Fast, simultaneous simulation of the integrated urban wastewater system using mechanistic surrogate models

J. Meirlaen*, B. Huyghebaert*, F. Sforzi, L. Benedetti*** and P. Vanrolleghem*.**

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

**Department of Systems and Computers, Università di Firenze, Via S. Marta 3, 50139 Firenze, Italy

***Enki s.r.l., Via Pagnini 21/r, 50134 Firenze, Italy

Water Science and Technology **Vol 43 No 7 pp 301–309** © 2001 IWA Publishing and the authors

Abstract The urban wastewater system components (sewer, treatment plant, and river) are often modelled using complex mechanistic models. Mechanistic surrogate models are introduced here as simplified models that still contain some physical knowledge. Surrogate models are faster, but are less but still sufficiently accurate, and require more data to be calibrated. The possibilities of replacing actual field data by virtual data generated with a complex mechanistic model for calibration of the surrogate model are examined. As an example, a series of tanks with variable volume is shown to approximate sufficiently well the flow propagation in the river Zwalm (Belgium) as predicted by the "de Saint-Venant" equations.

The three surrogate models can be implemented in the WEST® simulator, which makes a simultaneous simulation of the system possible. In this work a connection is made between the ASM1 and the new IWA River Model No. 1 (RWQM1) by using a translator between the models in such a way that both mass and elemental balances remain closed for the overall system. This approach is illustrated with a case study on the river Lambro (Italy). The dispersion process in this river with steady flow could be modelled by using a tanks in series model, while the water quality in the river was predicted to improve substantially with an increase in hydraulic capacity of the treatment plant. The simulation results with the upgraded plant still need to be checked by field data.

Keywords Integrated modelling; RWQM1; surrogate models; tanks-in-series; urban drainage system

Introduction

The integrated urban wastewater system (sewers, treatment plants and receiving water) has a major impact on the quality of the receiving water, due to discharges of combined sewer overflows (CSO) during storms and due to the effluents of the treatment plants. To minimise this impact in a given situation, real time control can be applied on the sewer system and/or on the treatment plants. The objectives of these control strategies are e.g. to minimise the total overflow volume or to avoid sludge washout during hydraulic overloading of the plant. Different studies show the beneficial effect of these control actions (among others, Nielsen et al., 1996; Entem et al., 1998; Petruck et al., 1998). However, Rauch and Harremoës (1999) showed that minimising the total overflow volume not necessarily results in the best water quality possible. It therefore seems necessary to take into account the resulting river water quality when determining the control actions, rather than starting from the traditional emission point of view.

One way to take the resulting river water quality into account is to use model based predictive control (MBPC) of the integrated urban wastewater system. This control method needs an integrated model, with the three subsystems running simultaneously. With a simultaneous simulation, the current and predicted states of the river water can be used to determine the control actions in e.g. the sewer system, whereas in sequential simulation this is not possible, since the water quality is only calculated after the simulation of the other systems is completed. An integrated model can also be used when designing or upgrading a system, where it allows to quickly quantify the effect of different design options on the resulting water quality. Because of these advantages, a simultaneous simulation of the integrated urban wastewater system has been asked for since the last few years (Rauch *et al.*, 1998; Schütze *et al.*, 1998).

Flow routing in sewer pipes or rivers can be described by the “de Saint-Venant” equations, which are based on the conservation of mass and momentum. These partial differential equations have to be solved using an advanced numerical integration algorithm with a high computational burden, which makes the model impractical for use in long-term simulations or in optimisation problems. However, this complex mechanistic model is able to accurately predict flood wave propagation in channels. Because design and model based predictive control are both examples of optimisation problems, typically a lot of simulations are necessary in order to find a (sub)optimal solution. Hence, short simulation times are required in order to find a solution within a reasonable time. For this reason, the complex mechanistic models as described above need to be substituted by faster models. In this paper the term “surrogate” models is used for the latter, as they form a surrogate (an approximate substitute) for the “real thing”, i.e. the complex mechanistic model that is approximating reality better. As with any model, two types of simplifications are possible, the empirical (black box) and the mechanistic (white box) approach. In this paper, a concept on how to go from reality to a mechanistic surrogate model is proposed and subsequently applied to the integrated urban wastewater system.

Another problem encountered in integrated modelling are the differences in variables used in the different state of the art models of the subsystems. The reason for this incompatibility is that these models have been developed independently during the last decades. An obvious example of these differences is the use of BOD in traditional sewer and river models to describe organic pollution, while in wastewater treatment plant models COD is typically used. The problem of combining sewer and treatment plant models was already discussed by Fronteau *et al.* (1997). Maryns and Bauwens (1997) have used an Activated Sludge Model No. 1 (ASM1, Henze *et al.*, 1987) model in riverine conditions in order to avoid translation of the states. In order to address this problem more fundamentally, an IWA Task Group has developed a novel river model, which is COD based (RWQM1, Reichert *et al.*, 2001). However, there still exist differences between the variables of RWQM1 and the variables of ASM1. In this work, a consistent translation of the states of ASM1 to the states of RWQM1 has been pursued. This connector between the treatment plant model and the river model closes both mass and elemental mass balances and is applied in a case study on the river Lambro (Italy).

Mechanistic surrogate models

The concept: from reality to mechanistic surrogate models

When modelling a system, there are typically three sources of information that can be used (Spriet and Vansteenkiste, 1982). The first one is the goal of the modelling exercise, the second one is the prior knowledge about the system under study, and the third one is the set of available data. In this work, the goal is independent of the type of model (complex or surrogate) chosen and is not considered further in the discussion. For the amounts of data and prior knowledge, however, these differences do exist.

When there is a lot of knowledge available (like e.g. in physical systems), a relatively small amount of data is needed to calibrate the few unknown parameters of the model (small data box in Figure1(a)). Indeed, many parameters have a physical meaning with a known value (e.g. the gravity constant *g*). When using a simplified model, in which less information is contained in the equations, more data are needed to compensate for the lack of information from the knowledge side (large data box in Figure1(b)). Many parameters no longer have a strict physical meaning, as they are the result of “lumping”. Rather than trying to recalculate them from the lumping equations, it is preferred to use data to

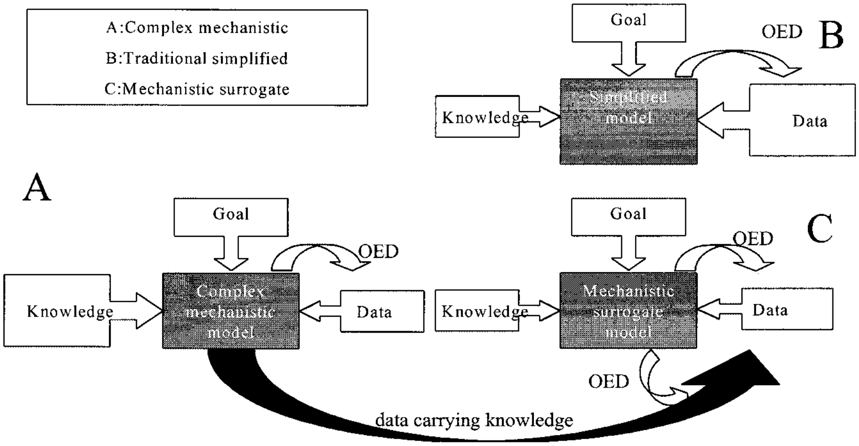

Figure 1 Sources of information during model building and surrogate modelling procedure

empirically find them. However, collecting data is an expensive and time-consuming task (Vanrolleghem *et al.*, 1999). Hence, collecting enough data to reliably calibrate a surrogate model may be an enormous or even impossible task. On the other hand, complex mechanistic models can be calibrated with a reasonable amount of effort, which is proven by numerous articles in the literature in a wide range of fields. These mechanistic models are accurate, but require too much calculation time for iterative use such as in design or MBPC, which is the future goal of this research project. The first problem to be tackled is the development of faster models.

The solution to this problem as proposed in this paper, is shown in Figure 1(c). Instead of trying to collect all the data from reality, one can also generate "data" by using the complex mechanistic model. After calibration of the complex model (with little data), simulation results yield virtual data that can be used to calibrate the surrogate model. This can be seen as "pumping" the knowledge summarised in the complex mechanistic model into the surrogate model by means of these virtual data. Once this surrogate model is calibrated, it can then be used within an optimisation procedure. To improve the reliability of using the model predictions one can, before use, check the solution suggested by optimisation with the surrogate model, with the predictions for that solution obtained with the more accurate complex model. When this last test is passed successfully, the solution can be applied to reality. The suggested procedure to go from reality to a surrogate model can be summarised as follows.

1. Determine the system under study, its boundaries and the problem to be solved.
2. Collect data on the system to calibrate the complex mechanistic model. This data collection may be assisted by optimal experimental design (OED) on reality by using the complex model to be calibrated.
3. Calibrate and validate the complex mechanistic model.
4. Generate data for the calibration of the mechanistic surrogate model. This data collection may be assisted by OED on the complex mechanistic model by using the mechanistic surrogate model. Note that, in contrast with real measurements (like in step 2), the number of degrees of freedom for experimental design optimisation is much higher and only limited by the creativity of the user.
5. Calibrate and validate the mechanistic surrogate model.

The application: from catchment to an integrated surrogate model of the urban wastewater system

Figure 2 shows the application of the previous concept on the urban wastewater system. Typical examples of complex mechanistic models of the different subsystems are given. In

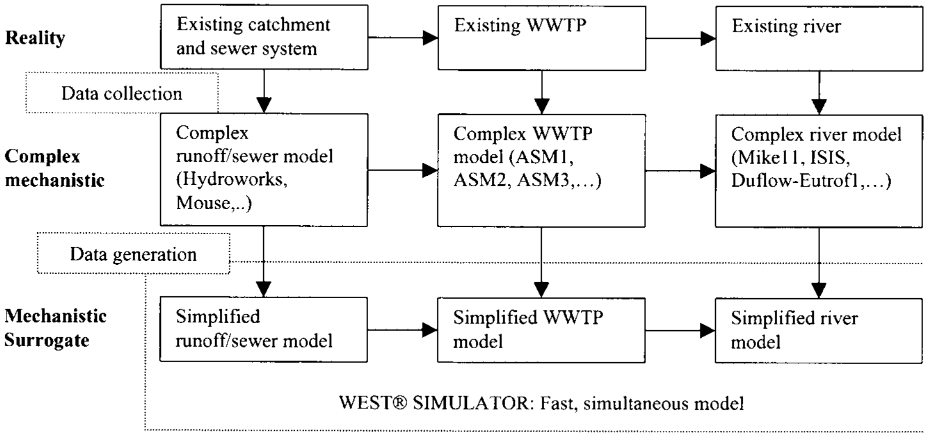

Figure 2 From reality to surrogate models using complex models

both the sewer models (HYDROWORKS, MOUSE, ...) and the river models (MIKE11, ISIS, DUFLOW-EUTROF1,...) flow propagation is described using the “de Saint-Venant” equations. These partial differential equations can be simplified in certain cases, but they still require advanced numerical integration and hence a long calculation time. Therefore alternatives for these models are suggested.

Fronteau (1999) compared the conceptual sewer model Kosim with the complex mechanistic Hydroworks model. Kosim describes run-off and flow propagation in the sewer system of each subcatchment by means of a Nash-cascade and a transportation time (Paulsen, 1986). When comparing this to the complex mechanistic models, where most of the pipes are calculated in detail, it is clear that the simulation time of the surrogate model will be much shorter. The long-term simulations of Kosim and Hydroworks gave similar results in Fronteau’s work, but Kosim was significantly faster. Since the Kosim model uses discrete time steps, it could not be implemented efficiently as such in the WEST® simulator (Hemmis NV, Kortrijk, Belgium) because this package is very effective in solving differential equations, but less so in solving discrete time equations. The corresponding differential equations have been worked out and implemented in the WEST® simulator.

Flow routing in rivers can be approximated by a series of tanks with variable volume (Beck and Young, 1975). Moreover, when these tanks are also used to describe the solutes in the river, the model is also able to describe dispersion in a reasonably accurate way (Reda, 1996). When, in addition, a biological conversion model is used to predict the conversion taking place in the river, a series of continuously stirred tanks can be easily used for modelling river water quality. The procedure of calibrating the flow propagating properties of a series of tanks on data generated by the “de Saint-Venant” equations will be illustrated in a case study performed on the river Zwalm (Belgium).

The WEST® simulator was originally used mainly for simulation of wastewater treatment plants and an extensive WWTP modelbase is available. Both the simplification of the runoff/sewer model and the tanks in series river model are now implemented in this package. Hence, the three parts of the integrated urban wastewater system (IUWS) are now available in a single software tool, making linking of the submodels straightforward. Moreover, problems with file or data transfer between different simulators are avoided and most importantly a simultaneous simulation is possible.

A case study on hydraulic surrogate models: a tanks in series model of the river Zwalm (Belgium)

The river Zwalm is a small river near Brakel in Flanders (Belgium). It has an unsteady flow with low flows in summer time (0.3 m^3/s) and high flows during rainy periods (up to 4.7 m^3/s). The studied reach of the Zwalm has a length of 5.5 km. A moveable weir forms the downstream boundary condition. In the reach studied, several subcatchments discharge into the Zwalm. A field survey of the reach yielded data about the longitudinal profile of the river, while also cross-sectional profiles were collected. These data were sufficient to define a model of the Zwalm in the simulation tool Aquasim (Reichert, 1994). Aquasim also uses the "de Saint-Venant" equations to calculate flow routing. The diffusive wave approximation of the "de Saint-Venant" equations was used to generate virtual data. The upstream inflows together with the side stream discharges were calculated using rain series from Ukkel (Belgium) which served as an input to a simple runoff model (Fronteau *et al.*, 1999). Subsequently, the water flows along the river could be calculated (For details, see Meirlaen *et al.*, in preparation b).

The physical inflows could be grouped into six lateral model inflows, which yielded a total of 7 tanks to describe the total reach under study. The first attempts to describe the flow with traditional rectangular tanks failed. This was explained by the fact that the river Zwalm is only a few meters wide, which means that its cross-section could not be approximated by a rectangle as used in the traditional tanks in series models. A parabolic relationship between the height and the cross-section turned out to be the most appropriate approximation (for details see Meirlaen *et al.*, in preparation b).

This yielded a surrogate tanks in series model with five parameters to be calibrated for each tank:

$$A_{\text{cross}} = ah^2 + bh$$

$$Q_e = \alpha(h - \beta)^{\gamma}$$

in which

A_{cross}	the cross-sectional area (m^2)
Q_e	the outflow rate of the tank (m^3/s)
h	water height of the tank
α, β, γ	outflow parameters
a, b	parameters characterising the shape of the cross-section

The parameters a and b are used to describe the relation between the height and the cross-sectional area and could be estimated using the field survey data. The parameters α, β and γ describe the relationship between the water height and the outflow rate of the tank, and were calibrated using the data generated by the complex mechanistic model. They were estimated for each tank separately by using an inflow series of 2 days during which one typical flood wave passes through the Zwalm. After calibration and validation of each tank, the tanks were connected to each other and the overall model was validated using a different inflow series of 7 days. The results of this validation are shown in Figure 3.

Figure 3(a) shows the validation results for a period of seven days. The Figure shows a very good fit of the two data series. When zooming in (Figure 3(b)) the fit appears not to be perfect. With the surrogate tanks in series model the peak flow is found 0.3 h earlier than

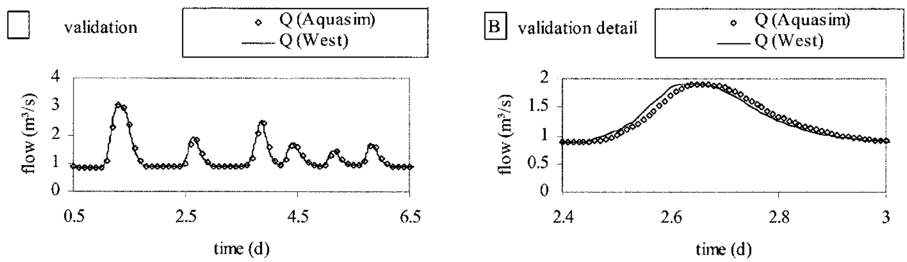

Figure 3 The validation results of the river Zwalm

with the complex model. This can be explained by the fact that in a tank the water level is always horizontal, so that when a flood wave enters the tank, this has an immediate effect on the outflow, whereas in reality this is not the case (Reda, 1996). Reda was able to reduce this error by introducing an extra parameter that artificially increases the height of the tank, thus resulting in a longer theoretical residence time and, consequently, in a slower propagation of the flood wave. In the case study of the Zwalm, this extension was not taken into account in the model because the simulation results proved sufficiently close to the data generated by the complex model. Consequently, it was concluded that the surrogate model was able to adequately describe the system. Moreover, the mechanistic surrogate model was found to be calculating three times faster than the complex mechanistic model.

Integrated modelling

Integrated modelling: connecting ASM1 and RWQM1

In order to simplify the connection between treatment plant models (typically COD based) and river water quality models (typically BOD based), a COD based river water quality model has been developed in which mass balances and elemental balances for C, N, O, P and H are closed (Reichert *et al.*, 2001; Shanahan *et al.*, 2001; Vanrolleghem *et al.*, 2001). In order to maintain this important feature for the integrated model, the proposed connector between the states of the ASM1 and the RWQM1 also has closed mass and elemental balances. To achieve this, it was necessary to assume a constant chemical composition of the ASM1 variables. Some states variables can be linked very easily, while others require some more attention. Particulate material that is biodegradable according to ASM1 (X_S) will also remain biodegradable in riverine conditions. Biomass translation is less evident. The fate of biomass coming out of the treatment plant is uncertain because totally different conditions occur in the receiving water. A fraction will probably remain active, another fraction is expected to die and has, hence, to be divided in a biodegradable and an inert part. Once the organic compounds are translated, elemental balances have to be closed. Dissolved carbon dioxide is used as a sink term for carbon, while also the other elements have inorganic representatives in the RWQM1 (H^+, HPO_4^-. NH_4^+, O_2) to act as compensation between the elemental contents in the ASM1 and the RWQM1. For more details, see Meirlaen *et al.* (in preparation b) and Benedetti and Sforzi (1999).

Integrated modelling: use of surrogate models on the river Lambro

The above connector and surrogate models have been used on a case study of the river Lambro, Italy. For a description of the catchment, see Whelan *et al.* (1999). The treatment plant in Merone has been focussed upon in this study. Before June 1998, this treatment plant had one primary clarifier, two oxidation tanks and two secondary settlers, but the hydraulic capacity was insufficient even to treat the dry weather flow, with a daily bypass of raw wastewater as a

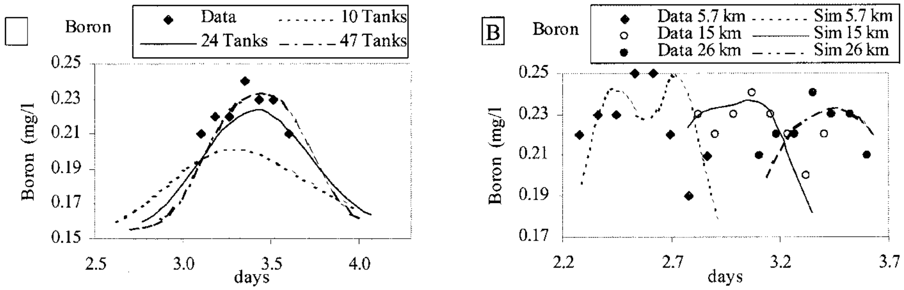

Figure 4 Effect of the number of tanks on dispersion after 26 km (a), longitudinal propagation of pollution wave (b)

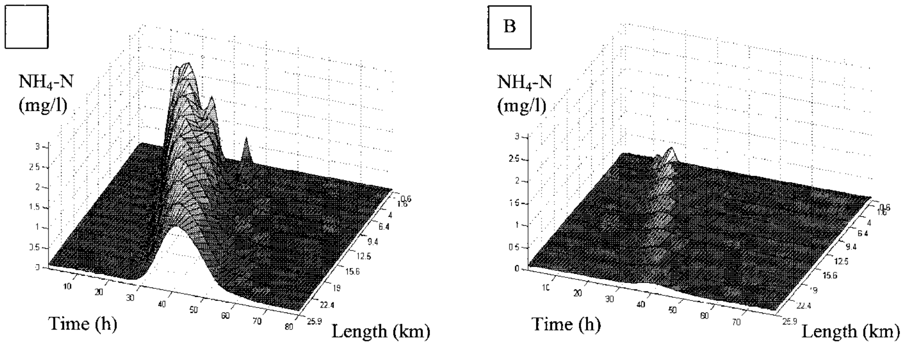

Figure 5 Ammonia concentrations in the Lambro during a storm, with old (a) and upgraded (b) plant capacity

consequence. The plant was extended with another primary settler, an extra oxidation tank and an extra secondary settler in June 1998 (see Meirlaen *et al.*, in preparation a).

In this study the river Lambro has been modelled as a series of completely mixed tanks with variable volume. Since the flow could be assumed to be rather steady (due to its origin in a lake further upstream), and since 2 rating curves along the river reach were available, it was not necessary to set up a complex hydraulic model. The variable volume tanks parameters were set on the basis of these rating curves. The available tracer data (Boron) were sufficient to determine the number of tanks needed to describe the behaviour of the river. Due to the steady flow, dispersion was the only important physical process to be taken into account. The effect of the number of tanks on dispersion is shown in Figure 4(a). It is clear that with an increasing number of tanks, dispersion is decreasing, while the peak concentration arrives later. By comparison with the field data, it was concluded that 47 tanks yielded a good approximation of the Boron concentration in the last section of the river, for peak time, peak height and dispersion rate. When looking at several places along the river (Figure 4(b)) and taking into account the uncertainty of the data, it was decided that 47 tanks described the hydraulic behaviour reasonably well (for details, see Meirlaen *et al.*, in preparation b).

Due to a lack of data, the treatment plant model was only calibrated on daily average effluent quality data, while using the parameters of the ASM1 as suggested by Henze *et al.* (1987) as a starting point. Simulations with both the previous and the current plant layout have been performed. The available data are from the former plant layout (Gandolfi and Facchi, 1998a, b). Since no data were available for the upgraded plant, identical biological parameters have been used in both cases, hence simulation results should be carefully interpreted.

After the development of both the river model and the treatment plant model were implemented in WEST®, and these models were linked to each other by the connector as described above. Some simulation results are shown in Figure 5. Because of the lower hydraulic capacity of the old plant, more water had to be bypassed, resulting in high pollution peak loads coming from the treatment plant (Figure 5a). The ammonium wave travels downstream with time while the peak concentration is lowered due to dispersion and nitrification in the river. With the upgraded plant, the overflow at the treatment plant is much lower, resulting in hardly elevated ammonium concentrations during the bypass under study. The effect of an increase in hydraulic capacity can be predicted as expected, but should be confirmed by measurements on the river with the upgraded plant in operation.

Conclusions

A new procedure to build and calibrate surrogate models with the aid of complex models has been outlined. This procedure solves both the problems of the long calculation times of the complex mechanistic models and the problem of the amount of field data required to calibrate models. Indeed, a calibrated complex mechanistic model is used to generate virtual data, which are subsequently used to calibrate a fast surrogate model. This approach was illustrated on the river Zwalm, where flood wave propagation could be approximated by a series of tanks with variable volume. This surrogate model to predict the hydraulic behaviour is easier to implement and can be used in combination with a biological conversion model. Moreover, it was three times faster than the complex mechanistic model. Hence, the use of surrogate models within optimisation problems seems to be promising.

A connector necessary to link the ASM1 and the RWQM1 models has been developed and applied in a case study on the river Lambro. Due to the steady flow regime and the availability of rating curves, the hydraulics of the Lambro could be approximated by a series of tanks with variable volume. It was not necessary to build a complex hydraulic model and generate virtual data with it. Because both the treatment plant and the river model could be implemented within the WEST® modelling and simulation software, linking of the two models is straightforward and simultaneous simulations can be conducted. Simulation results clearly show the detrimental effect of the pollution load of both overflow and treatment plant on the river. The integrated model predicts a substantial reduction in pollution load and an increase in the river water quality by the extension of the hydraulic capacity of the treatment plant. The simulation results for the upgraded plant still need validation by field data.

Acknowledgements

The author is Research Assistant of the Fund for Scientific Research - Flanders (Belgium). The authors would like to give special thanks to the Fund for Scientific Research (G.0102.97) for the financial support. The authors also thank members of the GREATER ECETOC task force (especially Prof. Gandolfi and his team from the university of Milan) for the supply of data.

References

Beck, M.B. and Young, P.C. (1975). A dynamic model for DO-BOD relationships in a non-tidal stream. *Wat. Res.*, **9**, 769–776.

Benedetti, L. and Sforzi, F. (1999). *Dynamic integrated modelling: a case study on the river Lambro*. Engineers thesis, Biomath Department, Ghent University.

Entem, S., Lahoud, A., Yde, L. and Bendsen, B. (1998). Real time control of the sewer system of Boulogne Billancourt – a contribution to improving the water quality of the Seine. *Wat. Sci. Tech.*, **37**(1), 327–332.

Fronteau, C. (1999). *Water quality management of river basins and evaluation of the impact of combined sewer overflows using an integrated modelling approach*. PhD. Thesis, University of Brussels, Brussels, Belgium, pp. 252.

Fronteau, C., Bauwens, W. and Vanrolleghem, P.A. (1997) Integrated modelling: Comparison of state variables, processes and parameters in sewer and waste water treatment models. *Wat. Sci. Tech.*, **36**(5), 373–380.

Fronteau, C., van Griensven, A. and Bauwens, W. (1999). Construction and calibration of an integrated urban drainage model. *Proc. For the 5th International Conference on Water Pollution 99,* May 24–26, Lemnos, Greece.

Gandolfi, C. and Facchi, A. (1998a). *Italian Subgroup Interim Activity Report,* January 1998–June 1998 – GREAT-ER Project.

Gandolfi, C. and Facchi, A. (1998b). *Italian Subgroup Interim Activity Report,* October 1998 – GREAT-ER Project.

Henze, M., Grady, C.P.L., Gujer, W., Marais, G.v.R. and Matsuo, T. (1987). *Activated Sludge Model No. 1*. IAWPRC, London, ISSN: 1010–707X.

Maryns, F. and Bauwens, W. (1997).The application of the activated sludge model no.1 to a river environment. *Wat. Sci. Tech.,* **36**(5), 201–208.

Meirlaen, J., Benedetti, L., Sforzi, F., Facchi, A., Gandolfi, C. and Vanrolleghem, P. (in preparation a). *Dynamic integrated modelling: a case study on the river Lambro.*

Meirlaen J., Sforzi F., Benedetti L. and Vanrolleghem P. (in preparation b). *The use of tanks in series models to describe hydraulic phenomena in rivers.*

Nielsen M.K., Carstensen J. and Harremoës P. (1996). Combined control of sewer and treatment plant during rainstorm. *Wat. Sci. Tech.*, **34**(3–4), 181–187.

Paulsen, O. (1986). Kontinuierliche Simulation von Abflüssen und Schmutzfrachten in der Trennentwässerung. Mitteilunges des Institutes für Wasserwirtschaft; Universität Hannover, Vol 62.

Petruck, A., Cassar, A. and Dettmar, J. (1998). Advanced real time control of a combined sewer system. *Wat. Sci. Tech.*,, **37**(1), 319–326.

Rauch, W., Aalderink, H., Krebs, P., Schilling, W. and Vanrolleghem, P. (1998) Requirements for integrated wastewater models – Driven by receiving water objectives. *Wat. Sci. Tech.*, **38**(11), 97–104.

Rauch, W. and Harremoës, P. (1999). Genetic algorithms in real time control applied to minimise transient pollution from urban wastewater systems. *Wat. Res.*, **33**, 1265–1277.

Reda, A. (1996). Simulation and control of stormwater impacts on river water quality. PhD. Thesis. Imperial College London, UK. pp. 512.

Reichert ,P. (1994). Aquasim – A tool for simulation and data analysis of aquatic systems. *Wat. Sci. Tech.*, **30**(2), 21–30.

Reichert, P., Borchardt, D., Henze, M., Rauch, W., Shanahan, P., Somlyódy, L. and Vanrolleghem, P.A. (2001). River water quality model no. 1 (RWQM1): II. Biochemical process equations. *Wat. Sci. Tech.*, **43**(6), 329–338 (this issue).

Schütze, M., Butler, D. and Beck, B.(1998). Optimisation of control strategies for the urban wastewater system – an integrated approach. *Wat. Sci. Tech.*, **39**(9), 209–216.

Shanahan, P., Borchardt, D., Henze, M., Rauch, W., Reichert, P., Somlyódy, L. and Vanrolleghem, P.A. (2001). River water quality model no. 1 (RWQM1): I. Modelling approach. *Wat. Sci. Tech.*, **43**(6), 329–338 (this issue).

Spriet, J.A. and Vansteenkiste, G.C. (1982). *Computer-aided Modelling and Simulation. Lecture Notes in Computer Science*. Academic Press, London.

Vanrolleghem, P.A., Schilling, W., Rauch, W., Krebs, P. and Aalderink, H. (1999) Setting up measuring campaigns for integrated wastewater modelling. *Wat. Sci. Tech.*, **39**(4), 257–268.

Vanrolleghem, P.A., Borchardt, D., Henze, M., Rauch, W., Reichert, P. Shanahan, P. and Somlyódy, L. (2001). River water quality model no. 1: III. Biochemical submodel selection. *Wat. Sci. Tech.*, **43**(6), xx–-xx (this issue).

Whelan, M.J., Gandolfi, C. and Bischetti, G.B. (1999). A simple stochastic model of point source solute transport in rivers based on gauging station data with implications for sampling requirements. *Wat. Res.*, **33**, 3171–3181.