Tools to support a model-based methodology for benefit/cost/risk analysis of wastewater treatment systems

Lorenzo Benedetti[a], **Davide Bixio**[b], **Filip Claeys**[a] **and Peter A. Vanrolleghem**[a,c]

[a] *BIOMATH, Ghent University, Coupure Links 653, 9000 Ghent, Belgium, lbene@biomath.ugent.be*

[b] *Aquafin nv, Technology Department, Dijkstraat 8, 2630 Aartselaar, Belgium*

[c] *modelEAU, Dept. Génie Civil, Pavillon Pouliot, Université Laval, Québec G1K 7P4, Canada*

Abstract: This paper presents a set of tools developed to support an innovative four-step methodology to design and upgrade wastewater treatment systems. For the first step of data collection and data reconstruction, two different tools have been developed, one for situations where data are available (using data reconstruction methods) and another for situations where no data are (yet) available (based on a simple draining catchment model driven by actual local rain series). The second step, i.e. model building, implied the development of a new simulation platform and of grid software to deal with the considerable simulation load generated by the third step, i.e. uncertainty analysis, which involves Monte Carlo simulations of one year time series with important dynamics and stiff behaviour. For the fourth and last step, evaluation of alternatives, the IWA/COST624 benchmarking evaluation approach has been implemented and expanded. The evaluator processes the results of the Monte Carlo simulations and plots the relevant information on the robustness of the process, as well as concentration-duration-frequency curves for the risk on non-compliance of emission limits. This paper illustrates the merits of each of these tools to make the innovative methodology of more practical interest for the design and upgrade of wastewater treatment infrastructure. Well-accepted models, risk assessment techniques and sufficient computational power (that can be tapped into thanks to adequate simulation software adaptations) are available. Therefore, the design practice should move from conventional procedures suited for a relatively fixed context as imposed by emission limits, to more advanced, transparent and cost-effective procedures appropriate to cope with the flexibility and complexity introduced by integrated water management approaches.

Keywords: Wastewater treatment plant design; Cost-benefit analysis; Risk; Modelling; Software tools

1. INTRODUCTION

Process choice and dimensioning of wastewater treatment plants (WWTPs) is a particularly sensitive step to cost-efficiently comply with regulatory standards. This step accounts only for a small fraction of the upfront costs, but it can lead to substantial savings.

With the new water quality based approach introduced with the European Water Framework Directive (WFD), the design of the systems is far less predetermined and the options to meet the goals become much more numerous [Vandenberghe et al., 2004]. Therefore, new design methodologies must be developed in order to be able to cope with such increased complexity in a cost-efficient way. The EU project CD4WC (www.cd4wc.org) is currently tackling these problems, focussing on the urban wastewater system.

A systematic methodology to evaluate system design/upgrade options has been introduced by Benedetti et al. [2005], showing its merits for the dimensioning of a selected process.

This methodology would be of limited practical use without the support of adequate software tools. This paper illustrates the tools developed for the steps constituting this methodology. The four steps are: (1) data collection and reconstruction, (2) model building, (3) uncertainty assessment and (4) evaluation of alternatives.

2. DATA COLLECTION AND RECONSTRUCTION

A weak point in many simulation studies is the limited availability of long time series representing realistic dynamic influent disturbance scenarios. There is a necessity to have adequate influent time series because the natural diurnal, weekly and seasonal variations and episodic events (e.g. "first flush") represent the main process disturbance.

In absence of these data, influent time series can be reconstructed, using available measurements and making assumptions on the influent properties (e.g. as in Bixio et al. [2002] and in Devisscher et al. [2006]) or, in case of the absence of data, generated by a phenomenological model of the sewer catchment [Gernaey et al., 2005].

A tool has been developed in Matlab/Simulink to account for these situations.

2. 1 Case with some data available

The approach taken in this case can be summarized as follows.

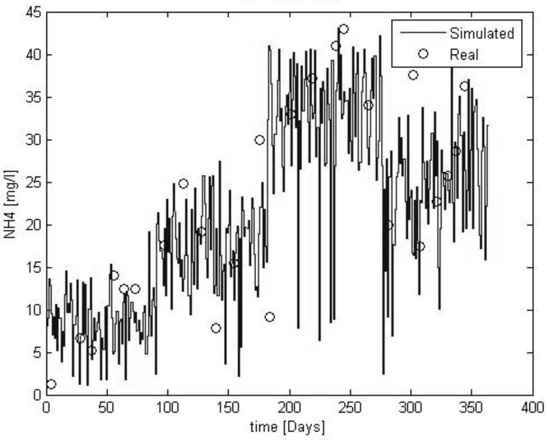

Figure 1. Example of a synthetically generated influent series and real data; from Devisscher et al. [2006].

If sufficient daily measured flow rate values are available, they are used directly, and classified into dry and wet days. Given the scarcity of water quality data, the pollutant load is computed from seasonal averages. To these synthetic daily sequences, factors are applied to account for weekends, or for first flush events (identified by checking whether a rain event appears after a number of dry days or another rain day). A daily pattern is applied to the flow rate, and concentrations are calculated from load and flow rate. An example is shown in Figure 1.

If daily flow rate values are not available, they are generated from a seasonal Poisson distribution, after which they undergo the same treatment as the flow rate values taken from data.

2. 2 Case with no data available

A simple phenomenological model was implemented, aimed at providing realistic WWTP influent dynamics without pretending at any point to provide a basis for studying urban drainage system mechanisms in detail.

Three basic modelling principles were applied: (1) model parsimony, limiting the number of model parameters as much as possible; (2) model transparency, for example by using model parameters that still have a physical meaning; (3) model flexibility, such that the proposed influent model can for example be extended easily for other applications where long influent time series are needed.

The proposed influent model produces dynamic influent flow rate and pollutant concentration trajectories. An example of the model structure relative to flow rate is shown in Figure 2.

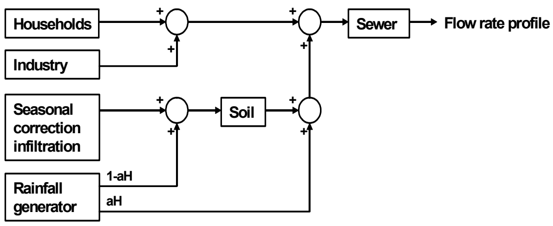

Figure 2. Schematic representation of the influent flow rate model; *aH* is the fraction of impervious surface; from Gernaey et al., [2005].

Water flows are generated adopting *per capita* discharges in households and industry (with daily, weekly and seasonal profiles), by rainfall-runoff on impervious surfaces connected to the sewer and by infiltration in the sewer from the soil compartment.

The soil is modelled as a variable volume tank with the level function of rainfall on pervious area and of an influent with seasonal variation representing the upstream aquifer. If the water level in the soil tank is higher than the invert level at which the sewer is placed, infiltration occurs at a rate function of such water level.

Rainfall can be either given as measured data (input file) or generated by a simple rainfall model.

The sewer is modelled as a series of tanks with variable volume. The size of the sewer system can be selected, assuming that a relatively small sewer system will result in sharp diurnal concentration peaks, whereas a large sewer system will result in smooth diurnal concentration variations.

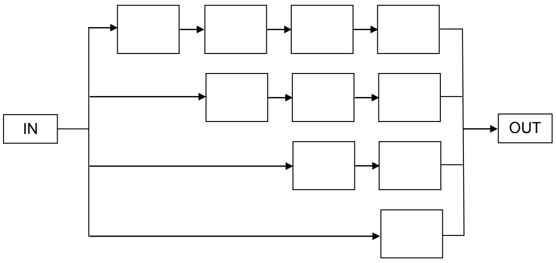

Figure 3. Schematic representation of the sewer model.

The example given in Figure 3 shows how that effect is achieved. In case of very small sewer networks, all the inflow passes only through the line with one block (each block consists of three tanks in series), while with a large sewer network the inflow will be evenly distributed to the four parallel lines. The model actually allows to choose to have up to eight parallel lines.

Pollutant loads are generated adopting *per capita* discharges in households and industry (with daily and weekly variations) for soluble COD, particulate COD, total nitrogen and total phosphorous.

Sedimentation and resuspension equations are included in the sewer model, to obtain a “first flush” effect under the appropriate conditions.

Noise is added to all generated quantities in order to reproduce the variability of the phenomena. Figure 4 shows an example of a generated influent time series.

The production of a yearly influent file with data every 15 minutes for a system with large sewer network takes less than 5 minutes on a Pentium 4 machine with a 3GHz processor.

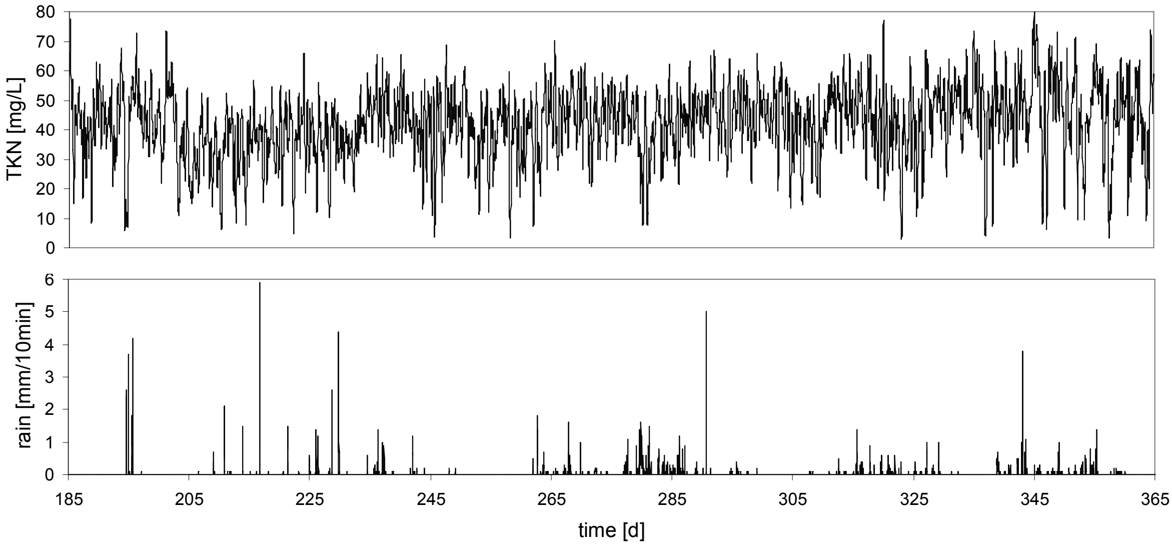

Figure 4. Example of input (rain) and output (Total Kjeldhal Nitrogen, TKN) of the influent generator.

3. MODEL BUILDING

3. 1 Models used

The most used models for biological treatment are the IWA activated sludge models (ASM) [Henze et al. 2000]. Other important parts of the plant that need to be modelled are secondary settlers, aeration systems and anaerobic digesters when present.

Such models are computationally intensive. The model of a low loaded activated sludge treatment system (LLAS) with chemical P-removal, several controllers and ASM2d in 8 tanks, has a total of 5109 parameters, 4210 algebraic variables and 276 derivatives.

3. 2 Simulation software

A new modelling and virtual experimentation kernel for water quality systems has been developed in order to be able to cope with the large computational load implied by the one-year simulations of complex WWTP layouts. This kernel was named “Tornado” and will be included in the new generation of the WEST® product family (HEMMIS, Kortrjik, Belgium), as well as in several other products and projects. Most important issues during development were versatility and efficiency. It is argued that classical approaches such as the adoption of Matlab/Simulink, custom FORTRAN codes and/or domain-specific simulators all have specific disadvantages. Therefore, a need arose for a kernel that offers a compromise between versatility and efficiency.

Tornado was developed in C++ using advanced language features, yielding a code base that offers fast execution, portability and increased readability. The software is composed of strictly separated environments for modelling and virtual experimentation. The modelling environment allows for the specification of complex ODE and DAE models in object-oriented, declarative languages such as MSL [Vanhooren et al., 2003] and Modelica. A model compiler translates these high-level models into efficient, flattened code. The experimentation environment allows for running atomic virtual experiments (such as simulations and steady-state analyses) as well as compound experiments (optimizations, scenario analyses …) on the basis of flattened models. Details on Tornado can be found in Claeys et al., [2006b].

As an example of performance: the LLAS model, running for 50 days in steady state and 415 days in dynamic conditions, with input and output data every 15 minutes, required only 31 minutes to execute on a Pentium 4 machine with a 3GHz processor. Using state of the art commercial software, it took 140 minutes.

4. UNCERTAINTY ASSESSMENT

An issue when dealing with deterministic models is the degree of uncertainty linked to their predictions. Probabilistic design, which is the combination of probabilistic modelling techniques with the currently available deterministic models, provides a solution to this issue [Bixio et al., 2002]. The adopted methodology makes use of Monte Carlo simulations to assess model parameters uncertainties and has been described in Benedetti et al. [2005]. Figure 5 shows two examples of uncertainty visualisation; on the left side of Figure 1 for all of the 10 configurations a cloud of 100 dots is plotted, each dot representing the yearly average of the effluent quality index (EQI) and of the total costs (TC) for one particular Monte Carlo simulation; on the right side, each cloud is summarised by a polygon joining the 5th and 95th percentiles for the two variables and by a marker for the 50th percentile. An important property of such graphs is that the larger the projection of a configuration's polygon on an axis is, the larger the uncertainty of that configuration for the variable associated to that axis is. For description of EQI, TC and data processing, please refer to section 5.

Another way to look at uncertainties is by means of concentration-duration curves. From Figure 6 it can be deduced that with the considered plant configuration, it is 95% sure that the threshold of 15mg/L of TN is not exceeded for more than 17% of the analysed period (one year), in case the effluent is measured by means of 2-hour composite samples.

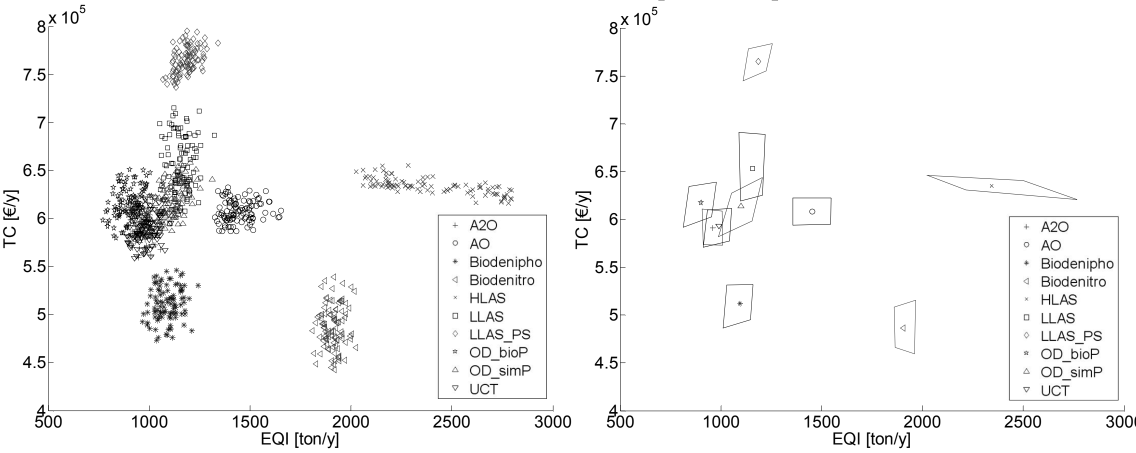

Figure 5. Two options to visualise Monte Carlo simulation results: all results as a cloud of markers (left) and polygons joining the 5th and 95th percentiles for the two variables and the 50th percentile as a marker (right); the data show yearly average effluent EQI and TC for 10 different plant configurations.

The number of necessary simulations for scenario analysis tends to be large, especially with Monte Carlo-based uncertainty assessment. To reduce this computational burden, tools that distribute simulations over idling PCs available in a local network are under development and were used in this study [Claeys et al., 2006a]. A framework for the distributed execution of simulations on a potentially heterogeneous pool of work nodes (Linux/Windows) has been implemented. It was named “Typhoon” and has been built on top of technologies such as C++, XML and SOAP. It was designed for stability, expandability, performance, platform-independence and ease of use.

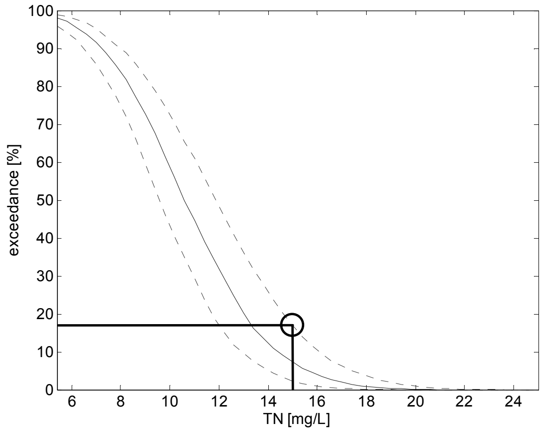

Figure 6. Concentration-exceedance curve for LLAS, 2-hour averages of TN effluent concentrations; full line: 50th percentile, dotted lines: 5th and 95th percentiles.

At BIOMATH, a cluster of 16 Linux machines with 3GHz processors is available, which allows to execute a batch of 100 Monte Carlo simulations of the LLAS model mentioned in section 3.1 in less than 4 hours. To execute the same simulations in series with commercial software on one machine with 3GHz processor, it would take almost 10 days.

5. EVALUATION OF ALTERNATIVES

The comparison of alternative scenarios can be based on performance criteria that can be grouped into two categories: environmental and economic criteria. The weight attributed to them in the decision making process depends on the specific situation of the project. For both categories uncertainties are computed, therefore risk issues can be analysed.

The proposed methodology adopts and extends the approach set out by IWA/COST [Copp et al., 2002]. It consists of the evaluation of three indicators: the effluent quality index (EQI), the time and number of effluent violations, and operating costs.

The EQI is meant to quantify the effluent pollution load to a receiving water body in a single variable. It is the weighted sum of the pollution loads due to (1) total suspended solids, (2) chemical oxygen demand, (3) biological oxygen demand after 5 days, (4) total Kjeldahl nitrogen, (5) nitrates and (6) total phosphorus over one complete year.

It is very difficult to calculate investment costs in order to compare different plant configurations and operational strategies. Detailed cost calculations should in general be preferred over the use of cost functions, which is feasible only for process options screening (Gillot et al., 1999), i.e. as is the case here.

The cost categories used in this paper are: aeration energy cost (AEC); energy cost (EC) including aeration, pumping and mixing costs; sludge cost (SC) which comprises sludge treatment and disposal; variable cost (VC) incorporating energy, sludge and chemicals cost; total cost (TC) which includes variable, personnel, maintenance and annualised capital costs.

An important criterion that has also been introduced is a measure to summarize the model output uncertainties – the relative reliability index (RRI) – which is inversely proportional to the sum of the standard deviations divided by the averages of all considered output variables. It gives a measure of how stable the performance is when the system is subjected to variations in model parameters. The RRI can also be computed for a subset of the output variables, e.g. in case of two variables the RRI is related to the perimeter of the rhombi in Figure 5.

Table 1. Plant configurations.

Short name	Description
A2O	anaerobic-anoxic-oxic – low loaded system, performs biological N and P removal
AO	anaerobic-oxic – high loaded system, performs biological P removal
BDNP	Biodenipho – low loaded system, performs biological N and P removal
BDN	Biodenitro – low loaded system, performs biological N removal
HLAS	high loaded activated sludge
LLAS	low loaded activated sludge – performs biological N and chemical P removal
LLAS_PS	LLAS with primary settler – performs biological N and chemical P removal
OD_bioP	oxidation ditch – low loaded system, performs biological N and P removal
OD_simP	oxidation ditch – low loaded system, performs biological N and chemical P removal
UCT	University Cape Town – low loaded system, performs biological N and P removal

Table 1 briefly introduces the plant configurations compared with the methodology described in this paper. They were all designed to treat 30000 PE in Oceanic climate conditions. It is an extract of a wider comparison involving all combinations of three plant sizes with four climatic conditions (see www.cd4wc.org). Further partial results of such work can be found in Benedetti et al. [2005; 2006].

Figure 7 shows the RRI for the plant layouts based on all output variables, while Figure 8 refers to the RRI calculated for EQI and TC, to be confronted with the rhombi in Figure 5; the values of RRI are normalised to have an average of 1. LLAS_PS has the best robustness of performance.

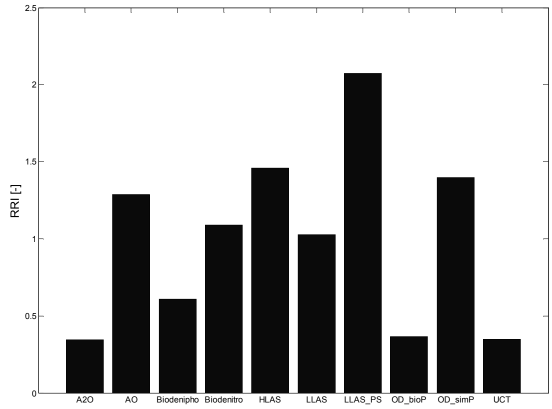

Figure 7. RRI for the 10 plant layouts based on all output variables.

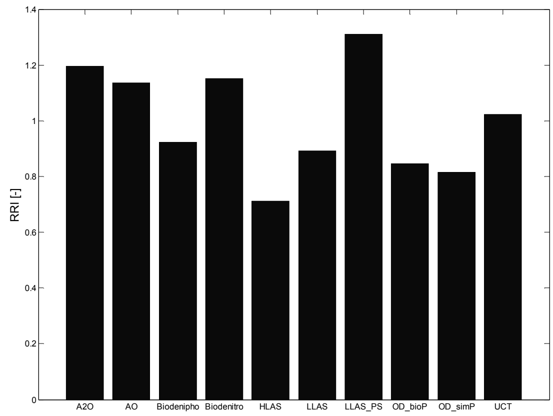

Figure 8. RRI for the 10 plant layouts based on EQI and TC.

All evaluations of alternatives concerning environmental and economic performance as well as associated uncertainties were performed by automatic post-processing of the output files generated by all simulations by using Matlab scripts created *ad hoc*. This allows to perform such evaluations without the need to re-run all simulations, but just performing again the post-processing with different values for costs and for the weights used to calculate indexes. The processing of a batch of 100 Monte Carlo output files of the mentioned LLAS model (each containing 12MB of data) required 10 minutes to execute on a Pentium 4 machine with a 3GHz processor.

6. CONCLUSIONS

With the use of the innovative tools introduced in this paper it was possible to generate the data to compare 10 plant layouts on their benefit/cost/risk in no longer than 2 days. It is therefore deemed to be of practical consideration for design in the wider decision-making context of river basin management introduced by the WFD.

7. ACKNOWLEDGEMENTS

The results presented in this publication have been elaborated in the frame of the EU projects CD4WC, contract no EVK1-CT-2002-00118 and Harmoni-CA, contract no. EVK1-CT-2002-20003. Peter A. Vanrolleghem is Canada Research Chair in Water Quality Modelling.

8. REFERENCES

Benedetti, L., Bixio, D. and Vanrolleghem, P. A., Assessment of WWTP design and upgrade options: balancing costs and risks of standards' exceedance, in: Proceedings of the 10th International Conference on Urban Drainage (ICUD), 21-26 August 2005, Copenhagen, Denmark, 2005.

Benedetti, L., Bixio, D. and Vanrolleghem, P. A., Benchmarking of WWTP design by assessing costs, effluent quality and process variability, accepted as long platform presentation for the IWA World Water Congress and Exhibition 10-14 September 2006, Beijing, China, 2006.

Bixio, D., Parmentier, G., Rousseau, D., Verdonck, F., Meirlaen, J., Vanrolleghem, P. A. and Thoeye, C., A quantitative risk analysis tool for design/simulation of wastewater treatment plants, Water Science and Technology 46(4-5), 301-307, 2002.

Claeys, F., Chtepen, M., Benedetti, L., Dhoedt, B. and Vanrolleghem, P. A., Distributed virtual experiments in water quality management. Water Science and Technology 53(1), 297-305, 2006a.

Claeys, F., De Pauw, D. J. W., Benedetti, L. and Vanrolleghem, P. A., Tornado: A versatile and efficient modelling & virtual experimentation kernel for water quality systems, accepted as platform presentation for the iEMSs 2006 conference, Burlington VE, USA, 2006b.

Copp, J., Spanjers, H. and Vanrolleghem, P. A., (eds.), Respirometry in control of the activated sludge process: Benchmarking Control Strategies, IWA Publishing, London UK, 2002.

Devisscher, M., Ciacci, G., Fé, L., Benedetti, L., Bixio, D., Thoeye, C., De Gueldre, G., Marsili-Libelli, S. and Vanrolleghem, P. A., Estimating costs and benefits of advanced control for wastewater treatment plants - the MAgIC methodology, Water Science and Technology 53(4-5), 215-223, 2006.

Gernaey, K., Rosen, C., Benedetti, L. and Jeppsson, U., Phenomenological modelling of wastewater treatment plant influent disturbance scenarios, in: Proceedings of the 10th International Conference on Urban Drainage (ICUD), 21-26 August 2005, Copenhagen, Denmark, 2005.

Gillot, S., De Clercq, B., Defour, D., Simoens, F., Gernaey, K. and Vanrolleghem, P. A. (1999). Optimization of wastewater treatment plant design and operation using simulation and cost analysis, in: Proceedings of the 72nd Annual WEF Conference and Exposition, 9-13 October 1999, New Orleans, USA, 1999.

Henze, M., Gujer, W., Mino, T. and van Loosdrecht, M. C. M., Activated sludge models ASM1, ASM2, ASM2d and ASM3, IWA Publishing, London, UK, 2000.

Vandenberghe, V., Benedetti, L. and Vanrolleghem, P.A. (2004) Synergetic use of the degrees of freedom to improve the urban waste water system, In: Proceedings of IWA World Water Congress 2004, Marrakech, Morocco, 19-24 September 2004.

Vanhooren, H., Meirlaen, J., Amerlink, Y., Claeys, F., Vangheluwe, H. and Vanrolleghem, P.A. (2003) WEST: Modelling biological wastewater treatment. J. Hydroinformatics 5(1), 27-50.