The Identification and Prediction of Urban Sewer Flows

A Preliminary Study

Beck, M. Bruce

1974

Document Version:
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Beck, M. B. (1974). *The Identification and Prediction of Urban Sewer Flows: A Preliminary Study*. (Research Reports TFRT-3118). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

THE IDENTIFICATION AND PREDICTION OF URBAN SEWER FLOWS

A PRELIMINARY STUDY

M.B. BECK

Report 7432 (C) November 1974
Lund Institute of Technology
Department of Automatic Control

THE IDENTIFICATION AND PREDICTION OF URBAN SEWER FLOWS

A PRELIMINARY STUDY

M.B. Beck

THE IDENTIFICATION AND PREDICTION OF URBAN SEWER FLOWS - A PRELIMINARY STUDY.

M.B. Beck †

ABSTRACT.

Dynamic models are presented for the relationships between rainfall-runoff and the influent flow to a wastewater treatment plant; the data are taken from the Käppala treatment plant, Lidingö, and the district surrounding the city of Stockholm. Previously, models for predicting sewer flows have been constructed upon a consideration of the many and complex physical characteristics that describe the nature of the system. The black box approach applied here to the identification of input/output relationships gives simpler models and yet it appears that the degree of empiricism inherent in this method is no more than that required to assign parameter values in the deterministic, mechanistic models.

The method of maximum likelihood is used for the identification and parameter estimation of stochastic models for the influent plant flow and recursive least-squares algorithms are employed in a suitable adaptive predictor for the same process. To have the facility for characterising the dynamic variations in the crude influx material to the plant is essential if the unit treatment processes, particularly those of sedimentation and activated-sludge, are

† Visitor, Division of Automatic Control, Lund Institute of Technology, and research fellow under the European Science Exchange Programme.

to be organised and controlled efficiently. Some discussion is made, therefore, on the practical implications for control given the possibility of using an on-line adaptive predictor for the influent flows.

TABLE OF CONTENTS

1. INTRODUCTION.

The analysis of flows in an urban sewer network is an integral link in the relationships between rainfall and the municipal use of water and the maintenance of quality in river and coastal waters (figure 1). Within these relationships three individual problems are recognised: (i) the determination of the input to a sewer network from surface runoff and municipal usage; (ii) the routing of flows across the network to outfalls and waste-water treatment facilities; (iii) the operation of treatment processes for the efficient removal of pollutant materials.

In the recent literature the dynamic properties of sewer network flows have been considered for the investigation of both the design and operational aspects of a sewer system and its adjacent treatment plants. It is relatively easy to characterise the variations in a dry-weather flow (DWF) resulting from municipal users' waste-water; these display observable daily, weekly, and seasonal fluctuations. On the other hand, the transient effects of a storm event, together with their attendant problems of untreated overflows to natural water courses, are more difficult to model: a factor reflected in the considerable attention given to the subject. Such events increase the loading on the treatment process in both a quantitative and qualitative manner and we may regard waste-water treatment as a process plant which receives a raw input material of an often widely-varying and largely ill-defined nature. For control synthesis it would be advantageous for treatment plant operation if the influent could be characterised more adequately and its major variations predicted in advance.

Hitherto models for the prediction sewer flows have assumed a large and purely deterministic nature resulting from the multitude of physical properties and phenomena which describe the system. In this paper we take a different approach using stochastic input/output models which are conceptually and in form simpler. The identification and prediction problems are tackled using the methods of maximum likelihood and recursive least-squares estimation, respectively; the data are taken from the Käppala treatment plant, Lidingö, and the district surrounding the city of Stockholm. It appears that the degree of empiricism inherent in this approach is no more than that required to assign parameter values to other models based on physical reasoning.

The study is part of a project supported by the Swedish Board for Technical Development (STU) whose aims are to examine improvements in the instrumentation and automatic control of waste-water treatment plants with particular reference to the activated sludge process [18]. Thus, the emphasis is placed upon constructing models of the influent flow for the purpose of plant control rather than the control and reduction of storm overflows from sewers. However, since the latter bears directly upon the former, and the two facets of the system are not really separable, it is given some consideration in the following. The discussion is by no means exhaustive and the results should be regarded as those of a preliminary investigation. It emerges that one particular limitation is the poor quality of the data; for this reason it is only feasible to consider the extensions to control analyses and application in a qualitative manner.

2. THE MODELLING AND CONTROL OF URBAN SEWER FLOWS.

For the purposes of the present study the survey of the available models for urban sewer flow dynamics is rather brief. Since it is our intention to use models based on a different conception of the system it is unnecessary to discuss here the details of previous approaches to the problem. Nevertheless, it should be possible to obtain a reasonable picture of the complex nature of deterministic models based on physical laws and arguments. The aspects of sewer flow and runoff modelling which are control-oriented are presented in Section 8.

Runoff from urbanised land surfaces.

An exemplary review and comparative study of methods for predicting inlet storm flows to an urban sewer network are given in Papadakis and Preul [19]. The several models developed over the past century are based upon considerations of the surface topography and physical characteristics, e.g. turf covering, asphalt covering, etc., together with the simulation of submodels for processes such as infiltration, surface retention, overland flow, and gutter flow. Verification results indicate that the models can reproduce successfully the storm water runoff transients in small drainage areas but there remain significant errors in the massively complex case of large urban watersheds. A complete account of a typical model for urban storm water runoff is presented by Chen and Shubinski [6]; the amount of physical detail included is quite remarkable. A more general treatment of land runoff in both rural and urban localities can be found, for example, in Offner [16].

The routing of flows through a sewer network.

Quite simply, as in figure 1, the output of a runoff model forms part of the input to a sewer network flow model. For the prediction of the influent flow to a treatment plant, theory requires that a set of partial differential equations be solved in the variables flow velocity and depth. Computationally, the method of characteristics may be employed, although for real-time applications to large and complex geometrical arrangements of a network the "progressive average-lag" method gives a faster but (acceptably) less accurate solution [9]. Factors to be supplied for the analysis include roughness coefficients, pipe-sizes and slopes, and some consideration of flow conditions at all points within the system.

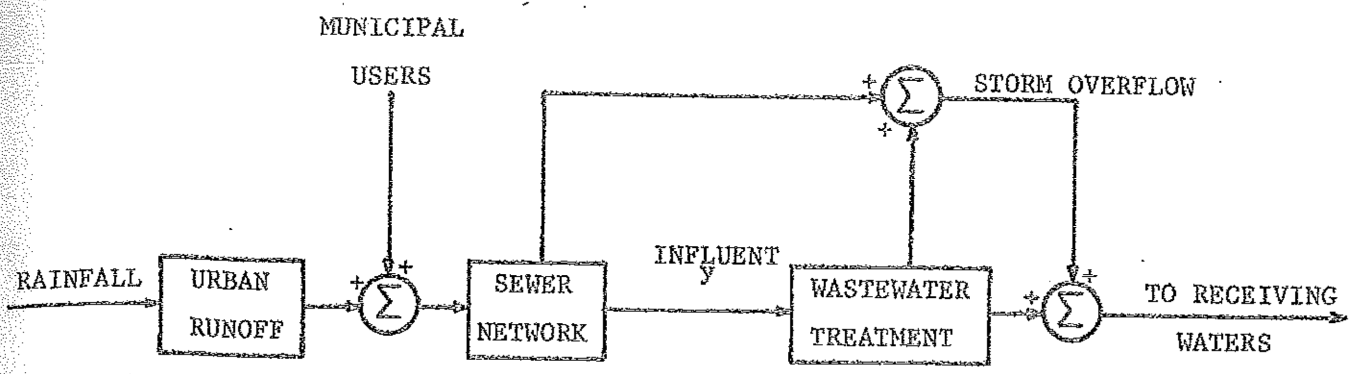


Figure 1 A schematic representation of a sewer network/wastewater treatment system

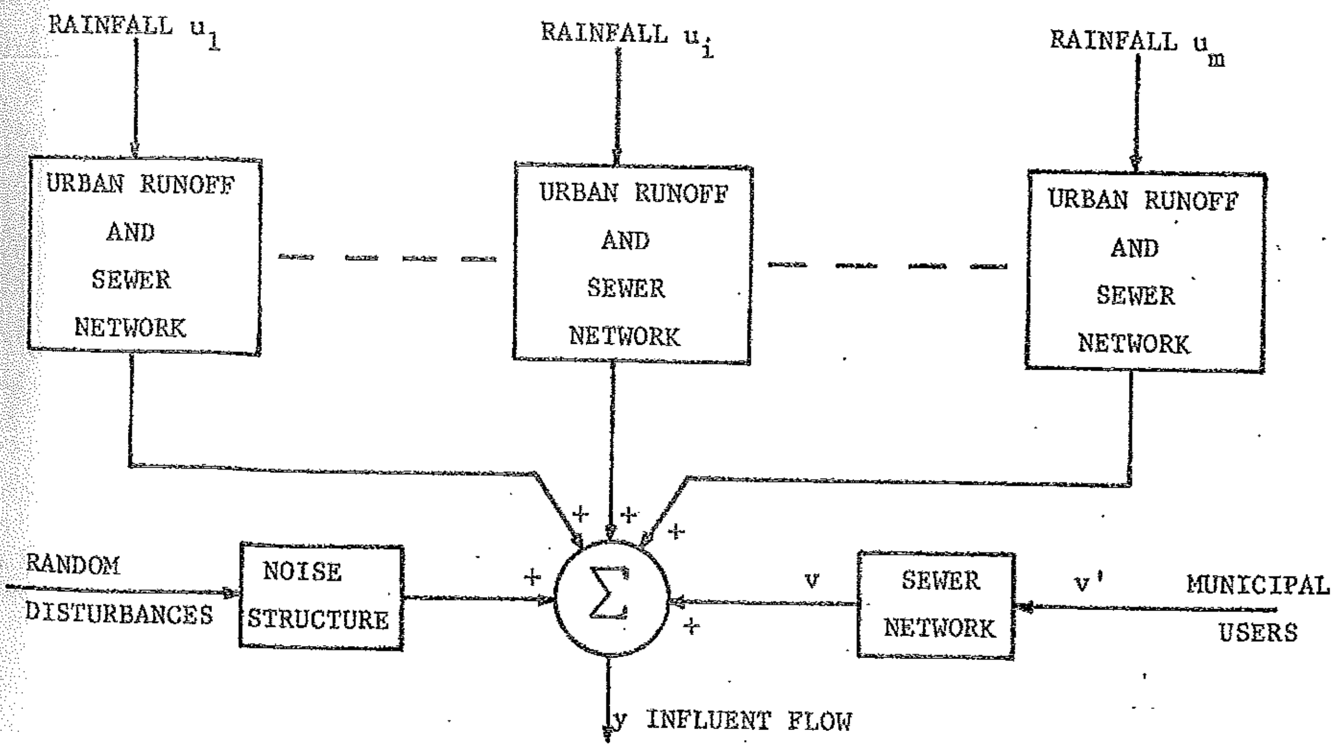


Figure 2 A schematic representation of the system for the identification of input/output flow models

3. THE PROBLEM FORMULATION.

The most striking feature of the currently available models for urban storm runoff and flow routing through sewer networks is their purely deterministic and complex structure. Of course, the problem is inherently complex since it is governed by considerations of highly diversified surface topographies and coverings together with rainfall events whose intensities vary spatially and temporarally and which travel spatially in time. Especially with the advent of the digital computer the approach has been to include as much detail as possible; this has led to a better understanding of the processes involved. Yet a theoretically complete analysis would produce an unwieldy and largely intractable mathematical model with a multitude of parameters required to be estimated or evaluated. Indeed, particularly in models for urban runoff it is acknowledged that "to a varying degree most of these methods rely upon empirical relationships and experience" [19].

It seems, therefore, that a stochastic model derived from time-series analysis, herein the maximum likelihood method [4], can yield equally usable results. Such an identification procedure assumes little a priori knowledge of the physical system and takes a relatively macroscopic view of the cause-effect relationships involved. As we have remarked earlier, many previous studies evolved from design problems of sewer networks or from a desire to reduce stormwater overflow. Our current emphasis, however, is upon the characterisation of the quality and flow of the influent to a waste-water treatment plant with a view to improved operational control of the unit processes for waste removal.

Thus, given the combined sewer network of the Käppala Reningsverk in the city of Stockholm (section 4) the problem is formulated as follows:

(i) To identify input-output models for the dynamic relationships between the input, rainfall u_i, at several spatial locations ($i = 1, 2, \ldots, m$) and the output, the influent flow to the Käppala plant, y (see figure 2); to examine whether any particular dynamical properties attach to individual locations of the input at i; or, alternatively, to examine the feasibility of using a single spatially-averaged rainfall input.

(ii) To construct an adaptive predictor for the influent y given an auxiliary signal v or v' for the flow resulting from municipal wastes; to evaluate the benefit of using additional measurements of u in such a predictor.

(iii) To investigate the use of models or predictors for the feedback of information on y in sewer network flow control synthesis and the feed forward of information on y in unit treatment process operation control.

As it is given, the problem formulation has been posed in a fairly general sense. Even so, we have not considered the rider problems of applying the same kind of analysis to quality variables or the identification of submodels (e.g. transfer functions) between intermediate points in the sewer network. It is quite likely (see section 8) that the dynamics and control of the influent flow to the plant are sensitive to those of certain sections of the system.

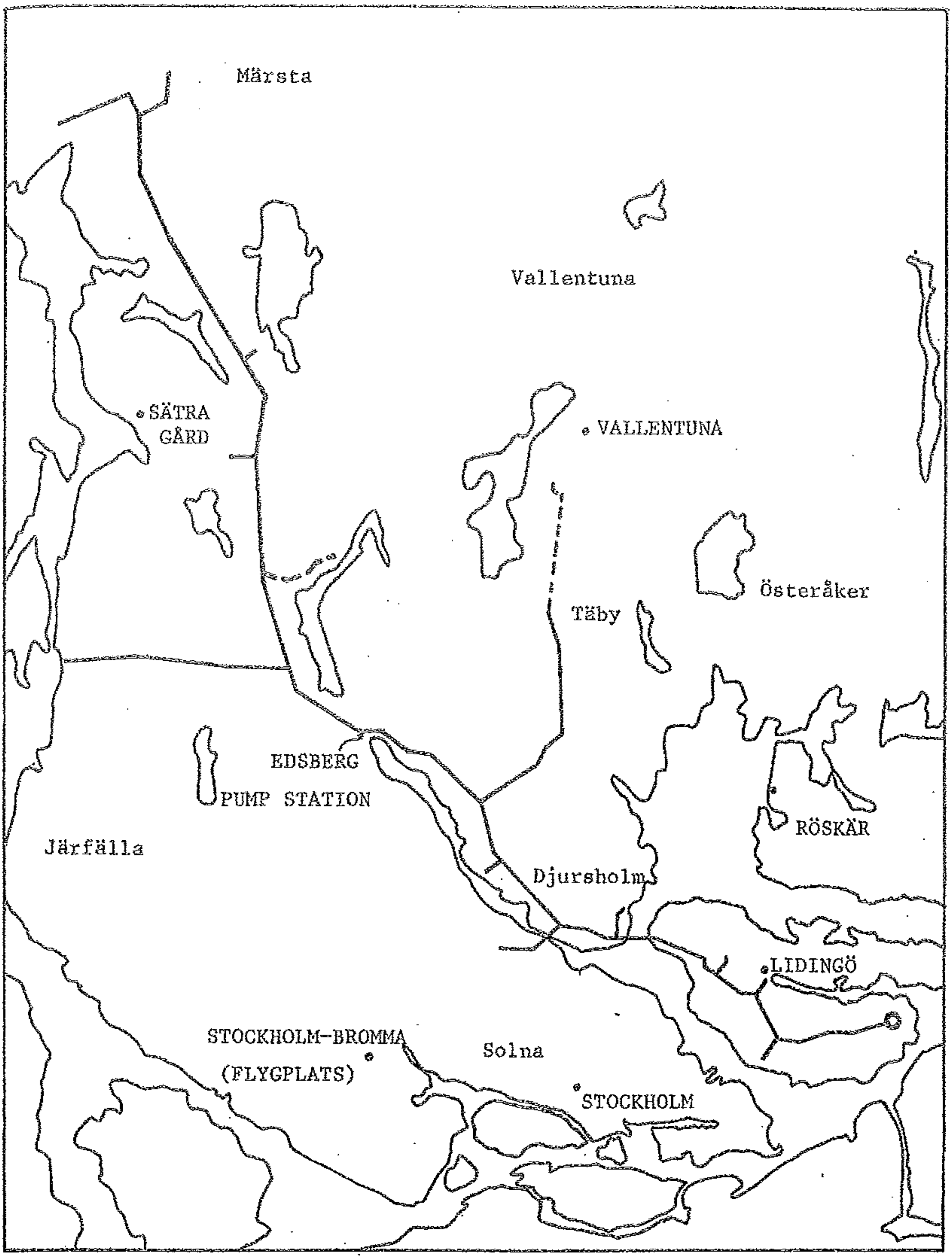
Furthermore, we should bear in mind that black box models are often specific to the data from which they are derived and are not necessarily amenable to a physical interpretation. However, experience shows that in many situations it is indeed possible to draw inferences on the nature of the physical system and it is encouraging to refer to the successful application of similar techniques in the analogous problem of predicting river flows in real time [11], [22].

4. THE KÄPPALA SEWER TUNNEL, STOCKHOLM.

Figure 3 shows the sewer tunnel system which collects waste-water from an area of some 1191 km^2 covering part of the city of Stockholm and its surrounding districts and serving a population of 290,000. For the purpose of this study it can be assumed that none of the flow is diverted from the Käppala tunnel after entry from subsidiary sewers. The Käppala treatment plant (design DWF 2 $m^3 s^{-1}$) includes several modern features, being commissioned in 1969 and employing a Siemens 304 process computer for data-logging and some control.

Within the sewer system continuous flow records are taken from 17 locations using Parshall flumes and the Käppala Works receives hourly sampled data on the mean influent flow to the plant. As with many other variables monitored by the plant, historical data on the influent flow are stored on the process computer for up to a period of three months behind the current time. 24-hr rainfall, and in some cases 12-hr and shorter, sampled observations are available from a group of ten stations within the Stockholm area.

An approximate assessment of transportation delays within the sewer system can be obtained from isolated tracer experiments; for instance, at a typical mean flow-rate of 1.32 $m^3 s^{-1}$ the mean residence time of the sewer flow between Edsberg and Käppala (20 km) is close to 8 hrs. This gives some idea of the time available for "feed forward" prediction of the influent flow if required for on-line purposes.



• = Meteorological station

o = Käppala treatment plant

Figure 3 The Käppala sewer tunnel system

5. IDENTIFICATION AND PREDICTION.

The class of models to be examined is one of parametric, linear, time-invariant models of a canonical form. They are black box models in the sense that they assume no knowledge of physical relationships between the system's inputs and output other than that the inputs should produce observable responses in the output. Our concern is with the identification of rainfall-urban sewer flow dynamics and the adaptive prediction of the influent flow to a treatment plant, see figure 2, assuming that the system is subject to stochastic disturbances and random errors of measurement.

5.1. Maximum Likelihood Identification.

In the general case, given the set of input/output data samples $\{u_i(t),\ i = 1, 2, \ldots, m;\ y(t);\ t = 1, 2, \ldots, N\}$, where $u_i(t)$, $i = 1, 2, \ldots, m$ are the m input signals, $y(t)$ is the output signal and t is the time of the t^{th} sampling instant, the identification problem is to find an estimate of the parameters of the system model [7]

$$A(q^{-1})y(t) = \sum_{i=1}^{m} B_i(q^{-1})u_i(t) + \lambda C(q^{-1})e(t) \qquad (1)$$

in which e(t) is a sequence of independent, normal (0,1) random variables and q denotes the shift operator

$$q\{y(t)\} = y(t+1) \quad \text{etc.} \qquad (2)$$

$A(q^{-1})$, $B_i(q^{-1})$, $i = 1, 2, \ldots, m$, and $C(q^{-1})$ are the polynomials

$$\left.\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \ldots + a_n q^{-n} \\ B_i(q^{-1}) &= b_{i0} + b_{i1} q^{-1} + \ldots + b_{in} q^{-n} \qquad i = 1,2,\ldots,m \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \ldots + c_n q^{-n} \end{aligned}\right\} \quad (3)$$

The residual errors of eqn. (1), $\{\varepsilon(t),\ t = 1, 2, \ldots, N\}$, defined by

$$C(q^{-1})\varepsilon(t) = A(q^{-1})y(t) - \sum_{i=1}^{m} B_i(q^{-1})u_i(t) \qquad (4)$$

are thus an independent and normal $(0,\lambda)$ sequence. The logarithm of the likelihood function is now

$$L = -\frac{1}{2\lambda^2}\sum_{t=1}^{N} \varepsilon^2(t) - N \log \lambda + \text{constant} \qquad (5)$$

and the maximisation of L is equivalent to minimising the loss function

$$V(\theta) = \frac{1}{2}\sum_{t=1}^{N} \varepsilon^2(t) \qquad (6)$$

where θ is the column vector of parameters in the model, eqn. (1),

$$\theta^T = [a_1,\ldots,a_n,\ b_{10},\ldots,b_{1n},\ b_{20},\ldots,b_{mn},\ c_1,\ldots,c_n] \quad (7)$$

(superfix T denotes the transpose of a vector or matrix). When $\hat{\theta}$ has been found, such that $V(\hat{\theta})$ is minimal, the maximum likelihood estimate of λ is given by

$$\hat{\lambda}^2 = \frac{2}{N} V(\hat{\theta}) \qquad (8)$$

Strictly speaking, the model of eqn. (1) applies only to stable, linear, time-invariant systems. A more complete discussion of the model structure, the minimisation of the loss function $V(\theta)$, and the conditions for the estimates to be consistent, asymptotically normal and efficient are given in the original source references [4], [5], [7].

Notice that in this application the inputs of the system correspond to the rainfall u_i measured at the locations $i = 1, 2, \ldots, m$ and the output y is the influent flow to the treatment plant.

5.2. Adaptive Prediction.

For the derivation of an adaptive predictor [23] let us consider the discrete-time stochastic process,

$$A(q^{-1}) = \lambda C(q^{-1})e(t) \tag{9}$$

with all variables defined as for the model of eqn (1); at this stage no knowledge is assumed for any deterministic signals u which may be related to the time-series y. Now denote the k-step ahead prediction of the output signal y based on the sampled observations $y(t), y(t-1), \ldots,$ by $\hat{y}(t+k|t)$. Introducing the loss function

$$V_k(t) = E\{\varepsilon(t+k)^2\} \tag{10}$$

where $E\{\ldots\}$ is the expectation operation and $\varepsilon(t+k)$ is the prediction error

$$\varepsilon(t+k) = y(t+k) - \hat{y}(t+k|t) \tag{11}$$

the prediction problem is solved in [3] for the process eqn. (9) having known A and C polynomials. Alternatively, if these parameters are unknown they can be estimated according to the procedure outlined above and then used in the construction of a predictor.

However, for an adaptive predictor of an unknown process it is not required that the parameters of that process, e.g. eqn. (9), be estimated explicitly; rather, the parameters of the predictor itself are estimated. Wittenmark [23] solves the problem by transforming it into the previously solved problem of an adaptive regulator; the adaptive predictor algorithms for an unknown process with constant parameters can then be derived by separation of the estimation and prediction steps.

In the particular application to be discussed here a slightly modified version of Wittenmark's algorithms are used since there is some indication that they have better stability properties, see [10].

Thus, for the adaptive predictor of the process, eqn. (9), we have:

Step 1: Estimation.
At time t, upon the receipt of a new observation y(t), the parameters $\alpha_1,\ldots,\alpha_p$, $\beta_1,\ldots,\beta_\ell$ are estimated in the model

$$y(t) = \alpha_1 y(t-k) + \ldots + \alpha_p y(t-k-p+1) - - \beta_1 \hat{y}(t-1|t-k-1) - \ldots - \beta_\ell \hat{y}(t-\ell|t-k-\ell) \qquad (12)$$

by the method of least squares;

Step 2: Prediction.

The estimates obtained in Step 1 are used to calculate the predicted output as

$$\hat{y}(t+k|t) = \alpha_1 y(t) + \ldots + \alpha_p y(t-p+1) -$$

$$- \beta_1 \hat{y}(t+k-1|t-1) + \ldots + \beta_\ell \hat{y}(t+k-\ell|t-\ell) \qquad (13)$$

More concisely, if we define the polynomials

$$A(q^{-1}) = \alpha_1 + \alpha_2 q^{-1} + \ldots + \alpha_p q^{-p+1}$$

$$B(q^{-1}) = \beta_1 q^{-1} + \beta_2 q^{-2} + \ldots + \beta_\ell q^{-\ell}$$

the adaptive predictor is given by

Estimation: $y(t) = A(q^{-1})y(t-k) - B(q^{-1})\hat{y}(t|t-k) \qquad (14)$

Prediction: $\hat{y}(t+k|t) = A(q^{-1})y(t) - B(q^{-1})\hat{y}(t+k|t) \qquad (15)$

In other words, the adaptive predictor is a function of both the past observations and the previous predictions of the output. Since a least-squares estimation is readily implemented in recursive form the predictor is well suited to real-time applications with each step of the procedure being repeated at each sampled instant of time. This is found to be feasible even if $C(q^{-1}) \neq 1$ in eqn. (9), in which case the least-squares parameter estimates of eqn. (12) may be biased.

The algorithms are fairly flexible such that, for example, trends in the process can be treated without difficulty. But of more importance here, time-varying parameters may be accounted for by introducing suitable exponential weighting of the data. Further, auxiliary variables $v_i(t)$ improve the prediction if periodic functions are present in the process $y(t)$ or additional measurements of another signal disturbing the system are available. Thus, defining the polynomials

$$C_i(q^{-1}) = \gamma_{i1} + \gamma_{i2}q^{-1} + \ldots + \gamma_{ir}q^{-r} \quad i = 1, 2, \ldots, m'$$

eqns. (14) and (15) become

$$\text{Estimation:} \quad y(t) = A(q^{-1})y(t-k) - B(q^{-1})\hat{y}(t|t-k) + \\ + \sum_{i=1}^{m'} C_i(q^{-1})v_i(t) \qquad (16)$$

$$\text{Prediction:} \quad \hat{y}(t+k|t) = A(q^{-1})y(t) - B(q^{-1})\hat{y}(t+k|t) + \\ + \sum_{i=1}^{m'} C_i(q^{-1})v_i(t+k) \qquad (17)$$

In this case $v_i(t)$ may be some representation of the periodic variation in the influent flow $y(t)$ caused by the waste from municipal users, or it may be a measurement of the rainfall (at a point location, or spatially averaged).

6. MAXIMUM LIKELIHOOD IDENTIFICATION OF RAINFALL-URBAN SEWER FLOW RELATIONSHIPS.

In this section part (1) of the problem formulation is addressed. The observed influent flow-rate to the Käppala plant is analysed for the period between October 1st (08.00 hrs) and October 31st (07.00 hrs), 1973, a total of 720 hourly samples. For the same interval data have been obtained from S.M.H.I. (Sveriges Meteorologiska och Hydrologiska Institut), Stockholm, and these relate to the rainfall measured at the four stations, Röskär, Stockholm-Bromma (flygplats), Stockholm, and Lidingö (see fig. 3). While these data reflect both typical rainfall events and dry-weather flows, they are not representative of any short and intense storm conditions.

Figure 4 shows the data for the defined period; the influent flows are computed from depth measurements and flow measurements downstream of the pumps to the Käppala plant. The rainfall is a "spatially"-averaged time-series for the four stations: at Röskär sampled values are calculated by distributing a 24 hr measurement equally among hourly observations on the timing of the precipitation (except for the interval 00.00 - 07.00 hrs); at Stockholm-Bromma, Stockholm, and Lidingö, the measured precipitation is distributed equally at each hourly sampling instant from observations taken at 12 hr, 12 hr, and 24 hr intervals, respectively.

In figure 5 a larger scale plot of the first week's data gives a good indication of the nature of typical dry-weather flows. The data display several features of an easily recognisable form. Firstly, there is the daily pattern with an approximately half sine-wave structure; working days display fluctuations of similar magnitudes, while the flows for Saturday and Sunday have a somewhat lower

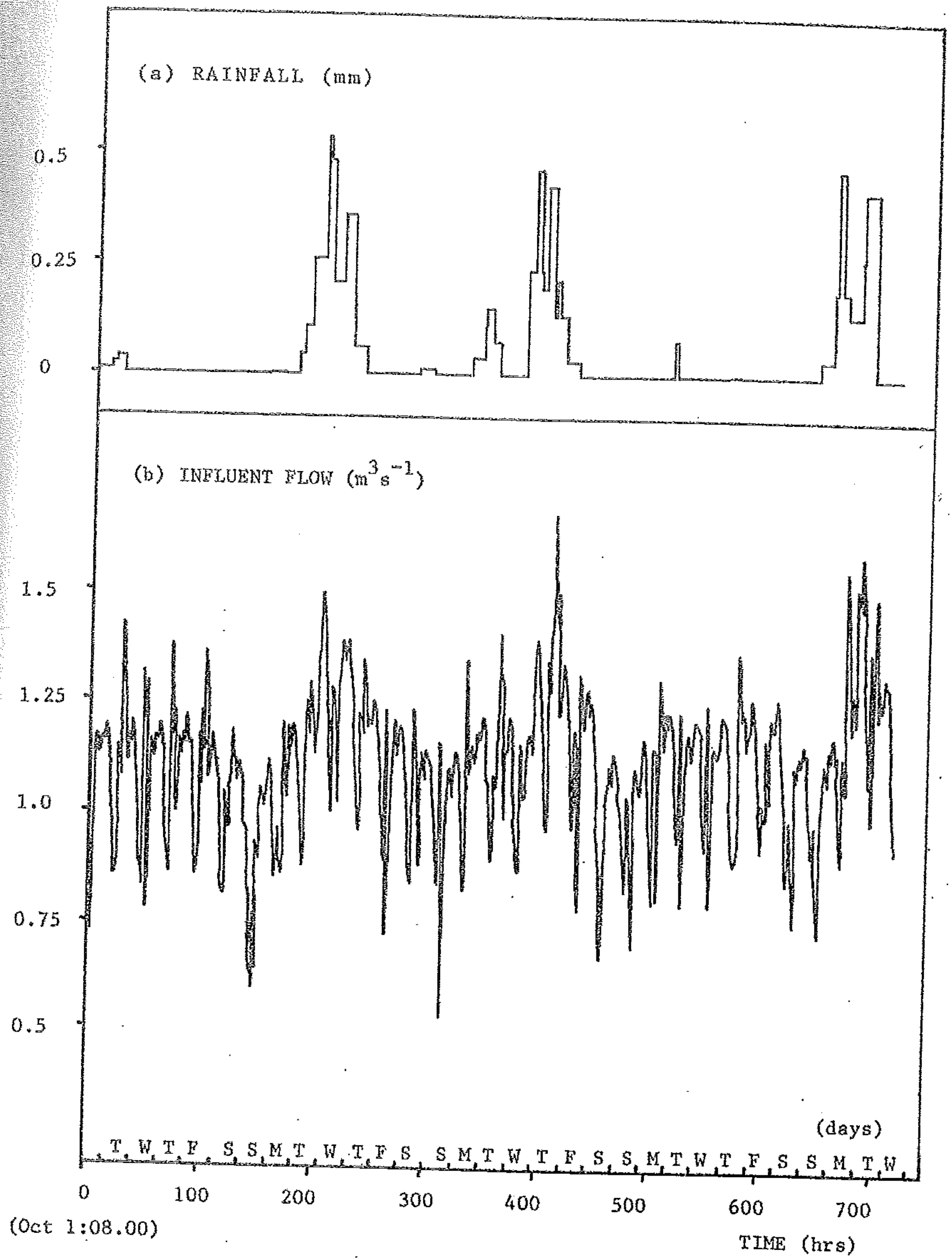


Figure 4 Rainfall (spatially averaged) and influent flow data for October 1973

amplitude and mean level. Referring again to figure 4 it is observable that this weekly process is repeated throughout the month with the runoff from rainfall events producing a temporary increase in flows as expected. The daily and weekly periodicities are correspondingly visible in the autocorrelation function of the influent flow data, figure 7.

Both figures 4 and 5 show the large "spikes" superimposed on the daily variations which result from pumping operations. Such a procedure enables the treatment plant to operate on a more or less constant flow-rate throughout the major part of the day by detaining the natural flow in a buffering well upstream of the influent gates and regularly (usually between 08.00 and 10.00 hrs) releasing the accumulated excess flow. Figure 6 gives the pattern of this manipulated flow to the treatment plant over the same interval as the natural flows of figure 5. That the pumping effects are still significant in the computed natural influent flow is a function of the sampled nature of the data and unavoidable inaccuracies in the relationship between the stored volume of sewage and depth measurements.

Indeed, among other factors, the large and relatively high-frequency disturbances imposed artificially by the pumping operations are a considerable drawback in the analysis of the data. Since lower-frequency oscillations of a deterministic nature, particularly the daily cycle, are recognisable in the time-series it might be appropriate to difference the data accordingly and apply the identification precedure to the resulting output process $\Delta y(t)$ given by

$$\Delta y(t) = y(t) - y(t-24)$$

However, it is well known that this tends to emphasise the level of noise in the data. Thus, we prefer to adopt alter-

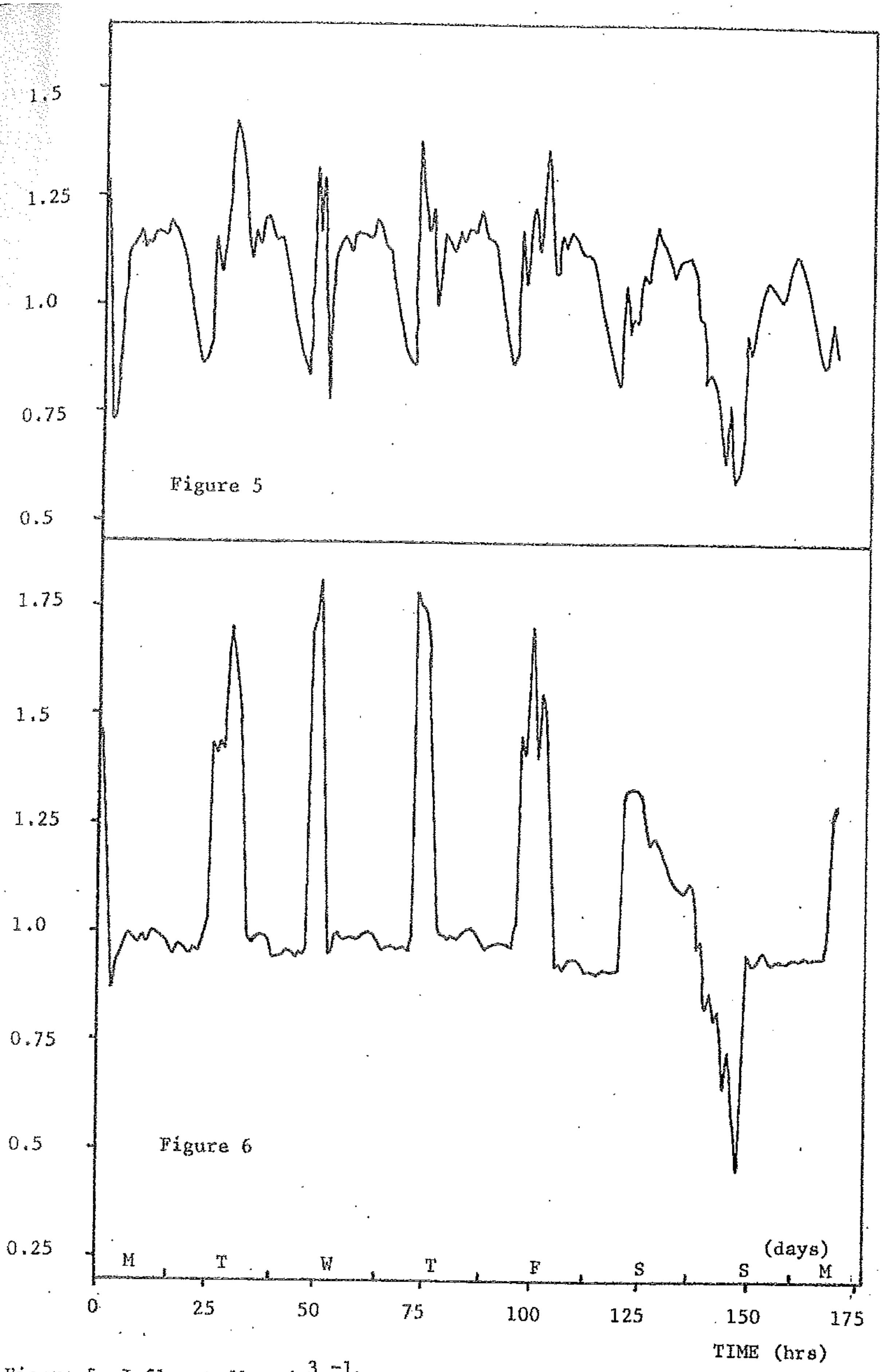


Figure 5 Influent flow ($m^3 s^{-1}$) to Käppala for first week of October

Figure 6 Pumped influent flow ($m^3 s^{-1}$) to Käppala for first week of October

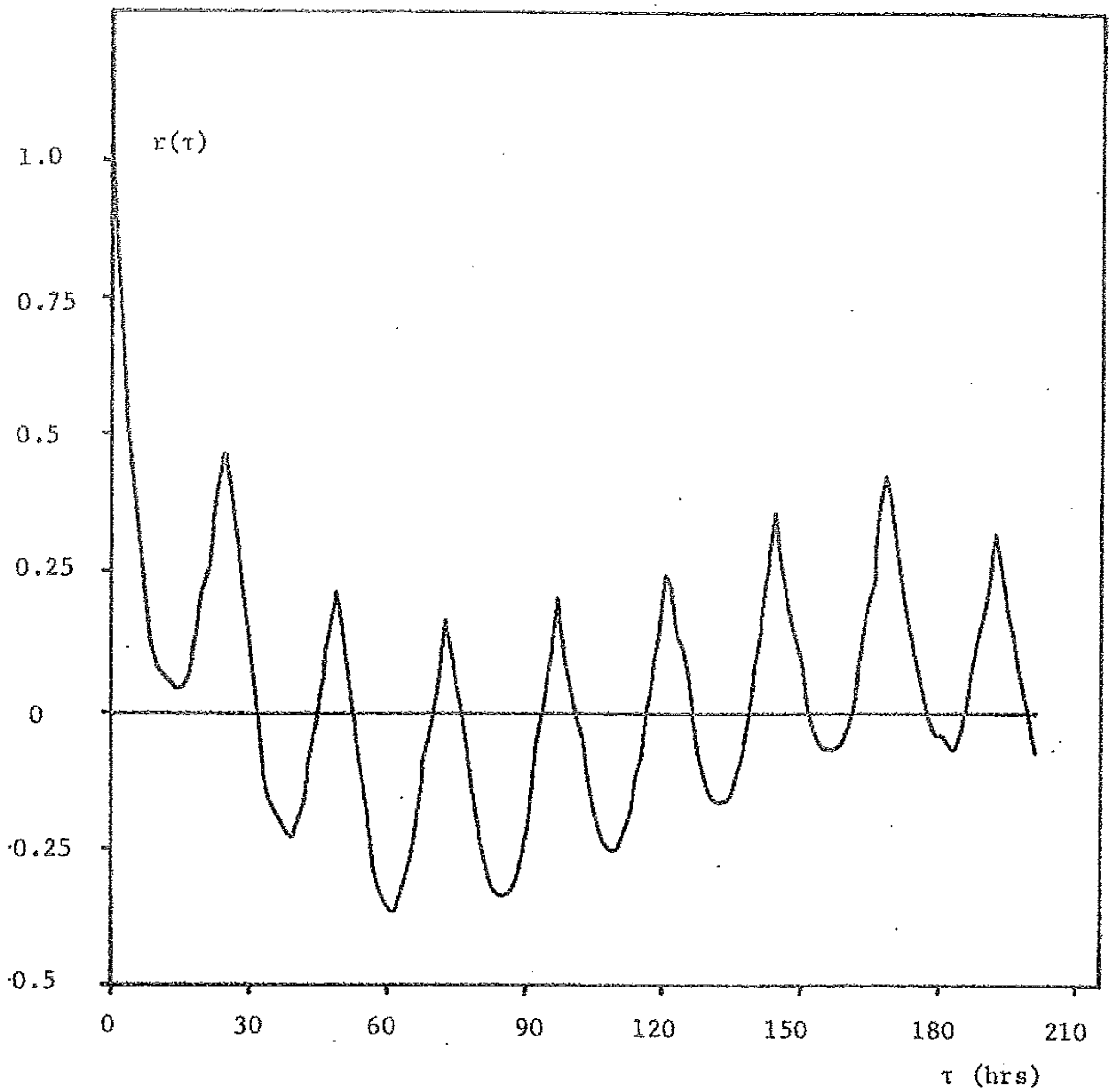


Figure 7 Sample autocorrelation function r(τ) for the influent flow data

alternative approaches using data which has been low-pass filtered in order to smooth the pumping disturbances; the rainfall time-series is filtered in an identical manner to preserve the original input-output relationships.

Given that the real problem of identification is the separation between the dry-weather flows and the excess flows from rainfall-runoff sources, two types of model are proposed. The first, called a general flow (GF) model, consists of a stochastic model for both kinds of flow; the second, a rainfall-runoff flow (RRF) model, uses a deterministic model to describe the dry-weather flows and a stochastic model to describe the additional effects of runoff.

6.1. A General Flow (GF) Model.

Table 1 and figure 8 give the identification results for a GF model. The output and input time-series are suitably smoothed using a discrete low-pass filter, see [8], of first order and with a cut-off frequency of 0.9 rad/hr. The d parameters of table 1 are estimates of initial values y(0), y(-1) of the output and their inclusion improves the general accuracy of the identification.

Table 1 - 2nd-order model: $y(t)\,|u_1(t),\ u_2(t)$

a_1	-0.916 ± 0.014	c_1	0
a_2	0	c_2	0
b_{11}	0.072 ± 0.018	d_1	0.880 ± 0.045
b_{12}	0	d_2	0
b_{21}	0.418 ± 0.030	λ	0.045 ± 0.001
b_{22}	-0.337 ± 0.030		

The notation $y(t)|u_1(t), u_2(t)$ indicates a single output/ two input system where $y(t)$ is the (low-pass filtered) influent flow-rate to the treatment plant (m^3s^{-1}), the input $u_1(t)$ is the spatially-averaged (and filtered) rainfall (mm) and the signal $u_2(t) = y(t-23)$, in this particular model, is an artificial input. This means that the output, nominally observed 24 hrs behind the current time, is fed forward to give a characterisation of the periodic (or approximately dry-weather flow) functions in the process $y(t)$. Consequently, differencing the data is avoided.

Hence, substituting the estimates of table 1 into the deterministic part of eqn. (1), we have, after rearrangement,

$$y(t) = 0.916y(t-1) + 0.418y(t-24) - $$

$$- 0.337y(t-25) + 0.072u_1(t-1) \qquad (18)$$

The model output of figure 8(c) is then given by eqn. (18) with $y(0) = d_1$ and the observed values of $y(t-24)$, $y(t-25)$ and $u_1(t-1)$ inserted at each discrete-time instant t. Clearly, the model errors do not exhibit a good correspondence between the purely deterministic model output and the observations of the influent flow-rate.

On the other hand the residuals, $\varepsilon(t)$, which can be interpreted as the prediction errors of the one-step ahead predictor

$$\varepsilon(t) = y(t) - \hat{y}(t|t-1) \qquad (19)$$

are quite small and have a standard deviation of ± 0.045 (m^3s^{-1}). Yet it is noticeable from figure 8(e) that the residuals contain periodic high frequency spikes and these are synchronous with the residues of the smoothed pumping effects and the "discontinuous" nature of the half sine-wave

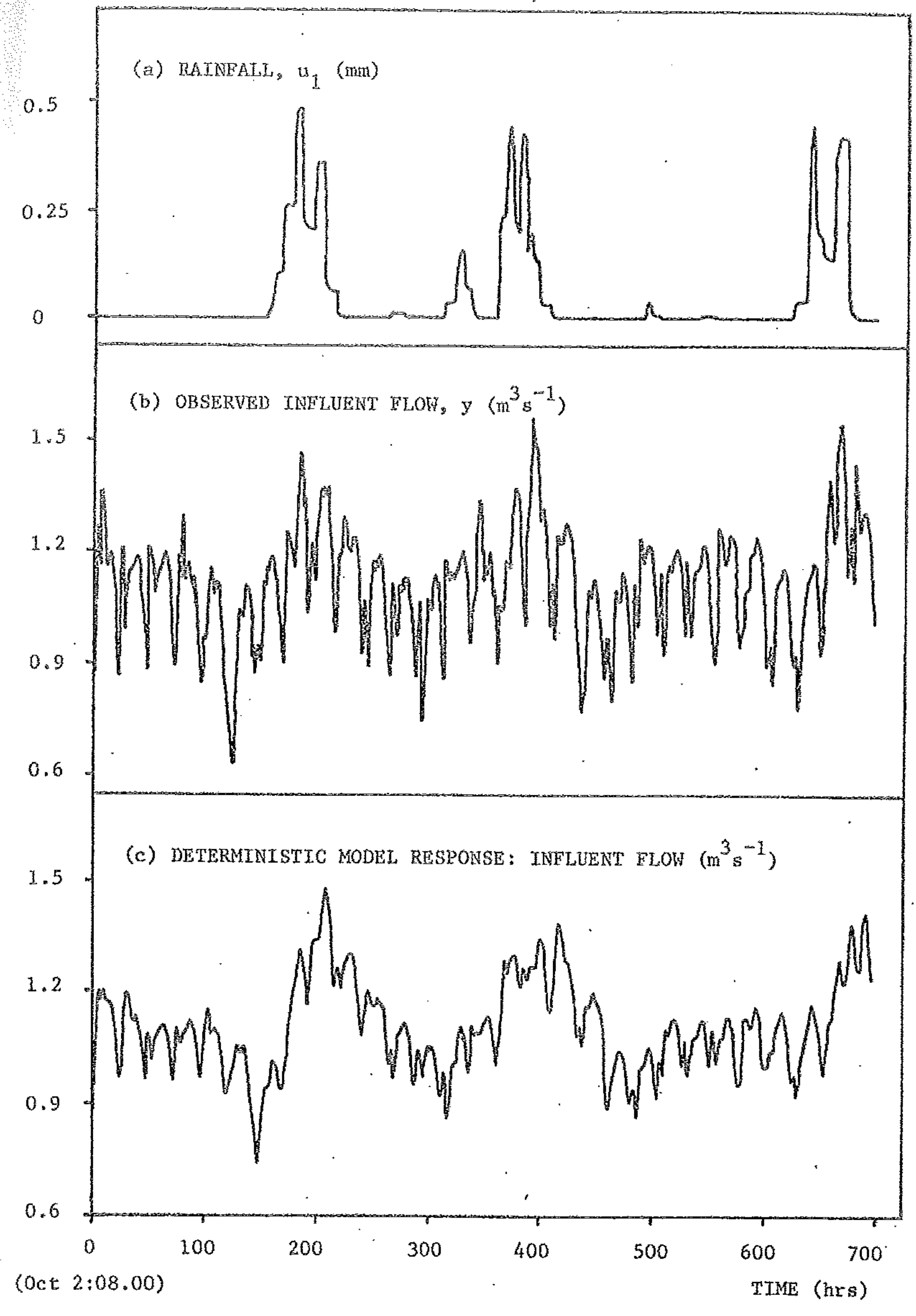


Figure 8 Maximum likelihood identification results for the GF model of table 1 (continued overleaf)

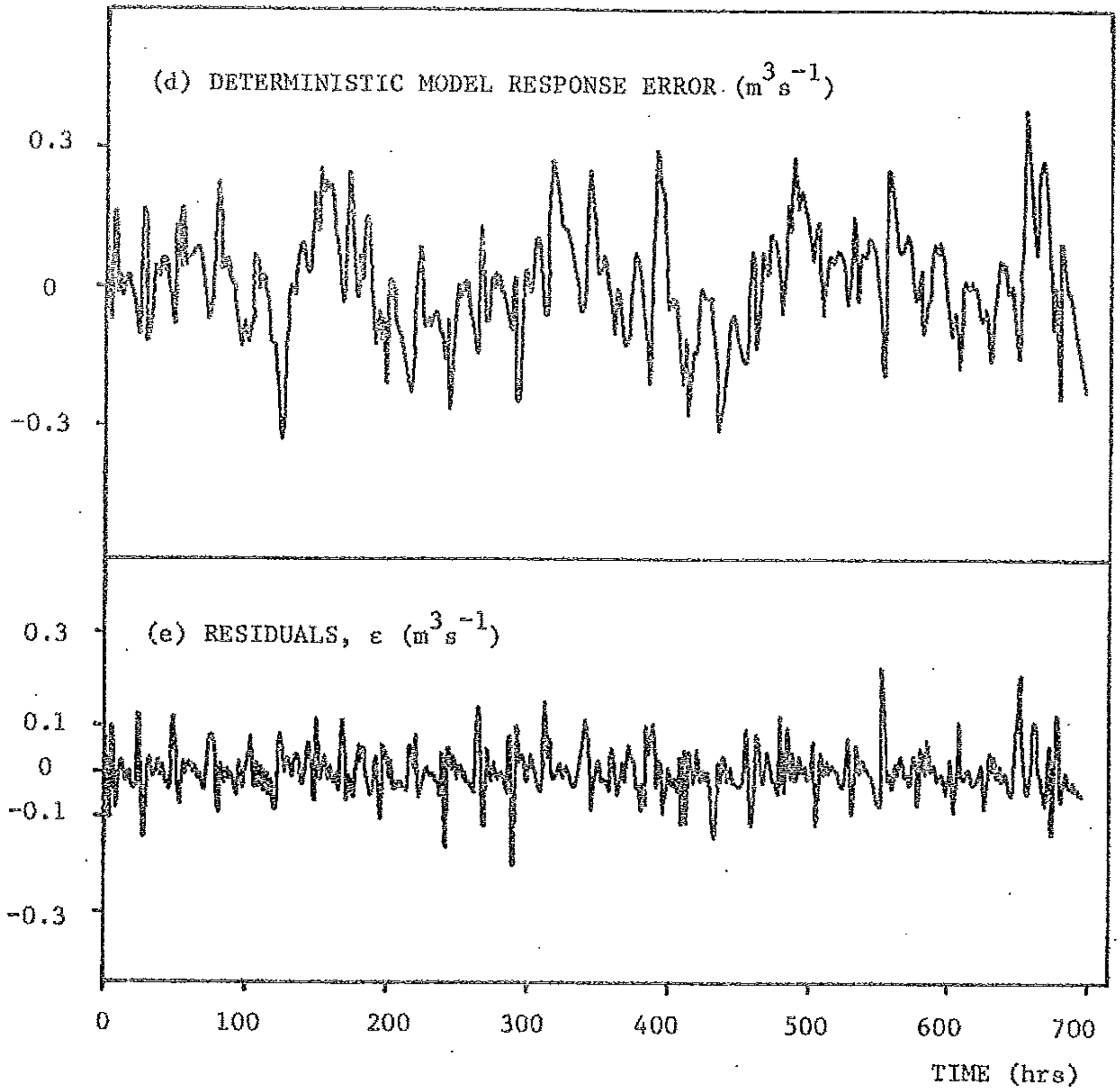


Figure 8(contd) Maximum likelihood identification results for the GF model of table 1 pattern of the flow data. More precise statistical tests of the residuals [8] confirm that they are significantly correlated in time, not normally distributed and correlated with the signal $u_2(t)$: factors which indicate that the identified model is not consistent with the assumptions for the validity of eqn. (1). Of course, some of these anomalies might be accounted for by the use of a least-squares estimation $\left(C(q^{-1}) = 1 \text{ in eqn. } (1)\right)$. However, any attempt to estimate $C(q^{-1}) \neq 1$ proves abortive since the C polynomial is found to be unstable.

The choice of two b_{2j} coefficients in the model of table 1 gives better properties of the residuals than the corresponding first-order model. But, owing to the low magnitudes of these "gains" in the feedforward process, the added complexity of estimating the coefficients of a third input, say

$$u_3(t) = -\,u_1(t-23) \qquad (20)$$

affords negligible improvement in the results. In theory, this should eliminate the spurious feedforward, through $u_2(t) = y(t-23)$, of excess flows resulting from rainfall runoff on the previous day.

Notice that the particular artificial signal $u_2(t)$ here is only one of several options

$$u_2(t) = \begin{cases} y(t-24) & \text{(i)} \\ y(t-168) & \text{(ii)} \\ \bar{y}_d(t) & \text{(iii)} \\ \bar{y}_w(t) & \text{(iv)} \end{cases} \qquad (21)$$

which might be used; $\bar{y}_d(t)$ and $\bar{y}_w(t)$ are mean daily and weekly periodic time-series, respectively, with

$\bar{\bar{y}}_d(t) = \bar{\bar{y}}_d(t-24)$

$\bar{\bar{y}}_w(t) = \bar{\bar{y}}_w(t-168)$

and can be computed a priori from the available data *. Different choices are used in other models (sections 6.2 and 7) but, for reasons which are discussed below, they are likely to be of limited value in a GF model.

6.2. A Rainfall-Runoff Flow (RRF) Model.

Evidently the structure of a GF model fails to describe adequately the periodic dry-weather flow variations in the data. However, observing in figure 4 that the first and fourth weeks of the data, samples 1 → 168 and 505 → → 672 inclusive, represent virtually dry-weather conditions, it is possible to compute a mean weekly dry-weather flow pattern $\bar{\bar{y}}_w(t)$. Operating again on the low-pass filtered observations, a time-series

$$y_r(t) = y(t) - \bar{\bar{y}}_w(t) \qquad (22)$$

is prepared, where $y_r(t)$ may be considered loosely as the excess flow resulting from rainfall-runoff sources.

We can now identify a single input/single output stochastic model for $y_r(t)|u_1(t)$ having removed a priori the deterministic dry-weather and periodic flow components. In fact, the results for this RRF model, given in table 2 and figure 9, apply to the model $y_r(t)|u_1'(t)$, $u_2'(t)$ in

---

* In longer term forecasting etc. these procedures may have to include some account of seasonal variations.

Table 2 - 1st-order model: $y_r(t)|u_1'(t)$, $u_2'(t)$

a_1	-0.739±0.028
b_{11}	0.063±0.032
b_{21}	0.086±0.031
c_1	0.984±0.007
λ	0.027±0.001

which $u_1'(t) = u_1(t-2)$ and $u_2'(t) = u_1(t-6)$; this is simply a more convenient way of estimating what would otherwise be a 6th-order system.

The choice of the model structure requires some explanation. Typically, the identification of first-order models $y_r(t)|u_1'(t)$ where

$$u_1'(t) = u_1(t-i) \qquad i = 0, 1, \ldots, 7$$

gives less than a 1.5% variation in the minimised loss function $V(\hat{\theta})$ of eqn. (6) and the largest difference in the estimates (of the b coefficient) is little more than 10% for any two values of i. Accordingly, the estimation of higher-order models, with low values of i, yields only one significantly non-zero b parameter. This is not surprising when we consider the nature of the rainfall data $u_1(t)$ which inherently distributes any precipitation event over a wide interval of one-hourly samples. The model of table 2, therefore, is no more than a counterbalance to the data and equally good results could have been obtained with other suitably separated inputs u_1' and u_2'.

The statistical properties of the residuals are similar to, but slightly better than, those of the GF model, al-

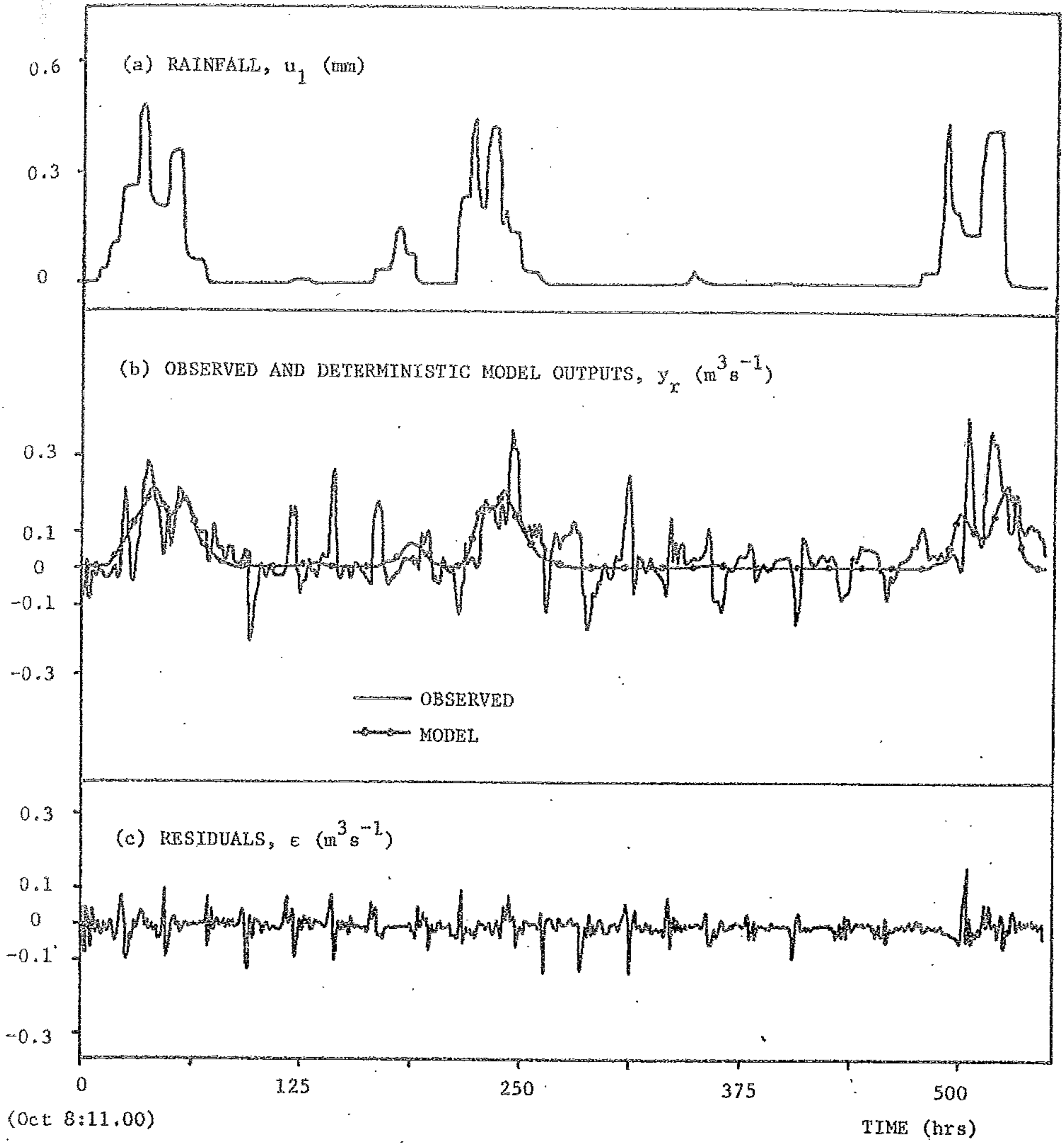


Figure 9 Maximum likelihood identification results for the RRF model of table 2

though they have a smaller standard deviation of ± 0.027 $(m^3 s^{-1})$. However, the estimate of c_1 indicates that the $C(q^{-1})$ polynomial is still only marginally stable. Substituting the deterministic output of the RRF model,

$$y_r(t) = -0.739 y_r(t-1) + 0.063 u_1(t-3) + 0.086 u_1(t-7) \quad (23)$$

into eqn. (22) allows more obvious comparison with the GF model and figure 10 shows the overall deterministic model response y(t) and model error for the period October 8th (11.00 hrs) - October 31st (07.00 hrs). In a deterministic sense the RRF model achieves a better characterisation of both the dry- and wet-weather dynamics of the influent flow to the treatment plant.

6.3. Comments.

It cannot faithfully be claimed that the identification process yields models that can be used with complete confidence in any forecasting or control synthesis context. Nevertheless, in a preliminary study where planned experimental work has not been possible we should try to analyse the failures and project a course for further investigation.

We may surmise that the identifiable models are only as good as the data permit and those available here leave much to be desired on several accounts. In the first instance, the method of constructing once-hourly sampled time-series from the rainfall observations means that the majority of precipitation events are distributed (necessarily) over a wide time interval. Hence, our knowledge of the timing of a particular event is very imprecise and it is not obvious which of the coefficients of the

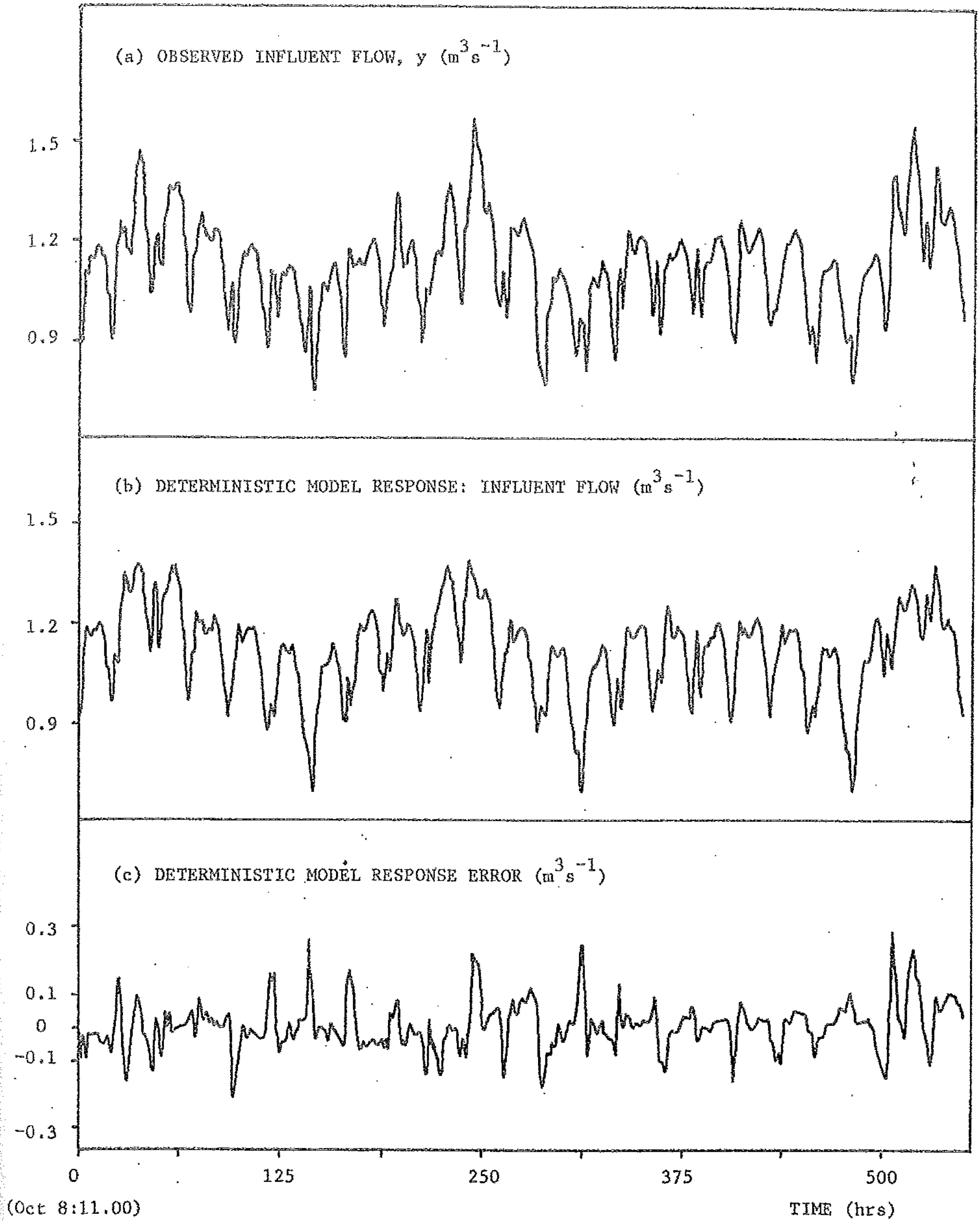


Figure 10 Comparison between the deterministic response of an RRF model, superimposed upon a mean weekly profile, and the observed influent flow

rainfall input $u_1(t)$ are most significant. This is clearly the case with the RRF model. Yet, while the nature of the rainfall data may be obscuring the identification of the true dynamics of urban runoff processes, it is also true that spatially-distributed runoff inputs are observed as temporally distributed in the influent flow to the plant. Thus, theoretically, an adaptive predictor which assumes no observations of rainfall should be capable of "recognising" such disturbances in advance of the time required for the peak response.

On the second account it is doubtful that the influent data describe the true flow at the downstream boundary of the Käppala tunnel. For example, we might have expected the natural flow to be reasonably free of pumping effects and to display smoother daily variations. The central problem here is the apparent discontinuity in the flow dynamics caused by the coincidence of the pump start-up with the minimum point in the 24hr cycle. A problem which is exacerbated by small differences in the timing of the pumping activity from one day to the next. Thus, whatever attempts are made at circumvention a satisfactory model cannot be identified without substantially modifying the data.

Notwithstanding the imperfections of the rainfall data, it is the nature of the flow data that dominates the identification. The algorithms for parameter estimation are required to treat the subsequent errors as statistically significant but are not provided with a model structure which accounts for their dynamics. Such conditions are more restrictive for the identification of a GF model. However, the structure of this type of model is more directly related to the probable structure of an adaptive predictor; the results presented here are a useful reference framework against which the discussion of the next section can be compared. On the other hand, if a strictly off-line analysis and control synthesis is to be undertaken, then an RRF model is preferable since it has the capacity to simulate more adequately the dynamic features of dry-weather flows and rainfall-runoff disturbances.

7. ADAPTIVE PREDICTION OF THE INFLUENT FLOW TO THE TREATMENT PLANT:

There are a considerable number of structures which could be hypothesised for an adaptive predictor described by eqns. (16) and (17). Ideally, the identification analysis should indicate a suitable choice for the orders p, ℓ, and r_i of the polynomials A, B, and C_i, respectively. However, in view of the uncertainty surrounding the validity of the GF model and the facility for parameter tracking afforded by the recursive algorithms of the predictor, it is advisable to reconsider and reassess several combinations of values for p, ℓ, and r_i. In addition, the auxiliary variables v_i may be any of the options given in eqn. (21), or measurements of the rainfall u_1, and we must choose k, the number of steps ahead for which prediction is required.

At this preliminary stage we shall restrict the discussion to simple and low-order structures of the predictor for $k = 1, 4$, and in the assumed absence of any measurements of the rainfall. Thus, several aspects of part (ii) of the problem formulation, particularly the use of rainfall data, are not examined. On the other hand, the resulting predictors give a first impression of the accuracy obtainable for short-term on-line prediction, one-step ahead, and longer-term prediction, four hours in advance of the current time. And, most important of all, since they require no on-line measurement of rainfall conditions, the predictors are of a form which can be more readily implemented in practice.

7.1. A One-Step Ahead Predictor.

An appropriate structure for the one-step ahead predictor is given in table 3. The a priori estimates of the parameters are based on those values to which the estimates converged after one run through the data in which $\hat{\theta}_0 = \underline{0}$ and $P_0 = (10)I$. Figures 11 and 12 show the prediction, prediction error, recursive parameter estimates and loss function

Table 3 - The one-step ahead predictor, k = 1, with 2 auxiliary variables, m' = 2.

Auxiliary variables: $v_1(t) = \tilde{y}_w(t)$
$v_2(t) = \tilde{y}_d(t)$

Polynomial orders: $A(q^{-1})$; $p = 1$
$B(q^{-1})$; $\ell = 1$
$C_1(q^{-1})$; $r_1 = 2$
$C_2(q^{-1})$; $r_2 = 2$

A priori estimates of the parameters, $\hat{\theta}_0$:

α_1	1.20	γ_{11}	0.47	γ_{21}	0.37
β_1	0.42	γ_{12}	-0.27	γ_{22}	-0.30

Exponential weighting factor, $\mu = 0.995$.
A priori covariance matrix of the estimation errors, $P_0 = (0.1)I^*$.

---

* Further definition of the use of this matrix in the algorithms for recursive least squares can be found in many texts, see e.g. [24]; I denotes the unit identity matrix.

$V_1(t)$ for this predictor, i.e. from eqn. (17),

$$\hat{y}(t+1|t) = \alpha_1 y(t) - \beta_1 \hat{y}(t|t-1) + \gamma_{11} v_1(t+1) +$$

$$+ \gamma_{12} v_1(t) + \gamma_{21} v_2(t+1) + \gamma_{22} v_2(t) \qquad (24)$$

Once more the results are given for the low-pass filtered time-series $y(t)$, from which the weekly profile $\bar{y}_w$, see section 6.2, and the daily profile $\bar{y}_d$, using all the observations, are computed, figure 13.

Despite the now almost inevitable errors coincident with the pumping operations, visible as spikes in the errors and giving the stepped form of the loss function, the one-step ahead prediction is very close to the observed data. In particular, the runoff from rainfall events is well described, although the errors are somewhat larger over these periods and the slope of the loss function is increased temporarily. Notice that the parameter estimates are relatively insensitive to these events, which are tantamount to an apparent change in the dynamic structure of the process. Thus, as suggested in section 6.3, the temporally-distributed observation of additional flows allows the predictor to recognise quickly this deterministic disturbance through $y(t)$ and $\hat{y}(t|t-1)$, in eqn. (24), and significant adjustment of the parameters, e.g. α_1 and β_1, becomes redundant.

The standard deviation of the prediction errors is ±0.029 ($m^3 s^{-1}$) and this compares favourably with that for the equivalent (residual) errors of the GF model, section 6.1. Since the structure of the GF model and the predictor are similar*, it can be shown through identity relationships

---

* With notable differences in the use of two auxiliary variables, v_1 and v_2, and the signal u_1 (the rainfall).

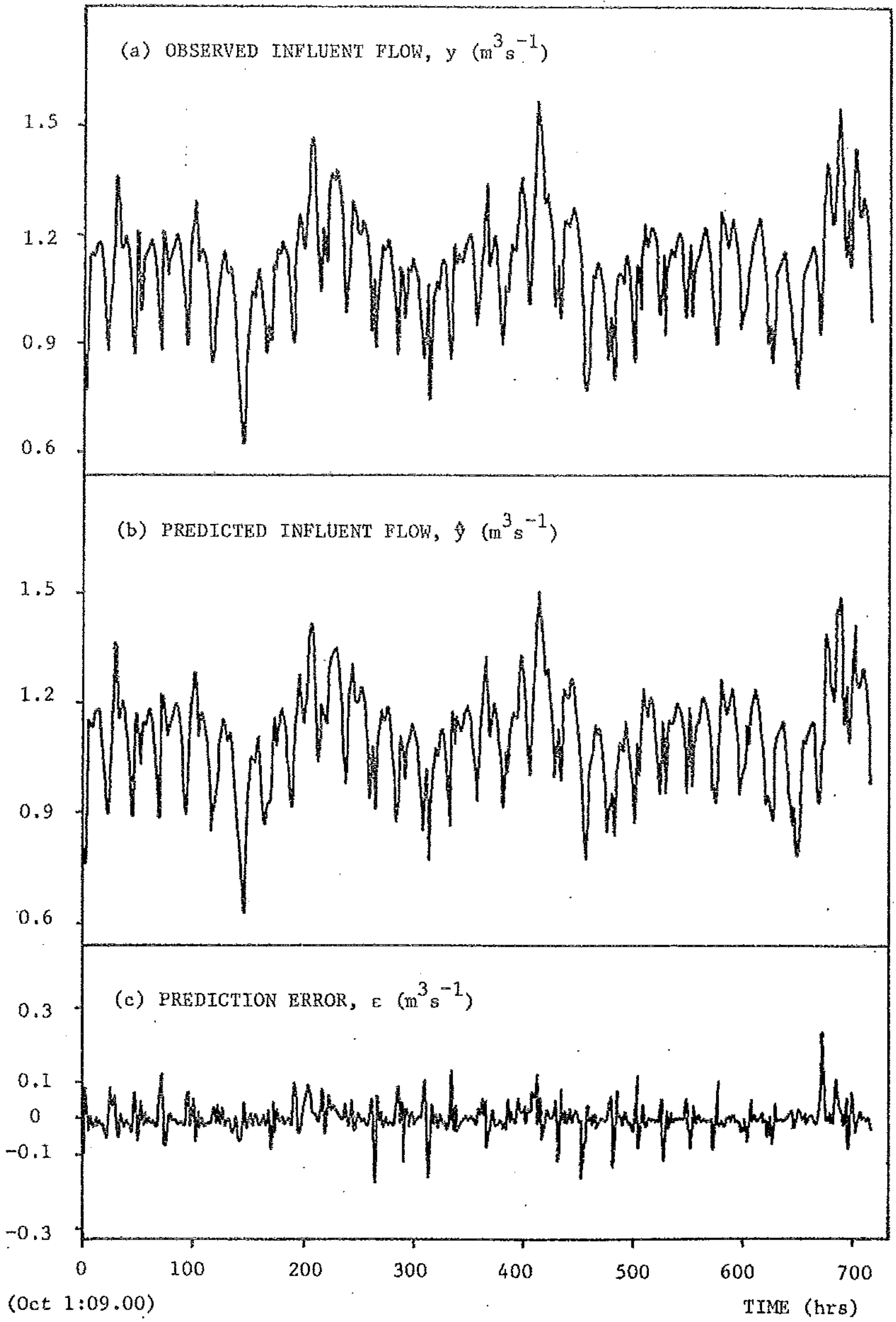


Figure 11 The results of the one-step ahead adaptive predictor of table 3

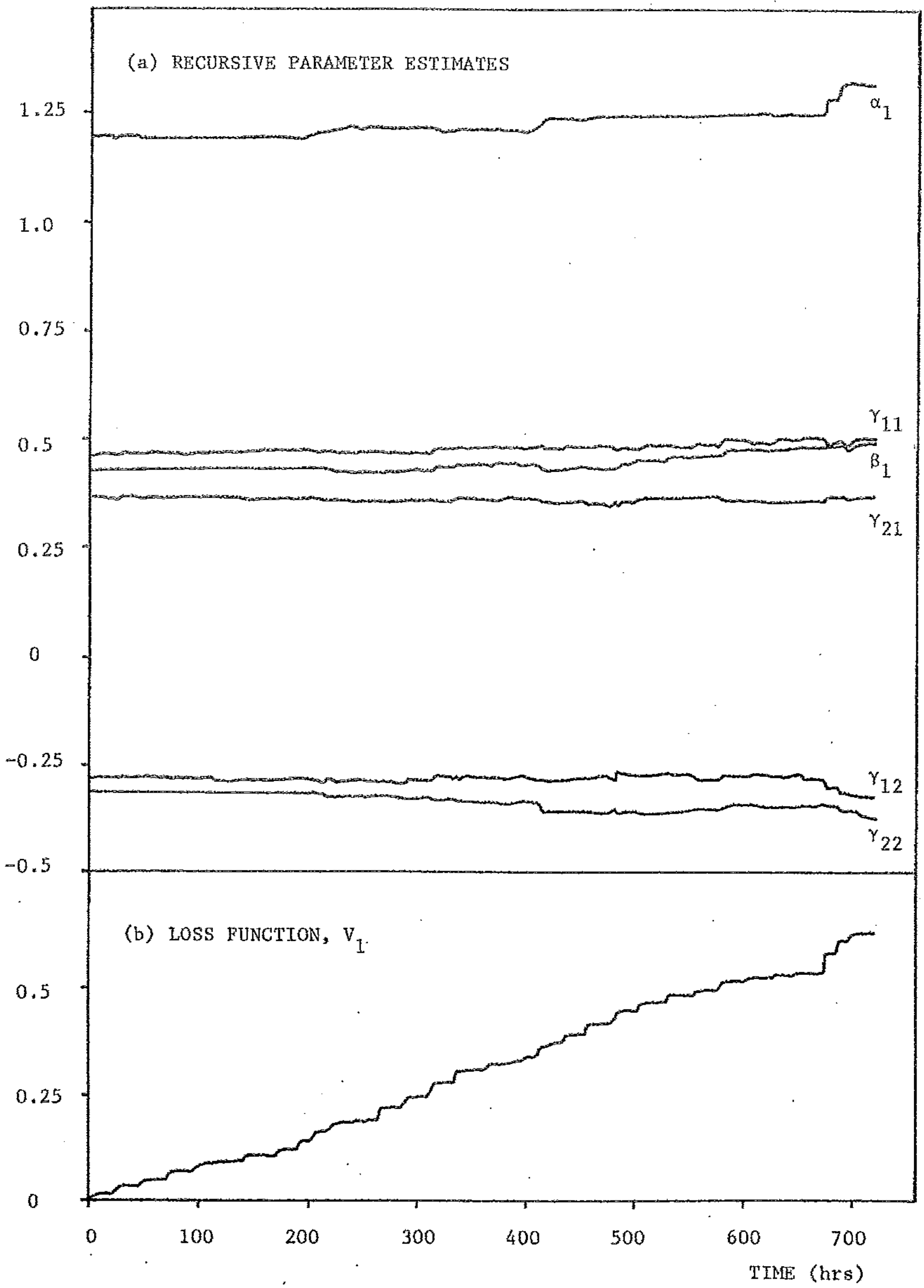
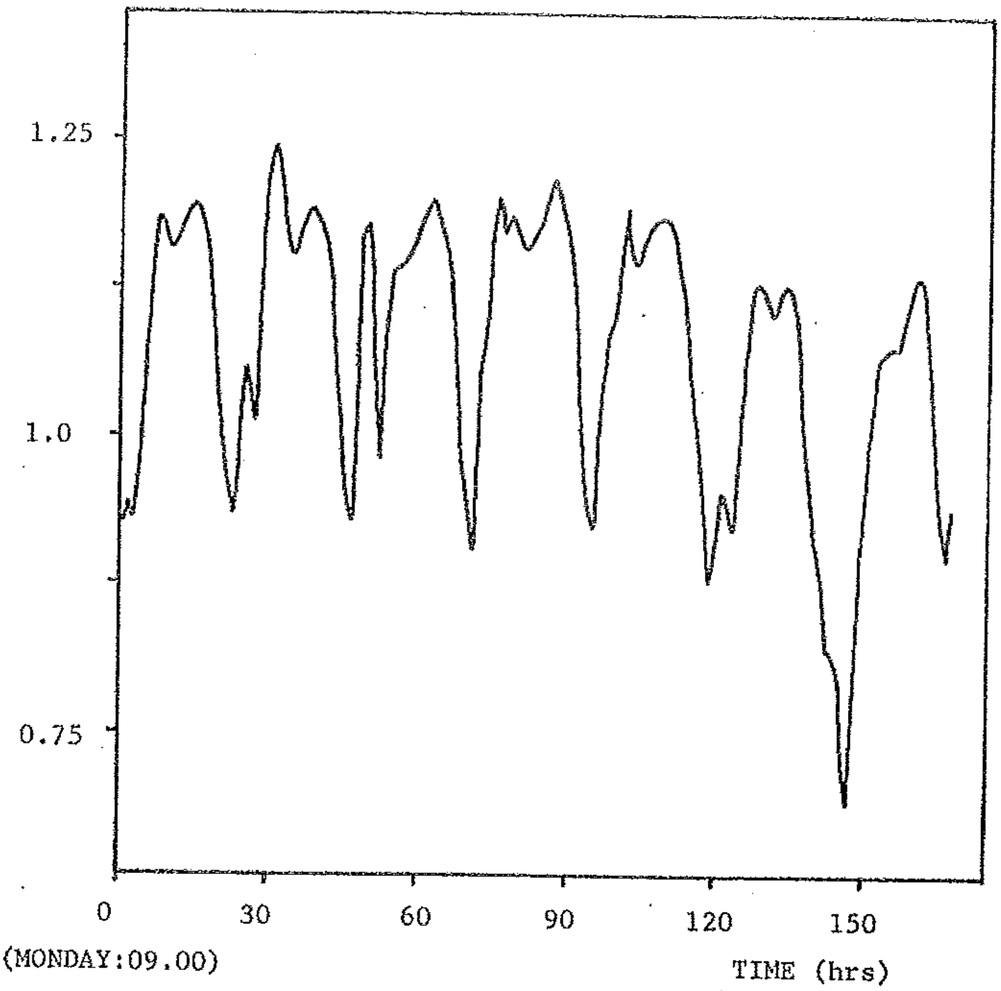
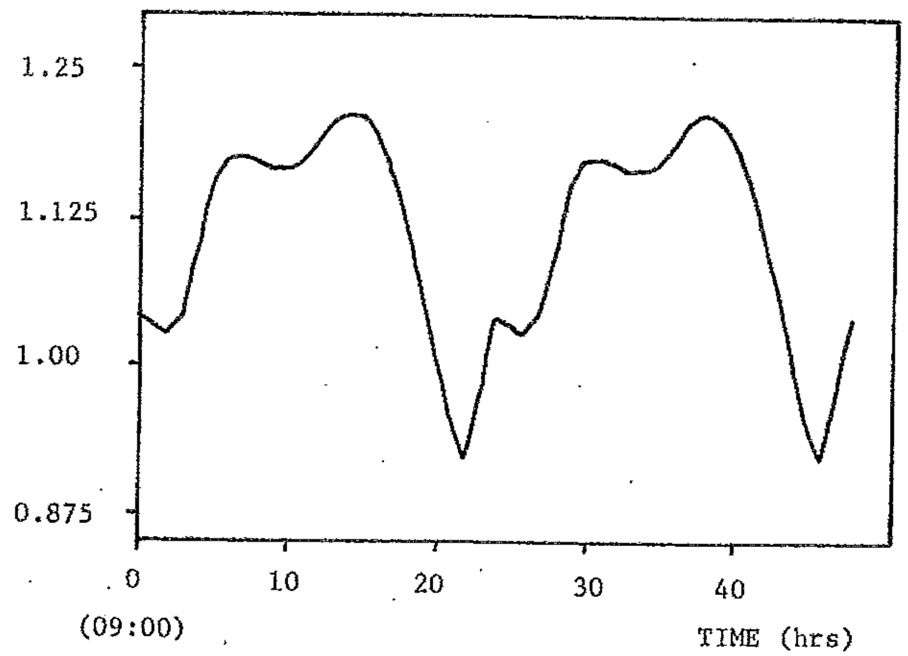


Figure 12 The results of the one-step ahead adaptive predictor of table 3



(a) Mean weekly influent flow profile, $\bar{y}_w$ (m^3s^{-1})



(b) Mean daily influent flow profile, $\bar{y}_d$ (m^3s^{-1})

Figure 13 The profiles of the auxiliary variables for adaptive prediction between the A, B, C, and A, B, C, polynomials, see [10], that the parameter estimates of the two descriptions of the process are in good agreement.

7.2. A Four-Step Ahead Adaptive Predictor.

Table 4 and figures 14 and 15 give the structure and results of a four-step ahead predictor; the recursive estimates are shown for those parameters which are significantly non-zero. The prediction errors have a standard deviation of ±0.074 (m^3s^{-1}). There is a marked decrease, with respect to the one-step ahead predictor, in the dependence of $\hat{y}(t+4|t)$ upon $y(t)$ and $\hat{y}(t+3|t-1)$ to the extent that the estimate of α_1 is much reduced and the $B(q^{-1})$ polynomial is found to be dispensable to all intents and purposes. The process of prediction is now more a function of the auxiliary variables and particularly the weekly profile $v_1(t)$. Thus, for the four-step ahead predictor the effects of rainfall-runoff are not quickly recognised through the signal $y(t)$; the prediction of the resultant peak flows are substantially attenuated and simultaneously the estimate of α_1, for example, is considerably "adapted" in order to describe the changes in the system dynamics. However, after the temporary additional flow α_1 returns slowly to its steady-state value for dry-weather conditions. Hence, in spite of the inaccuracies of the four-step ahead prediction we have a good example of the adaptability of the predictor.

Table 4 - The four-step ahead predictor, $k = 4$, with 2 auxiliary variables, $m' = 2$.

Auxiliary variables: $v_1(t) = \tilde{y}_w(t)$

$v_2(t) = \tilde{y}_d(t)$

Polynomial orders: $A(q^{-1})$; $p = 1$

$C_1(q^{-1})$; $r_1 = 5$

$C_2(q^{-1})$; $r_2 = 5$

A priori estimates of the parameters, $\hat{\theta}_0$:

α_1	0.49	γ_{11}	0.53	γ_{21}	0.32
		γ_{12}	0.16	γ_{22}	0.00
		γ_{13}	0.00	γ_{23}	0.00
		γ_{14}	0.00	γ_{24}	0.00
		γ_{15}	-0.26	γ_{25}	-0.22

Exponential weighting factor, $\mu = 0.995$

A priori covariance matrix of the estimation errors, $P_0 = (0.1)I$.

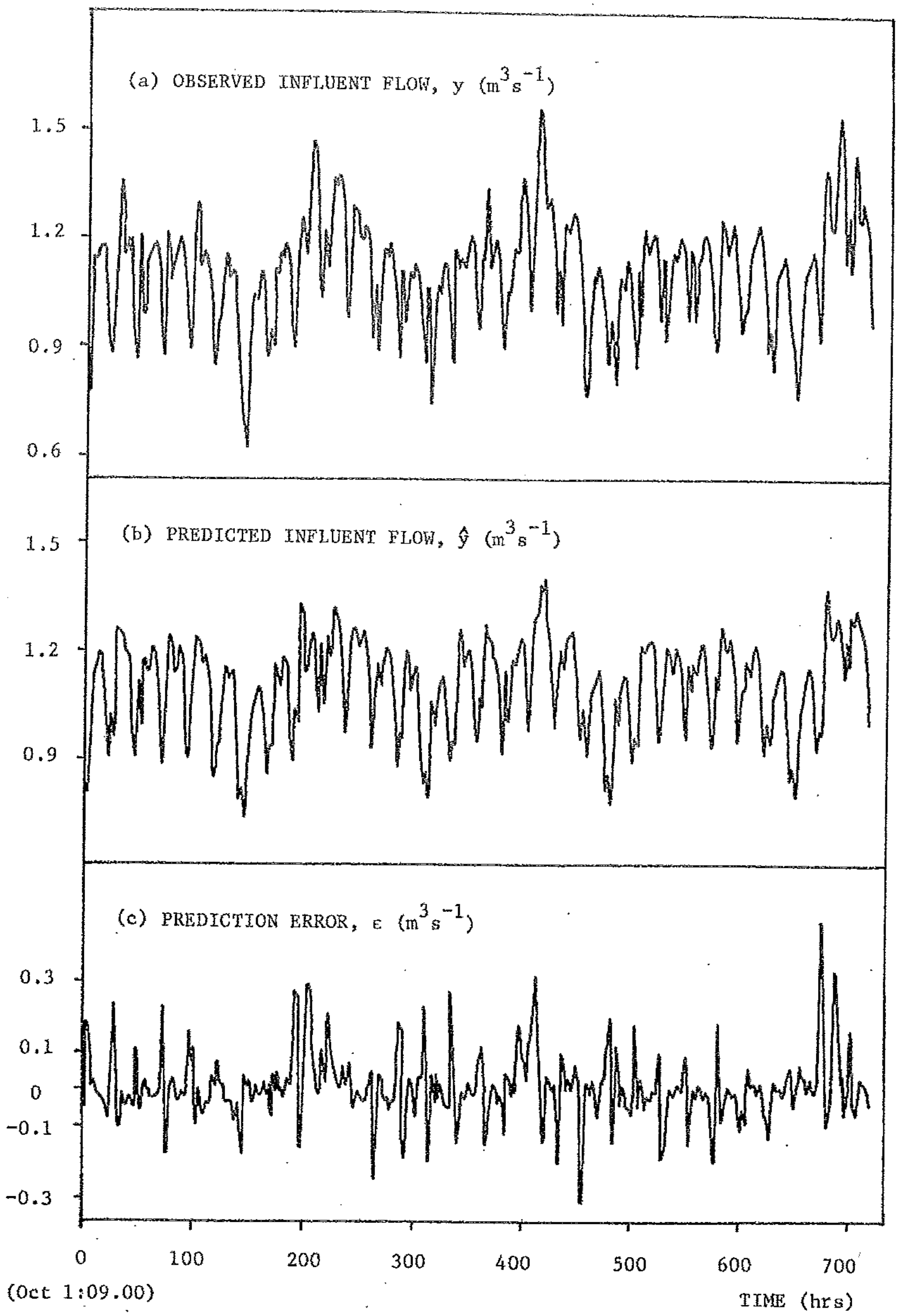


Figure 14 The results of the four-step ahead adaptive predictor of table 4

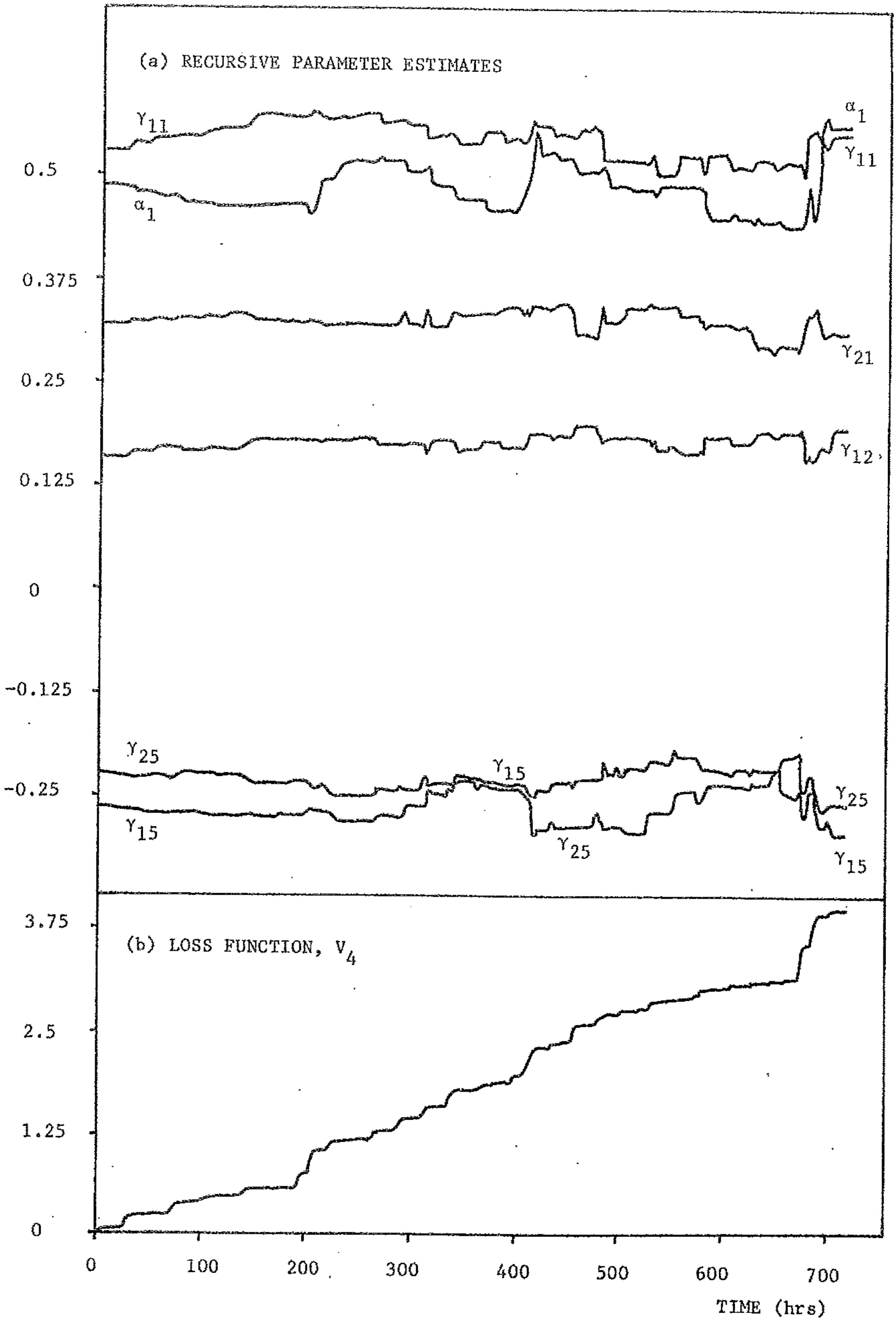


Figure 15 The results of the four-step ahead adaptive predictor of table 4

7.3. Comments.

From such a brief study of adaptive prediction it is a little dangerous to make generalisations on the method of evaluating predictor structures; for a more complete discussion the reader is referred to the work of Holst [10]. Bearing in mind that the choice of μ, P_0, $\hat{\theta}_0$, and the integers p, ℓ, r_i, are all factors which affect the initial rate of convergence of the parameter estimates, the criteria for predictor analysis should be based on the steady-state performance.

Here, we merely indicate that first-order A and B polynomials are preferred, partly because of the limitations of the data which emerged from the identification studies and partly since there is no immediately visible improvement when higher-order polynomials are used in the predictors. Of the possible profiles for v_i, i.e. those given in eqn. (21), the following ordering of goodness of prediction (with i = 1) is noted: (i) $\bar{y}_w(t)$; (ii) $\bar{y}_d(t)$; (iii) y(t-168); (iv) y(t-24). The inclusion of two profiles gives of the order of 5% reduction in the cumulative loss function V_1 (720) over the best single profile predictor. However, the important conclusion is that the mean profiles are seen to be consistently better than a feedforward process from the previous day or week. Furthermore, while only a short interval of observations for one month have been examined, it is likely that seasonal variations and longer-term trends should be recognised in any practical application. In this case it is not difficult to envisage a simple scheme for updating the profiles themselves, although the adaptability of the predictor may render this unnecessary.

The exponential weighting factor, μ, is found to give generally improved results. The added flexibility that this affords is not required for the one-step ahead predictor. In contrast, the value of μ might be reduced for the four-step ahead predictor where there is more difficulty in the speed of adaptation between dry- and wet-weather flows. For higher values of k the use of rainfall measurements would be advantageous. Consider the case for rainfall beginning at time t_0 such that we have the rainfall observations $u_1(t_0) \neq 0$, $u_1(t_0-1) = 0$. Due to a mean transportation delay of t_d in the sewer network the runoff input is first observed in $y(t_0+t_d)$ and the peak additional flow in $y(t_0+t_d+t_p)$, say. Let t_m be a pure time delay in the receipt of rainfall measurements, then

(i) we require $k < t_p$ if the predictor is to recognise runoff flow early enough to give reasonable prediction of $y(t_0+t_d+t_p)$ in the absence of rainfall measurement; or

(ii) we require $t_m < t_d$, if $k \geq t_p$, for similar objectives with the availability of rainfall measurements.

At the present time it is not easy to give any precise quantification of t_d or t_p; we may guess, however, that for the Käppala tunnel $2 < t_d < 4$ (hrs), while $1 < t_p < 4$ (hrs) as implied by the predictor results.

8. SOME IMPLICATIONS FOR SEWER SYSTEM AND WASTE-WATER TREATMENT PLANTS.

Thus far we have considered the solution of some aspects of parts (i) and (ii) of the problem formulation. The models and predictors which have been obtained express the dynamics of influent sewer flows to a treatment plant in a relatively simple and compact form. Such knowledge of the sewer network/waste-water treatment plant system is important in two senses, as stated in the third section of the problem formulation: firstly, in a "feedback" context it may be used in schemes for the regulation of flows within the sewer network; and secondly, it characterises the crude material influx to the treatment processes.

The manner in which these dynamic models are incorporated in control analyses and applications remains to be seen. However, it is a reasonable assumption that the initial synthesis of control laws will be constrained by the current practical limitations of automatic monitoring equipment and final control elements - the perennial stumbling block of water quality maintenance.

This section, therefore, takes a first look at the possibilities for implementing control of the system. As such it is an amalgam of factors absorbed from the available literature and matters arising from the preceding identification and predictor analyses. In some ways the following classification of sewer network/treatment plant control into two categories is a little arbitrary. However, for the time-being it fulfils our requirements of clarifying those aspects which are relevant to the overall objectives of improving plant operation.

8.1. Sewer Network Control.

Since transient flows should be estimated for both the design and operation of a sewer network system it is unnecessary here to make the usual distinction between the two problems. Nevertheless, the more specifically design-oriented studies of Rogers [21] and Lindholm [14], [15], for example, are only pertinent in so much as they impinge upon our present concern with the control of a given operational system.

Of more immediate interest are two major projects which have dealt with the on-line monitoring and operation of sewer flows although to a large extent they are also design studies. In a series of papers Anderson and co-workers [1], [2], [20], describe a computer-based real-time data acquisition system which has been installed and is in operation for the city of Cleveland, Ohio. Motivated by a desire to reduce the pollution caused by untreated storm-water overflows they summarise their objectives as follows: (i) the delivery of an essentially uniform flow and pollution load to the plant during dry-weather conditions, (ii) the reduction of high transient loads imposed on the plant by urban runoff, (iii) the elimination of by-pass of untreated sewage at the treatment plant, and (iv) the elimination of overflows from the sewer network during both dry and wet weather [2]. The second project, concerning the computer management of the sewer system of the city of Seattle, is reported in a comprehensive document by Leiser [13].

In the Cleveland study a mathematical flow model, drawn along the lines of those described in section 2, forms the basis of a dynamic programming routine for exercising automatic control throughout the system. Specific control objectives would be, for example, to organise flow detention and release* within the network during and after storm events; the final actuation of control is realised by an ingenious combination of gates and inflatable dams [20]. The Seattle scheme, in contrast, implements control according to so-called rule curves for storage in a trunk sewer and these are updated as the understanding of the system improves, i.e. a prescheduled, but "adaptive", type of control. Significantly, both studies suggest that downstream storage of the flow, especially near the treatment plant, is less effective than storage distributed at critical points throughout the sewer network.

It is beyond the scope of this discussion to comment further upon the details of sewer network flow control; clearly, a consideration of such factors is necessary for eventual regulation of the influent to the plant since the network has a sizeable and useful buffering capacity. It suffices to point out that, as these studies have demonstrated, the first steps towards the innovation of automatic control for sewer flows have yielded encouraging benefits in terms of economic investment, operational efficiency, and water quality maintenance. It is particularly satisfying to note the attention paid to the problems of monitoring and data retrieval and the referenced material provides an excellent background from which future applications may advance.

---

* Primarily through the use of storage tanks, but also, to a limited and closely monitored degree, within the sewers themselves.

8.2. Waste-Water Treatment Plant Control.

Unlike many process industries a waste-water treatment plant receives a raw input material whose variations with time are large and of a poorly-defined character. To be able to quantify that character and predict its time-variations in advance of the event may be regarded as the "raison d'être" of this study. However, even if the results of the identification and predictor analyses were conclusive, our task of discussing the implications for plant control is not an easy one: the reason being that the dynamics of the various unit processes of treatment are not well understood.

Nevertheless, some conjecture can be made on the problems to be tackled at a future stage. Recent and continuing investigations of the primary sedimentation [12] and activated-sludge processes [17], [18] indicate that a knowledge of flow dynamics is crucial to the efficiency of sedimentation and offers the possibility of flexible and tangible automatic control of the step-feed form of activated-sludge units. Thus, a good adaptive prediction of the influent flow permits in theory the <u>prior</u> organisation of flows through the plant such that diurnal fluctuations and the damaging shock loads of storm events may be attenuated. Especially in the case of lamella sedimentation processes and the secondary clarifiers efficiency is a function of maintaining the flow at acceptably low levels.

Indeed, this leads us to question the value of the Käppala treatment plant operating practice. Referring back to figure 6 it is evident that influent flow manipulation produces large pulse inputs to the plant. Such disturbances almost certainly mean that there is a transient period when a large portion of the suspended solids pass through sedimentation process without settling. Whether the introduction of this inefficiency is compensated by the increased efficiency of solids-settling over the remainder of the 24 hr period is difficult to state for there are other operational factors to be considered. As reported in Olsson et al [18] the crude sewage buffering basin itself acts as a sedimentation tank; subsequently, flow detention of more than one day is found to cause overloading of the influent screens when the accumulated volume is released.

Of course, the characterisation of the plant influent flow is only half the information required for describing the crude material influx; a knowledge of the quality variations is implied by the above. In this area the problems of instrumentation are more restrictive to on-line prediction and control. For instance, in order to obviate the difficulty, methods referred to in [14] propose a set of exponential-decay curves for the quality (biochemical oxygen demand) transient of storm runoff flows. Alternatively, while some quality indices can be monitored automatically, Pew et al [20] suggest the use of linear multiple regression models for the generation of runoff and sewer flow quality as a function of the more traditional (and less easily monitored) quantities of biochemical oxygen demand, suspended solids, and coliform organisms. It is precisely this kind of application where an adaptive predictor as presented here can be of most benefit.

But much remains to be done and, as so often occurs in this field, one has the feeling that we have only scratched at the surface of a complex problem. It emphasises the real need for further examination of the properties of treatment plant dynamics and the clarification of control objectives. Essentially, the synthesis of operational control laws for the individual units of treatment is a process of creating or discovering flexibility in designs which are at first sight rather inflexible. In this sense it is pleasing to note that, purely with respect to the Käppala plant, the buffering well adjacent to the influent gates has an (observed) capacity of up to 20,000 m^3. This creates a retention capability of some 5 hours (at typical existing flow-rates) with which to modulate the influent flow to be consistent with controlling the sedimentation and activated-sludge dynamics. A feature which is, therefore, not only significant in treatment plant control, but also reinforces the practicability of the flow detention schemes proposed by Young and Beck [25] in their studies of in-stream water quality control. And that, in the wider sense, is the ultimate aim of any examination of treatment plant operation.

9. CONCLUSIONS.

The major limitation of the study has been the nature of the observed rainfall and influent flow data and, therefore, it is recommended that for any future developments better data be obtained or a more accurate synthesis of the true influent flow be examined. However, if the errors of pumping disturbances are ignored the models and predictors presented here give quite reasonable results, and it may be assumed that the effects of pumping would be less or even negligible in the true flows.

In off-line applications of a model for rainfall-runoff/influent flow relationships the RRF model is more appropriate and for on-line purposes the 1-step ahead predictor approaches the satisfaction of those practical constraints to which the system is subject: namely that as little automated instrumentation as possible should be assumed. But the salient advantage of the black box methodology over the models based on physical laws and arguments is the simplicity and compactness of the relationships which are so derived. In whatever context the models are used they should be consistent with the overall objectives of synthesising control laws which are practicable. And at the present time the technology of waste-water treatment is more amenable to the simple rather than to the sophisticated.

10. REFERENCES.

[1] Anderson J.J., and Anderson D.J., "Computer control and modeling of sewer systems", A.I.Ch.E. Symposium Series, Volume 69, No. 129, pp. 432-438, American Institute of Chemical Engineers, New York, 1973.

[2] Anderson J.J., and Pew K.A., "Interfacing the sewer network with the treatment plant", Paper presented at IAWPR Workshop on Instrumentation, Control and Automation of Wastewater Treatment Systems, London, September 1973.

[3] Åström K.J., "Introduction to stochastic control theory", pp. 166-170, Academic Press, New York, 1970.

[4] Åström K.J., and Bohlin T., "Numerical identification of linear dynamic systems from normal operating records", Proc. IFAC Symposium on the Theory of Self-adaptive Control Systems, Teddington, England, 1965.

[5] Åström K.J., Bohlin T., and Wensmark S., "Automatic construction of linear stochastic dynamic models for stationary industrial processes with random disturbances using operating records", Technical paper TP 18.150, IBM Nordic Laboratory, Sweden 1965.

[6] Chen C.W., and Shubinski R.P., "Computer simulation of urban storm water runoff", Proc. A.S.C.E., J. Hydr. Division, Vol. 97, No. HY2, pp. 289-301, Febr. 1971.

[7] Gustavsson I., "Parametric identification of multiple input, single output linear dynamic systems", Report No. 6907, Division of Automatic Control, Lund Institute of Technology, Sweden, 1969.

[8] Gustavsson I., Selander S., and Wieslander J., "IDPAC - User's guide", Report No. 7331, Division of Automatic Control, Lund Institute of Technology, Sweden, 1973.

[9] Harris G.S., "Real time routing of flood hydrographs in storm sewers", Proc. A.S.C.E., J. Hydr. Division, Vol. 96, No. HY6, pp. 1247-1260, June 1970.

[10] Holst J., in preparation, Division of Automatic Control, Lund Institute of Technology, Sweden.

[11] Kashyap R.L., and Rao A.R., "Real time recursive prediction of river flows", Automatica, Vol. 9, No. 2, pp. 175-183, March 1973.

[12] Larsson R., and Schröder G., "Dynamiska modeller för primärsedimentering i reningsverk (Dynamic models of primary sedimentation in waste-water treatment)", Report No. RE-146, Division of Automatic Control, Lund Institute of Technology, (in Swedish), Sweden, September 1974.

[13] Leiser C.P., "Computer management of a combined sewer system", Environmental Protection Technology Series, EPA-670/2-74-022, U.S. Environmental Protection Agency, Cincinnati, Ohio 45268, July 1974.

[14] Lindholm O.G., "A pollutional analysis of the combined sewer system", Licentiate Thesis, Institutt for Vassbygging, Universitet i Trondheim, Norges Tekniske Høgskole, Norway, August 1974.

[15] Lindholm O.G., "Systemanalyse av avløpsanlegg (Systems analysis of a sewer network)", Vatten, No. 2, pp. 159-172, (in Norwegian), 1973.

[16] Offner F.F., "Computer simulation of storm water runoff", Proc. A.S.C.E., J. Hydr. Division, Vol. 99, No. HY12, pp. 2185-2194, December 1973.

[17] Olsson G., in preparation, Division of Automatic Control, Lund Institute of Technology, Sweden.

[18] Olsson G., Dahlqvist K-I., Eklund K., and Ulmgren L., "Control problems in waste-water treatment", Technical report, The Axel Johnson Institute for Industrial Research, Nynäshamn, Sweden, 1973.

[19] Papadakis C.N., and Preul H.C., "Testing of methods for determination of urban runoff", Proc. A.S.C.E., J. Hydr. Division, Vol. 99, No. HY9, pp. 1319-1335, September 1973.

[20] Pew K.A., Callery R.L., Brandsetter A., and Anderson J.J., "Data acquisition and combined sewer controls in Cleveland", J.W.P.C.F., Vol. 45, No. 11, pp. 2276-2289, November 1973.

[21] Rogers R.A., "Rational 'rational' method of storm drainage design", Proc. A.S.C.E., J. Irr. and Drain. Division, Vol. 94, No. IR4, December 1968.

[22] Whitehead P.G., and Young P.C., "A dynamic stochastic model for water quality in part of the Bedford-Ouse river system", Proc. IFIP Working Conference on Modelling and Simulation of Water Resources Systems, Ghent, Belgium, July/August 1974.

[23] Wittenmark B., "A self-tuning predictor", IEEE T-AC, Vol. 19, No. 6, December 1974.

[24] Young P.C., "A recursive approach to time-series analysis", Bulletin of the Institute of Mathematics and its Applications (IMA), Vol. 10, Nos. 5/6, pp. 209-224, 1974.

[25] Young P.C., and Beck M.B., "The modelling and control of water quality in a river system", Automatica, Vol. 10, No. 5, September 1974.

ACKNOWLEDGEMENTS.

It is my pleasure to acknowledge my appreciation of the help Professor Karl Johan Åström and his department have given me. I am deeply indebted to Jan Holst who has discussed with me many aspects of adaptive prediction. Further I should like to thank Björn Wittenmark and Ivar Gustavsson for their advice. However, in particular I am grateful to Gustaf Olsson for his continual encouragement and useful comments.

My thanks also go to Kjell-Ivar Dahlqvist of the Käppala treatment plant for the receipt of the data.