Pergamon

0273-1223(95)00186-7

Wat. Sci. Tech. Vol. 31, No. 2, pp. 125–133, 1995.

0273–1223/95 $9·50 + 0·00

STATISTICAL IDENTIFICATION OF MONOD-KINETIC PARAMETERS FROM ON-LINE MEASUREMENTS

Jacob Carstensen*, Poul Harremoës* and
Henrik Madsen**

**Institute of Environmental Science and Engineering, Building 115,
Technical University of Denmark, DK-2800 Lyngby, Denmark
**Institute of Mathematical Modelling, Building 321, Technical University of
Denmark, DK-2800 Lyngby, Denmark*

ABSTRACT

Time series analysis of on-line monitored ammonia and nitrate concentrations from full-scale wastewater treatment plants operated according to an alternating scheme makes the identification of Monod-kinetic expressions possible. The models presented in the present context only include kinetic parameters which have shown to be significant in a statistical sense. Estimates of kinetic parameters for the nitrification and denitrification processes are obtained by applying these models to the time series of ammonia and nitrate concentrations. In this paper, the concept of statistical identification which depends on the two conditions of theoretical and practical identification, is described. Experiences from estimating time series models of the nitrification and denitrification processes with data from two wastewater treatment plants are discussed. It appears that the dynamic of the biological processes on a full-scale plant is strongly varying. The proposed models are suitable for on-line control, because the states of the plant are continuously updated as new information from the on-line sensors becomes available.

KEYWORDS

Control; denitrification; estimation; grey box modelling; identifiability; Monod-kinetic expressions; nitrification; non-linear time series models; parameter identification; surveillance.

INTRODUCTION

Today, on-line sensors for measuring ammonia, nitrate and phosphate in the reaction tanks of a wastewater treatment plant are available. These instruments have been shown to give very reliable information on the wastewater processes for long periods of time when maintained regularly and properly. This opens up new world of information, and the most interesting parameters in the deterministic description of the biological processes (IAWQ model No. 1, Henze *et al.*, 1987) can be identified as suggested in Carstensen *et al.* (1994). The stochastic models used for identifying the significant parameters of the biological processes are called grey box models, since they incorporate the most important deterministic terms in the model formulation and stochastic terms to describe the residual variation. The objective of these simplified models is not to give a better description of the biological processes, but to formulate operational dynamic models capable of giving on-line information on the present state of the wastewater treatment plant. This is obtained by a built-in adaptivity of the models such that the model is updated for each new time step, when new information from the on-line measurements becomes available.

In the deterministic modelling of activated sludge processes the concept of model identification has been addressed in Jepsen & Olsson (1993) and Larrea *et al.* (1992). In both cases it was found that the IAWQ model No. 1 is overparameterised and hence, the order of the deterministic model has to be reduced in order to be able to estimate the parameters of the model. The alternative grey box modelling approach is primarily based on the available data from on-line sensors, where deterministic and stochastic terms are added to the model after their significance has been tested, as a part of the iterative modelling process.

One important feature of the grey box models is that the Monod-kinetic parameters of the nitrification and denitrification process can actually be identified and estimated by means of the prediction error decomposition (Harvey, 1989) and a maximum likelihood estimation method. In Holmberg & Ranta (1982) it is shown that the estimation of Monod-kinetic parameters may not always be possible due to lack of information in the data. This leads to the fundamental question of identifiability, i.e. is it possible to find a unique solution for each of the unknown parameters of the model, from data collected in experiments performed on a real system.

THE CONCEPT OF IDENTIFICATION

The question of identifiability is a fundamental one in stochastic modelling. Loosely speaking, the problem is whether the identification procedure will yield unique values of the parameters of the model structure given the data. In order to give the standard definition of identifiability of the statistical literature it is important to distinguish between a model and a structure. A model specifies a distribution for data while a structure specifies the parameters determining that distribution. Given this background, the following concepts may be defined.

(a) If two models based on different structures have the same joint density function they are said to be observationally equivalent.
(b) A structure is identifiable if there exist no other observationally equivalent structure.
(c) A model is identifiable if its structure is identifiable.

The concepts above are also known as *theoretical or structural identifiability*. The term deterministic identifiability is also widely used, since the problem of theoretical identifiability is an algebraic problem based on the deterministic equations. In Godfrey & DiStefano (1985) it is shown that the four parameters in a typical microbial growth model (maximum specific growth rate, half-saturation constant, yield coefficient and decay rate coefficient), consisting of two equations with Monod-kinetics, are theoretically identifiable.

Identifiability has an immediate bearing on estimation. If two structures are observationally equivalent (i.e. have the same joint density function), the probability of generating a particular set of observations is the same for both structures. Thus, there is no way of differentiating between them on the basis of data. Furthermore, it will often be the case that an attempt to estimate models which are not identifiable will run into practical difficulties.

This leads to the second condition of identifiability: *the practical or numerical identifiability*. This issue involves aspects on whether the data set contains sufficient information of the underlying data generating process to distinguish between different models (or could the data result from two observationally equivalent models). Thus, if the model is theoretically identifiable and the parameter estimates are found to be significant in a statistical sense, then the model is also practically identifiable. For the practical identification of Monod-kinetic expressions from on-line measurements of ammonia and nitrate, it is crucial that the wastewater treatment plant is persistently excited, i.e. the data should reflect abundant dynamics. Wastewater treatment plants of the SBR-type (Sequencing Batch Reactor) show this characteristic, while plants of the recirculating type may be operated by intermittent aeration in such a way that sufficient dynamic variation in data is obtained (see Nielsen *et al.*, 1994). Figure 1 shows the dynamics of ammonia and nitrate measurements in the aeration tank of a full-scale plant operated according to the BIO-DENIPHO scheme (BIO-DENIPHO is a patented process developed by I. Krüger Systems in cooperation with the Institute of Environmental Science and Engineering at the Technical University of Denmark).

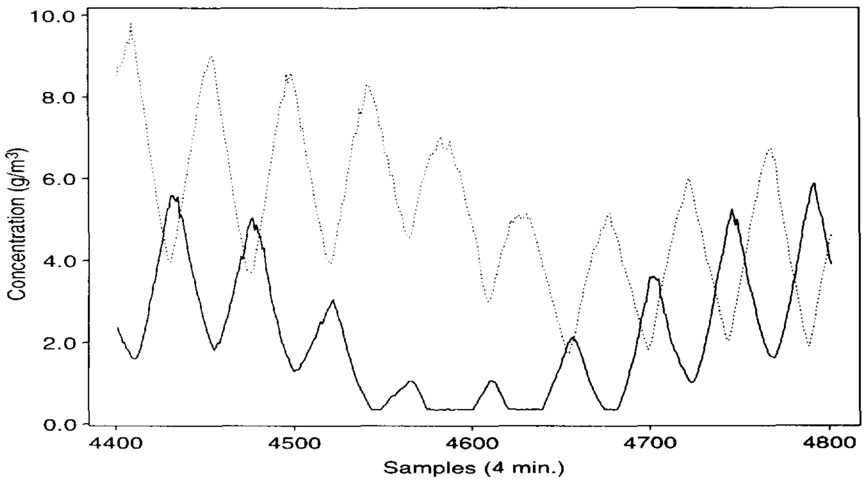

Fig. 1. On-line measurements of ammonia (solid curve) and nitrate (dotted curve) from Aalborg West WWTP.

ESTIMATING THE NITRIFICATION PROCESS

Nitrification is a two-step microbiological process transforming ammonia into nitrite and subsequently into nitrate – a process performed by autotrophic bacteria. Though, the intermediate formation and removal of nitrite is often neglected when modelling this process. For the modelling of the nitrification process, Monod-kinetic expressions are frequently used to describe the dependency of oxygen and ammonia on the process rate. In the deterministic world of activated sludge modelling these relationships are formulated by use of a non-linear differential equation for the removal of ammonia

$$\frac{dS_{NH4}}{dt} = -\frac{\mu_{max,A}}{Y_{obs,NH4}} \cdot \frac{S_{NH4}}{S_{NH4}+K_{NH4}} \cdot \frac{S_{O2}}{S_{O2}+K_{O2}} \cdot X_{B,A} \tag{1}$$

where

S_{NH4} = the concentration of ammonia
S_{O2} = the concentration of dissolved oxygen
$X_{B,A}$ = the concentration of active autotrophic biomass
$\mu_{max,A}$ = the maximum specific growth rate of autotrophic bacteria
$Y_{obs,NH4}$ = the observed biomass yield coefficient of ammonia
K_{NH4}, K_{O2} = the appropriate half-saturation constants

and a similar expression for the formulation of nitrate. An identifiability analysis of (1) quickly reveals that a simultaneous identification of both $\mu_{max,A}$ and $Y_{obs,NH4}$ is not theoretically feasible unless (1) is coupled with a differential equation for the growth of $X_{B,A}$. Furthermore, no methods for measuring the concentration of autotrophic biomass exist. This makes (1) inappropriate for the identification of the Monod-kinetic expressions used to describe the nitrification process. A more appropriate model is

$$\frac{\Delta S_{NH4,t}}{\Delta t}=-r_{nit,max}\cdot\frac{S_{NH4,t-1}}{S_{NH4,t-1}+K_{NH4}}\cdot\frac{S_{O2,t-1}}{S_{O2,t-1}+K_{O2}}\cdot X_{SS,t-1} \qquad (2)$$

for the removal of ammonia, and

$$\frac{\Delta S_{NO3,t}}{\Delta t}=r_{nit,max}\cdot\frac{S_{NH4,t-1}}{S_{NH4,t-1}+K_{NH4}}\cdot\frac{S_{O2,t-1}}{S_{O2,t-1}+K_{O2}}\cdot X_{SS,t-1} \qquad (3)$$

for the removal of ammonia, and for the formation of nitrate. $S_{NH4,t}$, $S_{O2,t}$, $X_{SS,t}$ are the on-line measured concentrations of ammonia, oxygen and suspended solids in the aeration tank at time t, and Δt is the time between samples. It is seen that $r_{nit,max}$ is a parameter incorporating the maximum specific growth rate of autotrophic bacteria, the observed biomass yield coefficient of ammonia and the fraction of autotrophic bacteria in the suspended solids. Comparing the discrete time formulation of the nitrification process with (1) it is seen that the use of lagged values in the Monod-kinetic expressions in (2) and (3) will result in biased estimates of K_{NH4} and K_{O2} because the process rate is adjusted one sample behind the present. This one sample time delay is required for the identification used – see Carstensen (1994). In order to compensate for this lack of continuous adjustment of the process rate higher values of K_{NH4} and K_{O2} will yield the same dynamics as the continuous time formulation. If the time between samples is small relative to the rate of the process the bias of the estimated half-saturation constants will also be small, and if the time between samples approaches zero, the estimates of K_{NH4} and K_{O2} are unbiased estimates of the half-saturation constants in (1).

In estimating the nitrification process by use of (2) and (3) and on-line measurements of ammonia, nitrate, oxygen and suspended solids concentrations, it should be stressed that these two equations will yield different results. This is due to the fact that during aerobic conditions in the aeration tank, ammonia is produced by hydrolysis of larger organic compounds, accumulated in the biomass, and removed by nitrification. Thus, the process rate in (2) actually is a conglomerate of the three simultaneous processes – ammonification, biomass assimilation and nitrification. On the other hand, the discrete time model based on nitrate (3) is an unbiased estimator of the nitrification process since the only significant reaction involving nitrate during aerobic conditions is the nitrification process, unless simultaneous denitrification takes place.

In Carstensen *et al.* (1993) and Carstensen *et al.* (1994) the nitrification process has been identified and estimated on data from the Lundtofte pilot-scale plant (data set covering 16 days) and the Aalborg West WWTP (data set covering 27 days), respectively. For both data sets the half-saturation constants are assumed to be constant throughout the total period, while $r_{nit,max}$ has been estimated for every single aerobic period in the alternating operation mode of the BIO-DENIPHO process. At the Aalborg West WWTP, expressions identical to (2) and (3) are identified, while the discrete time model formulation of the nitrification process at the Lundtofte pilot scale plant had to be reduced, due to two limitations in the data which seem to give non-identifiable parameters. Firstly, on-line measurements of the suspended solids concentrations are not available at the plant and hence, $X_{SS,t}$ is assumed to be constant. The suspended solids concentration in the aeration tanks is maintained approximately constant within the range 4–5 mg SS/l throughout intermittent removal of excess sludge. Secondly, throughout the given period the plant was operated with the same constant oxygen setpoint thereby making the practical identification of K_{O2} impossible, i.e. the effect of operating at different oxygen levels on the rate of the nitrification process could not be identified from the dynamics of the ammonia and nitrate measurements.

The results of estimating (2) and (3) based on the data sets from the Lundtofte pilot-scale plant and the Aalborg West WWTP are summarised in Table 1. At the Aalborg West WWTP both aeration tanks of the alternating process are equipped with on-line sensors so that redundant information on the nitrification process is obtained. The results of the parameter estimation appear to be very reproducible for each of the two plants and close to the suggested values in the literature, see e.g. Randall *et al.* (1992). The authors do not claim these parameters do exist. The keypoint is that (1) provides a good and well-established model formulation for the description of the nitrification process. Hence, by estimating (2) and (3) an operational model for on-line identification of the nitrification process incorporating some physical knowledge of the process is obtained. The standard deviations of the estimated half-saturation constants which are given with the parameter estimates in the table, are likely to be slightly too small due to the applied estimation routine.

TABLE 1. Parameter Estimates of the Nitrification Process

Parameter	Unit	Aalborg West WWTP - LT5	Aalborg West WWTP - LT6	Lundtofte pilot scale plant
$\hat{r}_{nit,max}$ in (2)	mg NH4-N/h/g SS	1.02 (±0.23)	1.28 (±0.26)	2.36 (±0.18)[1]
$\hat{k}_{NH4}$ in (2)	mg NH4-N/l	0.76 (±0.02)	0.56 (±0.03)	0.44 (±0.02)
$\hat{k}_{O2}$ in (2)	mg O2/l	0.71 (±0.01)	0.77 (±0.01)	-
$\hat{r}_{nit,max}$ in (3)	mg NO3-N/h/g SS	1.50 (±0.36)	1.47 (±0.36)	2.16 (±0.21)[1]
$\hat{k}_{NH4}$ in (3)	mg NH4-N/l	0.60 (±0.04)	0.49 (±0.05)	0.40 (±0.01)
$\hat{k}_{O2}$ in (3)	mg O2/l	0.67 (±0.02)	0.85 (±0.03)	-

[1] Estimates calculated with an average suspended solids concentration of 4.5 g SS/l

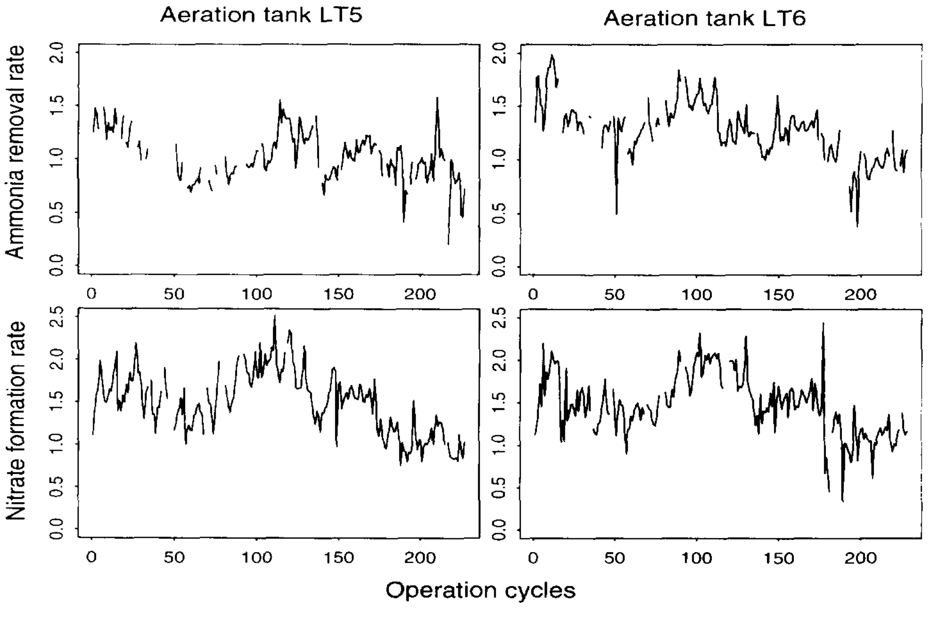

Fig. 2. Estimated maximum nitrification rates at the Aalborg West WWTP covering a period of 27 days.

Figure 2 shows the estimated maximum nitrification rates obtained from the aerobic phases of the BIO-DENIPHO process at the Aalborg West WWTP. It should be noted that the on-line ammonia sensor in aeration tank LT5 measured too low concentrations in the beginning of the data sampling period and was calibrated at operation cycle No. 112. The maximum nitrification rate appears to be highly fluctuant, even though some of the dynamic variation is due to uncertainty of the estimates. However, the four curves all contain the same underlying variation which could be caused by changes in e.g. temperature, wastewater composition or pH-value. Daily measurements of temperature cannot explain the underlying variations in Fig. 2. Detailed analyses of the measurements with respect to measuring technique and the method of interpretation have not revealed phenomena that can cause such fluctuations. The fact that they are observed on the basis of two independent estimation approaches: ammonia removal and nitrate production provides strong indications that the nitrification process is frequently affected by inhibiting materials in the wastewater. Similar results were obtained from estimating the nitrification process on data from the Lundtofte pilot scale plant. These are shown in Fig. 3. The gaps on the curves in Fig. 2 and Fig. 3 are due to aerobic periods with too few observations for a practical identification of $r_{nit,max}$. The mean and the standard deviation of the maximum nitrification rate are also given in Table 1. It should be noted that the estimate of $r_{nit,max}$ in (2) based on data from the Lundtofte pilot-scale plant does neither include the ammonification process nor biomass assimilation of ammonia, since a rate expression for the description of these two processes is practically identifiable and has been estimated separately as a joint process. Therefore, the two estimates of the mean of $r_{nit,max}$ in the last column of Table 1 and the curves in Fig. 3 do not incorporate the combined process of ammonification and biomass assimilation of ammonia, and may be implicitly compared. It is seen, that the estimates based on the ammonia and nitrate concentration have values of the same magnitude. It is also seen, that the maximum nitrification rates are generally higher at the Lundtofte pilot-scale plant. This is primarily due to higher temperatures in the activated sludge at the pilot plant, and secondarily due to chemical precipitation of phosphorus at the Aalborg West WWTP (lower bacterial activity per SS, which also includes the chemical sludge). Hence, estimating (2) and (3) provides an on-line information on the nitrification process, which is important for surveillance and optimizing the operation of nutrient removing plants.

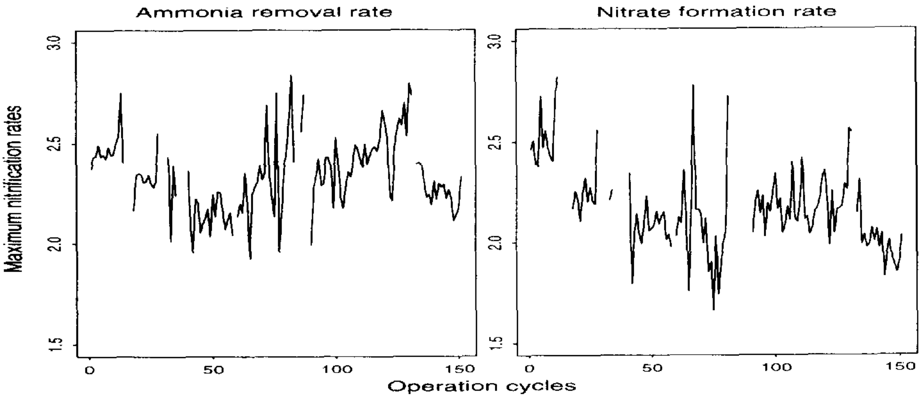

Fig. 3. Estimated maximum nitrification rates at the Lundtofte pilot scale plant.

ESTIMATING THE DENITRIFICATION PROCESS

Denitrification is a microbiological process performed by heterotrophic bacteria during anoxic conditions transforming nitrate into nitrogen gas by use of nitrate instead of oxygen as the oxidation agent. For the modelling of this process Monod-kinetic terms are frequently used to describe the dependency of oxygen and readily biodegradable substrate.

$$\frac{dS_{NO3}}{dt} = -\frac{\mu_{max,H}}{Y_{obs,NO3}} \cdot \frac{S_{NO3}}{S_{NO3}+K_{NO3}} \cdot \frac{S_S}{S_S+K_S} \cdot X_{B,H} \qquad (4)$$

where

S_{NO3} = the concentration of nitrate
S_S = the concentration of readily biodegradable substrate
$X_{B,H}$ = the concentration of heterotrophic biomass
$\mu_{max,H}$ = the maximum specific growth rate of heterotrophic bacteria
$Y_{obs,NO3}$ = the observed biomass yield coefficient of nitrate
K_{NO3}, K_S = the appropriate half-saturation constants.

As mentioned previously, a simultaneous identification of $\mu_{max,H}$ and $Y_{obs,NO3}$ is not feasible unless (4) is coupled with a differential equation for the growth of $X_{B,H}$. Furthermore, no method for measuring the concentration of heterotrophic biomass or readily biodegradable substrate exists. For modelling of the second Monod-kinetic term in (4), the load of ammonia to the plant, r_{load}, has proved useful by statistical means as a correlated measure of the readily biodegradable substrate concentration and the corresponding half- saturation constant is called K_{load}. Hence, it is better to use r_{load} to decribe the dependency of readily biodegradable substrate on the denitrification process than assuming the denitrification rate only to be limited by the nitrate concentration. Thus, the following rate expression can be identified for describing the denitrification process during anoxic conditions.

$$\frac{\Delta S_{NO3,t}}{\Delta t} = -r_{denit,max} \cdot \frac{S_{NO3,t-1}}{S_{NO3,t-1}+K_{NO3}} \cdot \frac{r_{load}}{r_{load}+K_{load}} \cdot X_{SS,t-1} \tag{5}$$

where $S_{NO3,t}$ is the measured nitrate concentration at time t and $r_{denit,max}$ is the maximum denitrification rate. The use of lagged values on the right-hand side of (5) will yield biased estimates of K_{NO3} and K_{load} as mentioned previously, but the rather crude approximation for the dependency of readily biodegradable substrate in (5) means that the estimates of the half-saturation constants found at the Aalborg West WWTP and the Lundtofte pilot-scale plant should not be unconditionally compared to suggested values of the deterministic model (4). At the Lundtofte pilot-scale plant the suspended solids concentration is assumed to be maintained approximately constant, that is on average 4.5 g SS/l.

TABLE 2. Parameter Estimates of the Denitrification Process

Parameter	Unit	Aalborg West WWTP - LT5	Aalborg West WWTP - LT6	Lundtofte pilot scale plant
$\hat{r}_{denit,max}$ in (5)	mg NO3-N/h/g SS	1.16 (±0.31)	1.18 (±0.34)	0.97 (±0.17)[1]
$\hat{K}_{NO3}$ in (5)	mg NO3-N/l	0.99 (±0.06)	1.08 (±0.13)	0.40 (±0.01)
$\hat{K}_{load}$ in (5)	mg NH4-N/l/h	1.55 (±0.05)	1.73 (±0.16)	0.43 (±0.05)

[1] Estimates calculated with an average suspended solids concentration of 4.5 g SS/l.

Estimates of the parameter in (5) based on data from the Aalborg West WWTP and the Lundtofte pilot-scale plant are found in Table 2. The results obtained from the two aeration tanks at the Aalborg West WWTP are very similar, but they deviate significantly from the results of the last column in the table. This is primarily due to the rather crude approximation of using r_{load} as a correlated measure of readily biodegradable substrate and a very different composition of the wastewater at the two plants. In fact, only the estimate of K_{NO3} at the Lundtofte pilot-scale plant is found to be close to the suggested values of the literature, see e.g. Randall *et al.* (1992).

Figure 4 shows the estimated maximum denitrification rates from every aerobic phase of the BIO-DENIPHO process at the Aalborg West WWTP. Similarly to the estimated maximum nitrification rates, the maximum denitrification rate also appears to be highly fluctuant, even though some of the dynamic is due to uncertainty of the estimates. However, the two curves contain the same underlying variation which could be caused by changes in temperature, wastewater composition or pH-value. In particular, changes in the COD/NH_4-N-ratio of the wastewater will result in changes in the estimates of the maximum denitrification rate. The keypoint is that (5) provides on-line information on the denitrification process, which is important for both surveillance and optimising the operation of nutrient removing plants.

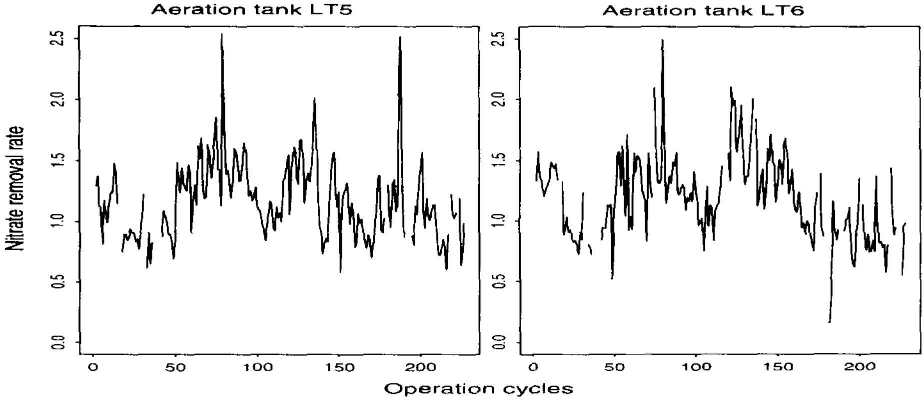

Fig. 4. Estimated maximum nitrification rates at the Aalborg West WWTP covering a period of 27 days.

DISCUSSION

A model-based method is proposed by which important information on the biological processes used for nitrogen removal can be deduced from the dynamics of on-line measured ammonia and nitrate concentrations. Such information is very valuable for plant surveillance and control, since the models provide on-line estimates of the nitrification and denitrification processes. In order to identify these processes it is crucial that the wastewater treatment plant is persistently excited – the data contains abundant dynamics. Such dynamics are obtained from plants of the SBR-type (e.g. the BIO-DENIPHO process). The parameter estimates involved in modelling the nitrification process and to some extent the denitrification processes of the two different plants are reproducible in the sense that similar results are obtained from different sets of data and the estimated half-saturation constants are close to the suggested values of the literature. In these studies large variations are found in the estimates of the maximum nitrification rate which indicates that the nitrification process is likely to be affected by inhibiting materials in the incoming wastewater. The models described in the present paper are only part of a major model system giving a description of all the significant processes in the aeration tanks. This system of models describing the hydraulic and biological processes is called a *grey box models* and is fully documented in Carstensen (1994).

Inhibition of the nitrifying and denitrifying bacteria may be detected if low parameter estimates of the maximum nitrification rate and maximum denitrification rate are encountered. Furthermore, the existing control of the oxygen set-point concentration and phase lengths of the BIO-DENIPHO process may be optimized by use of the information obtained from the parameter estimates. In Carstensen (1994) simulations of different control strategies based on the identified grey box models have shown large improvements measured by means of a cost function, which includes nutrient discharge and energy consumption costs. The simulations showed that improved control of the oxygen set-point concentration and phase lengths in an alternating plant would result in lower total nitrogen discharge as well as a lower cost of aeration. It is expected that the size of future plants can be reduced by 10–30% and that the complexity in reactor design for wastewater treatment plants with biological nutrient removal in the future will be replaced by more advanced control methods.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Ing. S.H. Isaacs, Department of Chemical Engineering, Technical University of Denmark, for providing the data from the Lundtofte pilot-scale plant and Henrik Bechmann M.Sc., I. Krüger Systems, for providing the data from the Aalborg West WWTP. Dr Marinus K. Nielsen, I. Krüger Systems is also acknowledged for giving valuable and critical comments of the results in this paper. This work was sponsored by The Danish Academy for Technical Sciences (in Danish: ATV) and I. Krüger Systems AS.

REFERENCES

Carstensen, J., Madsen, H., Poulsen, N.K., & Nielsen, M.K. (1993). Grey box modelling in two time domains of a wastewater pilot scale plant. *Environmetrics, 4(2)*, 187-208.

Carstensen, J. (1994). *Identification of wastewater processes.* Ph.D. thesis, Institute of Mathematical Modelling, Technical University of Denmark.

Carstensen, J., Madsen, H., Poulsen, N.K., & Nielsen, M.K. (1994). Identification of wastewater treatment processes for nutrient removal on a full-scale WWTP by statistical methods. *Water Research, 28*, 2055-2066.

Godfrey, K.R., & DiStefano, J.J. (1985). Identifiability of model parameters. *Proc. IFAC Identification and System Parameter Estimation*, York, UK.

Harvey, A. (1989). *Forecasting, structural time series models and the Kalman filter.* Cambridge University Press. 554 pp.

Henze, M., Grady, C., Gujer, W., Marais, G., & Matsuo, T. (1987). Activated sludge model No. 1. Technical Report, IAWPRC Science and Technology Reports No. 1, London.

Holmberg, A., & Ranta, J. (1982). Procedures for parameter and state estimation of a microbial growth process model. *Automatica, 18(2)*, 181-193.

Jepsen, U., & Olsson, G. (1993). Reduced order models for on-line parameter identification of the activated sludge process. *Wat. Sci. Tech.* 28(11-12), 173-183.

Larrea, L., García-Heras, J.L., Ayesa, E., & Florez, J. (1992). Designing experiments to determine the coefficients of activated sludge models by identification algorithms. *Wat. Sci. Tech.* 25(6), 149-165.

Nielsen, M.K., Madsen, H., & Carstensen, J. (1994). Identification and control of nutrient removing processes in wastewater treatment plants. *Submitted for the 3rd IEEE Conference on Control Applications*, University of Stratchclyde, Glasgow, Scotland.

Randall, C., Barnard, J., & Stensel, H. (Eds.) (1992). *Design and Retrofit of Wastewater Treatment Plants for Biological Nutrient Removal.* Technomic. 415 pp.