Nordic Hydrology, 1982, 233-246

Combination of a Conceptual Model and an Autoregressive Error Model for Improving Short Time Forecasting

Angela Lundberg
Div. of Water Resources Engineering,
WREL, Luleå, Sweden

An autoregressive error model has been tested on the residuals of the conceptual HBV-model for the Emån catchment. The autoregressive model gives considerable improvements for real shorttime forecasting, but for long range (10 days or more) forecasting no improvement is achieved compared to the conceptual model. Separation of the error functions for high and low discharges does not give any further improvement.

Introduction

Hydrologists working with runoff modelling can roughly be devided into two groups: those using the physical laws for modelling and those using the statistical laws. The discussion between these groups have often been hot-tempered and each group has rejected the methods of the opponents. The main argument of the statisticians are that the physical problems are too complicated and the physical parameters too unevenly distributed to be properly modelled. The "physicists" on the other hand claim that one must use one's knowledge of the physical processes. It is difficult to understand why it is so uncommon to utilize the advantages of both fields and to combine the two ways of looking upon the problem for achieving good results. An attempt to combine the two approaches is made in this paper.

Today the releasing plan is revised every fifth day. In the future with automatic data collection and improving computer capacity it would be possible to run conceptual models with automatic updating which would make it possible to change the plans with shorter intervals. To hydro power plants with low capacity due to small reservoirs or small installations this would mean that they could be used more efficiently.

The Conceptual HBV-Model

A conceptual model, i.e. a model based on physical arguments, developed by Bergström (1976) at the Swedish Meteorological and Hydrological Institution (SMHI), called the HBV-model has been shown to give good estimates of the runoff from several Scandinavian catchments. The regression coefficient is about 0.8-0.9 so there is still room for improvements, and a possible approach is to utilize a statistical error model in conjunction with the original HBV-model. This has been done in the present paper.

When simulated runoff is compared with observed runoff it is often found that there is a pronounced inertia in the errors (cf. Fig. 1). This means that if the model overestimates the runoff today, it will most likely do so tomorrow as well. The great inertia in the errors makes an autoregressive model suitable as an error model.

There is a restriction in the use of an error model; since the corrections are made on the residuals, i.e. the difference between the calculated and the observed runoff, the recorded runoff data have to be reliable. The quality of the recorded runoff data is today often rather poor, but it should be possible to improve those at a moderate cost.

The way the HBV-model is run today at the SMHI is that an experienced hydrologist with good knowledge of the model updates the model before a forecast is made (if updating is considered necessary). The updating is made by entering modified indata. A step towards automatic updating is to improve the forecast with a residual model.

Arma-Models

Autoregressive-Moving Average model is a combination of an autoregressive and a moving average model. Anderson (1976), Box-Jenkins (1976).

Autoregressive Processes

For an autoregressive proces of order *p*, and AR(*p*)-process, the current value of the process is expressed as a weighed sum of past values and a current shock

$$z_i = \phi_1 z_{i-1} + \ldots \qquad \phi_p z_{i-p} + a_i \tag{1}$$

where z_i = current value, ϕ = regression coefficients, a_i = current shock from a random process with zero mean and variance σ_a^2.

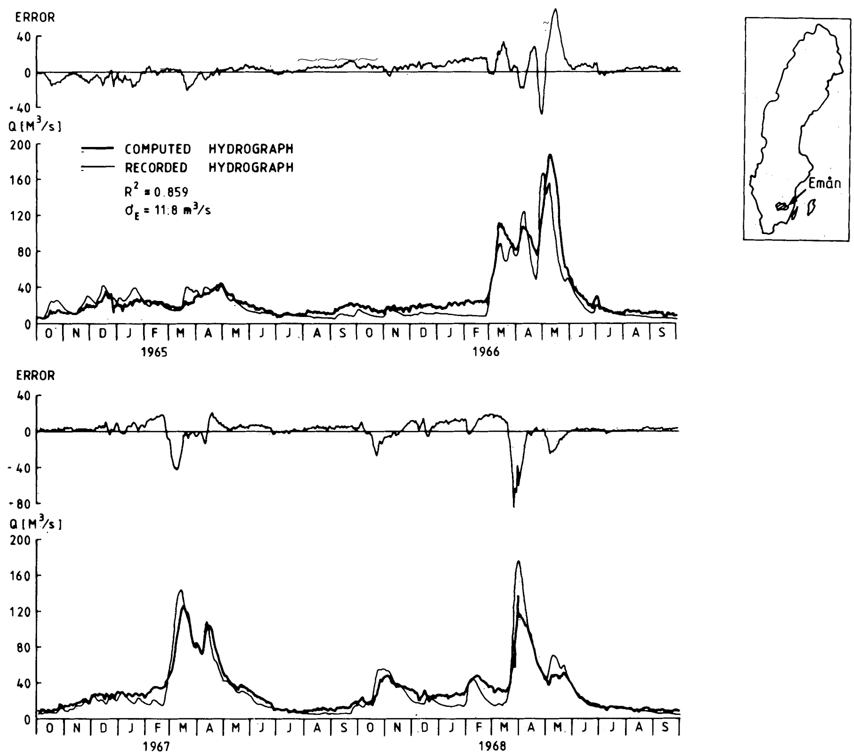

Fig. 1. Observed and by the HBV-model simulated runoff from the Emån catchment in southern Sweden. The residuals are plotted above. Regression coefficient = 0.859, residual standard deviation = 11.8 m^3/s. Emån catchment area 2.343 km^2.

The First Order AR-Process AR(1)

Consider the process

$$z_i = \phi_1 z_{i-1} + a_i \qquad (2)$$

and assume stationarity.

Taking expectations on Eq. (2) gives mean = 0 since $\phi_1 \neq 1$. The variance of z_i is

$$\mathrm{var}(z_i) = E\{(\phi_1 z_{i-1} + a_i)(\phi_1 z_{i-1} + a_i)\}$$

Since the current shock a_i is independent of z_{i-1} this gives

$$\sigma_z^2 = \phi_1^2 \sigma_z^2 + \sigma_a^2 \qquad \text{or} \qquad \sigma_z^2(1-\phi_1^2) = \sigma_a^2$$

For $\sigma_z{}^2$ to be finite and non-negative this requires $-1 < \phi_1 < 1$. This is the region of stationarity. Multiplying Eq. (2) by z_{i-k} and taking expectations

$$E[z_i z_{i-k}] = \phi_1 E[z_{i-1} z_{i-k}] + E[a_i z_{i-k}]$$

so for $k \geq 1$ to make sure that a_i and z_{i-k} are independent

$$\gamma(k) = \phi_1 \gamma(k-1) \qquad (3)$$

where $\gamma(k)$ the *autocoviariance* at lag k is defined

$$\gamma(k) = \text{cov}[z_i \, z_{i-k}]$$

The set of $\{\gamma(k) ; k = 0,1, ...\}$ is called the *autocovariance function.*
The *autocorrelation* at lag k $\varrho(k)$ is defined as

$$\rho(k) \equiv \frac{\gamma(k)}{\gamma(o)}$$

The set of $\{\varrho(k); k = 0,1 ...\}$ is called the *autocorrelation function ACF.*
Eq. (3) is a first order difference equation with solution

$$\gamma(k) = \phi_1^k \gamma(o) \quad \text{or} \quad \rho(k) \equiv \phi_1^k \quad \text{for} \quad k > 1$$

The partial autocorrelation function PACF denoted by $\{\phi_{kk}; k = 1,2 ...\}$ is the set of partial autocorrelations of various lags k. These are defined by

$$\phi_{kk} = |P_k^*| \,/\, |P_k|$$

where P_k is the $k \times k$ autocorrelations matrix.

$$P_k = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \rho(1) & & \\ \rho(2) & \rho(1) & 1 & & \\ \vdots & & & \ddots & \vdots \\ \vdots & & & & \rho(1) \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(1) & 1 \end{bmatrix} \quad \text{and } P_k^* \text{ is } P_k \text{ with the last column replaced by} \quad \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix}$$

so estimates of ϕ_{kk} can be obtained by replacing $\varrho(i)$ by $r(i)$ where $r(i)$ is the autocorrelation of lag i determined from the set. For the AR(1) process the theoretical PACF is given by a single term

$$\phi_{11} \equiv \rho(1) \equiv \phi_1$$

since

$$\phi_{22} = \frac{\rho(2)-\rho(1)^2}{1-\rho(1)^2} = \frac{\phi_1^2-\phi_1^2}{1-\phi_1^2} = 0 \quad \text{since} \quad \phi_1^2 \neq 1$$

General AR(p)

For the general AR(p)-process the region of stationarity, mean, variance, ACF and PACF, can be obtained in a similar way as for the one order process.

For an AR(p)-process the $\phi_{kk} = 0$ for $k > p$ so the PACF cuts off after p terms.

Mean Square Error of $r(k)$ and ϕ_{kk}

Barlett's formula, Bartlett (1946), states that

$$\text{var}[r(k)] \cong \frac{1}{N}(1 + 2\sum_{1}^{k} r(i)^2)$$

and the square root of this is the large-lag standard error of $r(k)$.

Quenouille's formula, Quenouille (1949), states that

$$\text{var}[\hat{\phi}_{kk}] \cong \frac{1}{N}$$

for lag large enough for the PACF to have died out.

Moving Average Processes MA(q)

The general form for a moving average process of order q, the MA(q), process is

$$z_i \equiv a_i + \theta_1 a_{i-1} + \theta_2 a_{i-2} + \quad \ldots \quad + \theta_q a_{i-q}$$

where θ_i are the moving average coefficients and a_i as before is a shock from a random process with zero mean and variance σ_a^2.

The First Order Moving Average Process MA(1)

$$z_i = a_i + \theta_1 a_{i-1} \qquad (4)$$

The MA(1) process is said to be *invertible,* a condition analogous to that of stationarity for an AR-process, if

$$-1 < \theta_1 < 1$$

A MA(∞)-process can be shown to be equivalent to an AR(p)-process and an AR(∞)-process to be equivalent to an MA(q).

The invertibility condition avoids certain model multiplications.

The ACF given by

$$\rho(k) = 0 \qquad k > 1$$

$$\rho(1) = \frac{\theta_1}{1+\theta_1^2}$$

cuts off after lag 1.

The PACF does not, however, cut off, but it can be shown to decay geometrically towards zero.

General Moving Average Processes MA(q)

The general MA(q)-process has complicated invertibility conditions but the ACF cuts off after q. The PACF does not cut off but decays towards zero.

Mixed Processes ARMA(p,q)

$$z_i = \theta_1 z_{i-1} + \ldots + \theta_p z_{i-p} + a_i + \theta_1 a_{i-1} + \ldots + \theta_q a_{i-q} \ldots \quad (5)$$

For mixed processes the stationarity and invertibility conditions are the same as for AR(p) and MA(q)-processes.

For mixed processes the ACF behaves like an AR(p) process after q-p lags and the PACF resembles that of an MA(q) process after p-q lags.

Identification and Estimation

The mean variance, ACF and PACF, is calculated and inspection of $r(k)$ and ϕ_{kk} indicates which model should be tested. If the process has not zero mean, it can be made so by replacing z_i with $\tilde{z}_i = z_i - \bar{z}$. Which model should be entered is decided by considering where the cutoffs, if any, occur in $\{\varrho(k)\}$ and $\{\phi_{kk}\}$ by comparing the estimated functions with their large-lag standard errors and see if the results fit into any of the theoretical patterns. Thus, the values of p and q are obtained (see Table 1).

The estimated autocorrelations are often highly correlated, thus $r(k)$ and ϕ_{kk} can only be used as general guidelines. If the patterns are not very obvious, several models might have to be fitted.

After having identified which model is to be entered one can find rough starting values in Table 2 (MA(2) has been excluded). First a check should be made to ensure that the values are in the admissible region.

After having identified a suitable model the next step is to obtain better estimates of the parameters using the rough values obtained at the identification as starting values. This can be done by a non-linear least square procedure and is rather complicated, but some statistical packages have procedures which give ARMA-estimates. The MINITAB (Ryan 1981) statistical package has been used for the calculations presented.

Calculations and Results

Autoregressive Process on the Residuals of the HBV-Model

A four year sequence of daily means show in Fig. 1 from the catchment Emån has been used as a test period. With the runoff simulated by the pure conceptual HBV-model the regression coefficient R^2 is 0.859, and the standard deviation error σ_E is 11.8 m^3/s. The difference between the observed and simulated runoff is called the error or the residual (i.e. the sequence part that still is not explained). The error in Fig. 1 was made zeromean and a check was made to ensure that there was no trend. The autocorrelation function ACF and the partial autocorrelation function PACF are plotted for different lags in Fig. 2.

Table 1

$\hat{\phi}_{kk} \cong \frac{1}{\sqrt{N}}$	$k > p \Rightarrow$ AR(p)
$r(k) \cong \left(1 + 2\sum_{1}^{q} r^2(i)\right)^{1/2}$	$k > q \Rightarrow$ MA(q)
neither $\hat{\phi}_{kk}$ nor $r(k)$ cuts off	$\Rightarrow$ ARMA

Table 2 -(From Anderson O.D. Time Series Analysis and Forecasting)

Process	Admissible region	Initial estimates
AR(1)	$-1 < r_1 < 1$	$\hat{\phi} = r_1$
AR(2)	$-1 < r_2 < 1$	$\hat{\phi}_1 = r_1(1-r_2)/(1-r_1^2)$
	$r_1^2 < \frac{1}{2}(r_2+1)$	$\hat{\phi}_2 = (r_2-r_1^2)/(1-r_1^2)$
MA(1)	$-0.5 < r_1 < 0.5$	$\hat{\theta}_1 = \left\{1-\sqrt{(1-4r_1^2)}\right\}/2r_1$
ARMA(1,1)	$2r_1^2 - \lvert r_1\rvert < r_2 < \lvert r_1\rvert$	$\hat{\phi}_1 = r_2/r_1$
		$\hat{\theta}_1 = \left\{b \pm \sqrt{(b^2-4)}\right\} /2$ *)
		where
		$b = (1-2r_2 + \hat{\phi}_1^2)/(r_1-\hat{\phi}_1)$

*) The sign being chosen to ensure $|\hat{\theta}_1| < 1$.

The standard deviation error for the PACF = $1/\sqrt{N}$. is also indicated. The ACF is slowly decreasing towards zero, while the PACF is oscillating around zero with decreasing amplitude. This indicates that an autoregressive-moving average (ARMA) model would be applicable. The PACF of lag 1 is much greater than the other lags. An AR(1) should therefore be sufficient. On account of the statistical principle of parsimonity (i.e. models prodigal in parameters are neglected in favour of those which are more economical) an AR(1) model is tested. Since for AR(1) $\phi_{11}=\varrho(1)$, and since the estimates achieved by the non-linear least square method differ very little from $\varrho(1)$, the $\varrho(1)$ values are used.

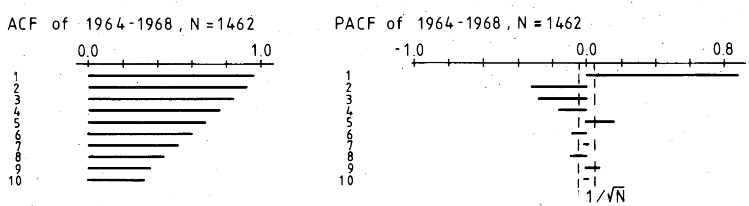

Fig. 2. Autocovariance function, ACF, and partial autocovariance function, PACF, for the Emån catchment 1964-1968. Standard deviation error of PACF ≡ $1/\sqrt{N}$.

The error model for the lags 1, 3, 5, and 7 becomes

$$z_{i+1} = 0.968 \times z_i + a_i(1) \qquad z_{i+5} = 0.685 \times z_i + a_i(5)$$
$$z_{i+3} = 0.848 \times z_i + a_i(3) \qquad z_{i+7} = 0.520 \times z_i + a_i(7)$$

where z_{i+n} is the remaining error of the n days ahead forecast after correcting for the error of day i of the conceptual model; i is the day number, z_i is error of the conceptual model for the i day, and $a_i(n)$ is a process, which should have mean = 0 and be random.

The simulated discharge from the combined conceptual and autoregressive model $Q_{\mathrm{AR},n(t)}$ can thus be calculated

$$Q_{\mathrm{AR},n(t)} = Q_{\mathrm{CONC}}(t) + \mathrm{CONST}(n)[Q_{\mathrm{CONC}}(t-n) - Q_{\mathrm{OBS}}(t-n)]$$

where index AR stands for autoregressive, n for number of days ahead of forecast, t is the day number, Q_{CONC} is the discharge calculated with the conceptual model, Q_{OBS} is the observed discharge. The function CONST(n) is the above determined regression coefficient.

In Figs. 3-4 the results of the combined conceptual and error models from the Emån catchment for 1964-1968 are plotted for different lags. In Table 3 the regression coefficient and the residual standard deviation are shown.

For forecasting up to nine days ahead an improved fitting is achieved, as can be seen from Table 3, with an error model. For forecasting up to five days the improvement is considerable.

Test on an Independent Period

For an independent period 1969-72 forecasts for different lags were also made. For the conceptual model the entire period 1964-72 is independent. Since only one parameter is introduced the risk of overparametrization should be small. The remaining residuals are plotted in Figs. 5 and 6. In Table 4 the regression coefficient and the standard deviation coefficient of the residuals for different lags are

Table 3 – Regression coefficient and standard deviation of residuals for the basic conceptual model and for the combined model for different numbers of days forecast. Calibration period 1964-1968. The Emån catchment.

	Regression coefficient	Residual standard deviation (m^3/s)
Basic conceptual HBV-model	0.859	11.8
Conceptual + AR(1) 1 day ahead	0.990	3.0
Conceptual + AR(1) 3 days ahead	0.958	6.4
Conceptual + AR(1) 5 days ahead	0.922	8.7
Conceptual + AR(1) 7 days ahead	0.893	10.3

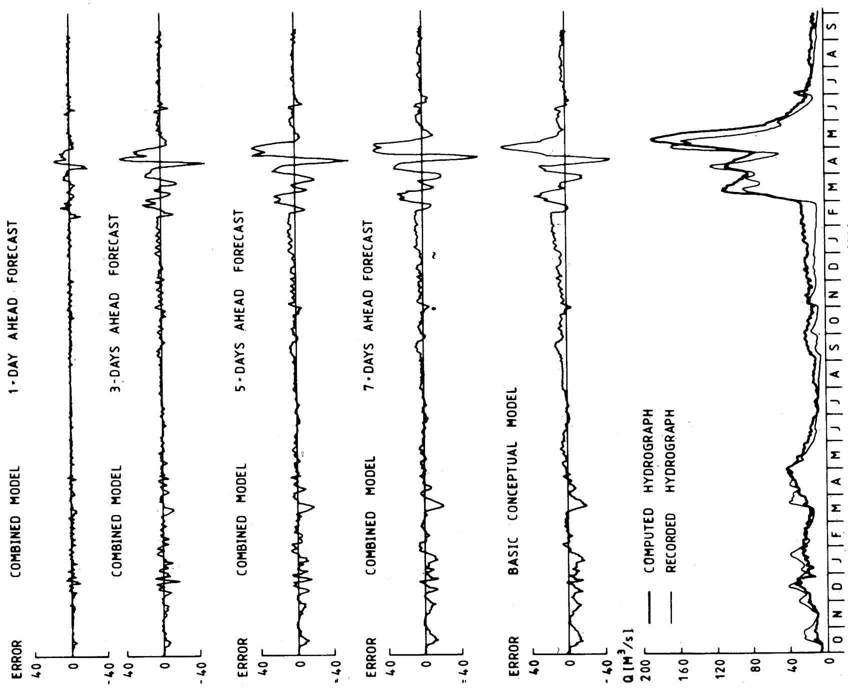

Fig. 3. Observed and simulated runoff with the conceptual model. Also plotted are the errors for different models. The Emån catchment. Calibration period 1965-66.

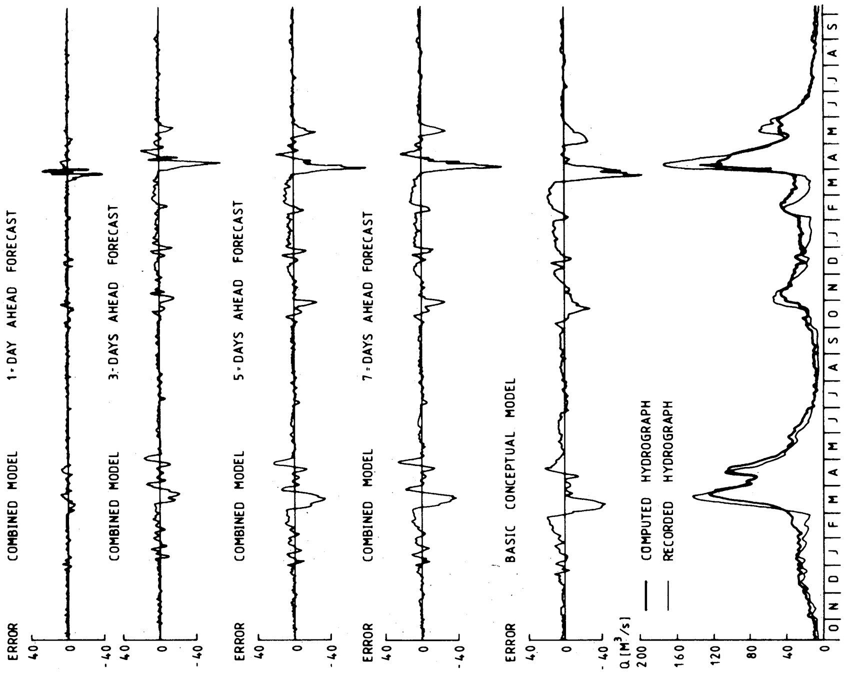

Fig. 4. Observed and simulated runoff with the conceptual model. Also plotted are the errors for different models. The Emån catchment. Calibration period 1967-68.

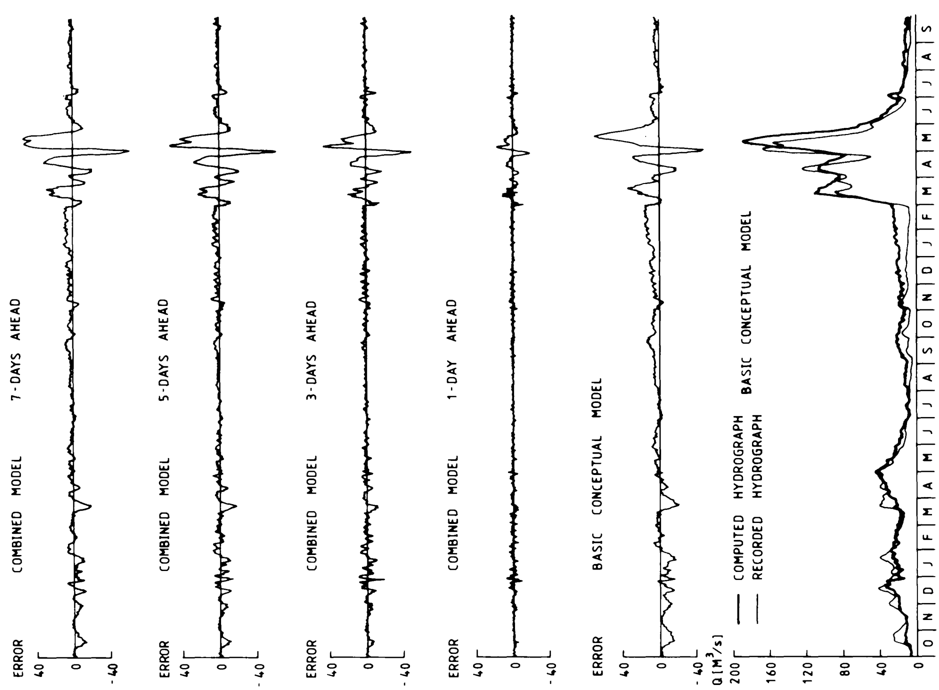

Fig. 5. Independent period 1969-70. Observed and simulated runoff with the conceptual model. Also plotted are the errors for different models. The Emån catchment.

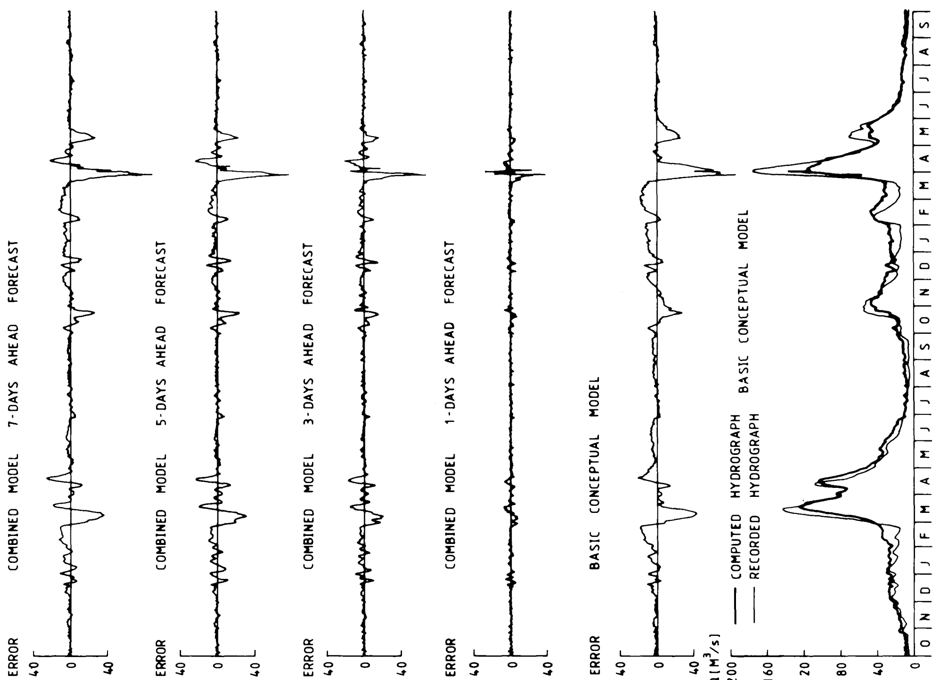

Fig. 6. Independent period 1971-72. Observed and simulated runoff with the conceptual model. Also plotted are the errors for different models. The Emån catchment.

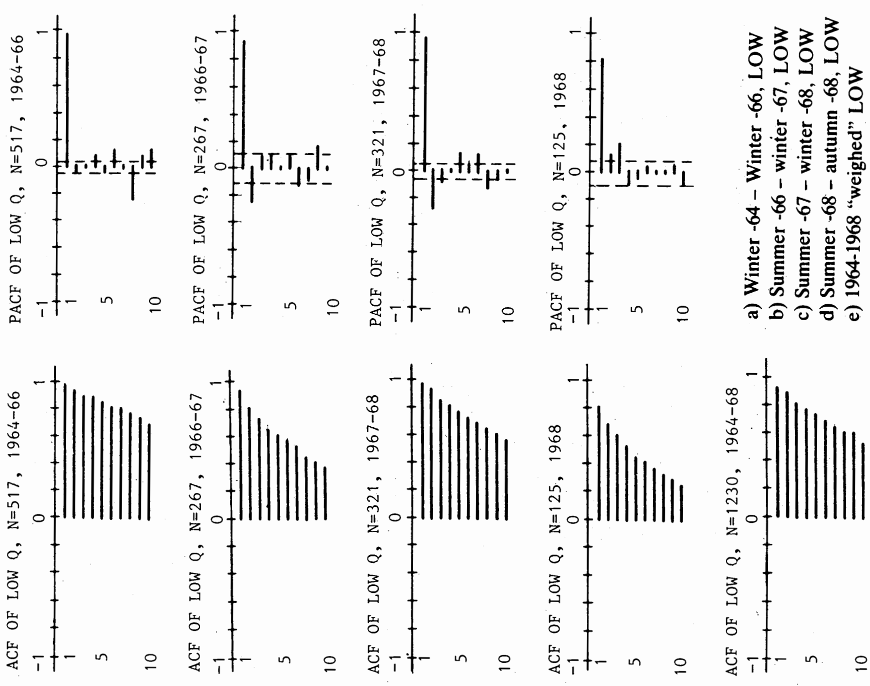

Fig. 7. Autocorrelation function, ACF, and partial autocor relation function, PACF, for discharges less than 40 m^3/s.

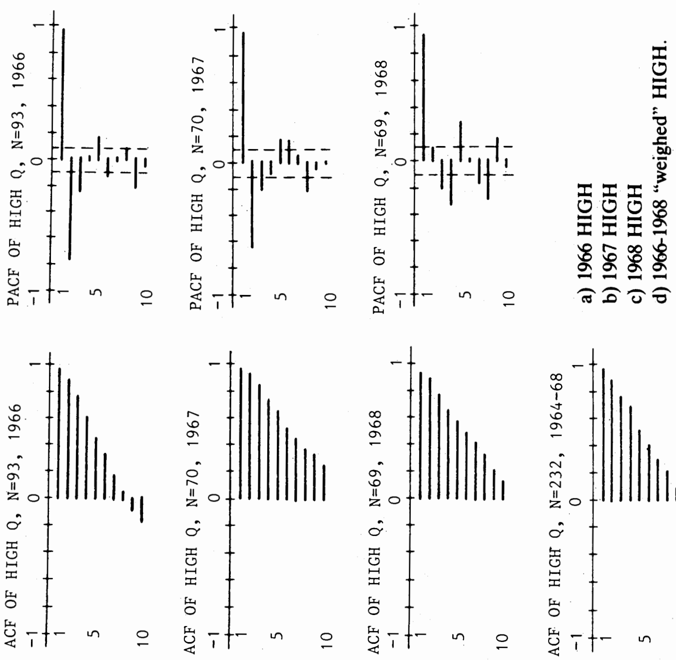

Fig. 8. Autocorrelation function, ACF, and partial autocor relation function, PACF, for discharges less than 40 m^3/s.

Table 4 – Regression coefficient and standard deviation of residuals for the basic conceptual model and for the combined model for different numbers of days forecast. Independent period 1969-72. The Emån catchment.

	Regression coefficient	Standard deviation residual (m^3/s)
Basic conceptual HBV-model	0.859	11.8
Conceptual + AR(1) 1 day ahead forecast	0.990	3.0
Conceptual + AR(1) 3 days ahead forecast	0.959	6.3
Conceptual + AR(1) 5 days ahead forecast	0.926	8.6
Conceptual + AR(1) 7 days ahead forecast	0.900	10.2

shown. As can be seen the numbers are almost the same as for the dependent period (cf. Table 3).

Separation of High and Low Discharges

Considering the error sequence of the conceptual model in Fig. 1 there seems to be a difference in the error propagation of the high discharges and the low discharges. The errors seem to decrease faster for the high discharges than for the low ones. To find out if this is the case the autocorrelations and the partial autocorrelations of the low and high discharges are plotted in Figs. 7 and 8. In general the ACF for the high discharges decreases faster than the one for the low discharges although the difference is not pronounced. An investigation is made to find out if even better fitting can be achieved if a separation of the regression coefficient for high and low discharges, respectively, is made.

Since there is no longer only one, but several series the problem arises how to derive *one* autocorrelation function for the high discharges and *one* for the low discharges. A simple weighing of the different ACF has been performed although what is obtained is not really an ACF. The expression for low discharges becomes

$$\text{"ACF}(K)\text{"}_{\text{TOTAL,LOW}} = \frac{\sum_{i=1}^{n} N_{i,\text{LOW}}\,\text{ACF}(K)_{i,\text{LOW}}}{\sum_{i=1}^{n} N_{i,\text{LOW}}}$$

where i is series number, k is the lag, N_i is the number of observations in series number i, n is the number of series.

The "ACF" for low and high discharges, respectively, obtained in the above mentioned way and the ACF for the total series are plotted in Fig. 9

The "ACF" for high discharges decreases somewhat faster than the one for the total series, and the "ACF" for the low discharges decreases somewhat slower than the one for the total series.

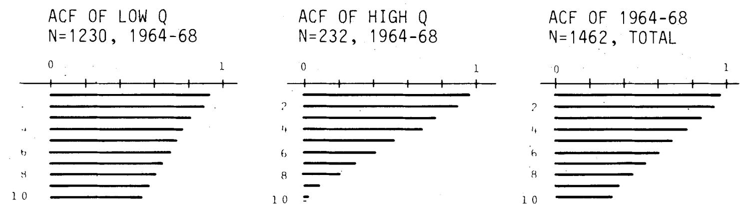

Fig. 9. Autocorrelation function ACF for the total series, for weighed high and weighed low discharges.

The error model for high and low discharges becomes for lags 1,3,5, and 7:

Recorded discharge greater than 40 m³/s	Recorded discharge less than 40 m³/s
$z_{i+} = 0.955 \times z_i + a_i(1)$	$= 0.938 \times z_i + a_i(1)$
$z_{i+} = 0.736 \times z_i + a_i(3)$	$= 0.816 \times z_i + a_i(3)$
$z_{i+} = 0.545 \times z_i + a_i(5)$	$= 0.716 \times z_i + a_i(5)$
$z_{i+} = 0.315 \times z_i + a_i(7)$	$= 0.652 \times z_i + a_i(7)$

The best improvement would be expected at large lags, since the difference in the autocorrelation coefficient is greatest there. In Fig. 10 the results of the seven days ahead forecast is shown. No improvement compared with the original model, where separation of the ACF for different discharges was not done, can be detected (cf. Figs. 3-4). Neither do the statistics of the errors show any improvement.

Conclusions

For a catchment where accurate discharges are available and the response of a rainfall or snowmelt is not very fast, e.g. a moderately large catchment, the short time forecast from a conceptual model can be considerably improved with an autoregressive error model. For the tested area, Emån catchment, the regression coefficient is improved from 0.86 for the original model to 0.99, 0.96, 0.93, and 0.90 for 1, 3, 5, and 7 days ahead forecast, respectively.

Acknowledgement

The research presented here was done while A. Sölløsi-Nagi, VITUKI, Budapest, was visiting WREL. Without his support this work would not have been made.

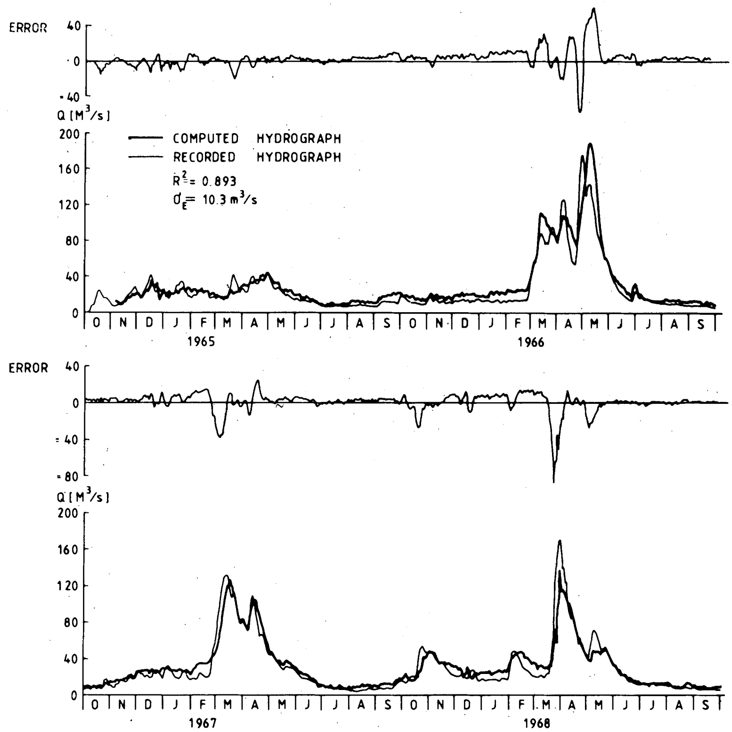

Fig. 10. Discharge from the Emån catchment, observed and calculated. Seven days ahead forecast. Conceptual model combined with an AR(1) with separated correlation-coefficient for high and low discharges. The residuals are plotted above. Regression coefficient ≡ 0.893, residual standard deviation 10.3. Calibration period 1964-1968.

References

Anderson, O. D. (1976) *Time Series Analysis and Forecasting,* The Box-Jenkins Approach. Butterworth & Co, London and Boston.

Bartlett, M. S. (1946) On the theoretical specification of sampling properties of autocorrelated time series. *Journ. Royal Stat. Soc. B8, 27.*

Bergström, S. (1976) Development and application of a conceptual runoff model for Scandinavian catchments, SMHI Nr RHO 7.

Box and Jenkins (1976) *Time Series Analysis, forecasting and control.* Holden Day, San Francisco, USA.

Ryan, T. (1981) Joiner Band Ryan B. Minitab reference manual, Duxbury Press, 20, Providence Street, Boston MA.

Quenouille, M. H. (1949) Approximate tests of correlation in time series, *Journ. Royal Stat. Soc. B11,* 68.

First received: February 5, 1982
Revised version received: June 2, 1982

Address: Div. Water Resources Engineering,
S-951 87 Luleå,
Sweden.