statistical properties:

(i) it has zero mean value, i.e. $E\{\varepsilon_i\} = 0$, and variance σ^2;

(ii) it is sequentially uncorrelated (serial correlation is zero), i.e.

$$E\{\varepsilon_i . \varepsilon_j\} = \sigma^2 . \delta_{ij};$$

$$\text{where } \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i=j \end{cases};$$

(iii) ε_i is independent of x_{ij}; i.e. $E\{\varepsilon_i . x_{ij}\} = 0$.

The method of least squares can be used to estimate the parameters $\underline{a} = (a_1, a_2, \dots a_n)^T$ of the problem mentioned above so that J, as defined below, is minimum.

$$J = \sum_{i=1}^{k} \left[\sum_{j=1}^{n} x_{ij} a_j - y_i \right]^2$$

$$\text{or} \quad J = \sum_{i=1}^{k} [X_i^T . \underline{a} - y_i]^2$$

$$\text{or} \quad J = [\chi^T \underline{a} - Y]^T [\chi^T \underline{a} - Y]$$

The derivative of J with respect to elements in vector $\underline{a}$ should be zero for minimum J:

$$\frac{1}{2} \nabla a(J) = \underline{a} \sum_{i=1}^{k} [X_i X_i^T] - \sum_{i=1}^{k} [X_i y_i] = \chi\chi^T . \underline{a} - \chi Y = 0 \quad \text{(A2.1.5)}$$

where $\nabla a(J)$ is the gradient with respect to every element of $\underline{a}$. If the matrix $[X_i X_i^T]$ is non-singular, the solution of equations becomes:

$$\underline{a}_k = \rho_k . b_k \quad \text{(A2.1.6)}$$

$$\text{where} \quad \rho_k = \left[\sum_{i=1}^{k} X_i X_i^T \right]^{-1} ; \quad b_k = \sum_{i=1}^{k} [X_i y_i] \quad \text{(A2.1.7)}$$

$$\text{or} \quad \rho_k = [\chi\chi^T]^{-1} ; \quad b_k = [\chi Y] \quad \text{(A2.1.8)}$$

In order to develop a recursive algorithm for the solution of Eqn(A2.1.6), the estimate of $\underline{a}$ after k-th instant $\hat{\underline{a}}_k$ is given by a linear addition of $\hat{\underline{a}}_{k-1}$ and a corrective term based on the information y_k and X_k received at k-th instant. From Eqns(A2.1.7) and (A2.1.8) we obtain:

$$\rho_k^{-1} = \rho_{k-1}^{-1} + X_k X_k^T \tag{A2.1.9}$$

$$b_k = b_{k-1} + X_k y_k \tag{A2.1.10}$$

We can invert the matrices in Eqn(A2.1.9) to obtain an expression for ρ_k^{-1} in terms of ρ_{k-1}^{-1} as follows (matrix inversion lemma):

$$\rho_k = \rho_{k-1} - \rho_{k-1} X_k (1 + X_k^T \rho_{k-1} X_k)^{-1} X_k^T \rho_{k-1} \tag{A2.1.11}$$

Substituting (A2.1.9) and (A2.1.10) in Eqn(A2.1.11), we get:

$$\underline{\hat{a}}_k = \underline{\hat{a}}_{k-1} + G_k (X_k^T \underline{\hat{a}}_{k-1} - y_k) \tag{A2.1.12}$$

where $$G_k = \rho_{k-1} X_k (1 + X_k^T \rho_{k-1} X_k)^{-1} \tag{A2.1.13}$$

So far, we have not used any of the assumptions made concerning the nature of the noise sequence. Let us consider the statistical properties of the prediction error $\underline{\tilde{a}}_k$ $(=(\underline{\hat{a}}_k - \underline{a}).)$:

(i) mean value of $\underline{\tilde{a}}_k$:

$$E\{\underline{\hat{a}} - \underline{a}\} = E\{[\chi\chi^T]^{-1} \chi E\} = 0$$

(ii) the covariance matrix of $\underline{\tilde{a}}_k$:

$$Var\{\underline{\tilde{a}}_k\} = E\{[\chi\chi^T]^{-1} \chi E E^T \chi^T \{[\chi\chi^T]^{-1}\}\}$$

Replacing $E\{E E^T\}$ by $\sigma^2 I$, we obtain:

$$Var\{\underline{\tilde{a}}_k\} = \sigma^2 [\chi\chi^T]^{-1} = \sigma^2 . \rho_k$$
$$P_k = \sigma^2 . \rho_k$$

Replacing ρ_k by $\frac{P_k}{\sigma^2}$ in Eqns(A2.1.11), (A2.1.12) and (A2.1.13), we get:

$$P_k = (I - G_k X_k^T) P_{k-1} \tag{A2.1.14}$$

where $$G_k = P_{k-1} X_k (\sigma^2 + X_k^T P_{k-1} X_k)^{-1} \tag{A2.1.15}$$

Thus the precision of the parameters $\hat{\underline{a}}_k$ is indicated by the covariance matrix P_k at any instant k.

2. Dynamic Least Squares Method

In the linear regression analysis, the parameters in the regression model are considered time invariant and are given by:

$$\underline{a}_k = \underline{a}_{k-1}$$

If this is not the case in reality, it is not wise to use this method. Let us assume that the parameters vary according to the stochastic difference equation given by:

$$\underline{a}_k = \Phi \, \underline{a}_{k-1} + \Gamma \, \underline{w}_k \qquad \text{(A2.2.1)}$$

where $\Phi = \Phi(k, k\text{-}1)$ transition matrix

$\Gamma = \Gamma(k,k\text{-}1)$ coefficient matrix of noise

and $\underline{w}_k =$ vector consisting of independent random variables $\sim N(0,Q)$,

i.e. $E\{\underline{w}_k\} = 0$

$E\{\underline{w}_k \, \underline{w}_{k-}^T\} = Q \, . \, \delta ij \, ;$

where $\delta ij = 1 \text{ if } i = j$

$ 0 \text{ if } i = j$

A simple example of such a model is the equation of random walk given by:

$$\underline{a}_k = \underline{a}_{k-1} + \underline{w}_k \qquad \text{(A2.2.2)}$$

Let us consider a linear model given by:

$$y_k = X_k^T \, . \, \underline{a} + \varepsilon_k \qquad \text{(A2.2.3)}$$

where the unknown parameter varies from (k-1)-th instant to k-th instant according to the stochastic difference equation, Eqn(A2.2.1). A step ahead prediction of $\underline{a}$.:

$$\underline{a}_{k/k-1} = \Phi \, \hat{\underline{a}}_{k-1} \qquad \text{(A2.2.4)}$$

From Eqns(A2.2.1) and (A2.2.4) error of prediction:

$$\tilde{\underline{a}}_{k/k-1} = \hat{\underline{a}}_{k/k-1} - \underline{a}_k$$

$$\tilde{\underline{a}}_{k/k-1} = \Phi \, \hat{\underline{a}}_{k-1} - \Phi \, \underline{a}_{k-1} - \Gamma \, \underline{w}_k = \Phi \, \tilde{\underline{a}}_{k-1} - \Gamma \, \underline{w}_k \qquad \text{(A2.2.5)}$$

The covariance matrix $P_{k/k-1}$:

$$P_{k/k-1} = E\{\underline{\tilde{a}}_{k/k-1}\ \underline{\tilde{a}}^T_{k/k-1-}\}$$
$$= E\{[\Phi\ \underline{\tilde{a}}_{k-1} - \Gamma\ \underline{w}_{\cdot k}]\ [\Phi\ \underline{\tilde{a}}_{k-1} - \Gamma\ \underline{w}_{\cdot k}]^T\}$$

Since $\underline{w}_{\cdot k}$ influences only $\underline{\tilde{a}}_k$, $E\{\underline{\tilde{a}}_{k-1} \cdot \underline{w}^T_{k-}\} = 0$, therefore:

$$P_{k/k-1} = \Phi\ P_{k-1}\ \Phi^T + \Gamma Q \Gamma^T \qquad \text{(A2.2.6)}$$

where $\quad P_{k/k-1} = E\{\underline{\tilde{a}}_{k-1}\ \underline{\tilde{a}}^T_{k-1-}\}$

A step ahead prediction of $\underline{a}_{\cdot k}$ ($\underline{\hat{a}}_{k/k-1}$) and P_k ($P_{k/k-1}$) is given by Eqns(A2.2.4) and (A2.2.6). Making use of these equations of predictions of variable parameters, we could modify the simple recursive least squares algorithm to sequentially estimate the variable parameter $\underline{a}_{\cdot k}$. Thus, in summary, we have the following equations:

Prediction:

$$\underline{a}_{\cdot k/k-1} = \Phi\ \underline{\hat{a}}_{k-1}$$

$$P_{k/k-1} = \Phi\ P_{k-1}\ \Phi^T + \Gamma Q \Gamma^T$$

Correction:

$$\underline{\hat{a}}_k = \underline{\hat{a}}_{k-1} + G_k\ (X^T_k\ \underline{a}_{k/k-1} - y_k)$$

$$P_k = (1 - G_k X^T_k)\ P_{k/k-1}$$

where $\quad G_k = P_{k/k-1}\ X_k\ (\sigma^2 + X^T_k\ P_{k/k-1}\ X_k)^{-1}$

Note: In practice, it is difficult to obtain a model to describe the variation of parameters given in Eqn(A2.2.1). The slowly varying parameters could be represented by the equation of random walk given by Eqn(A2.2.2). Thus the equations of prediction becomes:

$$\underline{a}_{\cdot k/k-1} = \underline{\hat{a}}_{k-1} \quad \text{and} \quad P_{k/k-1} = P_{k-1} + Q$$

3. Discrete Linear Optimal Estimation Technique (Kalman Filter)

Although there are several methods to establish a linear optimal estimator, we will consider only a simple approach. We shall seek to develop an estimator which is a linear function of measurements. This estimator should minimise the covariance of errors existing between estimation and the real value of the state. Initially, let us consider a discrete linear system given by:

$$X_{k+1} = \Phi_{k+1} X_k + \Gamma_{k+1} W_{k+1} \quad \text{(A2.3.1)}$$

The state X is observed by the measurement system given by:

$$Z_{k+1} = H X_{k+1} + V_{k+1} \quad \text{(A2.3.2)}$$

where
- X = state variable
- Z = measurement
- V = measurement error
- W = system error
- Φ = transition matrix
- H = measurement matrix
- Γ_{k+1} = coefficient matrix of system noise.

A step ahead prediction of the state at k-th instant is given by:

$$\hat{X}_{k+1/k} = \Phi_{k+1} X_{k/k} \quad \text{(A2.3.3)}$$

The prediction of measurement is:

$$\hat{Z}_{k+1} = HX_{k+1/k} \quad \text{(A2.3.4)}$$

Prediction error of the state:

$$\tilde{X}_{k+1/k} = X_{k+1} - \hat{X}_{k+1/k} \quad \text{(A2.3.5)}$$

Substituting from Eqns(A2.3.1) and (A2.3.3) we get:

$$\tilde{X}_{k+1/k} = \Phi_{k+1}(X_k - \hat{X}_{k/k}) + \Gamma_{k+1} W_{k+1}$$

or
$$\tilde{X}_{k+1/k} = \Phi_{k+1}\tilde{X}_k + \Gamma_{k+1} W_{k+1} \quad \text{(A2.3.6)}$$

The covariance matrix $P_{k/k-1}$ is given by:

$$P_{k+1/k} = E\left\{\tilde{X}_{k+1/k}\tilde{X}^T_{k+1/k}\right\}$$

Assuming no correlation between W_{k+1} and $\tilde{X}_k\left(\text{i.e.} E\left\{\tilde{X}_k W^T_{k+1}\right\} = 0\right)$, we obtain:

$$P_{k/k-1} = \Phi\ P_{k-1}\ \Phi^T + \Gamma\ Q\ \Gamma^T \quad \text{(A2.3.7)}$$

On receipt of $\underline{Z}_{k+1}$, this new information could be used for modifying the prediction. The deviation between the measurement and its prediction is:

$$\underline{\varepsilon} = \underline{Z}_{k+1} - H X_{k+1/k}$$

Since the estimation of state is a linear function of measurement, the corrected state can be given by:

$$\hat{\underline{X}}_{k+1/k+1} = \hat{\underline{X}}_{k+1/k} + \kappa_{k+1} (\underline{Z}_{k+1} - H \underline{X}_{k+1/k}) \qquad \text{(A2.3.8)}$$

Where κ_{k+1} is a gain matrix and its choice is such that it minimizes the covariance of error existing between estimation and the real value of the state variable X_{k+1}:

$$P_{k+1/k+1} \; [E\{ \tilde{X}_{k+1/k+1} \tilde{X}^T_{k+1/k+1} \}] .$$

$$\tilde{X}_{k+1/k+1} = X_{k+1} - \hat{X}_{k+1/k+1}$$

$$\tilde{X}_{k+1/k+1} = X_{k+1} - \hat{\underline{X}}_{k+1/k} - \kappa_{k+1} (\underline{Z}_{k+1} - H \underline{X}k+1/k)$$

By replacing $\underline{Z}_{k+1}$ by $(\underline{\varepsilon} + H X_{k+1})$ we obtain:

$$\tilde{X}_{k+1/k+1} = \tilde{X}_{k+1/k} - \kappa_{k+1} H \tilde{X}_{k+1/k} - \kappa_{k+1} (\underline{\varepsilon}_{k+1})$$

or $$\tilde{X}_{k+1/k+1} = (I - \kappa_{k+1} H) \tilde{X}_{k+1/k} - \kappa_{k+1} (\underline{\varepsilon}_{k+1})$$

Thus:

$$P_{k+1/k+1} = E\{ \tilde{X}_{k+1/k+1} \tilde{X}^T_{k+1/k+1} \}$$

$$= E\{[(1 - \kappa_{k+1} H) \tilde{X}_{k+1/k} - \kappa_{k+1} (\underline{\varepsilon}_{k+1})] [(1 - \kappa_{k+1} H) \tilde{X}_{k+1/k} - \kappa_{k+1} (\underline{\varepsilon}_{k+1})]^T\}$$

Assuming no correlation between the state estimation error and measurement error (i.e. $E[X_{k+1/k} \, \underline{\varepsilon}_{k+1}] = 0$) and the measurement error is serially independent, we can rewrite the expression for the covariance of state estimation error as:

$$P_{k+1/k+1} = (I - \kappa_{k+1} H) P_{k+1/k} (I - \kappa_{k+1} H)^T + \kappa_{k+1} R \kappa_{k+1} \qquad \text{(A2.3.9)}$$

When Eqn(A2.3.8) is expanded, it gives:

$$P_{k+1/k+1} = P_{k+1/k} - \kappa_{k+1} H P_{k+1/k} - P_{k+1/k} H^T \kappa_{k+1} + \kappa_{k+1} (H P_{k+1/k} H^T \kappa_{k+1} + R) K^T_{k+1} \qquad \text{(A2.3.10)}$$

The matrix $(H P_{k+1/k} H^T) + R$ in Eqn(A2.3.9) is symmetric, since $P_{k+1/k}$ and R are symmetric matrices. Thus it could be written as:

$$(H P_{k+1/k} H^T + R) = A_{k+1} A_{k+1}^T \qquad (A2.3.11)$$

The last three terms in Eqn(A2.3.10) have a form of quadratic matrix polynomial on κ_{k+1}. Now, assuming that there is a matrix B_{k+1}, such that:

$$B_{k+1} A_{k+1}^T = P_{k+1/k} H^T \qquad (A2.3.12)$$

eqn(A2.3.9) becomes:

$$P_{k+1/k+1} = P_{k+1/k} + (\kappa_{k+1} A_{k+1} - B_{k+1}) (\kappa_{k+1} A_{k+1} - B_{k+1})^T - B_{k+1} B_{k+1}^T$$

Supposing that the matrix $A_{k+1} A_{k+1}^T$ is positive definite, $P_{k+1/k+1}$ is minimum when $\kappa_{k+1} A_{k+1} = B_{k+1}$. Thus substituting for B_{k+1} in Eqn(A2.3.12), we obtain:

$$\kappa_{k+1} A_{k+1} A_{k+1}^T = P_{k+1/k} H^T \qquad (A2.3.13)$$

Thus the gain κ_{k+1} is given by:

$$\kappa_{k+1} = P_{k+1/k} H^T (H P_{k+1/k} H^T + R)^{-1} \qquad (A2.3.14)$$

Replacing κ_{k+1} by its value in the last term of Eqn(A2.3.10), we obtain:

$$P_{k+1/k+1} = (I - \kappa_{k+1} H^T) P_{k+1/k} \qquad (A2.3.15)$$

These expressions can be modified to take into account the deterministic inputs (U_{k+1}) given by:

$$X_{k+1/k} = \Phi_{k+1} X_k + \Omega_{k+1} U_{k+1} + \Gamma_{k+1} W_{k+1} \qquad (A2.3.16)$$

where U_{k+1} is the deterministic input
Ω_{k+1} is the input matrix.

Thus the optimal linear estimator commonly known as The Kalman Filter is given by the following equations:

Prediction:

$$X_{k+1/k} = \Phi_{k+1} X_{k/k} + \Omega_{k+1} U_{k+1}$$

$$P_{k+1/k} = \Phi P_{k-1} \Phi^T + \Gamma Q \Gamma^T$$

Correction (or filtering):

$$\underline{X}_{k+1/k+1} = \underline{X}_{k+1/k} + \kappa_{k+1} (\underline{Z}_{k+1} - H \underline{X}_{k+1/k})$$

$$P_{k+1/k+1} = (I - \kappa_{k+1} H^T) P_{k+1/k}$$

where $\kappa_{k+1} = P_{k+1/k} H^T (H P_{k+1/k} H^T + R)^{-1}$

APPENDIX 3

WORKED EXAMPLE

A lower order model [ARMAX(1,2,1)] is used to simulate runoff from rainfall data. The model is given by:

$$q_k = \delta_1 q_{k-1} + \omega_1 Rf_{(k-lg)} + \omega_2 Rf_{(k-1-lg)} + v_{(k)}$$

or $$q_k = H_k^T . X_k + v(k)$$

where
$$H_k^T = [q_{k-1}, Rf^{(k-1)}, Rf^{(k-2)}]$$
$$X_k = [\delta_1, \omega_1, \omega_2]^T$$
$$v(k) = \text{measurement noise } [\simeq N(O,R)]$$

The variation of parameters is represented by simple random walk:

$$\begin{bmatrix}\delta_1\\ \omega_1\\ \omega_2\end{bmatrix}_{k+1} = \begin{bmatrix}\delta_1\\ \omega_1\\ \omega_2\end{bmatrix}_{k} + \begin{bmatrix}w_1\\ w_2\\ w_3\end{bmatrix}_{k}$$

or $$X_{k+1} = X_k + W(k)$$

where $$W(k) \simeq N(O,Q)$$

$$Q = \begin{bmatrix}q^1\ 0\ 0\\ 0\ q^2\ 0\\ 0\ 0\ q^3\end{bmatrix}$$

The parameter values are simulated using generated independent random normal variates $Vn \simeq N(0,1)$, where $w_i = \sqrt{qi} . Vn$. The measurement noise is simulated likewise, where $V = \sqrt{R} . Vn$. The initial values of ARMAX model parameters are given by:

$$\begin{bmatrix}\delta_1\\ \omega_1\\ \omega_2\end{bmatrix}_{0} = \begin{bmatrix}0.62\\ 5.42\\ 5.25\end{bmatrix}$$

A numerical example of Recursive Algorithm (Kalman Filter):

Noise statistics:

$R = 100.0$

$$Q = \begin{bmatrix} 0.0001 & 0.0000 & 0.0000 \\ 0.0000 & 0.0100 & 0.0000 \\ 0.0000 & 0.0000 & 0.0100 \end{bmatrix}$$

State prediction:

$X_{k/k-1} = X_{k-1/k-1}$

$X^T_{k/k-1} = X^T_{k-1/k-1} = [0.598, 5.218, 4.581]$

$H^T_k = [327.0, 30.0, 21.0]$

Error covariance matrix:

$P_{k-1/k-1}$

$$\begin{bmatrix} 0.00099 & -.00609 & -.00634 \\ -.00609 & 0.16511 & -.05259 \\ -.00634 & -.05259 & 0.22532 \end{bmatrix}$$

Output prediction:

$q_{k/k-1} = H^T_k \cdot X_{k/k-1}$

$q_{k/k-1} = 448.3$; $q_k = 453$

Innovation:

$v_k = q_k - q_{k/k-1} = 4.7$

Predicted error covariance matrix:

$P_{k/k-1} = P_{k-1/k-1} + Q$

$$\begin{bmatrix} 0.00109 & -.00609 & -.00634 \\ -.00609 & 0.17511 & -.05259 \\ -.00634 & -.05259 & 0.23532 \end{bmatrix}$$

$P_{k/k-1} H_k = [0.04059, 2.1575, 1.2908]^T$

$\left[H^T_k \cdot P_{k/k-1} H_k + R\right] = 205.104$

Predictor gain:

$$G_k = P_{k/k-1}\ H_k \left[H_k^T . P_{k/k-1}\ H_k + R\right]^{-1}$$

$$G_k = [0.000198, 0.01050, 0.00629]^T$$

Correction:

$$G_k . v_k = [0.001, 0.049, 0.030]^T$$

State estimation using new measurements:

$$X_{k/k} = X_{k/k-1} + G_k\, v_k$$

$$X_{k/k}^T = [0.599, 5.268, 4.610]$$

Error covariance matrix after new measurement:

$$P_{k/k} = \left[I - G_k\, H_k^T\right] P_{k/k-1}$$

$$G_k\, H_k^T = \begin{bmatrix} 0.06475 & 0.00594 & 0.00416 \\ 3.4335 & 0.315 & 0.2205 \\ 2.0568 & 0.1887 & 0.13209 \end{bmatrix}$$

$$\left[I - G_k\, H_k^T\right] = \begin{bmatrix} 0.93525 & -.00594 & -.00416 \\ -3.4335 & 0.685 & -.2205 \\ -2.0568 & -.1887 & 0.86791 \end{bmatrix}$$

$$P_{k/k} = \begin{bmatrix} 0.00108 & -.00652 & -.00660 \\ -.00652 & 0.15246 & -.06614 \\ -.00660 & -.06616 & -.22720 \end{bmatrix}$$

The Cooperative Research Centre for Catchment Hydrology is a cooperative venture formed under the Commonwealth CRC Program between:

Bureau of Meteorology

CSIRO Division of Water Resources

Department of Conservation and Natural Resources

Melbourne Water

Monash University

Murray-Darling Basin Commission

Rural Water Corporation

The University of Melbourne

COOPERATIVE RESEARCH CENTRE FOR
CATCHMENT HYDROLOGY

Centre Office
Department of Civil Engineering, Monash University, Clayton, Victoria 3168, Australia
Telephone: (03) 9905 2704 Facsimile: (03) 9905 5033 International: +61 3 9905 2704