Development of a risk assessment based technique for design/retrofitting of WWTPs

D. Rousseau*, F. Verdonck*, O. Moerman*, R. Carrette, C. Thoeye**, J. Meirlaen* and P.A. Vanrolleghem***

*BIOMATH Department, Ghent University, Coupure Links 653, B-9000 Gent, Belgium

**Aquafin nv, Technology Department, Dijkstraat 8, B-2630 Aartselaar, Belgium

Abstract Up to now, within the design/retrofit of wastewater treatment plants (WWTPs), deterministic models were used to evaluate different scenarios on their merits in terms of effluent compliance. This paper describes an approach in which a Monte Carlo engine is coupled to a deterministic treatment plant model, followed by risk interpretation in the form of concentration – duration – frequency (cdf) curves of norm exceedance. The combination of probabilistic modelling techniques with the currently available deterministic models allows to determine the probability of exceeding the effluent limits of a WWTP. This percentage of exceedance is accompanied by confidence intervals resulting from the inherent uncertainty of influent characteristics and model parameters. The approach is illustrated for a hypothetical case study, consisting of a denitrifying plant model inspired by the benchmark model described by Spanjers *et al.*

Keywords Activated sludge performance optimisation; benchmarking; concentration-duration-frequency curves; dynamic simulation; Monte Carlo; risk assessment

Introduction

River water quality in Flanders (Belgium) has been dramatically bad during the past twenty years, because of the high degree of urbanisation, the industrial and agricultural pollution and insufficient basic treatment infrastructure. Almost no watercourses even met the lowest criteria, which were set out in river master plans. In 1990, the private company Aquafin was founded and assigned with the task of the design, construction, operation and financing of the necessary infrastructure for sewage treatment. From its inception, Aquafin has designed and built over 51 sewage treatment plants and in excess of 800 collector systems and pumping stations at a total construction cost of 750 million EURO. Aquafin currently operates a total of 167 sewage treatment plants, more than 500 pumping stations and 4000 km of collection systems.

One of the challenges Aquafin is facing now is to upgrade the patrimonium of old municipal wastewater treatment plants (WWTPs). These plants need to be retrofitted towards strict phosphorus and nitrogen removal consents. In 1991, when the European Directive for urban wastewater treatment 271/91 for sensitive areas to eutrophication was introduced, only one-quarter of the wastewater was treated in a WWTP. Moreover, the existing WWTPs did not comply to the present norms.

Within the currently followed design/retrofit-procedure, deterministic dynamic models are used to evaluate different renovation scenarios on their merits. One of the remaining issues when dealing with these deterministic models is the degree of uncertainty linked to their predictions. In other words, to what extent can the predictions of the model be taken for reality? The combination of probabilistic modelling techniques with the currently available deterministic models (steady state or dynamic models) could provide the answer needed. By building a probabilistic shell around the deterministic models one could quantify the uncertainty contained within the model predictions.

The concrete goal of this project is to determine the probability of exceeding the legal effluent standards of a WWTP. This percentage of exceedance should be accompanied by

Water Science and Technology **Vol 43 No 7 pp 287–294**

confidence intervals indicating the inherent uncertainty of influent characteristics and model parameters. Characterisation of uncertainty allows decision-makers to choose whether to actively take measures or to conduct additional research. The whole approach is then applied to a case study inspired by the benchmark work of Spanjers *et al.* (1998).

This paper describes the proposed approach of coupling a Monte Carlo engine with a deterministic WWTP model followed by risk interpretation in the form of concentration – duration – frequency (cdf) curves. Cdf curves are an excellent tool for risk assessment since they not only provide information about the number and the intensity of exceedances, but they also give detailed information concerning the duration of the exceedances. The whole approach is then applied to a nitrogen removal case study inspired by the benchmark work of Spanjers *et al.* (1998).

Time series analysis

The simplest way to evaluate the effluent quality is to compare the simulated effluent time series with the legal standards and look for exceedances. However, some preliminary calculations including the number of exceedances and the minimum, maximum and average time of exceedance already provide valuable extra information about the systems performance. To illustrate the concept, effluent TotN time series were generated using a simplified denitrification plant model based on the benchmark described by Spanjers *et al.* (1998) and with a 3-month influent file based on Bauwens *et al.* (1996) and influent quality data collected by Aquafin operators. Figure 1 clearly shows that the number of exceedances alone is a bad indicator. If the effluent concentrations vary in a narrow range around the effluent limit, the number of exceedances quickly increases. The additional information derived from the average duration, however, shows that the observed exceedances are not long-lasting events and, hence, that the environmental impact could be not as severe as expected.

Consequently, a concentration – duration – frequency (cdf) curve based on time series analysis is a more powerful tool (Fronteau *et al.*, 1995). These cdf curves are generated by dividing the norm exceedance times into a number of classes and by determining the number of exceedances for each class. Besides being a powerful tool, cdf curves also form a flexible tool since the number of classes, the width of one class and the limit all can be chosen.

Return period analysis, another tool for risk assessment, of combined sewer overflow (CSO) effects is often used for design decisions in urban storm water management. The return period of the exceedance of a variable is found as the reciprocal of its probability of exceedance and the mean number of events per year (Grum and Aalderink, 1999). For water quality objectives, this approach seems less useful since the duration of an exceedance is not taken into account.

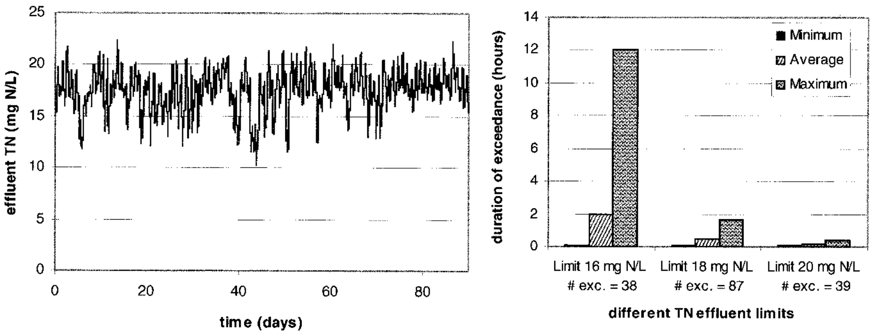

Figure 1 Effluent TotN time series from a simulated denitrifying plant (left); minimum, average and maximum duration of exceedance for different effluent limits and number of exceedances (right)

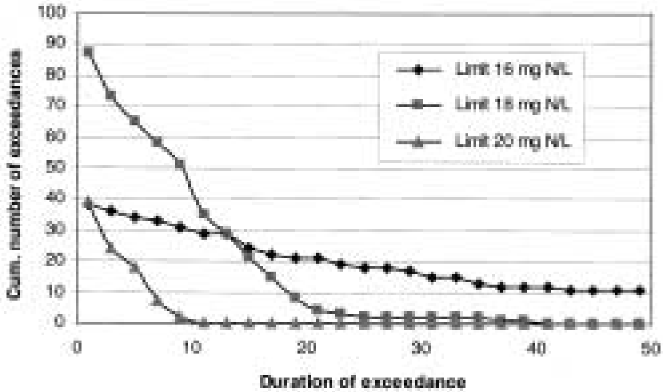

Figure 2 Cdf curves for different TotN effluent limits, based on the TotN effluent time series of Figure 1 with 2 hour averaged TotN-concentrations. Effluent limit exceedances were divided in 25 classes, each with a width of 2 hours and the last class containing all exceedances longer than 48 hours

The cdf tool can be used in three ways. First of all, several different effluent limits can be selected, resulting in as many cdf curves. These curves illustrate some characteristics of the effluent quality as can be seen in Figure 2. The curves in Figure 2 are cumulated and should be interpreted as follows: e.g. for the 18 mg N/L limit, one can see that there are 51 exceedances which last 10 hours **or longer**. Twenty-five classes were used to generate these cdf curves, each with a width of 2 hours and the last class counting all exceedances longer than 2 days. From Figure 2, it can for instance be seen that the effluent limit of 20 mg N/L is never exceeded more than 10 consecutive hours whereas the effluent limit of 16 mg N/L shows exceedances lasting longer than 48 hours. If there were an interest in these long-lasting phenomena, it could be appropriate to repeat the simulations with more classes or with bigger classes to determine the entire cdf-curve.

A second application of the cdf tool is a sensitivity analysis of particular model parameters. The more sensitive a parameter, the more the effluent time series and the resulting cdf curves will change if the value of the parameter is changed.

The third part of the time series analysis is that, for a fixed effluent limit, input variables as well as several model parameters can be varied, e.g. via Monte Carlo simulations. The combination of the Monte Carlo algorithm and the time series analysis results in a series of cdf curves. This set of cdf curves can be used to calculate a probability distribution of the cdf results. It is thus possible to determine the chance that effluent standards will be exceeded together with the uncertainty of this prediction. This application will be demonstrated further on by means of a case-study.

Overall approach

The probabilistic simulation takes into account both parameter and input uncertainty, in this way dealing with the difficulties to estimate model parameters and taking into account the inherent uncertainty in specific processes.

For each model input that is considered to be a random variable, a probability distribution is specified. Random samples are taken for each of the input distributions. One sample from each input distribution is selected, and the set of samples ("shot") is entered into the deterministic model. The model is then solved as it would be for any deterministic analysis. The model results are stored and the process is repeated until the specified number of model iterations is completed. Using Monte Carlo techniques, it is therefore possible to represent uncertainty in the output of a model by generating sample values for the model inputs, and

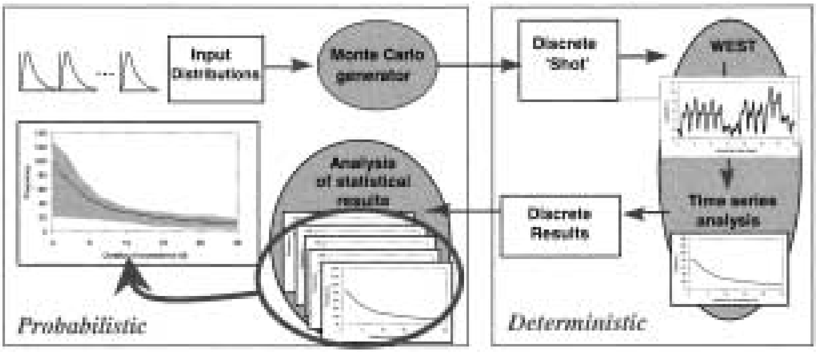

Figure 3 Monte Carlo methodology

running the model repetitively. Instead of obtaining a discrete number for model outputs as in a deterministic simulation, a set of output samples is obtained (Cullen and Frey, 1999).

In this case, the resulting model outputs are concentration-duration-frequency curves. After a large number of "shots", one obtains a large number of cdf curves, which can be used to construct an "uncertainty band" on the cdf curves (see Figure 3).

A distinction ought to be made between uncertainty and inherent variability. Variability represents heterogeneity or diversity, which is not reducible through further measurement or study. Uncertainty represents ignorance about a poorly characterised phenomenon which is sometimes reducible through further measurement or study. In the current status of the project, the variability is assumed to be completely captured via the dynamic simulations and uncertainty is captured via the Monte Carlo simulation. Therefore, there is no need for a second order Monte Carlo that would simulate variability and uncertainty in two loops, as illustrated in Grum and Aaldenberg (1999).

Case study with a denitrifying plant model

Simplified benchmark model

The above described method was implemented in the WEST modelling and simulation software (Hemmis NV, Kortrijk, Belgium) and tested on a case-study with a denitrifying WWTP model inspired by the benchmark model described by Spanjers *et al.* (1998). Two activated sludge units (ASUs) were placed in series, the first one being anoxic with a volume of 2000 m^3, the second one aerobic with a volume of 4000 m^3. The internal recirculation between the two ASUs was set at 66% of the flow leaving the second ASU. The biological treatment was simulated by means of the Activated Sludge Model No. 1 (ASM1) of Henze *et al.* (1987). Secondary sedimentation was simulated by a pointsettler assuming a non-settleable fraction of 0.005 and an underflow of 33% of the flow entering the pointsettler. The waste flow was set at about 385 m^3/day to obtain an SRT of 15 days. The overflow of the pointsettler was then submitted to time series analysis.

Influent variability and uncertainty

Measurements of the influent quality by Aquafin operators on several WWTPs provided an extensive dataset on which certain relationships between flow and wastewater components could be determined. The regression equations for medium-strength wastewater which were used in this case-study are indicated in Table 1. The relationships for COD versus flow are shown in Figure 4.

In order to be able to work with the ASM1 model, COD and KjN had to be fractionated into the specific components used by ASM1. These fractions are indicated in Table 2 and have been used to calculate the relationships between the ASM1 components in the influent and

Table 1 Relationships between certain wastewater components and flow. COD is expressed as g COD/m^3, KjN and NO_3–N are expressed as g N/m^3 and flow (Q) is expressed as m^3/day

	Minimum	Average	Maximum
COD	$y = 22066\ Q^{-0.6838}$	$y = 135478\ Q^{-0.8199}$	$y = 3.33 \times 10^6\ Q^{-1.0994}$
Kjeldahl N	$y = 103791\ Q^{-1.0836}$	$y = 34100\ Q^{-0.9071}$	$y = 54032\ Q^{-0.9045}$
Nitrate N	$y = 3 \times 10^{-5}\ Q^{-0.3266}$	$y = 0.00013\ Q^{+3.3301}$	$y = 0.0001\ Q^{+0.1388}$

Table 2 Influent fractionation

	% of COD	% of KjN
S_I	8	–
S_S	21	–
X_I	14	–
X_S	57	–
S_{NH}	–	64
S_{ND}	–	16
X_{ND}	–	20

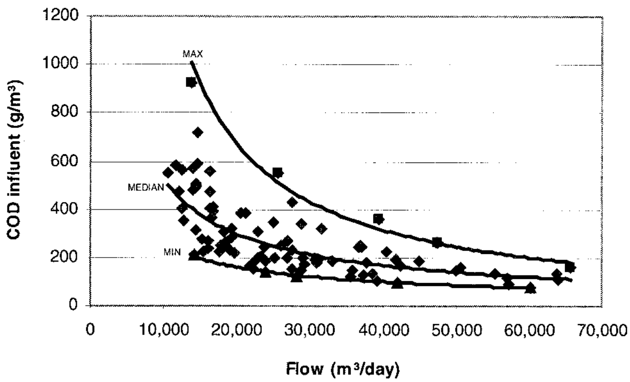

Figure 4 Relation between influent COD–concentrations and flow for medium strength wastewater + indication of minimum and maximum concentrations as given in Table 1

flow. Common practice dictates that X_{BH}, X_{BA}, X_P (heterotrophic and autotrophic biomass and particulate products of decay) and SO (dissolved oxygen) in the influent are set to zero.

For every component mentioned in Table 2, a triangular distribution was imposed between the minimum and maximum values calculated according to Table 1 and with the median on the average regression equations. Before every Monte Carlo shot, an influent file was built by the Monte Carlo engine based on a given flow series with DWF around 18000 m^3/day and by sampling from the specified distributions. The 6-month flow series was generated based on the work of Bauwens *et al.* (1996), after rescaling.

Parameter variability and uncertainty

Two types of variability were considered. First of all, the temperature dependency of certain parameters was implemented. This is the case for the heterotrophic and autotrophic growth rates μ_H and μ_A, the heterotrophic and autotrophic decay constants b_H and b_A, the hydrolysis rate k_h and the half saturation coefficient for hydrolysis of slowly biodegradable substrate K_X. Temperature dependencies were specified according to an Arrhenius equation using the values provided in the ASM1 of Henze *et al.* (1987). The temperature series itself was the same as used in Bauwens *et al.* (1996).

Table 3 Values and uncertainty ranges for the model parameters (for an explanation of these model parameters, see Henze et al., 1987)

Parameter	Unit	Distribution	Mean/median	Uncertainty
Y_H	gCOD/gCOD	triangular	0.67	5%
i_{XB}	gN/gCOD	triangular	0.08	5%
Y_A	gCOD/gN	triangular	0.24	5%
f_P	–	triangular	0.08	5%
i_{XP}	gN/gCOD	triangular	0.06	5%
K_S	gCOD/m^3	triangular	10	50%
K_{OH}	gO_2/m^3	triangular	0.2	50%
η_h	–	triangular	0.8	20%
η_g	–	triangular	0.8	20%
k_a	m^3/(gCOD*d)	triangular	0.05	50%
K_{NO}	gNO_3–N/m^3	triangular	0.5	50%
K_{NH}	gNH_4–N/m^3	triangular	1	50%
K_{OA}	gO_2/m^3	triangular	0.4	50%
f_{TSS}	g TSS/g COD	triangular	0.75	5%
μ_H	d^{-1}	truncated normal	3–6 (temp.dep)	20%
μ_A	d^{-1}	truncated normal	0.3–0.8 (temp.dep)	20%
b_H	d^{-1}	truncated normal	0.2–0.62 (temp.dep)	50%
b_A	d^{-1}	truncated normal	0.05–0.15 (temp.dep)	50%
k_h	gsbCOD/(gcellCOD ∗ d)	truncated normal	1–3 (temp.dep)	50%
K_X	gsbCOD/gcellCOD	truncated normal	0.01–0.03 (temp.dep)	50%

Secondly, all parameters were described by a distribution, as given in Table 3. The uncertainty ranges are based on Reichert and Vanrolleghem (2001). For the temperature-dependent parameters, a truncated normal distribution had to be used. The truncation was necessary to avoid negative values and was set at 0.00001.

Simulations

In order to cover the entire temperature range, simulations were done over a period of 180 days, starting in the winter period and ending in the summer period. For this case study, 300 Monte Carlo shots were simulated on a Pentium III – 650 MHz based PC. The effluent series were analysed for nitrate-N, ammonium-N and total-N with the effluent standards set to 10 mg N/L, 4 mg N/L and 18 mg N/L respectively. Concentrations were first time-averaged over a 2 hour period as imposed by environmental legislations in several countries.

Results

Norm compliance

The 300 cdf histograms resulting from the 300 Monte Carlo shots allowed to calculate the median and 5–95 percentiles for every class. The first class represents the total number of exceedances. Then, the percentage number of exceedances can be calculated. This can be transformed into% time exceedance. The results are shown in Figure 5. For ammonium-N for instance, the conclusion is that there is 95% certainty that the effluent limit will be exceeded less than 19% of the time. The nitrate-N limit will be exceeded 48% of the time (95% certainty) and the total-N limit will be exceeded 50% of the time (95% certainty). The European legal standards state that an installation may not exceed the effluent standards more than 5% of the time. We are only 43%, 25% and 5% certain that the effluent concentrations of respectively NH_4-N, NO_3-N and TotN comply to this standard.

Shape of uncertainty distributions

It is also possible to check the shape of the uncertainty distributions in the cdf-curves. Figure 4 shows that the nitrate-N limit exceedances show a lognormal distribution whilst the total-N limit exceedances show a totally different distribution.

Convergence of the Monte Carlo simulations

A preliminary study investigated the number of shots in Monte Carlo simulations that would be needed to obtain a sufficiently accurate cdf-distribution. Figure 6 shows that the

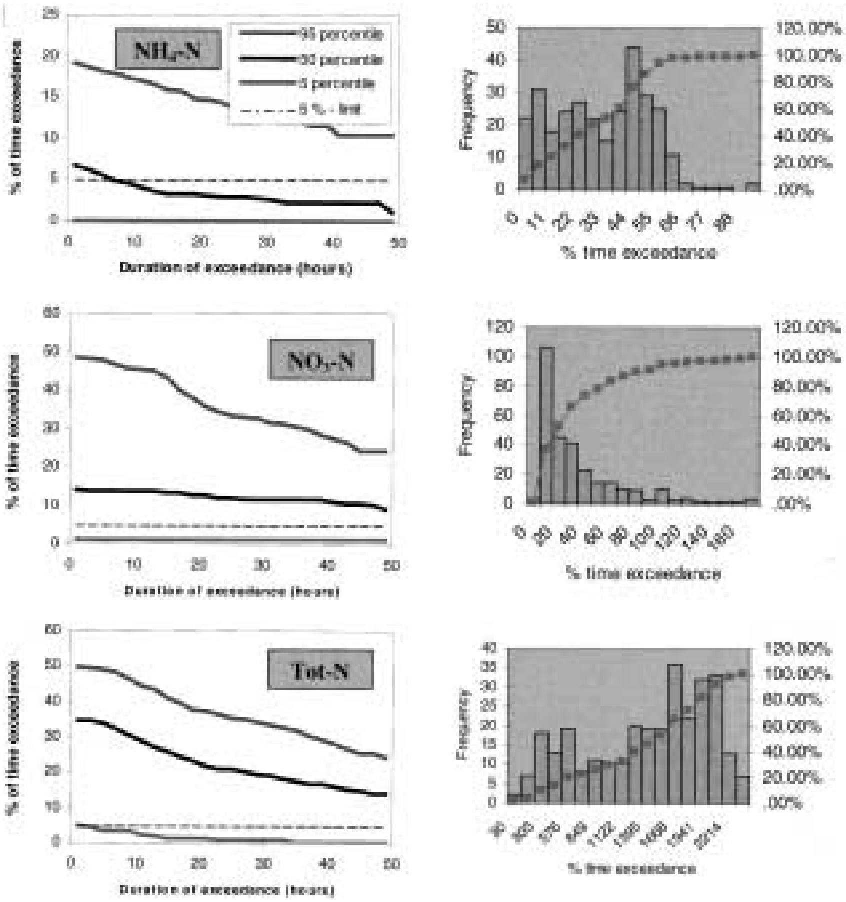

Figure 5 Cdf-curves and histograms of the first class based on 2-hour averaged effluent concentrations of ammonium-N, nitrate-N and total-N

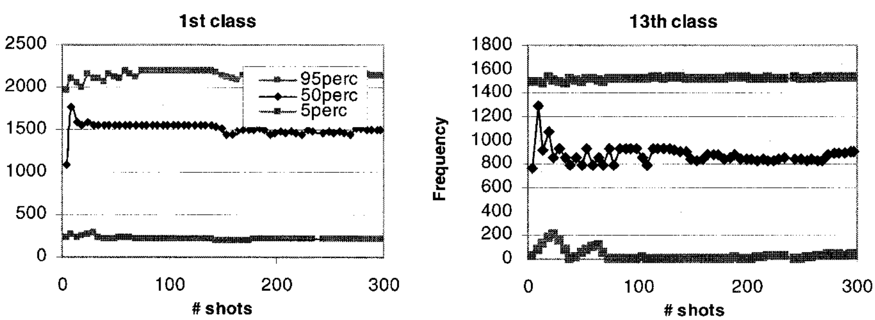

Figure 6 Convergence of the Monte Carlo shots for the 1st and 13th classes of the TotN cdf-curves

first class of the TotN cdf curve (total number of exceedances) converges within less than 50 shots. The 13th class looks less stable but this is partly due to the smaller number of exceedances. However, only about 100 shots are needed to stabilise the simulation outputs.

Conclusions

The combination of probabilistic modelling techniques with the currently available deterministic models (steady state or dynamic models) allows us to efficiently assess the uncertainty of model predictions.

A new tool was developed to determine the probability of exceeding the effluent limits of a WWTP. This percentage of exceedance is accompanied with confidence intervals indicating the inherent uncertainty of influent characteristics and model parameters. This characterisation of uncertainty allows decision-makers to choose whether to adjust the proposed design or to decide on another scenario.

Further research is still needed to investigate the influence of the choice of parameter distributions (normal, triangular, …) on the effluent distributions and to determine the appropriate number of Monte Carlo shots for other models. More simulations will also be done in which the interdependencies of certain parameters will be specified.

Acknowledgements

This research has been partly funded by a scholarship from the Flemish Institute for the Improvement of Scientific-Technological Research in the Industry (IWT). The authors would like to give special thanks to the Fund for Scientific Research (G.0102.97) for the financial support. We are also grateful to Ilse Smets, Jeroen Haegebaert and Jan Van Impe (BioTec, KULeuven) for their in-depth discussions.

References

Bauwens, W., Vanrolleghem, P.A. and Smeets, M. (1996). An evaluation of the efficiency of the combined sewer – wastewater treatment system under transient conditions. *Wat. Sci. Tech.*, **33**(2), 199-208.

Cullen, A.C. and Frey, H.C. (1999). *ProbabilisticTechniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. ISBN 0-306-45957-4. 335 p.

Fronteau, C., Bauwens, W., Smeets, M. and Vanrolleghem, P. (1995). Een evaluatie van de efficiëntie van het riool-waterzuivering-rivier systeem onder dynamische omstandigheden. *Water,* **84**, 203–210 (in dutch).

Grum, M. and Aalderink, R.H. (1999). Uncertainty in return period analysis of combined sewer overflow effects using embedded Monte Carlo simulations. *Wat. Sci. Tech.*, **39**(4), 233–240.

Henze, M., Grady, Jr C.P.L., Gujer, W., Marais, G.v.R. and Matsuo, T. (1987). *Activated Sludge Model No. 1,* IAWPRC Scientific and Technical Report No. 1, IAWPRC, London, Great-Britain.

Reichert, P. and Vanrolleghem, P. (2001). Identifiability and uncertainty analysis of the river water quality model no. 1 (RWQM1).*Wat. Sci. Tech.*, **43**(6), xx–xx (this issue).

Spanjers, H., Vanrolleghem, P.A., Nguyen, K., Vanhooren, H. and Patry, G.G. (1998). Towards a simulation benchmark for evaluating respirometry-based control strategies. *Wat. Sci. Tech.*, **37**(12), 219–226.