PERGAMON

Control Engineering Practice 7 (1999) 1489–1504

www.elsevier.com/locate/conengprac

Identification of the rainfall–runoff relationship in urban drainage networks

Fabio Previdi[a,*], Marco Lovera[a], Stefano Mambretti[b]

[a]*Dipartimento di Elettronica & Informazione, Politecnico di Milano, Via G. Ponzio, 34/5 – 20133 Milano, Italy*

[b]*Dipartimento di Ingegneria Idraulica, Ambientale e del Rilevamento, Politecnico di Milano, p.zza Leonardo da Vinci, 32 – 20133 Milano, Italy*

Received 14 May 1998; accepted 9 July 1999

Abstract

The calibration of conceptual models for the design of urban drainage networks is an important and well-known problem in hydraulic engineering. In this paper the problem is analysed and the use of black-box identification methods is proposed and applied to experimental data. Both linear (ARX and state space) and nonlinear (polynomial and neural NARX) models are considered and their performance in the simulation and prediction of the network flow from rainfall measurements is evaluated.

Keywords: System identification; Urban drainage systems; Prediction error methods; Subspace methods; Nonlinearity tests; Polynomial models; Neural networks

1. Introduction and problem statement

Black-box modelling of the rainfall-runoff relationship in urban drainage systems is not a new challenge for system identification (Boukhris, Giuliani, Mourot, Querelle & Frank, 1997; Keesman & Jakeman, 1997). In fact, the knowledge of the rainfall–runoff relationship is an essential tool in modelling and design of urban drainage networks (Harremoes, Jensen & Johansen, 1984; Dooge, 1977). Moreover, as increasing urbanization makes the flow in drainage networks larger and larger and urban catchments are usually designed to convey only medium intensity rainfall, it can be useful to perform fault detection and supervision and to design control systems for real-time storm effect management (Frank, 1990; Szafnicki & Graillot, 1996). In order to perform effective management of urban drainage networks, accurate models are needed, which must be capable of simulating and predicting the water flow in the main sections of the network. Very often attention is focused on quite large hydrologic basins often involving lakes and rivers, far away from hydraulic urban engineering interests. On the other hand, the identification procedures proposed in this paper are tailored on small urban catchments where the submerged component of the flow is not relevant to the model accuracy, and often the rain spatial distribution is also negligible.

In hydraulic practice the rainfall–runoff relationship is described through physical-based models which are completely characterised by a set of design parameters. Good estimates of parameters for these models should be obtained through long-term surveys on existing networks, which would be able to provide reliable parameter values in order to design networks on the basis of a previously assigned "design risk". The design risk is defined on the basis of the network maximum acceptable peak discharge Q as a function of the so-called "return period", i.e. the mean time interval T_r which must pass between the occurrence of two flow peaks of intensity Q (Chow, Maidment & Mays, 1988). Unfortunately, this approach is not always easily feasible, for three main reasons:

- It is expensive because of the required instrumentation and the need for good and frequent maintenance of the data acquisition set-up.
- If a design of a new sewer network is required, nothing but rainfall can be measured and a suitable model from measures on similar networks must be at hand.

* Corresponding author. Tel.: + 39-223-993650; fax: + 39-223-993412.

E-mail address: previdi@elet.polimi.it (F. Previdi)

0967-0661/99/$-see front matter
PII: S0967-0661(99)00118-5

- When the modelling aim is the rehabilitation of an existing sewer network, survey results may be not homogeneous because of changes that have occurred in the network due to urbanization.

Therefore, in network design simple models are used in order to obtain rainfall–runoff relationship models on the basis of a restricted data set. Usually, one of two ways is chosen (Boukhris et al., 1997; Dooge, 1977):

1. Representing the overall process behaviour through simple mathematical relationships (the so-called "*systemic approach*").
2. Modelling the process in a simplified way, in strict analogy with simple hydraulic systems (the so-called "*conceptual modelling approach*"). The main problem with these techniques is that only linear models are obtained, which can be useless for management and supervision purposes. Also, for use in network design and maintenance, parameter tuning must be performed carefully.

The aim of the present paper is to propose a novel calibration procedure for conceptual models, based on the estimation of black-box dynamic models for the rainfall–runoff relationship. In particular, it will be shown that the proposed procedure provides better estimates for the parameters of conceptual models than the calibration techniques that have been proposed in the hydraulic literature in recent years. Among other things, the use of black-box models for this task can be carried out in strict analogy with the Instantaneous Unit Hydrograph (IUH) estimation procedure, which is used in hydraulic practice for calibration (see Dooge (1977) for additional details).

As a further development, black-box models will allow real-time control techniques to be applied to network management since they can be easily and quickly calibrated, and since some of these models (the linear and nonlinear output error ones) do not need runoff measurements which are often difficult to obtain in real time.

In addition, black-box modelling can fulfil the need for accurate models of drainage in urban networks, provide information on the networks capability to cope with unexpected rainfall events, and give hints on the development of new networks. For this purpose, the prediction capability of the conceptual models obtained from ARX, OE and state-space black-box models are compared with the ones given by the models and the estimation methods more frequently used in hydraulic engineering. Finally, the performance of nonlinear NARX polynomial and neural models is compared with the results obtained from linear models.

The paper is organized as follows: in Section 2 the main conceptual models and parameter estimation methods are briefly outlined; in Section 3 the available data and the corresponding measurement set-up are described; Section 4 is devoted to the presentation of the results obtained in the identification of linear models, while in Section 5 the performance of the identified linear models for the tuning of hydraulic conceptual models is compared with that of classical tuning methods. In Section 6 results from nonlinearity tests are described, and finally in Section 7 the results obtained by using polynomial and neural nonlinear models are presented and the achievable improvements with respect to the linear case are described.

2. Classical *conceptual models* for drainage networks

In this section, a short presentation of the most widely used *conceptual models* is given, together with their parameter tuning procedures. In particular, the main step of the tuning procedure is the estimation of the IUH, i.e., the impulse response of the rainfall–runoff system.

2.1. Classical reservoir model

This model is based on the hypothesis that the urban drainage network can be effectively considered as a linear tank, i.e., it can be described as a linear system with the following impulse response $g(t)$:

$$g(t) = \frac{1}{k}\exp\left(-\frac{t}{k}\right). \tag{1}$$

In doing so, only one parameter is needed, the tank's characteristic time k, which can be defined as follows (Harremoes et al., 1984; Dooge, 1977):

$$k = \frac{W}{Q_r} = \frac{W_r + W_0}{Q_r}, \tag{2}$$

where W is the total storage volume which is the sum of W_r, the total water volume in the pipes when they are totally full and of W_0, which is the water volume stored on the catchment surface, Q_r is the flow at the catchment closure section when it is totally full. The tank's characteristic time cannot be trivially estimated from Eq. (2) since it is difficult to obtain measures of W_0; moreover, Q_r is frequently underestimated since the pipes are not always completely full. For this reason many empirical methods based on experimental data and network design parameters have been proposed in the literature for the estimation of the reservoir constant k. In particular, in this work, the following estimation formulas have been chosen for comparison with estimates obtained from black-box models:

- Pedersen formula (Pedersen, Peters & Helweg, 1980)

$$k = 3.458\,(L \cdot n)^{0.6}\, i^{-0.4} P^{-0.3}. \tag{3}$$

- Desbordes formula (Desbordes, 1984)

$$k = 5.07\, A^{0.18} (L)^{0.15} (1 + I)^{-1.9} (100P)^{-0.36} h^{-0.07}. \tag{4}$$

- Ciaponi–Papiri formula (Ciaponi & Papiri, 1992)

$$k = 0.50A^{0.351}S_r^{-0.29}d^{0.358}I^{-1.63}. \quad (5)$$

In all the preceding formulae L is the length of the main pipe in the network, n is Manning's roughness coefficient (Chow et al., 1988), i is the rainfall intensity, P is the mean catchment slope, A is the area of the catchment, I is the fraction of impervious ground on total catchment area, h is the height of the fallen rain, S_r is the percentual mean slope of the sewer network, and d is the density of drainage, defined as the ratio between the total network length and the catchment area.

Finally, the Italian Method (Asproni, Fasso & Putzu, 1981) and the Revised Italian Method (Mambretti & Paoletti, 1996a) have also been considered. The Italian Method gives an estimate of k on the basis of the physical characteristics of the catchment surface and of the drainage network pipes. The estimate is computed directly using the basic definition of the model in Eq. (2). The contribution of W_0 is tuned using the IUH estimated from rain-flow data and it has usually a value of about 30–50 m^3/ha.

In the Revised Italian Method the contribution of W_r is reduced in order to take into account the fact that the pipes are not always completely full, and the contribution of W_0 is computed to take into account only the impervious part of the catchment surface. The main disadvantage of these two methods is that accurate knowledge of the pipe positions and lengths is needed to obtain a precise estimate for W_r; their advantages are that fewer uncertain parameters are tuned directly by the available data than in the other methods.

2.2. Time-area model

This model (Dooge, 1977) is based on the hypothesis that flow phenomena are more relevant than accumulation ones, in contrast to the reservoir model. So, once the geometric characteristics of the catchment surface and of the network pipes are known, it is possible to compute a characteristic travel time (known as the *concentration time* T_c) that fully characterizes the flow dynamic in the whole catchment. This model has been conceived together with the Kinematic Method for tuning T_c, where the estimate is defined as the sum of two time intervals:

- the *network time* t_r, which is the time that a water drop takes to run along the longest pipe in the network, when it is completely full;
- the *entry time* t_e, which is the time that a water drop takes to enter the sewer network from outside starting from the most far point of the catchment.

The network time t_r can be analytically obtained knowing the water speed in full pipes and the length of the longest hydraulic path in the network. The travel speed during full pipe flow can be computed using the Chézy formula (6) (Chow et al., 1988):

$$V = K_s\left(\frac{D}{4}\right)^{2/3}\sqrt{s}\,, \quad (6)$$

where K_s is Strickler's roughness coefficient, D is the pipe diameter and s is the mean slope of the pipe. The entry time t_e should be estimated empirically.

An estimate for the concentration time T_c can be obtained directly by the IUH through a least-squares fitting of its integral (i.e., the system step response, called the "*S-shaped hydrograph*" from its typical shape) with a three-segment broken line. The concentration time estimate is the difference between the start–stop abscissa of the middle line segment (Mambretti & Paoletti, 1996b).

2.3. Nash model

In the Nash model (Chow et al., 1988) the catchment is modelled as a sequence of n linear tanks, all with the same characteristic time k. Using $n \in \Re^+$ a more general model is obtained with the following impulse response $g(t)$:

$$g(t) = \frac{1}{k\Gamma(n)}\left(\frac{t}{k}\right)^{n-1}\exp\left(-\frac{t}{k}\right), \quad (7)$$

where $\Gamma(n)$ is the well-known Euler Γ-function, i.e. $\Gamma(n) = \int_0^\infty \exp(-x^n)\,dx$.

The most used tuning procedure is due to Ramachanda (Dooge, 1977):

$$k = 0.0883A^{0.399}(1+I)^{-0.662}h^{-0.106}\vartheta^{0.222}, \quad (8)$$

$$n = \frac{0.1547A^{0.458}(1+I)^{1.662}h^{-0.267}\vartheta^{0.371}}{k}, \quad (9)$$

where ϑ is the duration of the storm.

3. Urban drainage networks and measurement set-up

The experimental data that was considered in this study was collected at Fort Lauderdale (USA) and at Munkerisparken (Denmark). The main characteristics of the two catchment areas are summarized in Table 1.

In the Fort Lauderdale catchment area, measurement of rain intensity was carried out with a tipping bucket-type gauge, each tip equal to 0.254 mm. The rainfall measuring component of the system has the ability to measure and record rainfall from three locations in the catchment. With regard to the discharge, the critical-depth flume-flow measurement principle with U-shaped constriction has been applied. If the flow was unsubmerged during all events, water level gauging upstream of the flume entrance would be sufficient. As the discharge increases, the water surface on the upstream side rises,

Table 1
Main characteristics of the examined catchments. The sampling time T_s is 60 s for both of them

Catchment	Total area (ha)	Impervious area in % of total area	Length of estimation data set	Length of validation data set
Munkerisparken	6.44	31.21	663	391
Fort Lauderdale	8.30	97.90	682	460

touching the top of the pipe, and filling the full flow area. Further increases in discharge trigger full pipe conditions downstream of the flume. For full pipe flows the flume may be considered a modified Venturi meter where headloss though the flume is used to calculate the discharge. Discharge ratings are available for each of these open channel conditions as well as for full pipe flow.

With regard to the Munkerisparken catchment area, for rainfall measurements a tipping bucket gauge was again used, with a 0.05 mm discrimination and a receiving funnel with an area of 649 cm^2. The gauge was calibrated in the laboratory, and found that the maximum measurable rain intensity was 144 mm/h. For discharge measurements, a constriction was again built to obtain a flume similar to the Parshall one (Chow et al., 1988).

In Fig. 1 the data available for the identification procedure (both the estimation and the validation sets) are shown. As can be seen from Fig. 1, the time histories for the rainfall input are characterized by sharp peaks; as a consequence, a similar behaviour is visible in the output flow measurements. It is useful to notice that in this framework an important criterion for the evaluation of the performance of rainfall–runoff models is the ability to predict the amplitude of flow peaks accurately.

4. Identification of linear models

Both input/output (of the ARX and OE type) and state-space models have been considered. The former have been identified by means of the classical least-squares (LS) method, while for the latter a subspace-based model identification (SMI) algorithm has been applied.

A unique model evaluation criterion has been used throughout the paper, i.e., the mean of the squares of the difference between the observed and the simulated flow, namely the mean square error (MSE) defined as

$$J = \frac{1}{N} \sum_{k=1}^{N} (y(k) - \hat{y}(k))^2, \tag{10}$$

where y is the measured flow, and $\hat{y}$ is the simulated flow and N is the total number of data.

4.1. Prediction error methods

The ARX[n_a, n_b, n_k] and OE[n_f, n_b, n_k] model structures are given in Eqs. (11)–(16), (Ljung, 1987):

$$A(q)y(t) = B(q)u(t - n_k) + e(t), \tag{11}$$

where

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, \tag{12}$$

$$B(q) = b_1 + b_2 q^{-1} + \cdots + b_{n_b} q^{-n_b+1}, \tag{13}$$

and

$$y(t) = \frac{B(q)}{F(q)} u(t - n_k) + e(t), \tag{14}$$

where

$$B(q) = b_1 + b_2 q^{-1} + \cdots + b_{n_b} q^{-n_b+1}, \tag{15}$$

$$F(q) = 1 + f_1 q^{-1} + \cdots + f_{n_f} q^{-n_f}. \tag{16}$$

The model structures have been chosen according to the following rationale: the experimental data clearly shows that the two catchments considered are characterized by an important time delay between input and output. Given the particular shape of the signals, it is easy to obtain an initial guess for the value of the delay by superimposing the input and output sequences and shifting them in order to obtain an accurate superposition of the rainfall and flow peaks. Using this initial estimate for the delay it is possible to estimate the values of n_a (or n_f) and n_b by means of cross-validation: the different models obtained in the different model structures have been compared by evaluating their performance when applied to a data set to which neither of them was adjusted; the model that shows the better performance (according to Eq. (10)) will be chosen. Unfortunately, J does not show a clear minimum as a function of n_a (or n_f) and n_b, so, this estimation has been done by choosing those values after which J does not decrease considerably. Finally, it is possible to refine the estimate of n_k. Moreover, it has been verified that the introduction of an MA component in the models does not give any significant improvement in the model performances.

For the Munkerisparken catchment, cross-validation suggests using an ARX[4 2 5]; unfortunately using this

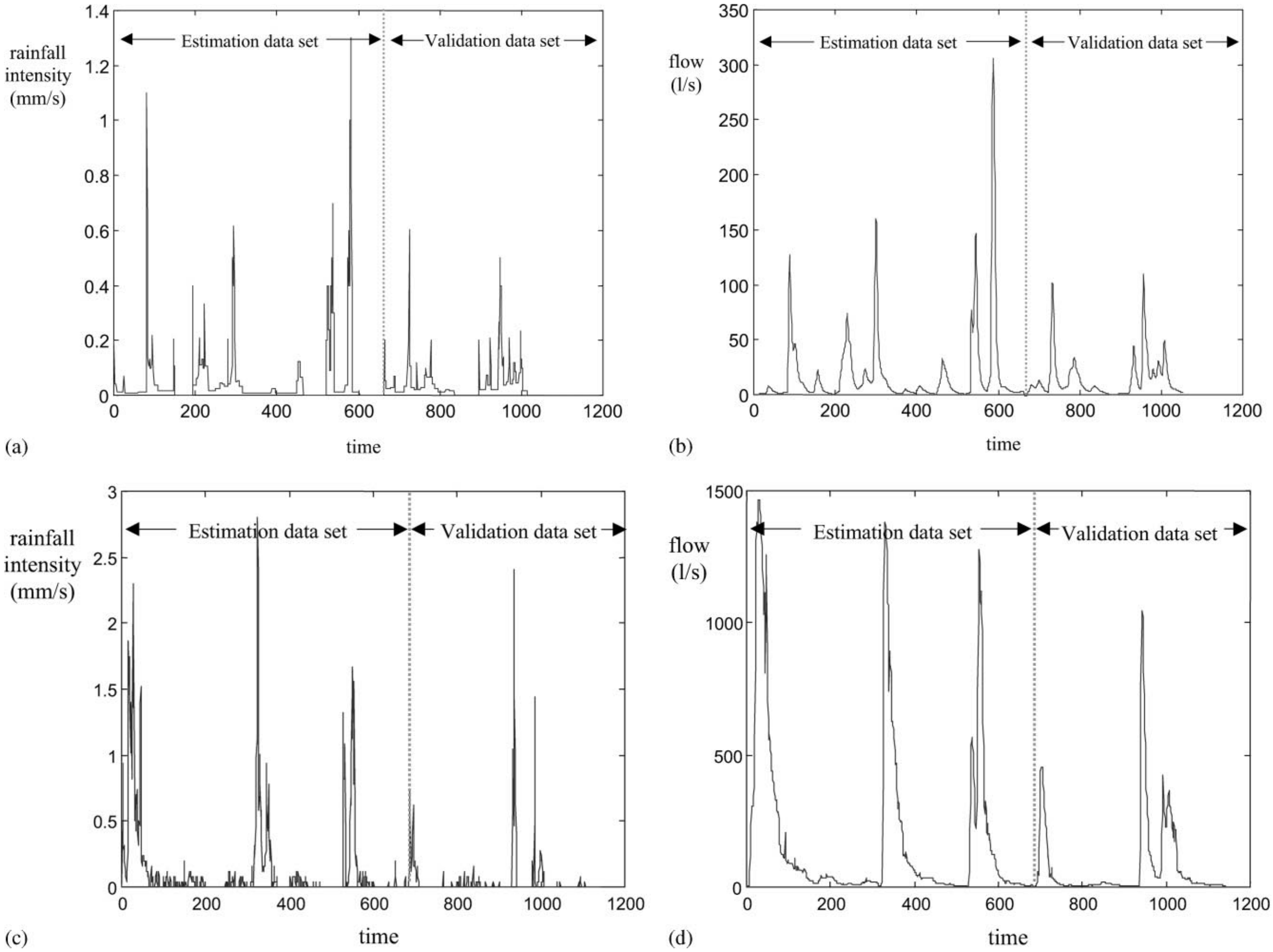


Fig. 1. Munkerisparken catchment: (a) input; (b) output. Fort Lauderdale catchment: (c) input; (d) output.

model order there is a risk of a pole-zero cancellation, and reducing the order of the model increases the value of the performance index considerably. So an ARX[4 3 5] has been chosen which eliminates the risk of a pole-zero cancellation and which has a performance index value similar to that of the initial model.

For the Fort Lauderdale catchment cross-validation suggests using an ARX[3 3 4]. A similar model order has been chosen for OE models. In Fig. 2 comparisons between ARX model simulations and the validation data sets are shown. No significant difference can be seen with the Output Error model simulation plots.

It should be mentioned that the performance of OE models is of particular interest, as their use, both in simulation and in prediction, does not require output (flow) past measurements that are difficult to obtain (a continuous supervision and a frequent maintenance of the measurement set-up is needed). For this reason the use of OE models should allow the real-time monitoring of networks in the presence of high-intensity peak discharge.

The value of criterion (10) is $J_1 = 40.39$ for the Munkerisparken ARX model and $J_2 = 3325.8$ for the Fort Lauderdale ARX model.

4.2. Subspace methods

The family of subspace identification algorithms provides methods for obtaining discrete-time state-space linear models from input/output measurements for MIMO systems of the form

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{w}_t,$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{u}_t + \mathbf{v}_t, \tag{17}$$

where $\mathbf{x} \in \Re^n$, $\mathbf{y} \in \Re^l$, $\mathbf{u} \in \Re^m$. As subspace methods represent a viable alternative to classical prediction error identification methods, their application to this problem has been considered a useful opportunity.

All methods are based on a geometric approach that involves subspaces spanned by rows and/or columns of matrices built with input/output data (Viberg, 1995).

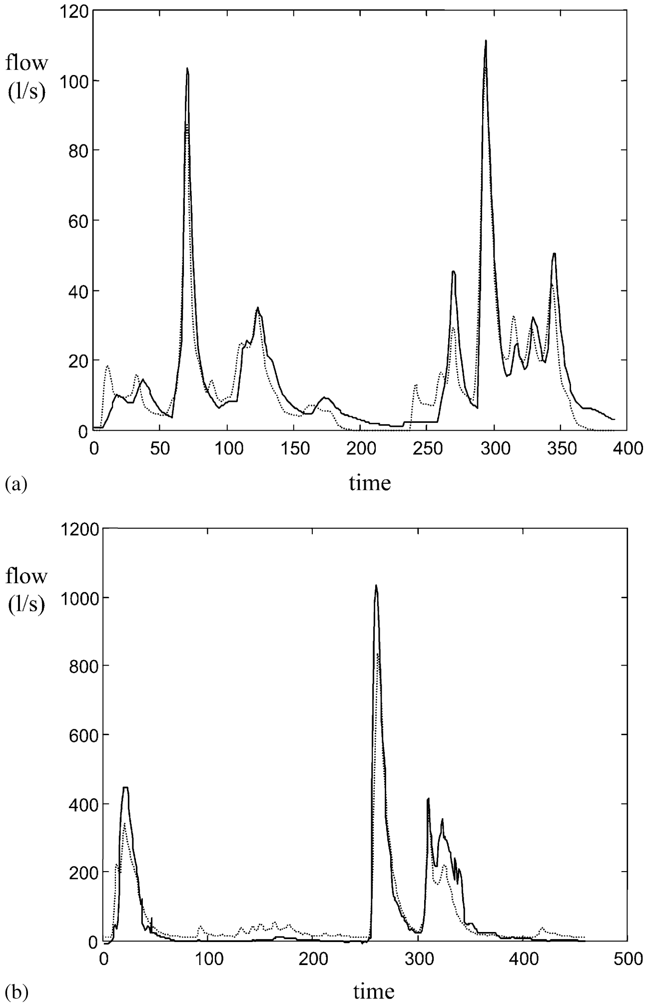


Fig. 2. ARX simulation (dashed) on validation data set (solid): (a) Munkerisparken catchment; (b) Fort Lauderdale catchment.

The main starting point of most subspace algorithms is the following equation:

$$\mathbf{Y} = \mathbf{\Gamma X} + \mathbf{HU} + \mathbf{EW} + \mathbf{V}, \tag{18}$$

where

- $\mathbf{Y}$, $\mathbf{U}$, $\mathbf{W}$ and $\mathbf{V}$ are Hankel matrices formed with the output signals, the input signals and the process and measurement noise, respectively:

$$\mathbf{Y}_{(li,\,j)} = \begin{bmatrix} \mathbf{y}_t & \mathbf{y}_{t+1} & \mathbf{y}_{t+2} & \mathbf{y}_{t+3} & \cdots & \mathbf{y}_{t+j-1} \\ \mathbf{y}_{t+1} & \mathbf{y}_{t+2} & \mathbf{y}_{t+3} & \cdots & & \mathbf{y}_{t+j} \\ \mathbf{y}_{t+2} & \mathbf{y}_{t+3} & \cdots & & & \vdots \\ \vdots & \vdots & & & & \vdots \\ \mathbf{y}_{t+i-1} & \mathbf{y}_{t+i} & \cdots & & & \mathbf{y}_{t+i+j-2} \end{bmatrix}_{(li,\,j)}$$

and similarly for $\mathbf{U}$, $\mathbf{W}$ and $\mathbf{V}$; i and j are parameters of the algorithm, to be chosen by the user.

- Γ is the extended observability matrix for the system:

$$\Gamma = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix}_{(li,\,n)}.$$

- $\mathbf{X}$ is a state sequence:

$$\mathbf{X}_{(n,\,j)} = [\mathbf{x}_t \;\; \mathbf{x}_{t+1} \;\; \mathbf{x}_{t+2} \cdots \;\; \mathbf{x}_{t+j+1}].$$

- $\mathbf{H}$ is a Toeplitz matrix formed with the Markov parameters of the model:

$$\mathbf{H}_{(li,\,mi)} = \begin{bmatrix} \mathbf{D} & & & & & \\ \mathbf{CB} & \mathbf{D} & & & 0 & \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & & & \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \\ \mathbf{CA}^{i-2}\mathbf{B} & & & \mathbf{CAB} & \mathbf{CB} & \mathbf{D} \end{bmatrix}_{(li,\,mi)}.$$

- $\mathbf{E}$ is a Toeplitz matrix:

$$\mathbf{E}_{(li,\,mi)} = \begin{bmatrix} 0 & & & & & \\ \mathbf{C} & 0 & & & 0 & \\ \mathbf{CA} & \mathbf{C} & 0 & & & \\ \mathbf{CA}^2 & \mathbf{CA} & \mathbf{C} & 0 & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & \\ \mathbf{CA}^{i-2} & & & \mathbf{CA} & \mathbf{C} & 0 \end{bmatrix}_{(li,\,mi)}.$$

Subspace methods can be divided into two main families: the first family (*one-step* methods) operates on (18) with the aim of calculating an estimate of state sequence matrix $\mathbf{X}$; this estimate allows the system matrices to be determined as the LS solution of a single, overdetermined set of linear equations.

On the other hand, the second family (*two-steps* methods) uses orthogonal projections of the rows of matrix $\mathbf{Y}$ to derive an estimate of the extended observability matrix Γ from which it is possible to compute matrices $\mathbf{A}$ and $\mathbf{C}$, in the following step, $\mathbf{B}$ and $\mathbf{D}$ are obtained by solving an LS problem.

The algorithms which have been used in this work belong to the MOESP (multivariable output error model state-space model realization, see Verhaegen, 1994; Lovera, 1998; Haverkamp & Verhaegen, 1998) and are known in the literature as MOESP past inputs (PI) and past outputs (PO).

These algorithms are two step methods, involving the use of instrumental variables (IV). The effect of the IVs is an asymptotic cancellation of noise, this is achieved by an orthogonal projection of the rows of **Y** on a subspace obtained from "past" input/output data.

The major advantage of the IV approach in the two-step family is the unbiased matrix estimates that are preserved even in the case of coloured noise; this feature has been considered important for the present application. Therefore, the problem is that of a "linear system" affected by process and measurement noise.

All subspace identification methods involve the calculation of the singular values of a matrix (which is different from method to method); in the ideal case of a noiseless, truly linear system generating the data, the number of nonzero singular values equals the system order n. When noise is present, and the S/N ratio is "reasonably" good, two sets of singular values can be detected: n "signal" singular values (significantly larger) and the remaining "noise" singular values. In this noisy case, an estimate of model order n can be obtained by splitting the singular values in these two sets.

As in the case of prediction error methods, a number of identification trials have been performed, in order to determine the best choice of time delay and model order.

In the overall, the best performance was achieved by making use of the PI algorithm, which led to the following models: in both cases a fourth-order model with a time delay of five samples was chosen, with a value of criterion (10) of $J_1 = 43.4$ for the Munkerisparken basin and of $J_2 = 3314$ for the Fort Lauderdale basin. The results are in line with those obtained by means of prediction error methods; the comparison between model simulations and measured outputs is omitted for brevity.

5. Comparison between classical and black-box based tuning of "*conceptual models*"

In this section the simulation results obtained from the hydraulic conceptual models described in Section 2 will be compared with the ones obtained from linear black-box models. In order to perform a fair comparison between black-box models and the existing models based on classical IUH theory (remember that the IUH is defined as the unit area response of the catchement to a rainfall impulse), there is a need for a procedure in order to obtain IUH-related parameters from black-box simulated data. Such a procedure will be outlined in the following.

5.1. Reservoir model tuning

The parameter k of the reservoir model can be tuned through many different procedures and, in the following, comparisons will be done between linear black-box models and some of the available empirical parameter-tuning techniques. In particular, a method has been developed to compare black-box models with existing storage constant estimation methods, which consists of the following steps.

The discrete impulse response of the identified model is normalized to unit area, thus giving an estimate of the IUH for the considered catchement. The estimate of k is taken to be the abscissa of the centre of mass of the estimated IUH; as a matter of fact, given the analytical expression of the IUH for the reservoir model:

$$g(t) = \frac{1}{k}\exp\left(-\frac{t}{k}\right), \tag{19}$$

it can be easily shown that the abscissa of its centre of mass is

$$\int_0^\infty \frac{1}{k}\, t \exp\left(-\frac{t}{k}\right) \mathrm{d}t = k_c. \tag{20}$$

So k_c is an estimate of the reservoir time constant k which can be easily obtained from black-box models. The reservoir time constant has been also estimated by least-squares fitting of the decreasing part of the normalized IUH, as obtained from black-box models. This estimate has been called k_{ls}. These quite unusual estimation procedures for the time constant from the IUH have been used since they are correlated in a straightforward way to the usual hydraulic techniques and they are easily understandable for an hydraulic technician directly involved in sewer design or maintenance.

In Tables 2 and 3 a comparison between the value of the storage constant obtained from empirical formulas and the values obtained from black-box models is shown. In particular k_c and k_{ls} are the two estimates obtained from the IUH given by the ARX and OE models, k_p is that obtained from the Pedersen estimate, k_d is given by the Desbordes estimate, k_i comes from the Italian tuning

Table 2
Munkerisparken network k storage constant (in s) as it is estimated from different methods

	k_c	k_{ls}	k_p	k_d	k_i	k_{ir}	k_{cp}
ARX[4 3 5]	4.81	3.78	8.21	13.57	16.34	4.90	7.43
OE[4 3 5]	4.52	4.19					

Table 3
Fort Lauderdale network k storage constant (in s) as it is estimated from different methods

	k_c	k_{ls}	k_p	k_d	k_i	k_{ir}	k_{cp}
ARX[3 3 4]	8.52	8.61	12.00	14.15	6.32	2.91	12.30
OE[3 3 4]	8.41	8.07					

method and k_{ir} from its revised version, and k_{cp} it is the Ciaponi–Papiri estimate.

Using the estimated k values in the reservoir model expression, it is possible to obtain simulations in time domain from input data of the validation data set. The black-box estimate of the time delay has been considered in the simulations, since the reservoir model does not include any delay. In this way it is possible to compare the different estimates of the reservoir constant, both visually and computing the performance index (10).

It is possible to note (Tables 2 and 3) that the estimates of k obtained from the ARX–OE models are similar, while the other tuning methods give results which differ considerably one from the other. Moreover, Fig. 3 shows the comparison between the simulation of the reservoir model tuned with the IUH estimated from the ARX model and the simulations obtained from the Desbordes tuning and the Italian method, respectively. So it is possible to argue that a correct value of k for the Munkerisparken catchment is $k \cong 4.5$–4.8, a value which is reliably obtained from the black-box tuning and not from other procedures, except for the Revised Italian Method.

Similar conclusions can be drawn looking at Table 3, which shows the Revised Italian Method does not give a correct estimate either.

Table 4
Estimates of the concentration time (in s) obtained using black-box linear models and classical methods

	ARX tuning	OE tuning	Kinematic method	Revised kinematic method
Munkerisparken	9.11	9.69	12.38	9.09
Fort Lauderdale	20.5	19.2	19.73	14.40

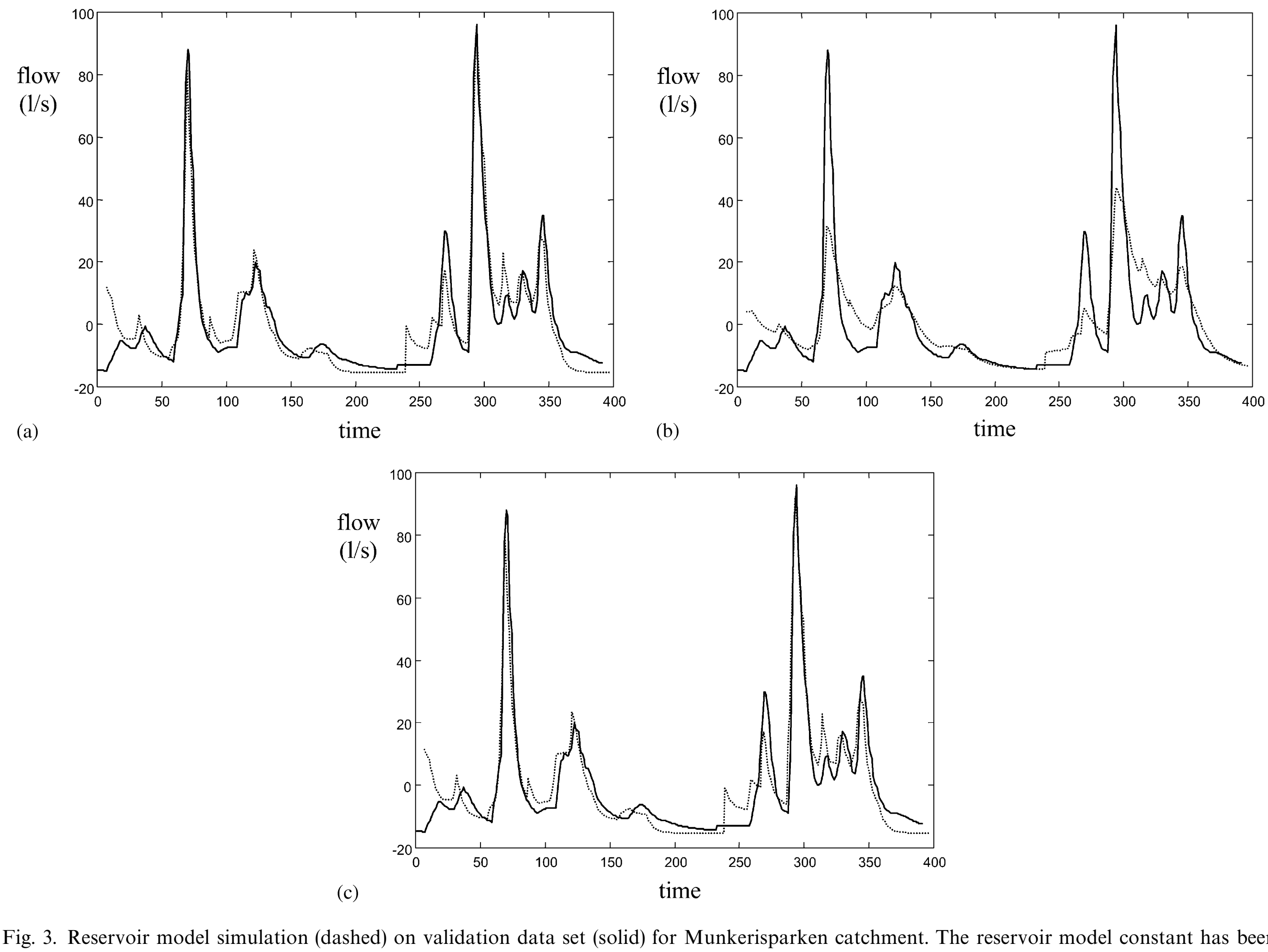


Fig. 3. Reservoir model simulation (dashed) on validation data set (solid) for Munkerisparken catchment. The reservoir model constant has been obtained as follows: (a) centre of mass of IUH from ARX model simulated data; (b) Desbordes formula; (c) Revised Italian method.

So it is possible to conclude that black-box model based tuning methods are effective and reliable for the reservoir model. This result is very important since the reservoir model is probably the most widely used for small sewer network modelling.

5.2. Time–area model tuning

In the hydraulic practice the concentration time T_c in the time-area model is estimated using the Kinematic method or its revised version. Mambretti and Paoletti (1996b) propose fitting the "*S-shaped hydrograph*" (which is the integral of the IUH and so it is strictly related to the step response of the system) with a three segment broken line by the least-squares method; the first and the third line segment are horizontal, while the middle line segment is a straight line to be fitted on the rising part of step response. The concentration time T_c is then given by the difference between the start–stop abscissa of the middle line segment. It is clear that the more precise the estimate of the step response of the system the more accurate will be the estimate of the concentration time T_c.

In the following the linear models estimated in Section 4 are used to obtain the step response of the system, and the three segment broken line procedure described above is used to obtain an estimate for the concentration time T_c.

In Table 4 the estimates of T_c are reported for the two examined catchments: it is shown a comparison between the black-box model-based estimates and the results from the kinematic method and its revised version.

Fig. 4 shows that the estimate for the concentration time T_c obtained from black-box models allows a faithful simulation of flow in the validation data set. It is worth noting that for the Munkerisparken catchment, an estimate in agreement with that obtained from the black-box models is also given by the kinematic method, but not from its revised version. In contrast, for the Fort Lauderdale catchment, the estimate obtained by the revised method is in agreement with the result obtained from the black-box models, which gives the best simulation with the validation data set in the sense of criterion (10). However, there is agreement between the estimates from different black-box models of the same catchment. So it is also possible to say that the black box models allow a precise and affordable estimate of the parameter T_c for the time–area model.

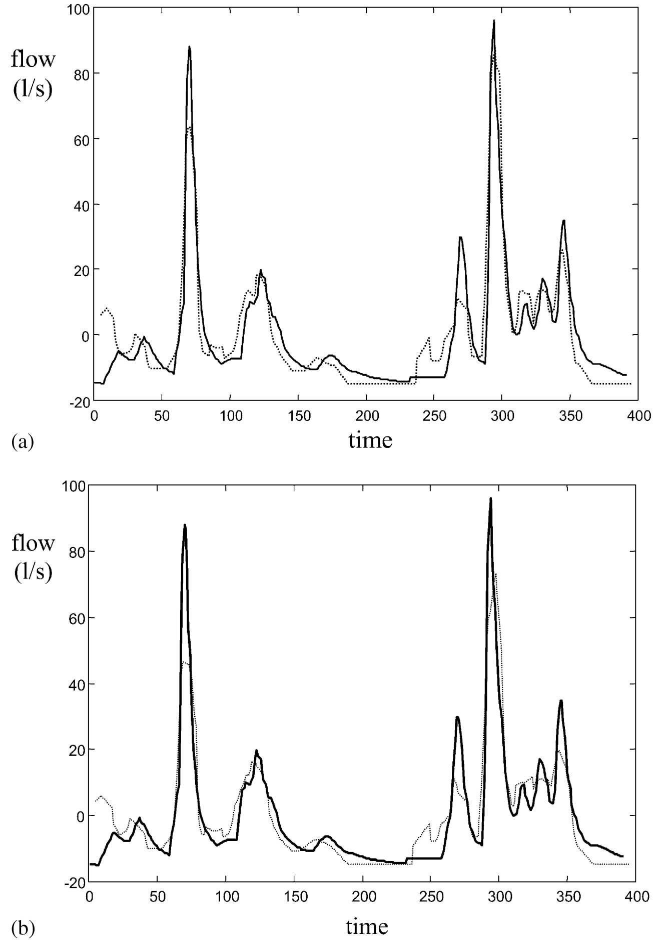


Fig. 4. Time–Area model simulation (dashed) on validation data set (solid) for Munkerisparken catchment using the concentration time as estimated by: (a) ARX[4 3 5]; (b) Kinematic method.

Table 5
Values for storage constant and tank number as obtained from linear models and Ramachanda estimation equations of Nash model parameters for Munkerisparken and Fort Lauderdale catchment

	k	n
Munkerisparken ARX tuning	1.54	3.31
Munkerisparken OE tuning	2.18	2.35
Munkerisparken Ramachanda tuning	4.01	4.72
Fort Lauderdale ARX tuning	5.13	1.65
Fort Lauderdale OE tuning	4.90	1.69
Fort Lauderdale Ramachanda tuning	7.42	6.06

5.3. Nash model tuning

Estimates of the storage constant k and of the optimal tank number n in the Nash model can be obtained through least-squares fitting of the IUH of the model Eq. (7) to the impulse response of ARX, OE linear black-box models. In particular, the obtained results will be compared to those from Ramachanda tuning Eqs. (8) and (9).

In Table 5 a comparison between the estimates obtained from the different methods is shown, while a comparison between simulations of the flow in the validation data set obtained from the estimated parameters is presented in Fig. 5.

As can be seen, Ramachanda formulae are not adequate to estimate the response characteristics of the analysed catchments. Moreover, Fig. 5 shows why correct parameter estimates are so important and how they can seriously affect the predictive capability of a model.

As shown in the previous paragraphs, linear models seem capable of providing accurate descriptions of the response of the considered drainage networks. As was previously mentioned, linearity in the response was expected, at least when considering "*critical*" storms, i.e., those that almost fill the network pipes. In fact, when a sewer pipe is completely filled, the velocity of the flow might be considered as constant and the system response almost linear.

In contrast, when the flow is small, the velocity is small as well, and the water's travelling time can also change considerably for small rainfall changes so, a nonlinear behaviour is expected in general. Finally, even though the data are collected in almost *critical* conditions (in the sense described above), some evidence of nonlinear behaviour can be pointed out, so the question arises whether it would be possible to derive even better models for the considered networks by resorting to nonlinear identification techniques. This issue will be the main one in the second part of this paper.

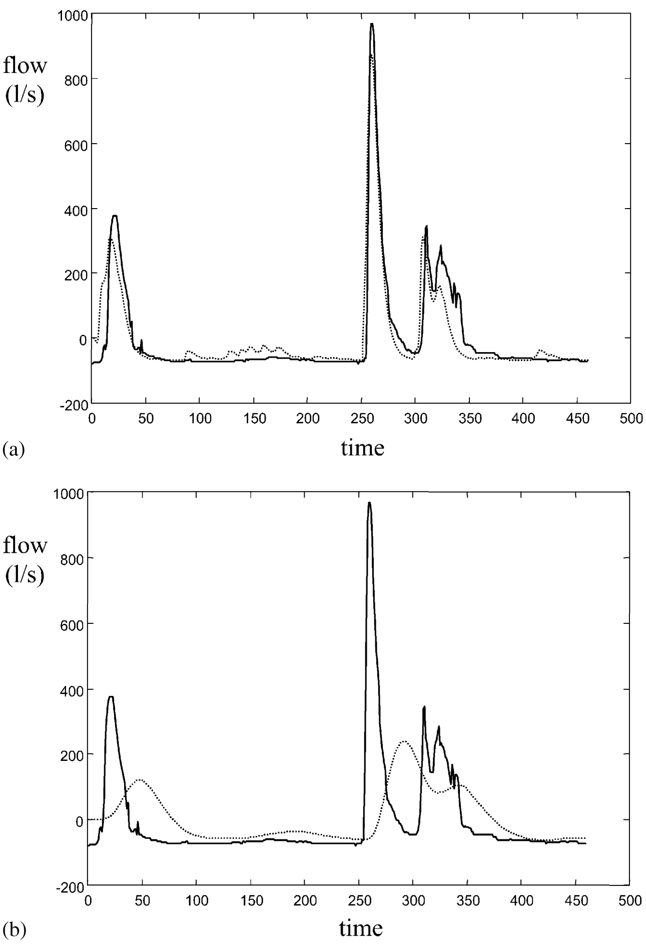


Fig. 5. Comparison between Fort Lauderdale validation data (solid) and Nash model simulation (dashed) tuned using: (a) ARX[4 3 5]; (b) Ramachanda formulae.

6. Testing for nonlinearity

Two classes of nonlinearity tests have been considered in this work:

1. Nonlinearity tests based on high-order correlation analysis (Billings & Voon, 1983,1986).
2. Tests based on frequency analysis and in particular on the so-called squared coherency function (Haber, 1985).

In order to apply correlation-based nonlinearity tests, a number of assumptions, mainly concerning the available input/output data, have to be checked. These hypotheses require the input $u(t)$ and the noise $e(t)$ to be independent zero mean processes, all the odd moments of $u(t)$ and $e(t)$ to be zero and the existence of all the even moments. Throughout this section all available data have been used.

If the previous conditions are satisfied the following holds if and only if the process is linear:

$$\phi_{y'y'^2}(\tau) = E[y'(t+\tau)\,(y'(t))^2] = 0 \quad \text{for every } \tau, \qquad (21)$$

where $y'(t) = y(t) - E[y(t)]$.

Third-order moments of the available input signals have been estimated and it has been verified that the zero value is not inside the confidence interval (see Billings and Zhu (1994) for the estimation of two-dimensional correlations). In the following, an alternative procedure to check the validity of these basic assumptions has been used. First, feed a linear system with the measured input data: if the nonlinearity test fails, then the input structure is not reliable for the application of the test; otherwise nothing can be said and more input realizations (together with some assumptions on the noise structure) should be needed.

Using the input data set from the Munkerisparken catchment as input for an ARX[4 3 5] model and hence performing test (21) on the corresponding output, the result given in Fig. 6 for the correlation function of interest was obtained. Notice that the ARX model has nothing to do with the real system: it is only a linear system that is useful in checking the hypotheses for (21). As can be seen from Fig. 6, the test fails even though the system is linear and so it can be concluded that the available input data do not match the assumptions for the tests.

Other tests have been therefore considered, namely, the one based on the squared coherency function

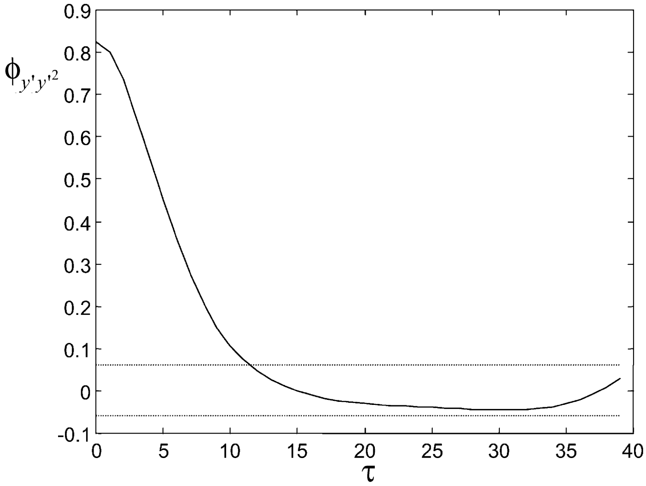


Fig. 6. High-order correlation function non linearity test for an ARX[4 3 5] linear system fed with input recorded on Munkerisparken catchment. The test evidently fails.

estimation (Haber, 1985; Jenkins & Watts, 1968). This procedure is based on mild assumptions about the data; its main limitation is found in its sensitivity to noise and in the bias of the estimate (Jenkins & Watts, 1968; Löhnberg, 1978).

An estimate of the squared coherency function is given by

$$\hat{C}(f) = \frac{\hat{G}_{u'y'}(f)\hat{G}_{y'u'}(f)}{\hat{G}_{u'u'}(f)\hat{G}_{y'y'}(f)}, \tag{22}$$

where $u'(t) = u(t) - E[u(t)]$ and $y'(t) = y(t) - E[y(t)]$, $G_{u'y'}$ is the cross spectral density between input and output, with complex conjugate $G_{y'u'}$; $G_{u'u'}$ and $G_{y'y'}$ are, respectively, the power spectral density of input and output.

A system is detected to be nonlinear if the following differs from zero:

$$\rho = 1 - \max_f \left\{ \frac{G_{u'y'}(f)G_{y'u'}(f)}{G_{u'u'}(f)G_{y'y'}(f)} \right\} = 1 - \max_f \{C(\omega)\}. \tag{23}$$

An estimate $\hat{\rho}$ of (23) is clearly dependent on the properties of estimate (22). When using test (23), in practice, essentially the coherency function can be less than unity for one or more of the following conditions (Bendat & Piersol, 1980):

(a) when extraneous noise is present in the input/output measurements;
(b) when bias errors are present in the spectral estimates;
(c) when the system relating input and output is not linear.

Noise does not have a great influence on the accuracy of the coherency function estimate when the signal-to-noise ratio is high, as in this case. In fact, the data has been carefully collected (see Section 3) and measurement errors are negligible both for input (rainfall intensity) and output (flow).

Bias in the estimate of the coherency function can cause serious errors and make test (23) useless. So, it is necessary to evaluate the bias of the squared coherency function estimate. In this paper, the approximation of bias $[\hat{C}]$ proposed in Löhnberg (1978) has been used. This approximation is better than the classical ones (see for instance Jenkins and Watts, 1968), since it is based on second-order Taylor expansion of $\hat{C}$ rather than on a first-order approximation.

Windows used in the power spectrum estimation play a key role in the estimation of the squared coherency. In particular, the knowledge of the dependence of bias$[\hat{C}]$ derived from the window characteristics can be used to choose a minimum bias window among a set, or even to design such a window.

In this work two families of windows have been considered for smoothed squared coherence estimation, namely the classical Parzen window and a family proposed by Papoulis (1973), which are minimum-bias windows in spectral estimates. At a first stage other window families have been considered, such as Hamming, Hanning and Blackman, but they were discarded since they give significantly higher values for bias$[\hat{C}]$ than the two considered families.

The estimate of bias$[\hat{C}]$ has been calculated following Löhnberg (1978) for different values of lag window width M, and the minimum bias window has been chosen. For the Munkerisparken data $M = 9$ has been chosen for both Parzen and Papoulis windows; for the Fort Lauderdale data $M = 5$ has been chosen for both window families (Fig. 7).

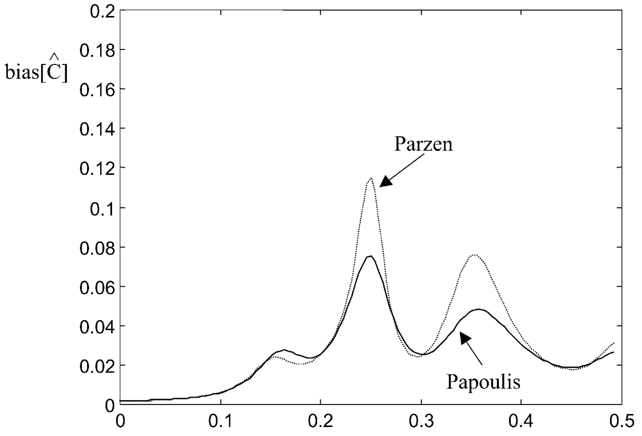
Then, the smoothed-squared coherency estimate was computed together with its 95% confidence limits according to Jenkins and Watts (1968, see Fig. 8). The confidence interval has not been corrected for the bias effects since the bias (like the variance) is inherently low at the frequency at which the estimate of C reaches its maximum value, which is the one used for the computation of $\hat{\rho}$. Finally, an estimate $\hat{\rho}$ of Eq. (23) has been obtained.

The estimated values for ρ are shown in Table 6.

Finally, it is possible to conclude that the two available data sets indicate that the relationship between rainfall and runoff shows nonlinear behaviour in the two considered basins. This is in agreement with the previous comments regarding *critical* rains and their capability of emphasizing nonlinear dynamics.

7. Identification of NARX models

In this section nonlinear models are used for the identification of the dynamic response of the considered



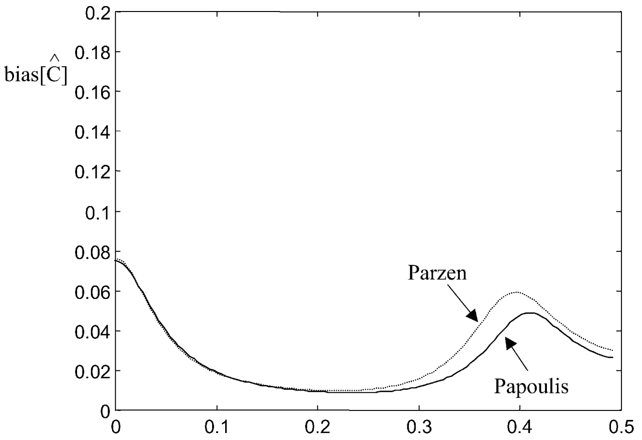


Fig. 7. Estimate of bias[$\hat{C}$] as a function of normalized frequency using Parzen and Papoulis windows: (a) Munkerisparken catchment; (b) Fort Lauderdale catchment.

Table 6
Estimates of ρ for the two catchments together with their 95% confidence interval

	$\hat{\rho}$	95% confidence interval
Munkerisparken (Parzen $M=9$)	0.4601	[0.3992; 0.5260]
Fort Lauderdale (Parzen $M=5$)	0.8194	[0.7710; 0.8644]
Munkerisparken (Papoulis $M=9$)	0.4558	[0.3954; 0.5212]
Fort Lauderdale (Papoulis $M=5$)	0.8348	[0.7876; 0.8783]

catchments. NARX models have been considered, i.e. nonlinear relationships between past observations $[u(t-1), u(t-2), \dots, y(t-1), y(t-2), \dots]$ and future outputs $y(t)$ of the form (Sjoberg et al., 1995; Sjoberg, 1995; Juditsky et al., 1995):

$$y(t) = f(u(t-1), u(t-2), \dots, y(t-1), y(t-2), \dots, \theta) + e(t), \tag{24}$$

where $f(\cdot)$ is the nonlinear map to be estimated, $e(t)$ is the modelling error and θ is a finite-dimensional vector which parametrizes the class of nonlinear functions $f(\cdot)$.

The identification problem for NARX models consists of two separate stages:

- Choosing the dimension of the regression vector, i.e. how many past observations of input and output are needed (structural identification).
- Estimating the map $f(\cdot)$. In most cases, this amounts to finding a parameterization $\hat{\theta}$ that minimizes a model evaluation criterion (parameter estimation).

The parametrization of the nonlinear mapping will be obtained using the following function expansion (Sjoberg, 1995):

$$f(u(t-1), \dots, y(t-1), \dots, \theta) = \sum_k \alpha_k f_k(u(t-1), \dots, y(t-1), \dots, \beta_k), \tag{25}$$

where $f_k(\cdot)$ is a parametrization of a single family of functions, θ is built with the α_k and with the (possible) $f_k(\cdot)$ family parametrization β_k. Also, the dimension of the k-expansion has to be chosen and it can be a further problem or it can be included in the regressor choice problem.

Two different cases will be considered:

- the polynomial expansion case, where the $f_k(\cdot)$ will be monomials built using $[u(t-1), \dots, y(t-1), \dots]$, according to (Chen, Billings & Luo, 1989);
- the neural network case, where the $f_k(\cdot)$ will be a parametrized sigmoidal function.

The model evaluation criterion will be (10) for both model families and the minimization of it will be achieved using well established techniques.

7.1. Polynomial NARX models

The NARX model family obtained through polynomial expansion of the unknown nonlinear map $f(\cdot)$ is well known and there are many successful applications of this methodology (see Thomson, Schooling & Soufian, 1996 for example). Expanding (24) as a polynomial of degree l the representation in (25) is

$$y(t) = \sum_{i=0}^{M} \theta_i p_i(t) + e(t), \tag{26}$$

where

$$p_i(t) = \prod_{j=1}^{p} y(t - N_{y_j}) \prod_{k=1}^{q} u(t - N_{u_q}) \quad \text{with}$$

$p, q \geqslant 0,\ 1 \leqslant p + q \leqslant l,$

$1 \leqslant N_{y_j} \leqslant N_o,\ 1 \leqslant N_{u_q} \leqslant N_i,$

N_o, N_i is the regressor dimension;

θ_i is the ith model parameter;

$$M = \sum_{s=0}^{l} n_s \quad \text{with} \quad \begin{cases} n_0 = 1, \\ n_s = n_{s-1} \dfrac{N_0 + N_i + s - 1}{s}. \end{cases}$$

Model (26) is linear in the parameters; this is clearly an advantage for parameter estimation, even though the estimation of the parameters is far from being a trivial problem, as the information matrix is often badly ill-conditioned. Moreover, the number of parameters in (26) can be very large even for small values of l, N_o, N_i, and a structure selection procedure is needed, since the model structure is rarely known a priori.

Parameter estimation can be performed using the well-known algorithm based on orthogonal regression estimation, known as Modified Gram–Schmidt (Chen et al., 1989). Structure selection can be performed by choosing the term (among all the candidates) that gives the highest decrease in the output error variance at each step of the estimation procedure, i.e., by estimating the percentage reduction due to the ith parameter with respect to the maximum mean-squared prediction error, called error reduction ratio (ERR):

$$\text{ERR}_i = \frac{\sum_{t=1}^{N} g_i^2 w_i^2(t)}{\sum_{t=1}^{N} y^2(t)} 100 \tag{27}$$

where g_i is the parameter estimate in the orthogonal form of the model (analogous to θ_i), $w_i(t)$ is the expansion term in the orthogonal model (analogous to $p_i(t)$), and N is the length of the data set.

The main problem in this procedure is to establish a stopping criterion, i.e., to decide when the introduction of a new term in the expansion is useless since the ERR given by that term is too small and the presence of that term in the expansion will not improve the model quality. The selection procedure can be stopped when the contribution to the output variance explanation is below an a priori chosen threshold ρ:

$$1 - \sum_{i=1}^{M_S} \text{ERR}_i < \rho, \tag{28}$$

where M_S is the number of the selected parameters. Moreover a working procedure for the choice of ρ based on trial and error is proposed in Thomson et al. (1996).

These two procedures seem to be alternative; however they can be fruitfully combined, thus obtaining a new trial and error procedure which allows models that are both parsimonious and precise to be obtained and to avoid overfitting, as shown by simulation on validation data.

The combined approach which has been used in this paper involves the following (Previdi, 1999):

1. Choosing N_o and N_i (regressor dimension) on the basis of the best linear model. The use of $l = 2$ is compulsory, since few data are available for parameter estimation.
2. Choosing a priori a maximum number of parameters to be estimated. This choice can be at least partially guided by the availability of experimental data.
3. Estimating an initial NARX model and computing the variance of the residual for the obtained model. In this way, it is possible to give a first estimate for the value of ρ.
4. Re-estimating the model using the above value of the residual covariance as a stopping criterion in the Billings algorithm.
5. Eventually modifying the regressor structure on the basis of the selected parameters in steps 3 and 4.

For the Munkerisparken basin a polynomial NARX model with $l = 2$, $N_o = 4$, $N_i = 4$, and with a four-step time delay between input and output has been chosen. The estimation procedure was stopped by trial and error in the way explained above; 13 parameters were selected and estimated.

For the Fort Lauderdale basin a polynomial NARX model with $l = 2$, $N_o = 5$, $N_i = 4$, and with a three-step time delay between input and output was chosen. Sixteen parameters were estimated using the procedure described above.

In Fig. 9 the comparisons between model simulation and validation data set for both systems is shown.

The obtained results on validation data set are clearly better than those of linear models. In particular, criterion (10) values are $J_1 = 28.58$ for the Munkerisparken basin and $J_1 = 2608.0$ for the Fort Lauderdale one. The estimates of the flow peaks of the validation data set are also as precise as those obtained from linear models.

7.2. Neural network NARX models

Feedforward neural networks of the multi-layer perceptron family have been used for the identification of the rainfall–runoff relationship. Feedforward neural networks have an input layer, one or more hidden layers and an output layer of neurons. In this work only one hidden layer and only one neuron in the output layer have been used. The network is called feedforward because there are no connections between neurons in the same layer, and the connections between two layers are unidirectional.

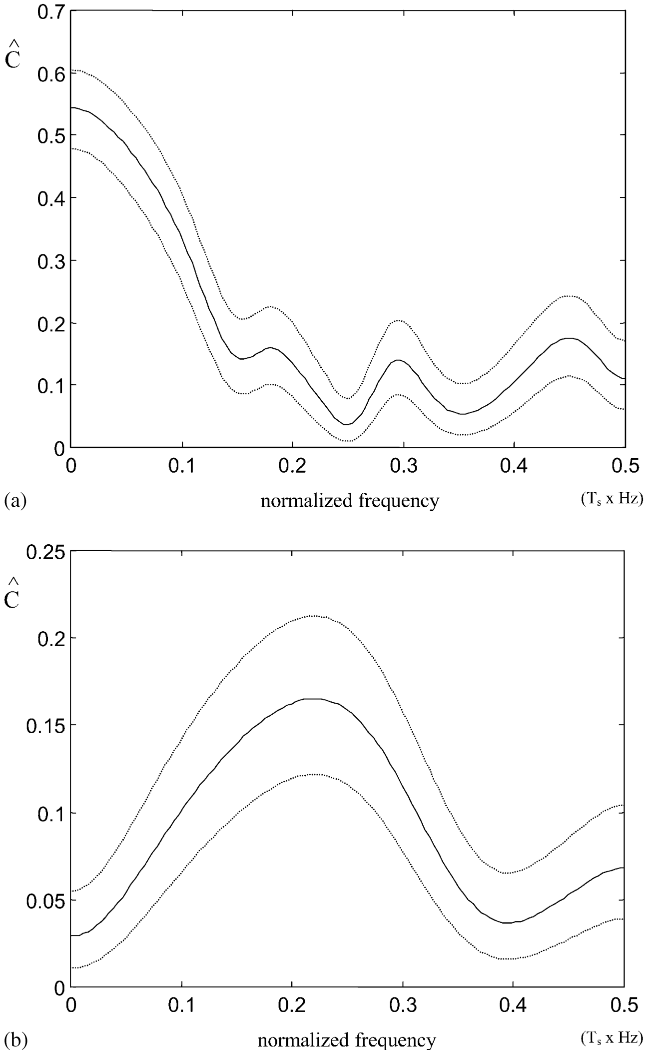


Fig. 8. Squared coherency function estimates using the selected Papoulis windows: (a) Munkerisparken catchment; (b) Fort Lauderdale catchment.

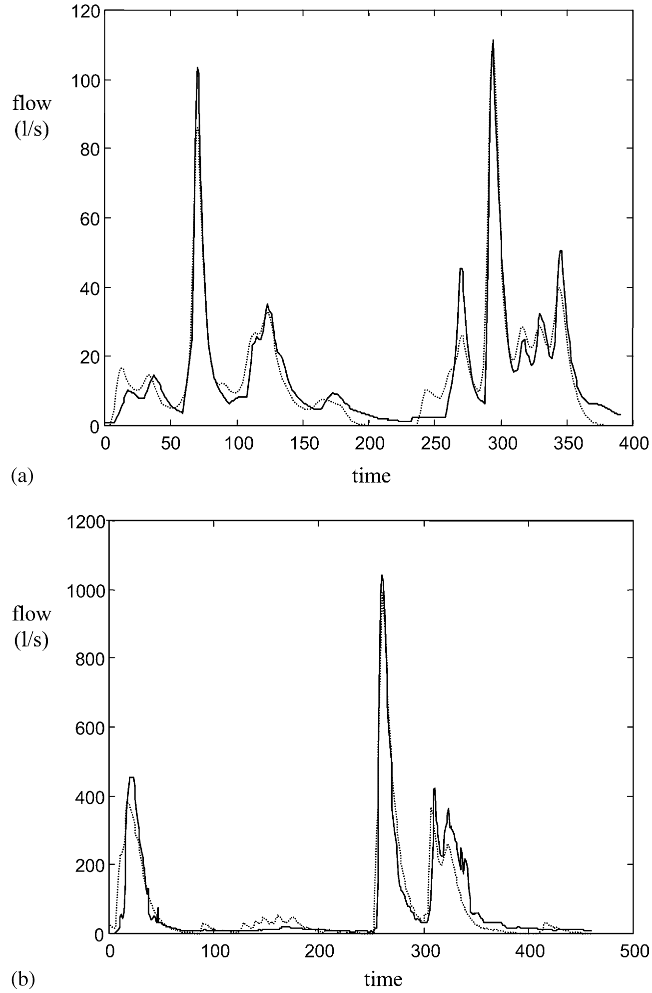


Fig. 9. Comparison between polynomial model simulation (dashed) and validation data set (solid): (a) Munkerisparken catchment; (b) Fort Lauderdale catchment.

The output of each neuron is a nonlinear function of its input (the hyperbolic tangent has been used), except for the output neuron (which gives the predicted model output). The input to a neuron is a linear combination of the output of the neurons in the previous layer summed with a bias.

The model parameters have been tuned using a standard gradient-descent back-propagation algorithm in which criterion (10) is minimized. The number of inputs of the neural network is chosen according to the desired model order: n_a being the number of past samples of the output, n_b the number of past samples of the input, n_k the number of samples of delay between input and output, and the number of neural network inputs is $n_a + n_b$.

During the training procedure the network is fed with input and output from the estimation data set using a pseudo-random extraction procedure in order to generate a large number of training patterns from the available data. The parameters are estimated with the network in a NARX configuration and the prediction error is minimized (Narendra & Parthasarathy, 1990). Training is stopped after a fixed number of iterations; in fact, it is known that terminating the search for the minimum is a kind of regularization and helps in avoiding overtraining (Sjoberg & Ljung, 1995). During simulation the neural network is used in a nonlinear output error (NOE) configuration, i.e., the network predicted output is passed through a suitable delay line and taken to the input.

The regressor order (i.e., the values of n_a, n_b and n_k) has been selected according to the following cross-validation procedure:

- the initial value of n_k is that of the best available linear model;
- different values for n_a and n_b (close to those of the linear model) are tried; a first choice for n_a and n_b is made according to the value of criterion (10) on the validation data set;

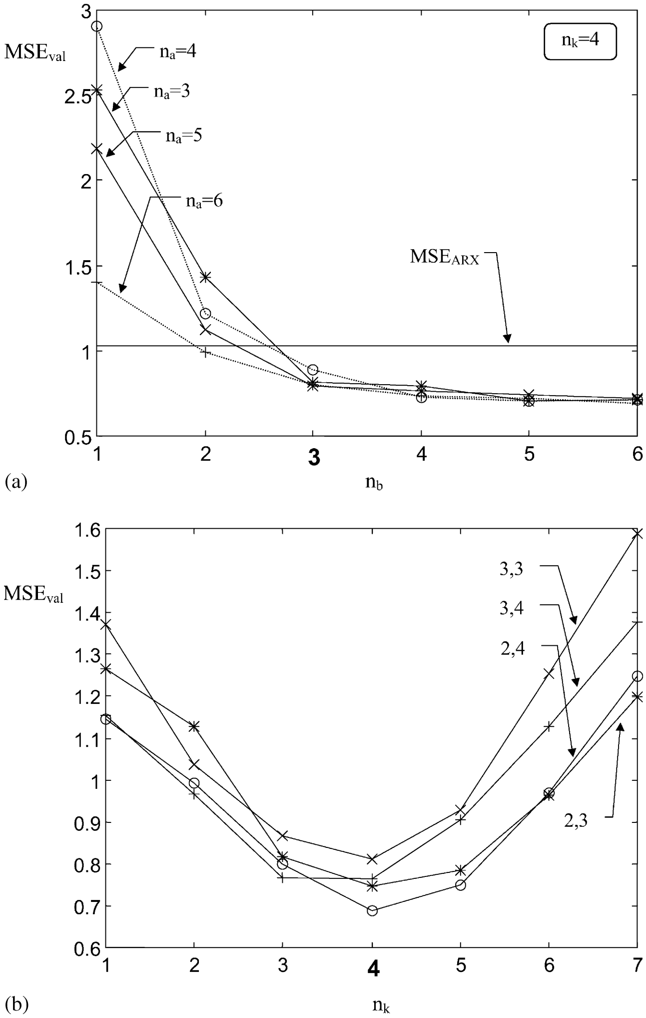


Fig. 10. (a) MSE on validation data vs. number of past inputs n_b at different values of number of past outputs n_a for the Fort Lauderdale models. In the plot the normalized criterion value for the linear ARX model is also shown. (b) MSE criterion vs. time delay n_k between input and output for some of the best fitting models.

- different values for n_k are tried using n_a and n_b obtained from the previous step; the best value for n_k is selected according to criterion (10) on the validation data set;
- the best values for n_a and n_b are selected according to criterion (10); in this selection the parsimony principle is applied (Ljung, 1987) in order to avoid overfitting models.

For the sake of brevity only the results about model identification for the Fort Lauderdale catchment are presented. In Fig. 10 the dependence of the normalized criterion (10) (computed on the validation data set) upon the number of past inputs and the value of the time delay is shown; each curve of Fig. 10a is obtained at fixed number of past outputs.

The procedure used for model order selection suggests the choice of three past outputs, three past inputs and a four time step delay. Ten repetitions of the training procedure have been performed and the mean value of criterion (10) on the validation data set is $J_1 = 2534.8$ (Fig. 11).

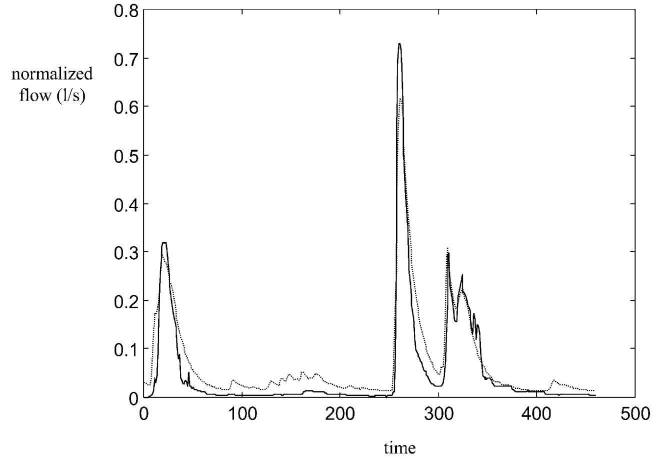


Fig. 11. Fort Lauderdale neural model (dashed) vs. validation data (solid).

8. Conclusions

The problem of deriving accurate models for the rainfall–runoff relation in urban drainage networks has been considered and a new procedure for the tuning of hydraulic conceptual models, based on black-box identification techniques has been developed and applied to experimental data.

It has been shown that, for the basins considered, satisfactory results can already be obtained by making use of linear models. Further performance improvements can be achieved by considering nonlinear models. In particular, two practical procedures for the structural identification of NARX polynomial and neural models have been proposed and applied to experimental data.

Acknowledgements

The authors would like to thank Prof. S. A. Billings for his useful feedback on some of the more technical issues in this paper.

References

Asproni, G. R., Fassò, C. A., & Putzu, G. (1981). Italian method for design of storms sewers. *Proceedings of 2nd international conference on urban storm drainage*, Urbana, Illinois, USA (pp. 475–481).

Bendat, J. S., & Piersol, A. G. (1980). *Engineering applications of correlation and spectral analysis*. New York: Wiley.

Billings, S. A., & Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear systems. *IEE Proceedings*, *130 D*, 193–199.

Billings, S. A., & Voon, W. S. F. (1986). Correlation based model validity tests for non-linear models. *International Journal of Control*, *44*, 235–244.

Billings, S. A., & Zhu, Q. M. (1994). Nonlinear model validation using correlation tests. *International Journal of Control*, *60*, 1107–1120.

Boukhris, A., Giuliani, S., Mourot, G., Querelle, R., Frank, P. M. (1997). Comparison of two fuzzy modelling approaches applied to the rainfall-runoff relationship. *Proceedings of the IFAC symposium on system identification SYSID '97*, vol. 3, Fukuoka, Japan (pp. 1513–1518).

Chen, S., Billings, S. A., & Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of Control*, *50*, 1873–1896.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). *Applied hydrology*. New York: McGraw-Hill.

Ciaponi, C., & Papiri, S. (1992). Una taratura del modello dell'invaso lineare per i bacini di drenaggio urbano. *Ingegneria Ambientale*, *21*(6), 331–343.

Desbordes, M. (1984). Evacuation des eaux pluviales urbaines: les orientaton francaises, researches et applications de la modelisation, perspectives d'evolution. *Proceedings of Seminar on Urban Drainage*, Cosenza, Italy (pp. 792–800).

Dooge, J. C. I. (1977). Problems and methods of rainfall-runoff modelling. In T. A. Ciriani, & J. R. Wallis, *Mathematical models for surface water hydrology*. New York: Wiley.

Frank, P. M. (1990). Fault detection in dynamic systems using analytical and knowledge-based redundancy — A survey and some new results. *Automatica*, *26*, 450–472.

Haber, R. (1985). Nonlinearity tests for dynamic processes. *Proceedings of the IFAC symposium on identification and system parameter estimation*, York, UK (pp. 409–414).

Harremoes, P., Jensen, M., & Johansen, N. B. (1984). A staged approach to application of rainfall data to urban runoff calculations. *Water Science and Technology*, *16*, 8.

Haverkamp, B., & Verhaegen, M. (1998). *SMI toolbox: State space model identification software for multivariable dynamical systems*, vol. 1.0. Delft, The Netherlands: Delft University of Technology. http://lcewww.et.tudelft.nl ~ haver/smi.html.

Jenkins, G. M., & Watts, D. G. (1968). *Spectral analysis and its application*. San Francisco, USA: Holden-Day.

Juditsky, A., Hjalmarsson, H., Benveniste, A., Delyon, B., Ljung, L., Sjoberg, J., & Zhang, Q. (1995). Nonlinear black-box models in system identification: Mathematical foundations. *Automatica*, *31*, 1725–1750.

Keesman, K., & Jakeman, T. (1997). Identification for long-term prediction of rainfall–streamflow system. *Proceedings of the IFAC symposium on system identification SYSID '97*, vol. 3, Fukuoka, Japan (pp. 1519–1523).

Ljung, L. (1987). *System identification: Theory for the user*. New Jersey, USA: Prentice Hall.

Löhnberg, P. (1978). Improved approximation of bias in squared coherence estimates for weakly smooth spectra. *IEEE Transactions On Acoustic, Speech, and Signal Processing*, *26*, 172–174.

Lovera, M. (1998). *Subspace identification methods: theory and applications*. Ph.D. Thesis, Politecnico di Milano, Italy.

Mambretti S., & Paoletti, A. (1996a). A new approach in overland flow simulation in urban catchment. *Proceedings of 7th international conference on urban storm drainage*, Hannover, Germany (pp. 965–970).

Mambretti, S., & Paoletti, A. (1996b). Il modello del condotto equivalente nella simulazione del deflusso superficiale in ambiente urbano. *Proceedings of Seminar on Urban Drainage*, Sorrento, Italy (pp. 269–306).

Narendra, K., & Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, *1*, 4–27.

Papoulis, A. (1973). Minimum-bias windows for high-resolution spectral estimates. *IEEE Transactions On Information Theory*, *19*, 9–12.

Pedersen, J. T., Peters, J. C., & Helweg, O. J. (1980). Hydrograph by single linear reservoir model. *ASCE Journal of the Hydraulic Division*, *106*(HY5), 837–852.

Previdi, F. (1999). *Identification and control using local linear models*, Ph.D. Thesis, Politecnico di Milano, Italy.

Sjoberg, J. (1995). *Non-linear system identification with neural networks*, Ph.D. Thesis, University of Linkoping, Sweden.

Sjoberg, J., & Ljung, L. (1995). Overtraining, regularization, and searching for minimum with application to neural networks. *International Journal of Control*, *62*, 1391–1407.

Sjoberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P. -Y., Hjalmarsson, H., & Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, *31*, 1691–1724.

Szafnicki, K., & Graillot, D. (1996). Knowledge-based real-time fault detection and supervision of urban drainage systems. *Automatica*, *32*, 1043–1047.

Thomson, M., Schooling, S. P., & Soufian, M. (1996). The practical application of a nonlinear identification methodology. *Control Engineering Practice*, *4*, 295–306.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input–output data. *Automatica*, *30*, 61–74.

Viberg, M. (1995). Subspace-based methods for the identification of linear time-invariant systems. *Automatica*, *31*, 1835–1852.