Pergamon

0043-1354(94)E0035-5

Wat. Res. Vol. 28, No. 10, pp. 2055–2066, 1994

0043-1354/94 $7.00 + 0.00

IDENTIFICATION OF WASTEWATER TREATMENT PROCESSES FOR NUTRIENT REMOVAL ON A FULL-SCALE WWTP BY STATISTICAL METHODS

JACOB CARSTENSEN[1], HENRIK MADSEN[1], NIELS KJØLSTAD POULSEN[1] and MARINUS K. NIELSEN[2]

[1]Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Bygning 101, 2800 Lyngby, Copenhagen, [2]I. Krüger Systems AS, Research & Development Division, Gladsaxevej 363, 2860 Søborg, Denmark

(First received August 1993; accepted in revised form January 1994)

Abstract—The introduction of on-line sensors of nutrient salt concentrations on wastewater treatment plants opens a wide new area of modelling wastewater processes. Time series models of these processes are very useful for gaining insight in real time operation of wastewater treatment systems which deal with variable influent flows and pollution loads. In this paper nonlinear time series models describing the variations of the ammonia and nitrate concentrations in the aeration tanks of a biological nutrient removal WWTP are established. The models proposed herein are identified by combining well-known theory of the processes, i.e. including prior knowledge, with the significant effects found in data by using statistical identification methods. Rates of the biochemical and hydraulic processes are identified by statistical methods and the related constants for the biochemical processes are estimated assuming Monod kinetics. The models only include those hydraulic and kinetic parameters, which have shown to be significant in a statistical sense, and hence they can be quantified. The application potential of these models is on-line control, because the present state of the plant is given by the variables of the models which are continuously updated as new information from the on-line sensors becomes available.

Key words—nonlinear time series models, grey box modelling, estimation Monod-kinetic expressions, ammonia load estimation, nitrification, denitrification, surveillance, control

1. INTRODUCTION

In the last decade reliable on-line sensors for monitoring of nutrient salt concentrations on wastewater treatment plants have been developed. Experiences with locating the sampling positions, filtering of the samples and maintenance of the sensors have been obtained on several Danish full-scale plants (Thomsen and Nielsen, 1992). The main objective of these sensors has been surveillance of plant performance and alarm handling, but the potential of using on-line sensors for controlling wastewater treatment plants is big and needs to be investigated.

Today many plants are operated according to predetermined schemes with very little considerations to the variations of the material loads. Use of sensors in on-line control of the operation of the plant may enhance the ability to comply with the assigned effluent standards and reduce the sizes of plants to be designed in the future.

The acquired information from the on-line sensors can be handled in several ways. These methods range from simple ruled based control to calibration of complex models, characterized by a large number of parameters, like IAWPRC model No. 1 (Henze *et al.*, 1987), which is used for simulating different operational strategies. Other methods for on-line control of activated sludge processes include ARMAX-models (Olsson *et al.*, 1989), fuzzy logic (Couillard and Zhu, 1992), and neural networks (Bhat and McAvoy, 1990), of which the last two methods are capable of handling nonlinearities.

In this paper, stochastic models incorporating physical knowledge of the system are proposed. These models are denoted grey box models, since they are not as detailed as the "white" deterministic models and still contain some physical insight of the system as opposed to black box models, e.g. neural networks and ARMAX-models. The data from the plants should contain some dynamic variations in order to identify physical phenomena by statistical methods.

2. PLANT AND DATA DESCRIPTION

The Aalborg West WWTP is designed for 265,000 p.e. with an average daily load of 60,000 m^3 wastewater. The biological part of the plant consists of 18 anaerobic pretreatment tanks, 6 aeration tanks and 15 clarifiers, and the plant is operated according to the BIO-DENIPHO* scheme (Bundgaard, 1988),

*BIO-DENIPHO is patented process developed by I. Krüger Systems in cooperation with the Department of Environmental Engineering at the Technical University of Denmark.

Table 1. Summary of available measurements and registered controlling signals

	Variable	Unit
Weir (inlet LT5)	$W_{i,\mathrm{LT5},t}$	0/1 (closed/open)
Weir (inlet LT6)	$W_{i,\mathrm{LT6},t}$	0/1 (closed/open)
Weir (outlet LT5)	$W_{o,\mathrm{LT5},t}$	0/1 (closed/open)
Weir (outlet LT6)	$W_{o,\mathrm{LT6},t}$	0/1 (closed/open)
Oxygen Supply Rate LT5	$\mathrm{OSR}_{\mathrm{LT5},t}$	g O_2/h/m^3
Oxygen Supply Rate LT6	$\mathrm{OSR}_{\mathrm{LT6},t}$	g O_2/h/m^3
Oxygen concentration LT5, sensor 1	$O_{21,\mathrm{LT5},t}$	g O_2/m^3
Oxygen concentration LT5, sensor 2	$O_{22,\mathrm{LT5},t}$	g O_2/m^3
Oxygen concentration LT6, sensor 1	$O_{21,\mathrm{LT6},t}$	g O_2/m^3
Oxygen concentration LT6, sensor 2	$O_{22,\mathrm{LT6},t}$	g O_2/m^3
Ammonia concentration LT5	$C_{\mathrm{NH4,LT5},t}$	g NH4-N/m^3
Ammonia concentration LT6	$C_{\mathrm{NH4,LT6},t}$	g NH4-N/m^3
Nitrate concentration LT5	$C_{\mathrm{NO3,LT5},t}$	g NO3-N/m^3
Nitrate concentration LT6	$C_{\mathrm{NO3,LT6},t}$	g NO3-N/m^3
Phosphate concentration LT6	$C_{\mathrm{PO4,LT6},t}$	g PO4-N/m^3
SS concentration LT6	$S_{\mathrm{SS,LT6},t}$	kg/m^3
Flow biological part of plant	$Q_{\mathrm{bio},t}$	m^3/h

where the nutrient salts are removed biologically. However, a major part of the phosphate in the wastewater is removed chemically by adding ferrosulphate.

The 6 aeration tanks are operated in three parallel couples of two alternating tanks, and two alternating tanks (denoted LT5 and LT6) are equipped with sensors for measuring ammonia and nitrate concentrations. Furthermore, two sensors for measuring oxygen concentration (one in each end of the tank) are placed in every aeration tank, one sensor for measuring phosphate and one for measuring the suspended solids (SS) concentrations are placed in LT6, and the oxygen supply from the rotors in all the aeration tanks is registered. Finally, the hydraulic load of wastewater to the biological part of the plant is measured. For the models in this paper only measurements from LT5 and LT6 were considered together with the flow to the biological part of the plant. Table 1 summarizes the variables used for modelling.

In October and November 1992 the variables given in Table 1 were recorded. A data set of approximately 27 days without large interrupts in the sampling was obtained. A few outliers and measurements taken during calibration of the sensors were removed from the data set and replaced with missing values, which can handled by the methods considered in this paper. A total of 9507 observations with 4 min between subsequent samples were recorded. For all the variables in Table 1, $t = 1 \ldots 9507$, denotes the sample number. In Fig. 1, a small part of the time series of ammonia and nitrate is shown, and the alternating operation of the plant is clearly recognized. The

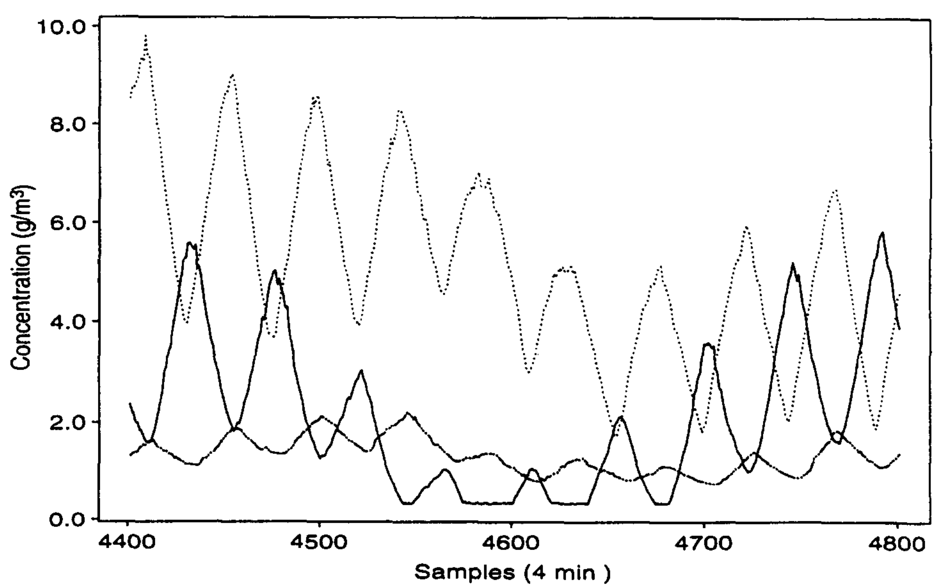

Fig. 1. On-line measurements of ammonia (solid curve) and nitrate (dotted curve) concentrations from Aalborg West WWTP. The operation of the plant follows the BIO-DENIPHO scheme.

detection limits of the ammonia and nitrate sensors are in the range 0.3–0.4 g/m^3.

3. MODELLING OF THE PROCESSES

The variety of biochemical processes and bacteria types of a full scale WWTP is big. In the IAWPRC model No. 1 the many biochemical processes are divided into separate classes of processes, which are used for modelling. However, with the data available from on-line sensors today, it is still not possible to identify and estimate the parameters of the IAWPRC model No. 1, except a major part of the parameters are fixed. Thus, in order to extract information about the biochemical processes from the on-line measurements by statistical methods, simpler models need to be derived.

The models proposed in this paper are found by combining well-known theory of the processes with significant effects found by statistical methods in the data. By this approach, terms for the most important processes are extracted and lack of these terms are compensated for by adding noise, i.e. by formulating stochastic models. The estimates of the models are found by maximum likelihood estimation, and the significance of the parameters is determined using the likelihood ratio test (Goodwin and Payne, 1977) and BIC* (Schwarz, 1978). The grey box models presented in this paper only include the measured variables of Table 1 and may appear to be very simple compared to the deterministic models. However, the fact that we are actually able to estimate the parameters of these models makes them complicated compared to traditional stochastic models, e.g. ARMAX-models.

In this paper the variations of the ammonia and nitrate concentrations in LT5 and LT6 are modelled using the information available from previous samples of all the variables listed in Table 1. However, the fluctuations of some of the explanatory variables are such that a low-pass filtering is required. A weighted moving average filter with the filter weights given below was found to be adequate for filtering. Thus, the following variables are defined

$$\bar{O}_{2,\mathrm{LT5},t} = (3 \cdot O_{21,\mathrm{LT5},t} + 3 \cdot O_{22,\mathrm{LT5},t} + 2 \cdot O_{21,\mathrm{LT5},t-1} + 2 \cdot O_{22,\mathrm{LT5},t-1} + O_{21,\mathrm{LT5},t-2} + O_{22,\mathrm{LT5},t-2})/12$$

$$\bar{O}_{2,\mathrm{LT6},t} = (3 \cdot O_{21,\mathrm{LT6},t} + 3 \cdot O_{22,\mathrm{LT6},t} + 2 \cdot O_{21,\mathrm{LT6},t-1} + 2 \cdot O_{22,\mathrm{LT6},t-1} + O_{21,\mathrm{LT6},t-2} + O_{22,\mathrm{LT6},t-2})/12$$

$$\bar{Q}_{\mathrm{bio},t} = (Q_{\mathrm{bio},t} + Q_{\mathrm{bio},t-1} + Q_{\mathrm{bio},t-2})/3$$

$$\overline{\mathrm{OSR}}_{\mathrm{LT5},t} = (\mathrm{OSR}_{\mathrm{LT5},t} + \mathrm{OSR}_{\mathrm{LT5},t-1} + \mathrm{OSR}_{\mathrm{LT5},t-2})/3$$

$$\overline{\mathrm{OSR}}_{\mathrm{LT6},t} = (\mathrm{OSR}_{\mathrm{LT6},t} + \mathrm{OSR}_{\mathrm{LT6},t-1} + \mathrm{OSR}_{\mathrm{LT6},t-2})/3$$

$$\overline{\mathrm{SS}}_{\mathrm{LT6},t} = (\mathrm{SS}_{\mathrm{LT6},t} + \mathrm{SS}_{\mathrm{LT6},t-1} + \mathrm{SS}_{\mathrm{LT6},t-2})/3. \quad (1)$$

*Bayes Information Criterion.

For the present data set (which does not include measurements of phosphate concentrations), four biochemical and hydraulic processes were found to be significant:

- load to the WWTP;
- the nitrification process;
- the denitrification process;
- transport of nutrients.

Each of these processes is dealt with separately in the following, though it should be stressed that all the processes are estimated simultaneously, and that there are periods where several processes are active. The rates of the processes are all treated numerically in order to make them graphically more comparable, even though they are related to both formation and removal of nutrients. Some of the parameters used in describing these processes are assumed to be constant within one operation cycle, but it is statistically shown that they vary from one operation cycle to another. Therefore, to indicate that these parameters are time varying, the operation cycle number $f = 1 \ldots 227$ is used. One operation cycle covers approximately 45 samples. Thus, the time between two subsequent operation cycles is approximately 3 h, but it may vary within the range of 1.5–4.5 h.

3.1. *Load to the WWTP*

The raw wastewater to the biological part of the Aalborg West WWTP mainly consists of organic materials, ammonia and phosphate. During the anoxic phase in the aeration tank when the inlet weir ($W_{i,\mathrm{LT5},t}$ and $W_{i,\mathrm{LT6},t}$) is open, the ammonia concentration increases due to the load of ammonia in the raw wastewater. The ammonia concentration increases approximately at a constant rate, $r_{\mathrm{load},f}$, during one operation cycle of the BIO-DENIPHO scheme, when the load to the WWTP is the only significant process influencing the ammonia concentration. This can also be verified by inspecting Fig. 1.

If the aeration tank is ideal mixed with an inlet concentration of ammonia, C_{inlet}, running into the tank, the ammonia concentration in the tank, C_t, is given by the differential equation

$$\frac{\mathrm{d}C_t}{\mathrm{d}t} = \frac{C_{\mathrm{inlet}} - C_t}{T_h} \quad (2)$$

where T_h is the hydraulic retention time of the aeration tank. If the inlet weir is opened at $t = 0$ with the ammonia concentration in the tank being C_0, the solution to (2) is

$$C_t = C_{\mathrm{inlet}} - (C_{\mathrm{inlet}} - C_0) \cdot e^{-t/T_h} \quad (3)$$

Because the period with load to the aeration tank is relatively short compared to the hydraulic retention time, C_t increases approximately linear. Furthermore, the ammonia concentration in the aeration tank at

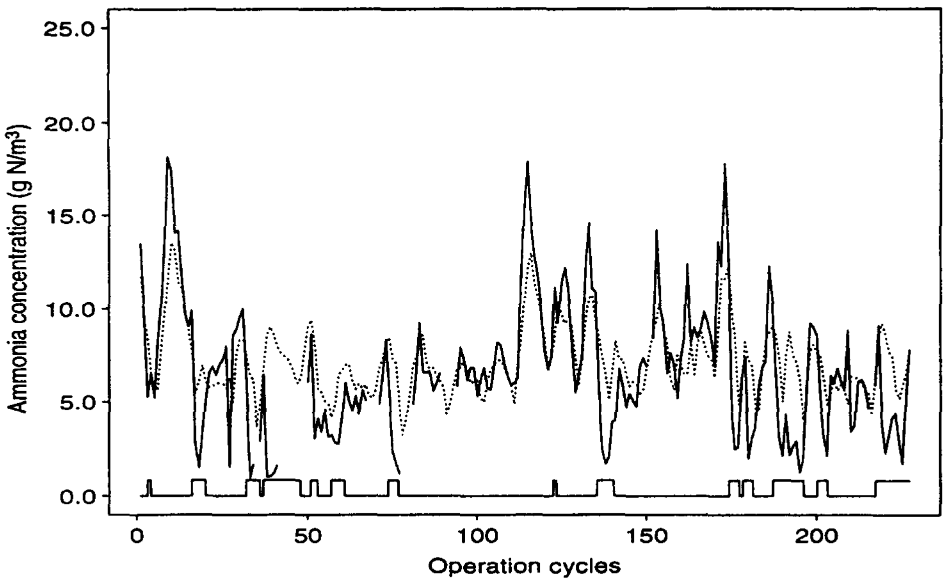

Fig. 2. Estimates (solid line) and one-step predictions (dotted line) of inlet concentrations of ammonia running into aeration tank LT5. The lower curve indicates the rainy periods (low during dry weather periods and high during rainy weather periods).

$t = 0$ is in general small relative to the inlet concentration, giving the following approximation

$$r_{\text{load},f} \approx \frac{C_{\text{inlet}}}{T_h} \tag{4}$$

Thus, by estimating $r_{\text{load},f}$ for every operation cycle and knowing T_h, an estimate of the average inlet concentration of ammonia to the aeration tank in that period is also found. However, the estimates of C_{inlet} are biased, because the estimates of $r_{\text{load},f}$ incorporates the simultaneous ammonia influent process and removal of ammonia due to denitrification. In Carstensen *et al.* (1993), it is shown by comparing the estimated influent ammonia concentrations with measured influent concentrations, that approximately 32% of the influent ammonia is removed during the denitrification. The estimated inlet concentrations of ammonia to LT5 and LT6 are shown in Fig. 2 and Fig. 3, respectively. Estimated load rates and inlet concentrations based on less than 4 observations are

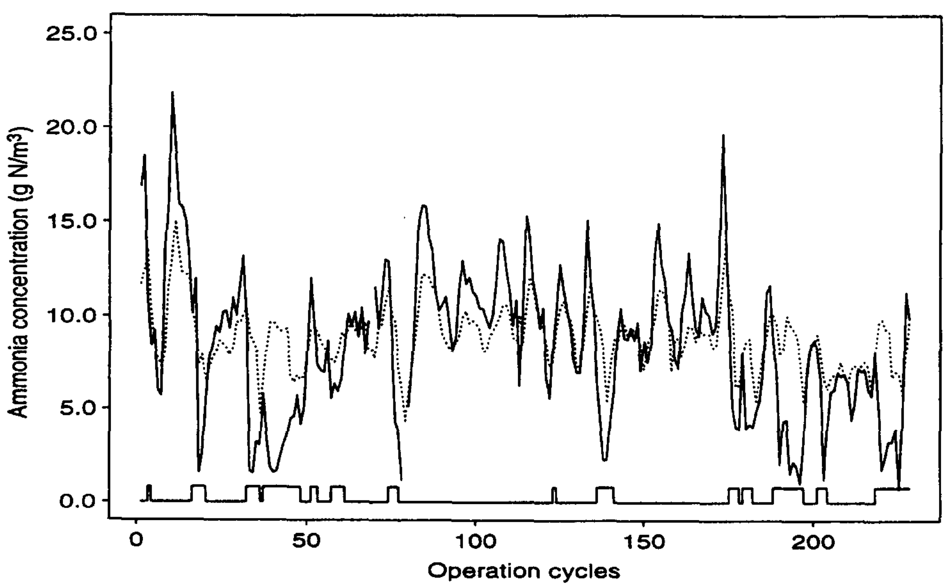

Fig. 3. Estimates (solid line) and one-step predictions (dotted line) of inlet concentrations of ammonia running into aeration tank LT6. The lower curve indicates the rainy periods (low during dry weather periods and high during rainy weather periods).

considered as missing values, which is indicated by the gaps on the solid curves in the figures. In general, the two figures correspond well with each other, except that the estimated and predicted (see later) inlet concentrations in Fig. 3 are higher than that in Fig. 2 for the first 112 operation cycles. This is caused by a drift in the ammonia sensor in LT5, which was re-calibrated in the 112th operation cycle of the time series. After the calibration the two curves agree very well in both scaling and dynamics. However, it should be emphasized that the two figures should not be identical, because $W_{i,\mathrm{LT6},t}$ is half an operation cycle displaced compared to $W_{i,\mathrm{LT5},t}$. Therefore, short term high or low ammonia loads may only show up on one of the figures.

Statistical analysis have shown that the variations of the ammonia load rate can be modelled using an AR (AutoRegressive) model (Box & Jenkins, 1976)

$$(1 + \phi_{\mathrm{load}} B)(r_{\mathrm{load},f} - \mu_{\mathrm{load},t}) = \epsilon_{\mathrm{load},f} \quad (5)$$

where B is the back-shift operator ($B_{r\mathrm{load},f} = r_{\mathrm{load},f-1}$), the noise $\epsilon_{\mathrm{load},f}$ is assumed to be N_{iid} $(0, \sigma^2_{\epsilon,\mathrm{load}})$†. Analysis of data has shown that the mean load, $\mu_{\mathrm{load},t}$, is adequately described by a diurnal profile, ρ_t, and one load parameter for trading days, μ_{td}, and one load parameter for weekends, μ_{weekend}, i.e.

$$\mu_{\mathrm{load},t} = [\mu_{\mathrm{td}} \cdot I_{\mathrm{td},t} + \mu_{\mathrm{weekend}} \cdot (1 - I_{\mathrm{td},t})] \cdot \rho_t$$

$$\rho_t = 1 + \alpha_1 \cos\left(\frac{2\pi t}{360}\right) + \beta_1 \sin\left(\frac{2\pi t}{360}\right) + \alpha_2 \cos\left(\frac{4\pi t}{360}\right) + \beta_2 \sin\left(\frac{4\pi t}{360}\right) \quad (6)$$

where $I_{\mathrm{td},t}$ is an indicator function equal to 1 if the samples of the given operation cycle f are taken on a trading day and 0 if the samples of the given operation cycle are taken on a weekend. The diurnal variation of the ammonia load is modelled using a second order Fourier expansion, which is identical on trading days and weekends. The AR(1)-process in (5) describes the correlated deviations from the mean load $\mu_{\mathrm{load},t}$ from one operation cycle to another, i.e. if the load of the last operation cycle was lower than the average mean load at the given time, given by (6), it is likely that the load of the present operation cycle will also be lower. The one step (operation cycle) predictions of (5) are shown in Figs 2 and 3 by dotted curves. The one-step predictions of the inlet concentrations in the figures are in general good except for two cases. Firstly, short term peak ammonia loads are difficult to predict as they cannot be described as simple diurnal variations, and the occurrences of peak loads are not found to be correlated with other measurements at the WWTP. Secondly, low inlet concentrations of ammonia due to rainfall in the catchment area is a different case from the dry weather situations for which the formulation of a model is given later.

Table 2. Estimates from the ammonia load process

	Aeration tank LT5	Aeration tank LT6	Unit
$\hat{\mu}_{\mathrm{td}}$	8.0 (±0.1)	9.4 (±0.1)	g/m³
$\hat{\mu}_{\mathrm{weekend}}$	6.2 (±0.5)	8.1 (±0.5)	g/m³
$\hat{C}_{\mathrm{rain}}$	2.9 (±0.1)	2.7 (±0.1)	g/m³
$\hat{T}_{h,\mathrm{sewer,total}}$	2.3 (±0.1)	2.3 (±0.1)	h
$\hat{\phi}_{\mathrm{load}}$	0.68 (±0.01)	0.50 (±0.02)	—
$\hat{\alpha}_1$	−0.018 (±0.002)	0.026 (±0.007)	—
$\hat{\beta}_1$	−0.185 (±0.011)	−0.102 (±0.010)	—
$\hat{\alpha}_2$	0.001 (±5%)	0.012 (±0.005)	—
$\hat{\beta}_2$	0.075 (±0.004)	0.061 (±0.010)	—
$\hat{\sigma}_{\epsilon,\mathrm{load}}$	1.8	2.7	g/m³
$\hat{\sigma}_{\epsilon,\mathrm{load}}$	1.6	1.1	g/m³

It is assumed, that the ammonia concentration in the wastewater from households and industries can be characterized by a stationary model like (5) and that this process is not influenced by rainy weather. In the beginning of a rainy period the sewer system contains a large buffer of household and industry wastewater, which is partly mixed with rainwater and flushed to the WWTP. Due to mixing with rainwater the ammonia concentration in the anaerobic tanks will slowly decrease towards a level C_{rain}, which is the average ammonia concentration in the sewer during a long term rainy weather. After the rain has stopped the ammonia concentration will slowly increase to the normal level given by (5). A rainfall event is defined when the flow to the biological part of the plant exceeds 4000 m³/h, and it is considered to be finished when the flow has been below 3500 m³/h for 2 h.

The large buffer of the sewer and all the pretreatment tanks that the raw wastewater has to pass before entering the aeration tanks, can neither be considered as one big ideal mixed tank nor as one large pipe with no mixing. In order to describe the partial mixing of raw wastewater with rainwater and the large buffer capacity of the sewer and the pretreatment tanks, a simple model of N ideal mixed identical tanks in series was found adequate (Carstensen, 1994). For the given data $N = 2$ ideal mixed tanks in series with an hydraulic retention time, $T_{h,\mathrm{sewer}}$ of approximately 1.2 h each gave the best fitting of data during rainy weather. Though, using $N = 3$ ideal mixed tanks in series with an hydraulic retention time of approximately 0.8 h each, a fit nearly as good as using $N = 2$ was obtained. The dotted curves on Figs 2 and 3 show the predicted ammonia concentrations in the anaerobic tanks assuming that the rainfall events do not occur. Thus, during rainy weather (5) is not updated with the estimated inlet concentrations, but instead replaced with the rainy weather model described above. This is seen on the figures where the estimated inlet concentrations of ammonia (solid curves) approaches C_{rain} found in Table 2 and the dotted curves approaches the mean dry weather influent ammonia load, $\mu_{\mathrm{load},t}$ given by (6).

†Normally independent identical distributed.

The estimates of $r_{\mathrm{load},f}$ based on data from one operation cycle are naturally encumbered with some uncertainty

$$\hat{r}_{\mathrm{load},f} = r_{\mathrm{load},f} + e_{\mathrm{load},f} \tag{7}$$

where $e_{\mathrm{load},f}$ is $N_{iid}(0,\sigma^2_{e,\mathrm{load}})$. This problem of two noise components can be handled by applying a Kalman filter to the model (5) (Harvey, 1989). Estimating the parameters relevant to the ammonia load process yields the estimates in Table 2. The SD of the estimates is also given, and approximate 95%-confidence limits are found by multiplying the SD by two.

3.2. *The nitrification process*

During the aerobic phase of the operation cycle ammonia is removed and nitrate is produced by the nitrification process. The intermediate formation and removal of nitrite is not considered in the models proposed in this paper. The simultaneous dilution of nitrogen from the aeration tanks is a separate process which is described later. A Monod-kinetic type of expression is often used to describe the dependency of the oxygen and ammonia concentration using two half-saturation constants. The removal of ammonia by nitrification is according to Henze *et al.* (1987) given by the following differential equation

$$\frac{\mathrm{d}C_{\mathrm{NH4},t}}{\mathrm{d}t} = k_{\mathrm{nit,max}} \cdot \frac{C_{\mathrm{NH4},t}}{C_{\mathrm{NH4},t} + K_{\mathrm{NH4}}} \cdot \frac{C_{\mathrm{O_2},t}}{C_{\mathrm{O_2},t} + K_{\mathrm{O_2}}} \cdot X_{\mathrm{B,nit},t} \tag{8}$$

where $k_{\mathrm{nit,max}}$ is the maximum nitrification rate, K_{NH4} and $K_{\mathrm{O_2}}$ are the half saturation constants for the Monod-kinetic expressions, and $X_{\mathrm{B,nit},t}$ is the suspended sludge concentration of nitrifying bacteria. A similar expression holds for the formation of nitrate.

From the declining curves of ammonia concentrations and the inclining curves of nitrate concentrations a rate expression in discrete time similar to (8) can be identified.

$$r_{\mathrm{nit},t} = r_{\mathrm{nit,max},f} \cdot \frac{C_{\mathrm{NH4},t-1}}{C_{\mathrm{NH4},t-1} + K_{\mathrm{NH4}}} \cdot \frac{\bar{O}_{2,t-2}}{\bar{O}_{2,t-2} + K_{\mathrm{O_2}}} \cdot \overline{\mathrm{SS}}_{t-1}. \tag{9}$$

For the rates estimated on nitrate concentrations the average delay from the measured oxygen concentrations to $r_{\mathrm{nit},t}$ was only one sample. Note that the rate expression only depends on lagged values of the ammonia, oxygen, and suspended sludge concentration. The use of lagged values in the Monod-kinetic expressions in (9) is justifiable, because of the persistence of the process between samples. It should be stressed that the expression above (9) also incorporates the use of ammonia for heterotrophic growth.

Thus, with the slopes of the ammonia and nitrate curves during the aerobic phase of the operation cycle given by $r_{\mathrm{nit},t}$ an estimate of $r_{\mathrm{nit,max},f}$ can be derived. The two half-saturation constants K_{NH4} and $K_{\mathrm{O_2}}$ are also estimated, but they are assumed to be constant throughout the given dataset while $r_{\mathrm{nit,max},f}$ may vary from one operation cycle to another as indicated by the index f. Holmberg and Ranta (1982) show that a practical identification of the half-saturation constants in an expression like (8) is not applicable based on a single decay of ammonia concentrations or increase of nitrate concentrations, since the parameter estimates will be very correlated. However, the fact that the saturation constants are estimated on 227 operation cycles with wide ranges of ammonia and oxygen concentrations results in a low correlation between the estimates. The estimates of the half-saturation constants are, however, still more mutual correlated than the other estimates are. Comparing the differential equation (8) with the discrete time rate expression (9) the following approximation can be found

$$r_{\mathrm{nit,max},f} \approx k \cdot \frac{X_{\mathrm{B,nit},f}}{SS_f} \tag{10}$$

i.e. $r_{\mathrm{nit,max},f}$ shows the variations of the relative concentration of active nitrifying bacteria.

The estimated maximum nitrification rates for each specific operation cycle based on the ammonia concentrations in LT5 and LT6 are shown in Figs 4 and 5, and in Figs 6 and 7 the analogous estimates based on the nitrate concentrations are shown. Estimated maximum nitrification rates based on less than 4 observations are regarded as missing values, which is indicated by the gaps on the solid curves in the figures. The variation of the relative amount of active nitrifying bacteria in the activated sludge is larger than the removal of excess sludge can account for (the average sludge age is approximately 20 days). However, there seems to be a correlation between the estimated maximum nitrification rates and past load of ammonia such that low inlet concentrations of ammonia result in a low amount of active nitrifiers. The scaling of the figures is not identical, since the parameters are estimated separately with different half-saturation constants, but in general the trends of the figures correspond well. The correlation of the 4 time series of estimated maximum nitrification rates is in the range from 36% to 79% with the smallest correlation on the estimates from the ammonia sensor in LT5 which was recalibrated in the middle of the period. The estimates based on measurements of ammonia are biased because ammonia is accumulated in the biomass and simultaneously produced by hydrolysis, which are phenomenon not incorporated in the estimates. However, the apparent daily variation in the figures cannot be explained from (8).

The variations of the maximum nitrification rate can be modelled using an AR(1)-model

$$(1 + \phi_{\mathrm{nit}} B)(r_{\mathrm{nit,max},f} - \mu_{\mathrm{nit,max},f}) = \epsilon_{\mathrm{nit},f} \tag{11}$$

where B is the back-shift operator, $\epsilon_{\mathrm{nit},f}$ is $N_{iid}(0,\sigma^2_{\epsilon,\mathrm{nit}})$ and $\mu_{\mathrm{nit,max},f}$ is the mean maximum nitrification rate for the given operation cycle. The correlation between the maximum nitrification rate and the

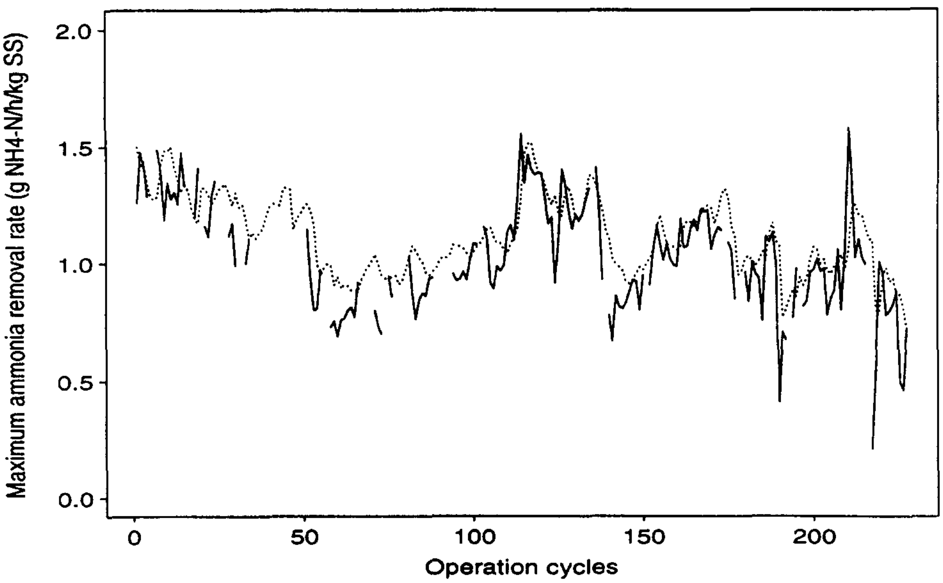

Fig. 4. Estimates (solid line) and one-step predictions (dotted line) of maximum nitrification rates, $r_{\text{nit,max},f}$, $f = 1 \ldots 227$, based on ammonia concentrations from aeration tank LT5.

ammonia load rate can be expressed in the following way

$$\mu_{\text{nit,max},f} + \mu^0_{\text{nit,max}} + \gamma \cdot r_{\text{load},f-1} \quad (12)$$

i.e. the mean maximum nitrification rate is a linear function of the previous ammonia load rate. Because perfect knowledge of $r_{\text{load},f-1}$ cannot be obtained, the estimator of $r_{\text{load},f-1}$ as the mean of the posteriori distribution after Kalman filtering has been applied (Harvey, 1989). The estimator of $r_{\text{load},f-1}$ also incorporates the rainy weather model described above to give an estimate of the ammonia load during rainy weather. The linear relationship in (12) is rather empirical and further investigation needs to be done to find a more reasonable relationship.

The estimates of $r_{\text{nit,max},f}$ based on data from one operation cycle are encumbered with some uncertainty

$$\hat{r}_{\text{nit,max},f} = r_{\text{nit,max},f} + e_{\text{nit},f} \quad (13)$$

where $e_{\text{nit},f}$ is $N_{iid}(0,\sigma^2_{e,\text{nit}})$. A Kalman filter was used to distinguish the two noise components in (11) and

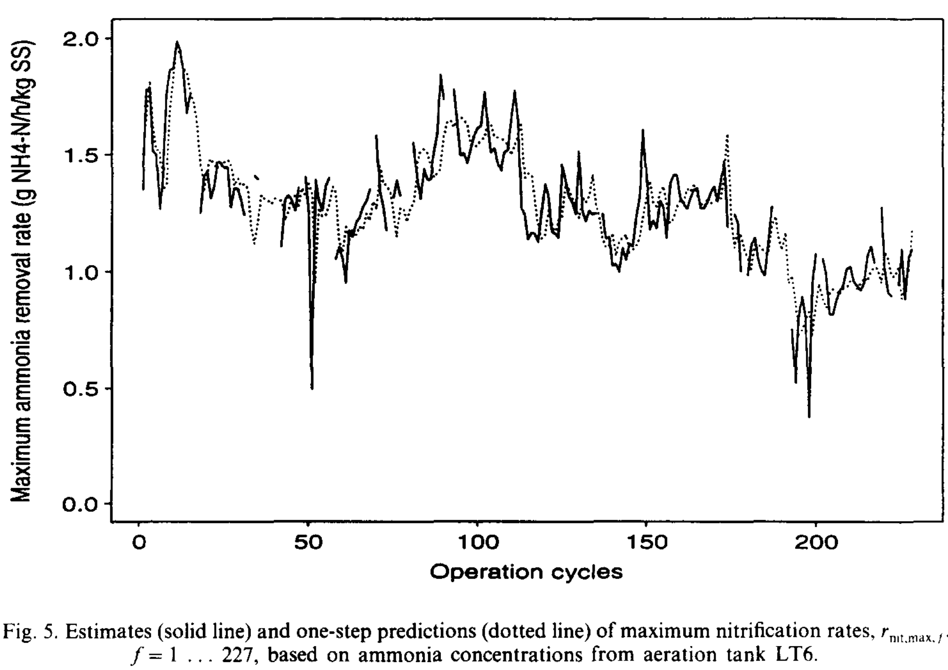

Fig. 5. Estimates (solid line) and one-step predictions (dotted line) of maximum nitrification rates, $r_{\text{nit,max},f}$, $f = 1 \ldots 227$, based on ammonia concentrations from aeration tank LT6.

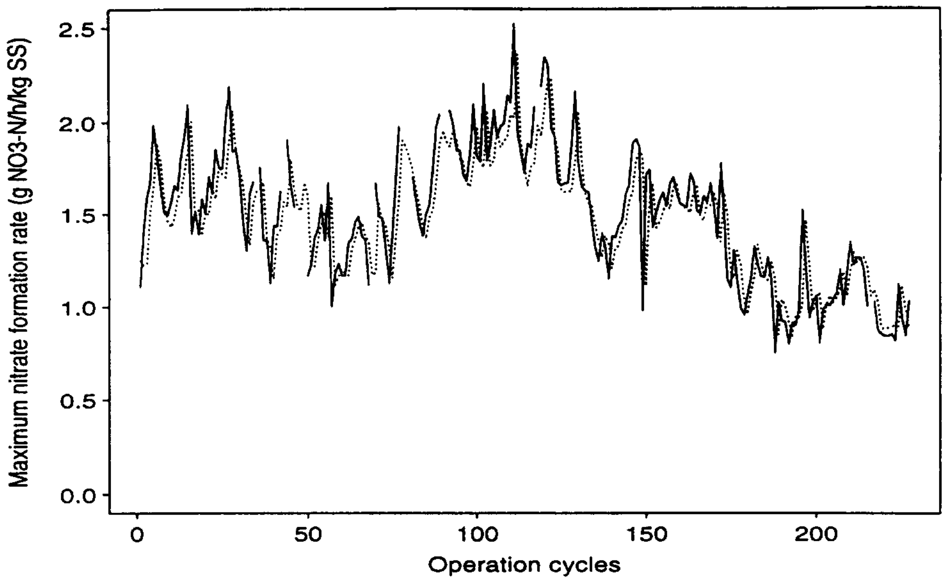

Fig. 6. Estimates (solid line) and one-step predictions (dotted line) of maximum nitrification rates, $r_{nit,max,f}$, $f = 1 \ldots 227$, based on nitrate concentrations from aeration tank LT5.

(13). The estimated parameters of the nitrification process are shown in Table 3. The SD of the estimates based on nitrate concentrations is larger than the uncertainty of the estimates based on ammonia concentrations. This is due to a built-in filtering of the ammonia concentration sensors, such that the uncertainty of the ammonia measurements is smaller. Because the filtering removes some of the dynamics of the ammonia concentration in the tank, the estimates in the first two columns of Table 3 are somewhat biased.

The true ammonia removal rate is actually larger than estimated in the models proposed here, because ammonia is produced by hydrolysis simultaneously with the nitrification process. However, the hydrolysis process cannot be identified from the given dataset, and as a result the estimated nitrification rate is larger when estimated from the nitrate concentrations. The estimates of the half-saturation constants are based only on the available data with possible calibration offset errors, and should as though not be considered as true theoretical values.

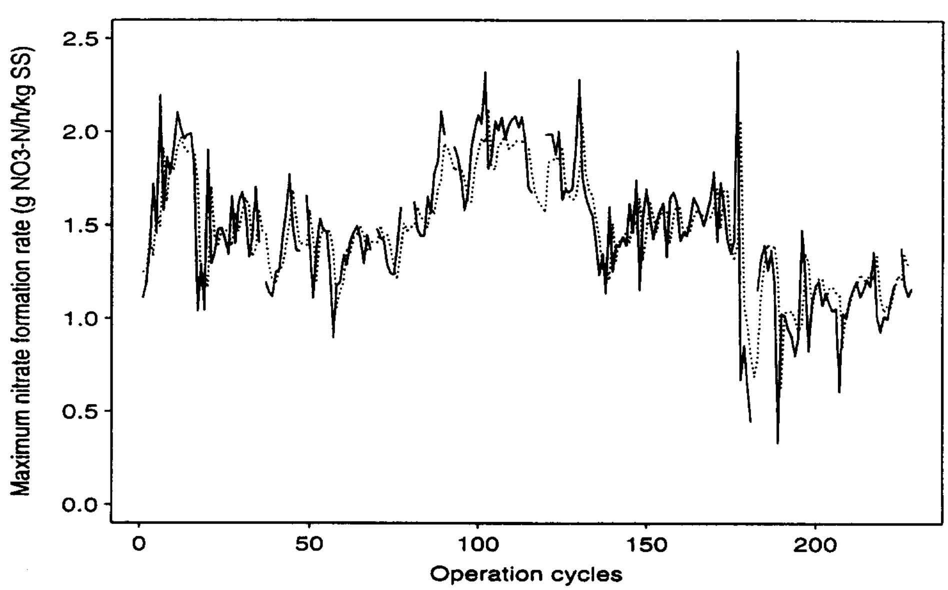

Fig. 7. Estimates (solid line) and one-step predictions (dotted line) of maximum nitrification rates, $r_{nit,max,f}$, $f = 1 \ldots 227$, based on nitrate concentrations from aeration tank LT6.

Table 3. Estimates from the nitrification process

	$C_{NH4,LT5}$	$C_{NH4,LT6}$	$C_{NO3,LT5}$	$C_{NO3,LT6}$	Unit
$\bar{r}_{nit,max,f}$	1.11	1.28	1.48	1.48	g/h/kg SS
$\hat{K}_{NH4}$	0.76 (±0.02)	0.56 (±0.03)	0.60 (±0.04)	0.49 (±0.05)	g/m³
$\hat{K}_{O_2}$	0.71 (±0.01)	0.77 (±0.01)	0.67 (±0.02)	0.86 (±0.03)	g/m³
$\hat{\phi}_{nit}$	0.93 (±0.02)	1.00 (±0.02)	0.91 (±0.04)	0.86 (±0.05)	—
$\hat{\sigma}_{\iota,nit}$	0.06	0.06	0.18	0.20	g/h/kg SS
$\hat{\sigma}_{\epsilon,nit}$	0.12	0.12	0.08	0.12	g/h/kg SS

However, the 4 sets of estimates of the half-saturation constants for the nitrification process are close to the expected values. The high AR(1)-parameter estimates (close to 1) show that the maximum nitrification rate contains slow variations which are not modelled in (9). The slow variations could be caused by variations in the temperature, but unfortunately no data to investigate this hypothesis have been available. While reasonable explanations to the slow variation can be found, the short term variations of up to 25% of the maximum nitrification rate is quite surprising. This topic needs to be further investigated.

3.3. *The denitrification process*

During the anoxic phase of the operation cycle nitrate is transformed into nitrogen gas by the denitrification process. The simultaneous dilution of nitrate from the aeration tanks is a separate process which is described in the next section. In Henze *et al.* (1987) the removal of nitrate by denitrification is described by the following differential equation

$$\frac{dC_{NO3,t}}{dt} = k_{denit,max} \cdot \frac{C_{NO3,t}}{C_{NO3,t} + K_{NO3}} \cdot \frac{S}{S + K_S} \cdot X_{B,denit,t} \tag{14}$$

where $k_{denit,max}$ is the maximum denitrification rate, S is the readily bio-degradable substrate concentration, $X_{B,denit,t}$ is the suspended sludge concentration of denitrifying bacteria in the aeration tank and K_{NO3} and K_S are the half-saturation constants for the two rate limiting concentrations, the nitrate and the readily bio-degradable substrate concentration.

At present no measuring equipment for on-line monitoring of the readily bio-degradable substrate concentration is available at Aalborg West WWTP. Thus, an identification of a rate expression with both Monod kinetic expressions as in (14) is not possible, and some basic assumptions about the readily bio-degradable substrate concentration are necessary in order to identify the effect of the readily bio-degradable substrate concentration on the denitrification process.

Primarily, during the anoxic phase of the operation cycle with raw wastewater running into the aeration tank, it is assumed that the readily bio-degradable substrate concentration will remain approximately constant through simultaneous supply from the raw wastewater and removal by the denitrification process. Secondly, the readily bio-degradable substrate concentration in the process tank during the anoxic phase of the operation cycle is assumed to be proportional to the previous load of ammonia. The first assumption is crude because low loads of readily bio-degradable substrate may not be able to maintain a constant concentration in the aeration tank, and high loads of readily bio-degradable substrate may result in a rising concentration in the aeration tanks during the anoxic phase. However, for average loads of readily bio-degradable substrate this assumption may prove to be right. The second assumption is based on well-known result that the ammonia and organic material loads are highly correlated (Dupont & Sinkjær, 1994).

Using the assumption above and the data from the declining curves of nitrate concentrations a rate expression similar to (14) can be identified

$$r_{denit,t} = r_{denit,max,f} \times \cdot \frac{C_{NO3,t-1}}{C_{NO3,t-1} + K_{NO3}} \cdot \frac{r_{load,f-1}}{r_{load,f-1} + K_{load}} \cdot \overline{SS}_{t-1}. \tag{15}$$

The half-saturation constants are assumed to be constant throughout the given data set, while the maximum denitrification rate may vary from one operation cycle to another. Note that the readily bio-degradable substrate concentration in (14) is replaced by $r_{load,f-1}$ and $X_{B,denit,t}$ is replaced by $\overline{SS}_{t-1}$, which leads to the following approximation given the two assumptios above

$$r_{denit,max,f} \approx k \cdot \frac{X_{B,denit,f}}{SS_f} \tag{16}$$

i.e. $r_{denit,max,f}$ shows the variations of the relative concentration of active denitrifying bacteria.

The estimated maximum denitrification rates for each specific operation cycle based on the nitrate concentrations in LT5 and LT6 are shown in Fig. 8 and Fig. 9, respectively. Estimated maximum denitrification rates based on less than 4 observations are regarded as missing values, which is indicated by the gaps on the solid curves in the figures. The trends of the two figures are similar, but the influence of the readily bio-degradable substrate concentration on the denitrification rate makes the estimates on the figures differ for a number of operation cycles. The correlation between the two time series of estimated maximum denitrification rates is 54%.

The variation of the maximum denitrification rate can be modelled using an AR(1)-model

$$(1 + \phi_{denit} B)(r_{denit,max,f} - \mu_{denit,max}) = \epsilon_{denit,f} \tag{17}$$

where B is the back-shift operator, $\epsilon_{denit,f}$ is $N_{iid}(0, \sigma^2_{\epsilon,denit})$ and $\mu_{denit,max}$ is the mean maximum

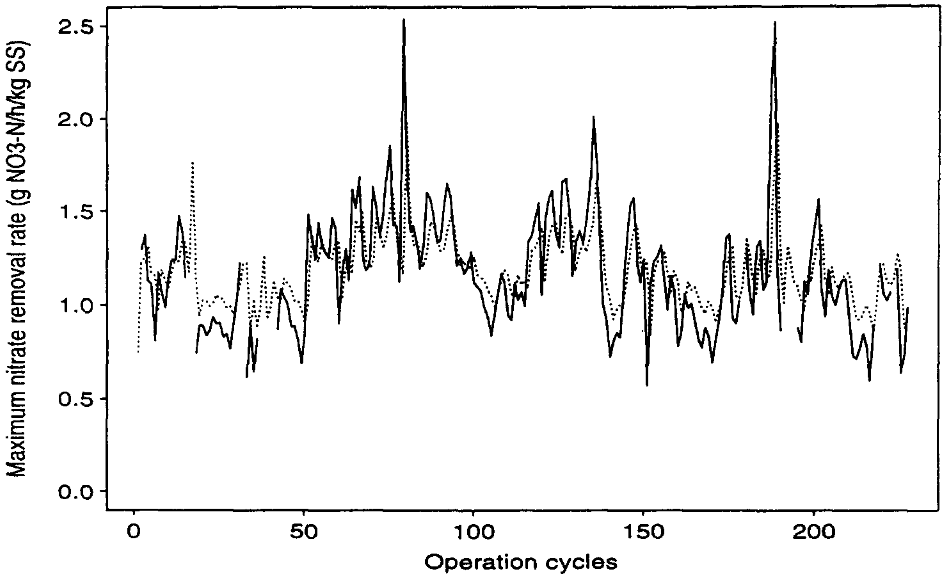

Fig. 8. Estimates (solid line) and one-step predictions (dotted line) of maximum denitrification rates based on nitrate concentrations from aeration tank LT5.

denitrification rate. The influence of the load process on the denitrification process has already been modelled in (15).

The estimates of $r_{\text{denit,max},f}$ based on data from one operation cycle are encumbered with some uncertainty which is modelled as follows

$$\hat{r}_{\text{denit,max},f} = r_{\text{denit,max},f} + e_{\text{denit},f} \quad (18)$$

where $e_{\text{denit},f}$ is $N_{iid}(0,\sigma^2_{\text{e,denit}})$. As for the nitrification process a Kalman filter was used to distinguish the two noise component in (17) and (18). The parameter estimates of the denitrification process are shown in Table 4.

The estimates of the half-saturation constant for the nitrate concentration are most likely too high, due to the shortage of readily bio-degradable substrate at the end of the anoxic phase. The estimates of K_{load} cannot be related to K_S, because the relationship between r_{load} and the readily bio-degradable substrate concentration is not known, but the use of r_{load} in the Monod-kinetic expression gives a good description of the dependency of the readily bio-degradable substrate.

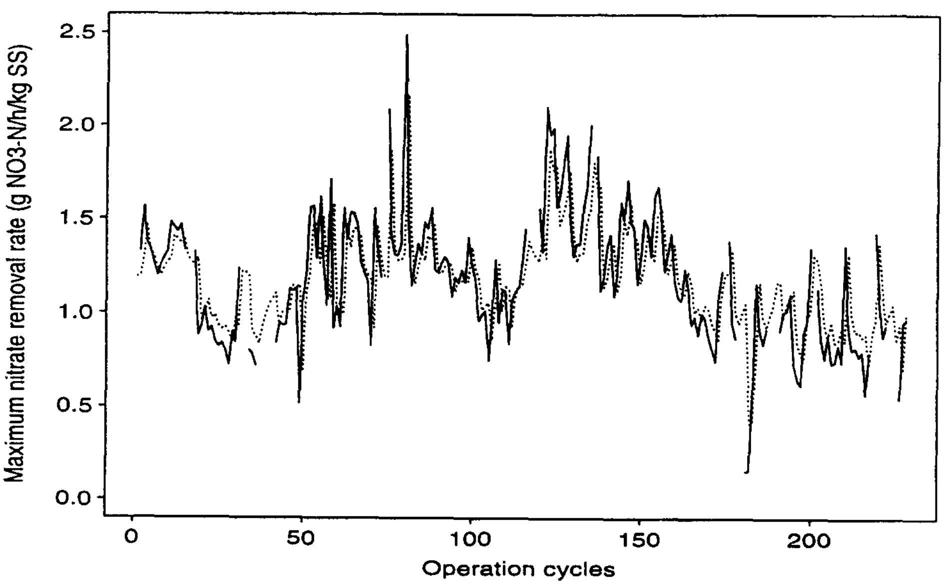

Fig. 9. Estimates (solid line) and one-step predictions (dotted line) of maximum denitrification rates based on nitrate concentrations from aeration tank LT6.

Table 4. Estimates from the denitrification process

	Aeration tank LT5	Aeration tank LT6	Unit
$\bar{r}_{denit,max,f}$	1.19	1.18	g/h/kg SS
$\hat{K}_{NO3}$	0.99 ($\pm$0.06)	1.08 ($\pm$0.13)	g/m^3
$\hat{K}_{load}$	1.55 ($\pm$0.05)	1.73 ($\pm$0.16)	g/m^3/h
$\hat{\phi}_{denit}$	0.59 ($\pm$0.04)	0.75 ($\pm$0.06)	—
$\hat{\sigma}_{\varepsilon,denit}$	0.26	0.28	g/h/kg SS
$\hat{\sigma}_{c,denit}$	0.00	0.00	g/h/kg SS

3.4. Transport of nutrients

In a BIO-DENIPHO WWTP the two alternating aeration tanks are connected by a spill-way between the two tanks, such that a continuous flow through the plant is ensured. This flow between the two aeration tanks naturally affects the ammonia and nitrate concentrations in the two tanks, because the nutrients are also transported by the flow. Identification of this process is necessary for the mass balance of the plant, while the results of estimating this process have been left out. Information about estimating this process and the related results can be found in Carstensen (1994) and Carstensen *et al.* (1993).

4. CONCLUSION

In this study, the complex processes of biological nutrient removal on a full scale WWTP operated according to the BIO-DENIPHO scheme have been modelled using nonlinear grey box models, which incorporates physical knowledge of the plant in the model formulation. Examination of data covering 27 days of on-line monitoring of ammonia and nitrate concentrations in two alternating aeration tanks have shown that four significant processes can be revealed from the data:

ammonia load to the plant;
the nitrification process;
the denitrification process;
nutrient transport.

Rates of these processes are estimated for every operation cycle and stochastic models describing the variation of the rates from one operation cycle to another are proposed and identified. The ammonia load rate is found to correlate strongly to both the nitrification and denitrification process rates. Thus, further research is needed to explain these correlations.

From a physical interpretation of the parameters in the rate expressions and using simple assumptions, estimates of the relative number of active nitrifying and denitrifying bacteria in the suspended sludge are obtained. The estimates of these parameters, which are essential to biological nutrient removal, have a larger variation than the removal of excess sludge can actually account for. The reason for this phenomena is not known, so whether the variations are controllable or forced on the plant from load or factors not included in the models (e.g. temperature, alkalinity, wastewater composition) needs to be further investigated. The physical parameters from the models have all been estimated to values close to the expected physical parameters values, and on-line estimates of the process rates gives a clear indication of the state of the plant at any time. Furthermore, information concerning other parts of the plant and the sewer system are obtained using data from the aeration tanks only.

The models include the significant processes, which can be quantified from the given time series of ammonia and nitrate concentrations, and as though not suitable for giving a detailed picture of the total biochemical and hydraulic processes in the plant. The objective of the models is to make it possible to extract as much information as possible from the given data about the processes by statistical methods. Experimental planning of the operation of the plant might increase the number of processes, that can be identified and quantified, and such planning will most likely improve the parameters estimates.

The models are useful for on-line surveillance of load and nutrient removal rates, and for control of the plant, because the states of the plant will be quantified by the parameters of the proposed models. Application of the models at a plant will potentially result in lower energy costs and effluent discharges, because the plant can be operated according to the given material loads and removal rates. Changes in the state of the plant are detected in the parameter estimates of the models. In that way the necessary on-line controlling actions can be taken, and alarms for inhibitions and extreme loads can be obtained.

Acknowledgements—This work was sponsored by The Danish Academy for Technical Sciences (in Danish: ATV) and I. Krüger Systems AS. Henrik Bechmann M.Sc. kindly provided the data from Aalborg West WWTP and Dr Dines Thornberg gave valuable comments.

REFERENCES

Bhat N. and McAvoy T. (1990) Use of neural nets for dynamic modeling and control of chemical process systems. *Comput. Chem. Engng* **14,** 573–583.

Box G. and Jenkins G. (1976) *Time Series Analysis, Forecasting and Control.* Holden-Day, Calif.

Bundgaard E. (1988) Nitrogen and phosphorus removal by the BIO-DENITRO and BIO-DENIPHO processes. In *Proc. Int. Workshop on Wastewater Treat. Technol.,* Copenhagen.

Carstensen J. (1994) Identification of wastewater processes. Ph.D. thesis, Technical Univ. of Denmark, Lyngby.

Carstensen J., Madsen H., Poulsen N. and Nielsen M. (1993) Grey box modelling in two time domains of wastewater pilot scale plant. *Environmetrics* **4,** 187–208.

Couillard D. and Zhu S. (1992) Control strategy for the activated sludge process under shock loading. *Wat. Res.* **26,** 649–655.

Dupont R. and Sinkjær O. (1994) Optimisation of wastewater treatment plants by means of computer models. Submitted for *IAWQ Biannual Conf.,* Budapest, 1994.

Goodwin G. and Payne R. (1977) *Dynamic System Identification—Experiment Design and Data Analysis.* Academic Press, N.Y.

Harvey A. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge Univ. Press.

Henze M., Grady C., Gujer W., Marais G. and Matsuo T. (1987) Activated sludge model No. 1. IAWPRC Sci. Technol. Rep. No. 1, London.

Holmberg A. and Ranta J. (1982) Procedures for parameter and state estimation of a microbial growth process model. *Automatica* **18,** 181–193.

Olsson G., Andersson B., Hellström B., Holmström H., Reinius L. and Vopatek P. (1989) Measurements, data analysis and control methods in wastewater treatment plants—state of the art and future trends. *Wat. Sci. Technol.* **21,** 1333–1345.

Schwarz G. (1978) Estimating the dimension of a model. *Ann. Statist.* **6,** 461–464.

Thomsen H. and Nielsen M. (1992) Practical experience with on-line measurements of NH4, NO3, PO4, redox, MLSS, and SS in advanced activated sludge plants. In *HYDROTOP92, The City and the Water*, Marseilles, France, Vol. 2, pp. 378–388.