Climate Variability and Change—Hydrological Impacts (Proceedings of the Fifth FRIEND World Conference held at Havana, Cuba, November 2006), IAHS Publ. 308, 2006.

Multi-model technique for low flow forecasting

MARY-JEANNE ADLER & VALENTINA UNGUREANU

National Administration "Apele Romane" – National Institute of Hydrology and Water Management, 97 Sos Bucuresti-Ploiesti, Sector 2, Bucharest, Romania

adler@hidro.ro

Abstract In forecasting low flows, four models, two determinist-statistic models (the VIDRA Model and a modified Stanford Model), a statistical model based on hydrometeorological analogues and an ARIMA class of model, were used to simulate the non-stationary and seasonal nature of the monthly flows. The result obtained was a large range of forecasted values which required a procedure in order to reach a decision on the most appropriate forecast. A solution was found in the development of a packet of procedures including both forecasting models and a decision procedure. This procedure is weight-based and uses the model errors on the last time step in its computing, so that with the help of model-weighting it is possible to incorporate in the procedure those models which best match the actual data (multi-model results optimization procedure).

Key words low flow forecast; multi-model procedure; forecasting decision support

INTRODUCTION

In hydrological forecasting, some of the problems that can occur in real time are a lack of information necessary to the model (problems of data transmission), errors of measurement or errors of representation of the variables, and a particular situation that cannot be simulated with good results by the model.

So the hydrologist has to perform several operations, each with some hazards, and so he needs good software which should include both forecasting models and a decision procedure. Given these tools, the forecaster will have more time to look for supplementary data and make a better analysis of the existing situation. The reduction of subjectivity and a better understanding of the situation will lead to improved forecasting.

MULTI-MODEL DECISION PROCEDURE

A simple description of the problem is given below.

- There are several forecast models available.
- It is presumed that each model checks a group of hypotheses.
- The forecast quality of each model is only guaranteed if it checks these hypotheses.
- In real time it is not possible to check all the hypotheses.
- The relative quality of the results of various models depends on the situation (operating configuration) of the system at the time.

The simple real-time estimating procedure for the weight given to each model presumes that the errors registered on the last time steps allow construction of an operating configuration available for the given moment and different from the normal operating configuration of the system. This procedure uses the model errors on the last time step in its computing so that, with the help of model-weighting, it is possible to incorporate into the procedure those models which best match the actual data (Roche & Tamin, 1986a,b; Roche & Torterotot, 1987).

The formula to compute the weighting is:

$$W_k = (e_k^2)^{-1}/\Sigma_{j=1}^{m}(e_j^2)^{-1}$$

where m is the number of models used for low flow forecast and e_k is the forecast error of the kth model. The sum of all models' weights is $\Sigma_{k=1}^{m} W_k = 1$.

LOW FLOW FORECASTING

The major objective in modelling the hydrological behaviour of a watershed is to simulate its streamflow hydrograph in response to an input of precipitation. To accomplish this, the hydrological cycle is analysed and expressed as a collection of mathematical formulations based on rational parameters that may be adjusted after trial simulations with known input and output. This may be continued until the model is judged to be an adequate representation of the hydrological cycle for a study area – for low flows in our case. The test gauging station was Bals on the Oltet River, in an endemically droughty area of the southeastern part of Romania. The applied models were calibrated using the climatic series of monthly data 1960–1990 and validated on the 1991–2004 period. An excessively droughty period during the vegetation growth period was identified in 1994. Using this critical period, imposing high precision in forecasting low flows, the multi-model methodology was tested.

Deterministic-statistical model

The deterministic-statistical model simulates separately the two main components of the monthly mean discharge: the mean pluvial discharge and the base flow (Q_b).

The base flow is forecast using the recession curve during low flows:

$$Q_b = \frac{1}{T}\int_0^T Q_0 \times \mathrm{e}^{-\alpha\times t}dt = \frac{Q_0}{\alpha\times T}\left(1-\mathrm{e}^{-\alpha\times T}\right)$$

where T is the number of days of the considered month, Q_0 is the initial value at the beginning of each month and α the recession coefficient. The α parameter is a function depending on Q_0:

$$\alpha = \frac{\ln Q_0 - \ln Q(t)}{t}$$

For each basin the $\alpha = f(Q_0)$ relations are calculated for each month or groups of months. In this case, two distinct cases were found, see Table 1.

Table 1 The recession parameters for each characteristic interval.

Q_o (m^3 s^{-1}) Month	a (10^{-2} day^{-1}) VI–VIII	IX–XI	Q_o (m^3 s^{-1}) Month:	a (10^{-2} day^{-1}) VI–VIII	IX–XI
1	5.6	3.2	6	8.1	5.1
2	6.3	3.9	7	8.3	5.3
3	6.9	4.5	8	8.6	5.4
4	7.3	4.7	9	8.8	-
5	7.8	4.9	10	8.9	-

For a forecasting condition Q_0 is known and the coefficient α is determined from Fig. 1 and using the equation for Q_b, the monthly mean discharge of the base flow is calculated.

In order to determine the contribution of precipitation to the mean discharge (Q_p) a multiple regression was used for each month giving: $Q_P = f(Q_0,P)$; Fig. 1(a) and (b).

A long-term meteorological forecast is usually a qualitative one (excessively droughty period, drought, normal droughty period, and similar for rainy periods. The transformation into numeric values was obtained using the frequency of the mean precipitation in the basin, and associating a probabilistic interval to each forecast indication:

- excessive drought p = 90–95%
- drought p = 80–90%
- normal drought p = 60–80%
- normal rainy p = 20–40%
- rainy p = 10–20%
- normal rainy p = 5–10%

The rainy interval is obtained for each meteorological forecast from the probability curve of each basin; an example is given for Bals - Table 2.

Finally, the forecast discharge (Q_F):

$$Q_F = Q_b + Q_P$$

is calculated taking the superposition of the effects of the mean monthly discharge into consideration. An application of this model is presented in Table 3.

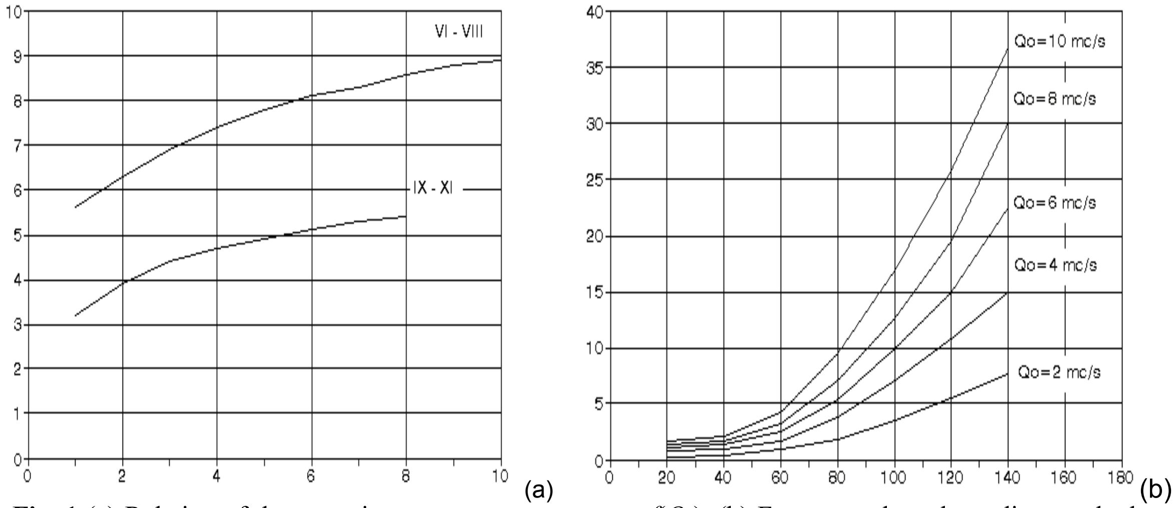

Fig. 1 (a) Relation of the recession curve parameters, $a = f(Q_0)$. (b) Forecast values depending on the base discharge and precipitation $Q_p = f(Q_0,P)$ (Bals application).

Table 2 Monthly precipitation for Bals.

No.	P (%)	VI	VII	VIII	IX	X	XI
1	10	155	130	117	100	100	108
2	20	135	104	88	76	63	86.5
3	50	94	65	50	39	42	55
4	80	60	38	27	35	19.5	29
5	90	45	28	17.8	6.4	9	19.5

Table 3 The mean monthly forecasted discharges during the summer-autumn period for the excessively droughty year 1994.

Month	P (mm)	Q_o ($m^3 s^{-1}$)	a ($10^{-2} zi^{-1}$)	Q_b ($m^3 s^{-1}$)	Q_P ($m^3 s^{-1}$)	Q_F ($m^3 s^{-1}$)	Q_{obs} ($m^3 s^{-1}$)
VI	20.0	3.36	7.5	1.03	2.0	3.03	3.61
VII	17.1	3.36	6.9	1.86	1.5	3.36	3.18
VIII	59.8	1.86	6.9	0.784	0.25	1.09	1.19
IX	4.00	0.03	3.0	0.020	0.20	0.22	0.474
X	68.4	1.86	6.9	0.784	1.00	1.78	2.98
XI	7.84	1.70	4.5	0.820	0.20	1.02	1.71

Model using hydrometeorological analogues

The forecasting model based on the hydrometeorological analogues presumes the follow stages:

- the definition of the characteristic scale, of the quantitative and/or qualitative variables, determining the analogy with an anterior stage;
- the definition of the position in the space of characteristics in the forecasting moment;
- the search of an analogue position using the theory of shapes.

The forecasting algorithm estimates two tendencies:

- a regime tendency using the hydrological information only,
- a regime tendency using the information from the analogue years from the meteorological point of view furnished by the meteorological prevision (Matreata, 1997).

The first tendency (C_h) is obtained using the percentage from the mean multi-year discharge (Q_{mean}) of the forecasted discharge (Q_F): $C_h = Q_F^j/Q_{mean}^j$.

The mean monthly forecast discharge Q^j_F is obtained from the annual hydrograph of droughty type (dimensioned by the mean annual discharge: Q^j_{mean}) or excessively droughty type for the physico-geographical space of interest and re-dimensioned using the mean annual forecast discharge Q^j_{meanF}, Fig. 2.

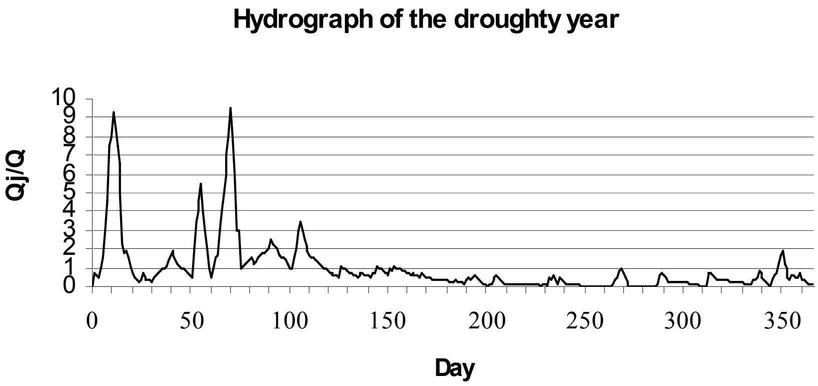

Fig. 2 Hydrograph of the droughty year for the Oltet Basin (regional analysis).

The estimation of the second tendency is based on the evolution of the mean monthly discharges from the meteorological analogue year ($N_{analogs}$), for each month, using the relation:

$$C_m = \frac{\dfrac{1}{N_{analogs}\sum_{i=1}^{N} Q_i^j}}{Q_{mean}^j}$$

where Q^j_i is the monthly mean discharge of the *j*th month from the *i*th analogue year from the meteorological point of view. The droughty analogue characteristic years were determined using the deciles method.

The final forecast is obtained using the equation:

$$Q_F = (0.6C_h + 0.4C_m)\cdot Q_{mean}^j$$

An example of such a calculation is presented below. The mean annual discharge forecast using ARMA (1,1) was 2.86 $m^3 s^{-1}$ and the computed monthly low flows are presented in Table 4.

Table 4 The forecasted low flows using the hydrometeorological analogue model.

Year	Month	Q^j_{analog}	Q^j_{mean}	Cm	Q^j_F	C_h	Q_F	Q_{obs}
1994	VI	5.77	4.53	0.676	4.03	0.473	3.12	3.61
	VII	1.66	4.28	0.200	1.23	0.148	3.80	3.18
	VIII	1.02	3.38	0.302	0.532	0.157	0.727	1.19
	IX	0.922	2.02	0.456	0.406	0.201	0.624	0.474
	X	1.67	2.06	0.811	0.669	0.325	1.50	2.98
	XI	2.30	3.66	0.628	0.947	0.259	1.49	1.71

Stochastic models used for the mean monthly discharges during low flow period

The non-stationary and seasonal nature of the monthly flows is modelled using the ARIMA class of models. A step-by-step procedure was used to obtain valid models through proper model identification, parameter estimation, performance evaluation, model parsimony and validation of the residuals. In building the model it was especially important to obtain the parsimonious models with white residuals. The procedure was applied to the monthly flows of the River Oltet at Bals and a valid model capable of acceptable prediction results was obtained: ARIMA (2,1,1). The forecast monthly mean discharges (Q_F) and the observed ones (Q_{obs}) are presented in Table 5.

Deterministic model for the mean monthly low flows forecast

The model used was VIDRA, a reservoir model similar to the Stanford Watershed Simulation Model (Crawford & Linsley, 1966; Corbus & Adler, 2004), a mathematical model programmed for a PC, synthesizing a continuous hydrograph (watershed outflow *vs* time) of streamflow from climatic data (precipitation and evaporation) and watershed parameters (soil surface moisture and

Table 5 The mean monthly forecasted discharges using the ARMA Model.

Year	Month	Q_F (m^3 s^{-1})	Q_{obs} (m^3 s^{-1})
1994	VI	3.40	3.80
	VII	3.13	3.26
	VIII	1.25	1.26
	IX	0.65	0.550
	X	2.80	3.20
	XI	1.98	2.02

retention properties, interflow storage and flow conditions, groundwater storage and flow conditions, and the physical state and geomorphic properties of the basin).

The computer program was carefully studied, flow diagrammed in detail and an adequate calibration for low flows was obtained, using mainly the facilities of the multiple recession constants: CB, infiltration index; LZSN, soil moisture storage index; UZSN, soil surface moisture index; K24L, parameter indicating groundwater flow leaving the basin; KK24, daily base recession constant; and GWF, base flow, determined for low flow period calibration of the model.

The 1990–1993 period was used for calibration and the low flow period June–November 1994 was forecast (Table 6).

Table 6 The mean monthly forecasted discharges using the Stanford Model.

Year	Month	Q_F (m^3 s^{-1})	Q_{obs} (m^3 s^{-1})
1994	VI	3.80	3.50
	VII	3.26	3.20
	VIII	1.26	1.50
	IX	0.55	0.900
	X	3.20	3.00
	XI	2.02	1.59

RESULTS FOR THE MULTI-MODEL REAL-TIME DECISION PROCEDURE IN THE HYDROLOGICAL FORECAST

The multi-model procedure for low flow forecasts was applied for the four models presented. As the forecast errors of the four models for May are unknown, the value for June, the first month of the forecasted interval, was considered to be the same for all models (e_k = 1.00, where k = 1,2,3,4). In this case, the weight of each k model was calculated with the weight formula w_k = (1/(1+1+1+1)) = ¼ = 0.25 and the forecast mean monthly discharge for June is:

$$Q_F^{\mathrm{VI}} = 3.03 \times 0.25 + 3.12 \times 0.25 + 3.80 \times 0.25 + 3.50 \times 0.25 = 3.36$$

In July, the forecast errors of the four models for June was calculated and those values have been used to compute the weight of each model as in the following example: measured discharge in June – 3.61 m^3 s^{-1}, deterministic-statistical model – 3.03 m^3 s^{-1}, hydrometeorological analogues method – 3.12 m^3 s^{-1}, stochastic model ARIMA – 3.80 m^3 s^{-1}, Stanford model – 3.50. The forecast errors in June were as follows: –16.066%, –13.573%, 5.263%, –3.047%. The weights of the models computed for July were:

$$w_1 = [(-16.066)^2]^{-1}/\{[(-16.066)^2]^{-1} + [(-13.573)^2]^{-1} + [(5.263)^2]^{-1} + [(-3.047)^2]^{-1}\} = 0.0253$$

and similarly the rest of the weights were obtained, w_2 = 0.0354, w_3 = 0.2358, w_4 = 0.7035. The forecasted value using the multi-model procedure for July, is:

$$Q_F^{\mathrm{VII}} = 3.36 \times 0.0253 + 3.80 \times 0.0354 + 3.26 \times 0.2358 + 3.20 \times 0.0253 = 3.24.$$

A similar procedure was applied for August, September, October and November. The results are presented in Table 7 and in Figs 3 and 4.

The results obtained show clearly the usefulness of the application of the multi-model procedure in providing an increase in the precision of forecast values.

Table 7 Observed discharge, calculated discharges (Q_C), forecast values (Q_F) and errors (*Er*) provided by each of the four models and by the multi-model procedure, for the six month period, June–November 1994.

Month	Observed discharge	Determinist-statistic model (1)		Hydromet. analogues method (2)		Stochastic model ARMA type (3)		Stanford type model (4)		Multi-model procedure (mm) (5)	
	($m^3 s^{-1}$)	Q_C ($m^3 s^{-1}$)	*Er* (%)	Q_C ($m^3 s^{-1}$)	*Er* (%)	Q_C ($m^3 s^{-1}$)	*Er (%)*	Q_C ($m^3 s^{-1}$)	*Er* (%)	Q_F ($m^3 s^{-1}$)	*Er* (%)
VI	3.61	3.03	–16.07	3.12	–13.57	3.80	5.26	3.50	–3.05	3.36	–6.86
VII	3.18	3.36	5.66	3.8	19.50	3.26	2.52	3.20	0.63	3.41	7.08
VIII	1.19	1.09	–8.40	0.727	–38.91	1.26	5.88	1.50	26.05	1.48	24.42
IX	0.474	0.220	–53.59	0.624	31.65	0.550	16.03	0.900	89.87	0.46	–3.14
X	2.98	1.78	–40.27	2.5	–16.11	3.20	7.38	3.00	0.67	2.79	–6.50
XI	1.71	1.02	–40.35	1.49	–12.87	2.02	18.13	1.59	–7.02	1.59	–6.82

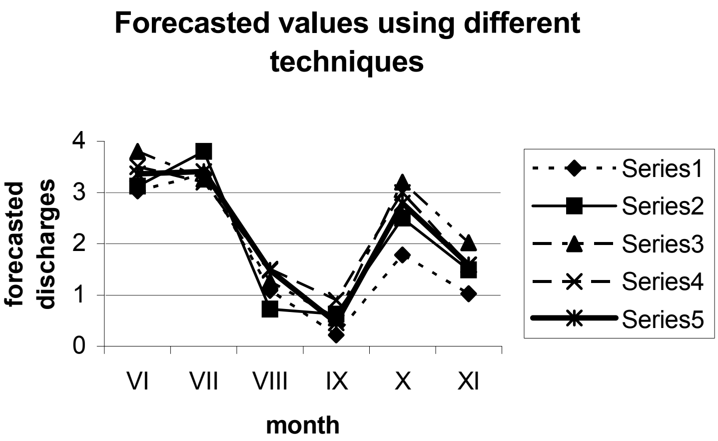

Fig. 3 Forecast of discharges using four mathematical models and the multi-model procedure.

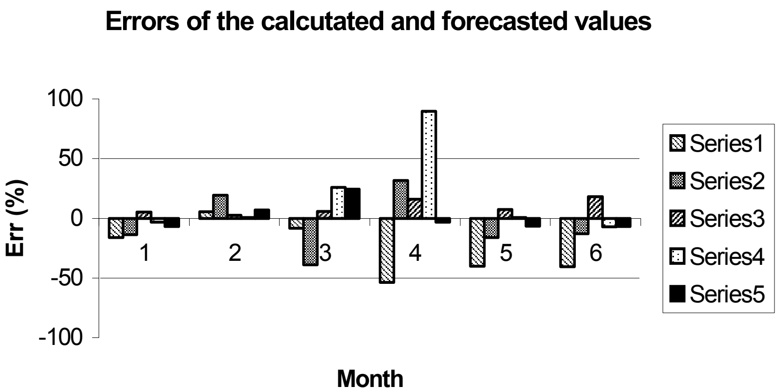

Fig. 4 Representation of the forecasting errors of the four mathematical models and of the multi-model procedure.

CONCLUSIONS

Four models were used for low flow forecasting for the extreme droughty period June–November 1994. The models which best match the data are the stochastical model and the Stanford model. The multi-model approach ameliorates the forecast; in these five months the forecasted errors are less then 10%. The multi-model procedure is very efficient. The weights of the model which had provided the worst result in a given month diminish the influence of that forecasted value on the following month, while for those giving better results, the weights increase their contribution to the forecasted value for the following month. The multi-model procedure increased the performance of the forecasts and it is for this reason that we recommend it for application in operational activity.

REFERENCES

Crawford, N. H. & Linsley, R. K. (1966) *Digital simulation in hydrology: Stanford Watershed Model IV*. Technical Report 39, Department of Civil Engineering, Stanford University, Palo Alto, California, USA.

Cavadias, G. & Morin, G. (1985) Amélioration des performances des modèles hydrologiques par combinaison des débits simulés. *Revue internationale des sciences de l'eau* **1**(1/4), 12–28.

Makridakis, S. & Wincler, R. L. (1983) Average of forecast: some empirical results. *Manage. Sci*. **29**, 42–46.

Matreata, M. (1997) Dinamico-statistic model for low flows. *Studii de hidrologie*, INMH, 24–37.

Newbold, P. & Granger, C. W. J. (1974) Experience with forecasting univariate time series and the combination of forecasts. *J. Royal Statist. Soc*. A **137**, 57–69.

Roche, P. A. & Tamin, R.(1986) Procédures de décision en temps réel pour la prévision des crues. CERGRENE.

Roche, P.A., Torterotot, J. P. (1987) Prévision des crues sur la Haute-Garonne. Utilisation d'une procédure Multi-Modèles. Méthodologie et critique des données. CERGRENE.