Dissolved Loads of Rivers and Surface Water Quantity/Quality Relationships
(Proceedings of the Hamburg Symposium, August 1983). IAHS Publ. no. 141.

Real time water quality forecasting models based on the water quantity/quality relationship

J. L. MARIVOET
Laboratory of Hydrology, Free University Brussels (VUB), Brussels, Belgium

ABSTRACT For the real time management of a safety reservoir used in drinking water supply, short term forecasts of the water quality in the River Meuse are needed. In this paper four methods for forecasting fluoride and chloride concentrations are compared. The first model consists of a simple ARMA model forecasting future concentrations from the last measured concentrations. The second and third methods are based on dilution relationships. A transfer function-noise model, with the reciprocal of the river flow as input and the concentration as output, is proposed as a fourth method. A simple rainfall-runoff model is introduced to forecast future runoffs. The transfer function-noise model yields the best results for short term forecasts and the dilution models for long term forecasts.

Modèles de prévision en temps réel de la qualité de l'eau basés sur la relation quantité/qualité
RESUME Pour la gestion en temps réel d'un réservoir de sécurité pour la distribution d'eau potable, nous avons besoin de prévisions de la qualité de l'eau de la Meuse. Quatre méthodes pour la prévision de la concentration en fluorure et en chlorure sont comparées. Le premier modèle est un modèle ARMA permettant d'évaluer les concentrations futures à partir des dernières concentrations mesurées. La deuxième et troisième méthode sont basées sur une relation de dilution. Le quatrième modèle est un modèle "fonction de transfert" utilisant la réciproque de débit comme donnée d'entrée et la concentration comme donnée de sortie. Un modèle simple "pluie-débit" est introduit pour la prévision des débits. Le modèle "fonction de transfert" donne les meilleurs résultats de prévision à court terme et les modèles de dilution à long terme.

INTRODUCTION

In order to produce drinking water for the metropolis of Antwerp, the Antwerp Water Works take water from a canal system that connects the River Meuse to Antwerp. The water can flow by different canals from the River Meuse to Antwerp. The River Meuse is polluted by waste waters from chemical plants and by the untreated sewage from some large cities. In periods of low flow conditions the river water no longer satisfies the European standards for drinking water (EEC directive No.75/440) and on some occasions the fluoride concentrations exceed the 1.5 mg l^{-1} level.

A safety reservoir has been constructed to avoid violations of drinking water quality standards and a second reservoir is planned near the water intake points. The capacity of the reservoirs will allow the needs of the water treatment plants to be covered for a period of three weeks. For the real time management of these safety reservoirs (Marivoet & Van Craenenbroeck, 1983), a mathematical model describing the transport and the major processes taking place in the canal system has been constructed and calibrated (De Smedt *et al.*, 1981; Marivoet, 1983). If the concentrations in the River Meuse at the entrance of the canals are known, this model allows prediction of the concentration at the water intake points. In order to extend the forecasting period of this water quality model, forecasts of the future concentrations in the River Meuse are needed.

THE ARMA MODEL

ARMA models, as proposed by Box & Jenkins (1976), enable future values of a time series to be easily forecast. From the analysis of the autocorrelation function and the partial autocorrelation function it is concluded that a first order autoregressive model can be used to model the time series of daily fluoride and chloride concentrations. This model is described by the formula:

$$\tilde{C}_t = \phi_1 \tilde{C}_{t-1} + a_t \tag{1}$$

where

$$\tilde{C}_t = C_t - C_o \tag{2}$$

and ϕ_1 is the autoregressive parameter, a_t is a noise term and C_o is the background concentration. Box & Jenkins (1976) define C_t as the deviation from the mean concentration. Since the mean concentration has no physical meaning, the use of deviations from the background or "natural" concentrations of the investigated solute species is preferred in the present study. These background concentrations are estimated from concentration measurements at extremely high river flow.

Runoff and concentration data for the period May 1979-October 1980 have been used for model calibration, and the calibrated parameters with their 95% confidence limits and the estimated standard deviation s_a are given in Table 1.

THE DILUTION MODELS

A simple deterministic model can be built up by considering the river concentration as the sum of a natural concentration C_o and a concentration resulting from the dilution of the river load by the river flow. The concentration C_t is now described by the equation:

$$C_t = C_o + L/Q_{t-b} \tag{3}$$

TABLE 1 Calibrated parameters for the ARMA model (equation 1)

Parameter	*Fluoride*	*Chloride*
C_o (mg l^{-1})	0.13	15.9
ϕ_1	0.981 ± 0.017	0.991 ± 0.013
s_a (mg l^{-1})	0.191	7.80

where L is the river load for the studied solute species and b is time delay or lag. A time delay between river flow and concentration is considered in order to take into account the travel time from the injection to the measurement point. Time delays of 3 days for fluoride and 4 days for chloride have been estimated from the cross-correlation function between the reciprocal of the river flow and the concentration. Equation (3) is calibrated by linear regression and the results are summarized in Table 2.

TABLE 2 Calibrated parameters for equation (3)

Parameter	*Fluoride*	*Chloride*
b (days)	3	4
C_o (mg l^{-1})	0.113 ± 0.050	23.2 ± 2.0
L (g s^{-1})	86.0 ± 5.4	3139 ± 221
s (mg l^{-1})	0.330	13.4

Since this time delay varies with the river flow, a second model has been developed. This model assumes that the injection of the load takes place at a fixed point, which may be different for every solute species, and that the volume V of the river segment between the injection point and the measuring point is constant. This last assumption can be justified for a wide range of river flows since the River Meuse is canalized and the locks are operated to maintain a constant level.

The time delay T_t is estimated as:

$$T_t = V/Q_t \tag{4}$$

and a linear interpolation:

$$C_t = C_o + (1 - d_t)\, L/Q_{t-b_t} + d_t\, L/Q_{t-b_t-1} \tag{5}$$

is applied where

$$b_t = \text{greatest integer} \leq T_t \tag{6}$$

$$d_t = T_t - b_t \tag{7}$$

TABLE 3 Calibrated parameters for equation (5)

Parameter	Fluoride	Chloride
V ($10^6 m^3$)	17.97 ± 1.08	30.84 ± 1.31
C_o ($mg\ l^{-1}$)	0.090 ± 0.047	23.93 ± 2.11
L ($g\ s^{-1}$)	90.3 ± 5.1	3106 ± 227
s ($mg\ l^{-1}$)	0.314	13.7

This model has three parameters V, C_o and L. In order to calibrate these parameters, the objective function S_1:

$$S_1 = \Sigma_{t=T_o}^{T_1} N_t^2 \tag{8}$$

is optimized where:

$$N_t = C_t - \hat{C}_t \tag{9}$$

and T_o is the start of the measurement period, T_1 is the end of the measurement period, C_t is measured concentration and $\hat{C}_t$ is calculated concentration. The results of the calibration are given in Table 3. This model yields better results for fluoride but in the case of chloride the results are comparable with those from equation (3).

Improvement of the forecasts obtained with the model is possible using a noise model for the residuals. From the analysis of the autocorrelation function and the partial autocorrelation function of the residuals N_t it is concluded that a first order autoregressive process can be used to model the residuals. This gives:

$$N_t = \psi_1 N_{t-1} + a_t \tag{10}$$

where ψ_1 is the autoregressive parameter and a_t is a white noise term. The model is now formed by equations (4-7), (9) and (10). Calibration of this model is achieved by optimizing the objective function S_2:

$$S_2 = \Sigma_{t=T_o}^{T_1} a_t^2 \tag{11}$$

The results of this calibration are given in Table 4.

TABLE 4 Calibrated parameters for the dilution model (5) with a noise model (10)

Parameter	Fluoride	Chloride
V ($10^6 m^3$)	19.62 ± 0.61	32.14 ± 0.97
C_o ($mg\ l^{-1}$)	0.113	23.2
L ($g\ s^{-1}$)	71.4 ± 6.8	2383 ± 295
ψ_1	0.852 ± 0.047	0.848 ± 0.051
s_a ($mg\ l^{-1}$)	0.187	8.72

THE TRANSFER FUNCTION MODEL

The combined transfer function-noise model, as proposed by Box & Jenkins (1976), can be used to model the concentrations. Model input is taken as the reciprocal X_t of the river flow:

$$X_t = 1/Q_t \tag{12}$$

The transfer function can be written as:

$$\tilde{C}_t = \Delta^{-1}(z)\ \Omega(z)\ X_{t-b} + N_t \tag{13}$$

where

$$\Delta(z) = 1 - \delta_1 z - \delta_2 z^2 - \dots - \delta_r z^r$$

$$\Omega(z) = \omega_1 z - \omega_2 z^2 - \dots - \omega_s z^s$$

and z is a backward shift operator. The noise model can be written as:

$$N_t = \Phi^{-1}(z)\ \Theta(z)\ a_t \tag{14}$$

where

$$\Phi(z) = 1 - \psi_1 z - \psi_2 z^2 - \dots - \psi_p z^p$$

$$\Theta(z) = 1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_q z^q$$

The model identification, based on the prewhitening method, shows the following model structure for both fluoride and chloride:

$$(r, s, b) = (1, 1, 0)$$

and

$$(p, q) = (1, 0)$$

The model calibration, however, indicated that some model parameters are not significantly different from zero. The transfer function model has therefore been simplified to a (1, 0, 0) model for fluoride and to a (1, 0, 1) model for chloride. The results of the model calibrations are given in Table 5.

RIVER FLOW FORECASTING MODELS

In the former models the river flow is used as a leading indicator, but rainfall is now introduced as a leading indicator for the river flow in order to improve the short term forecasts. Rainfall measured at Rochefort is used as model input, and a combined transfer function-noise model is to forecast river flows (Ledolter, 1978;

TABLE 5 Calibrated parameters for the combined transfer function-noise model

Parameter	*Fluoride*	*Chloride*
b *(days)*	*0*	*1*
δ_1	*0.8456 ± 0.0758*	*0.8724 ± 0.0625*
ω_o $(g\ s^{-1})$	*13.35 ± 6.40*	*462 ± 194*
ψ_1	*0.8242 ± 0.0487*	*0.7956 ± 0.0545*
C_o $(mg\ l^{-1})$	*0.113*	*23.2*
s_a $(mg\ l^{-1})$	*0.182*	*7.35*

Anselmo & Ubertini, 1979). This model becomes:

$$\tilde{Q}_t = \Delta^{-1}(z)\ \Omega(z)\ P_{t-b} + N_t \tag{15}$$

where P_t is rainfall. The model identification shows the following model structure:

$$(r, s, b) = (1, 1, 1)$$

and

$$(p, q) = (2, 0)$$

The response of a river basin to rainfall depends on the antecedent precipitation and, therefore, the use of an antecedent precipitation index (API) in modelling has been proposed by Kohler & Linsley (1951). API_t is defined as:

$$API_t = \alpha API_{t-1} + P_t/100 \tag{16}$$

where α is an autoregressive parameter. A gain coefficient G_t is given by:

$$G_t = 1 - \exp\ (-API_{t-1}) \tag{17}$$

The model now becomes:

$$\tilde{Q}_t = \delta_1\tilde{Q}_{t-1} + \omega_o G_t P_{t-1} - \omega_1 G_t P_{t-2} + N_t \tag{18}$$

with a first order autoregressive noise model. The results of the calibration of these two rainfall-runoff models are given in Table 6. In further calculations model (18) is preferred, since it gives better results than model (15). If no rainfall information is available, an ARMA model may be used to forecast future river flows. The analysis of the autocorrelation function and the partial autocorrelation function indicates that a third order autoregressive model should be used in this case.

TABLE 6 Calibrated parameters for the rainfall-runoff models

Parameter	*Model (15)*	*Model (18)*
δ_1	0.750 ± 0.054	0.635 ± 0.048
$\omega_0 (10^3 m^2 s^{-1})$	12.63 ± 1.42	25.88 ± 4.03
$\omega_1 (10^3 m^2 s^{-1})$	-10.58 ± 1.50	-25.54 ± 4.78
ψ_1	1.221 ± 0.100	0.9793 ± 0.0211
ψ_2	-0.282 ± 0.101	-
α	-	0.9065 ± 0.0394
$Q_0 (m^3 s^{-1})$	11	11
$s_a (m^3 s^{-1})$	63.5	53.4

APPLICATION OF THE MODELS

The different calibrated models were applied to forecast the concentrations for the period November-December 1980. This test period shows a very typical autumn pattern. At the end of the summer drought the river flow is very low and the concentrations are high. The autumn storms produce a series of hydrographs and the baseflow increases. The performance of the different models are compared by the following parameter:

$$R^2 = (F_o - F_1)/F_o$$

where

$$F_o = \Sigma_{t=T_o}^{T_1} C_i^2$$

and

$$F_1 = \Sigma_{t=T_o}^{T_1} (C_i - \hat{C}_i)^2$$

The parameters R^2 obtained with the different forecasting models for the test period are given in Tables 7 and 8 for different step-ahead forecasts. Runoff is forecasted assuming future rainfalls to be zero. If rainfall information is not available the ARMA model is used to forecast future runoffs. The measured flows and concentrations together with their three days ahead forecasts are shown in Figs 1-3.

DISCUSSION

The results obtained for the test period show that the ARMA model gives good forecasting results for very short forecasting periods. A great advantage of this model is its simplicity. The deterministic dilution models, without noise components, give inferior results for the first two step-ahead forecasts but they give the best results for longer forecasting periods. The addition of a noise model

TABLE 7 Parameters R^2 and s_a obtained for fluoride with different forecasting models for the test period

Runoff model	Quality model		Step-ahead forecasts (days): 1	2	3	4	5
	(1)	R^2	0.940	0.870	0.812	0.783	0.735
		s_a	0.146	0.216	0.261	0.282	0.314
(18)	(3)	R^2	0.898	0.895	0.893	0.881	0.841
		s_a	0.190	0.194	0.198	0.209	0.244
	(5) + (10)	R^2	0.923	0.916	0.878	0.860	0.840
		s_a	0.165	0.174	0.210	0.227	0.244
	(13) + (14)	R^2	0.950	0.912	0.874	0.829	0.755
		s_a	0.133	0.178	0.214	0.250	0.302
ARMA	(5) + (10)	R^2	0.923	0.907	0.852	0.842	0.811
		s_a	0.165	0.183	0.232	0.241	0.265
	(13) + (14)	R^2	0.940	0.872	0.809	0.755	0.695
		s_a	0.145	0.214	0.262	0.300	0.337

TABLE 8 Parameters R^2 and s_a obtained for chloride with the different forecasting models for the test period

Runoff model	Quality model		Step-ahead forecasts (days): 1	2	3	4	5
	(1)	R^2	0.977	0.943	0.916	0.897	0.885
		s_a	6.58	10.51	12.85	14.42	15.37
(18)	(3)	R^2	0.956	0.956	0.955	0.954	0.950
		s_a	9.10	9.25	9.41	9.57	10.15
	(5) + (10)	R^2	0.972	0.960	0.953	0.943	0.936
		s_a	7.29	8.76	9.67	10.68	11.42
	(13) + (14)	R^2	0.982	0.964	0.957	0.953	0.944
		s_a	5.87	8.37	9.17	9.75	10.71
ARMA	(5) + (10)	R^2	0.972	0.960	0.953	0.940	0.934
		s_a	7.29	8.76	9.66	10.95	11.67
	(13) + (14)	R^2	0.982	0.965	0.959	0.950	0.939
		s_a	5.87	8.21	9.03	10.00	11.21

improves the quality of the short term forecasts.

Excellent short term results are obtained with the combined transfer function-noise model. The proposed model is not a pure black box model but a grey box model, because the model input of river runoff is for deterministic reasons transformed to its reciprocal. This model is the best adapted to describe the decrease of the concentration after a rapid rise of the river flow.

The ARMA runoff model gives in most cases results comparable with

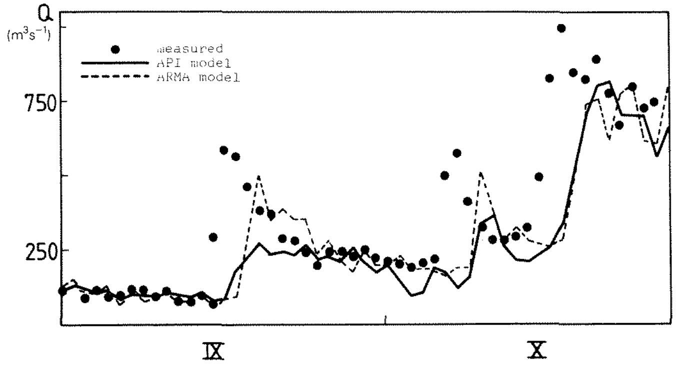

FIG.1 Measured flows and their three-day-ahead forecasts (test period).

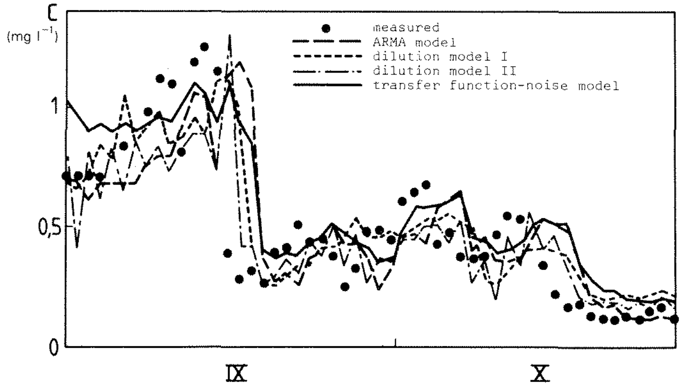

FIG.2 Measured fluoride concentrations and their three-day-ahead forecasts (test period).

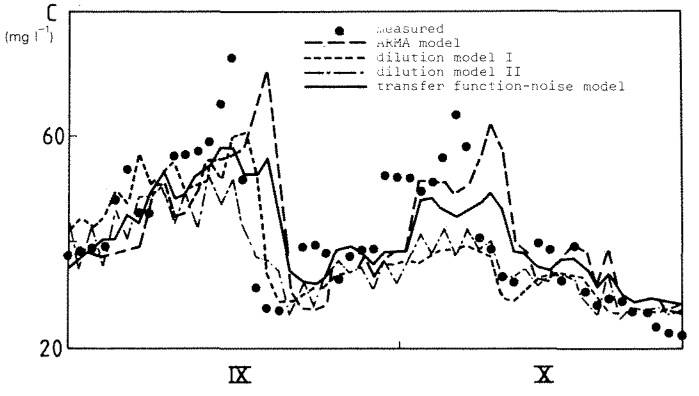

FIG.3 Measured chloride concentrations and their three-day-ahead forecasts (test period).

the rainfall-runoff model, but after intensive rainfalls the latter model permits better forecasts of the hydrographs.

CONCLUSIONS

The combined transfer function-noise model, with the reciprocal of the runoff as model input, gives satisfactory results for short term forecasts of fluoride and chloride concentrations. For long term forecasts dilution models are preferred. If rainfall information is available a simple rainfall-runoff model enables improvement in the performance of the forecasting model.

ACKNOWLEDGEMENTS This study is supported by the I.V.Antwerpse Waterwerken NV. The author is grateful to Prof.Dr A.Van der Beken for supervising this work and to Dr W.Van Craenenbroeck for many fruitful discussions. The help of the Rijksinstituut voor Zuivering van Afvalwater (The Netherlands) in providing the concentration data is highly appreciated.

REFERENCES

Anselmo, V. & Ubertini, L. (1979) Transfer function-noise model applied to flow forecasting. *Hydrol. Sci. Bull.* 24 (3), 353-359.

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

De Smedt, F., Ideler, W. & Van der Beken, A. (1981) A water quality transport model for a channel network of Northern Belgium. In: *Proc. Internat. Conf. Numerical Modelling of River, Channel and Overland Flow for Water Resources and Environmental Applications* (Bratislava May 1981). Section 2.2, 1-12.

Kohler, M.A. & Linsley, R.K. (1951) Predicting the runoff from storm rainfall *US Weather Bureau, Research Pap. 34.*

Ledolter, J. (1978) The analysis of multivariate time series applied to problems in hydrology. *J. Hydrol.* 36, 327-352.

Marivoet, J.L. (1983) A water quality model for a canal system in Northern Belgium. (in preparation).

Marivoet, J.L. & Van Craenenbroeck, W. (1983) The use of a water quality forecasting model for the control of drinking water intake points. In: *Security in Water Supply* (IWSA Conf., Brussels, September 1983).