Environmental Modelling & Software 26 (2011) 1255–1267

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach

Krist V. Gernaey [a,*], Xavier Flores-Alsina [a,b], Christian Rosen [b,c], Lorenzo Benedetti [d,e], Ulf Jeppsson [b]

[a] Center for Process Engineering and Technology (PROCESS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby, Denmark
[b] Department of Measurement Technology and Industrial Electrical Engineering (MIE), Division of Industrial Electrical Engineering and Automation (IEA), Lund University, Box 118, SE-221 00 Lund, Sweden
[c] AnoxKaldnes AB, Klosterängsvägen 11A, SE-226 47 Lund, Sweden
[d] Department of Applied Mathematics, Biometrics and Process Control (BIOMATH), Ghent University, Coupure Links 653, B-9000 Ghent, Belgium
[e] Waterways srl, Via del Ferrone 88, 50023 Impruneta (FI), Italy

ARTICLE INFO

Article history:
Received 3 November 2010
Received in revised form
31 May 2011
Accepted 1 June 2011
Available online 2 July 2011

Keywords:
Dynamic disturbances
Influent modelling
Urban drainage
Wastewater treatment plant
BSM2

ABSTRACT

Activated Sludge Models are widely used for simulation-based evaluation of wastewater treatment plant (WWTP) performance. However, due to the high workload and cost of a measuring campaign on a full-scale WWTP, many simulation studies suffer from lack of sufficiently long influent flow rate and concentration time series representing realistic wastewater influent dynamics. In this paper, a simple phenomenological modelling approach is proposed as an alternative to generate dynamic influent pollutant disturbance scenarios. The presented set of models is constructed following the principles of parsimony (limiting the number of parameters as much as possible), transparency (using parameters with physical meaning where possible) and flexibility (easily extendable to other applications where long dynamic influent time series are needed). The proposed approach is sub-divided in four main model blocks: 1) model block for flow rate generation, 2) model block for pollutants generation (carbon, nitrogen and phosphorus), 3) model block for temperature generation and 4) model block for transport of water and pollutants. The paper is illustrated with the results obtained during the development of the dynamic influent of the Benchmark Simulation Model no. 2 (BSM2). The series of simulations show that it is possible to generate a dry weather influent describing diurnal flow rate dynamics (low rate at night, high rate during day time), weekend effects (with different flow rate during weekends, compared to weekdays), holiday effects (where the wastewater production is assumed to be different for a number of weeks) and seasonal effects (with variations in the infiltration and thus also the flow rate to the WWTP). In addition, the dry weather model can be extended with a rain and storm weather generator, where the proposed phenomenological model can also mimic the "first flush" effect from the sewer network and the influent dilution phenomena that are typically observed at full-scale WWTPs following a rain event. Finally, the extension of the sewer system can be incorporated in the influent dynamics as well: the larger the simulated sewer network, the smoother the simulated diurnal flow rate and concentration variations. In the discussion, it is pointed out how the proposed phenomenological models can be expanded to other applications, for example to represent heavy metal or organic micro-pollutant loads entering the treatment plant.

Software availability

Name of the software: WWTP Influent Generator Model
Developer: Krist V. Gernaey, Xavier Flores-Alsina, Lorenzo Benedetti, Christian Rosen, Ulf Jeppsson
Programming language: Matlab 7.0
Availability: The source code for generation of WWTP influent files can be obtained for free. Contact Krist V. Gernaey, Center for Process Engineering and Technology (PROCESS), Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby, Denmark. E-mail: kvg@kt.dtu.dk

* Corresponding author. Tel.: +45 45 25 29 70; fax: +45 45 93 29 06.
E-mail address: kvg@kt.dtu.dk (K.V. Gernaey).

1364-8152/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2011.06.001

Nomenclature

A_1 Surface area of the variable volume tank ('Soil' model block) [m^2]
A_2 Surface area of the sewer variable volume tank ('Sewer' model block) [m^2]
aH A parameter determining the direct contribution of rainfall falling on impermeable surfaces in the catchment area to the flow rate in the sewer ('Rain generator' model block) [%]
ASMs Activated Sludge Models
BSM1 Benchmark Simulation Model no.1
BSM2 Benchmark Simulation Model no.2
C Gain relating the outlet flow rate Q_{out} to the liquid level in the variable volume tank ('Sewer' model block) [$m^{1.5}$ d^{-1}]
COD Chemical Oxygen Demand
COD_{part} Particulate COD
CODpart_gperPEperd COD_{part} load per person equivalent ('Households pollutants' model block) [(g COD pe^{-1}) d^{-1}]
CODpart_Ind_kgperd COD_{part} load from industry ('Industry pollutants' model block) [kg COD d^{-1}]
COD_{sol} Soluble COD
CODsol_gperPEperd COD_{sol} load per person equivalent ('Households pollutants' model block) [(g COD pe^{-1}) d^{-1}]
CODsol_Ind_kgperd COD_{sol} load from industry ('Industry pollutants' model block) [kg COD d^{-1}]
COD_{tot} Total COD
FF Gain for first flush effect ('First flush effect' model block) [d^{-1}]
FFfraction Fraction of *TSS* that can settle in the sewer ('First flush effect' model block) [-]
h_1 Actual height of the water level in the variable volume tank ('Soil' model block) [m]
h_2 Actual height of the water level in the sewer variable volume tank ('Sewer' model block) [m]
H_{inf} Height of the water level above the invert level ('Soil' model block) [m]
H_{inv} Height of the invert level ('Soil' model block) [m]
H_{max} Maximum height of the water level in the variable volume tank ('Soil' model block) [m]
H_{min} The minimum liquid level that needs to be exceeded before there is a flow rate out of the variable volume tank model ('Sewer' model block) [m]
InfAmp Amplitude of the sine wave for generating seasonal effects due to infiltration ('Seasonal correction factor' model block) (m^3 d^{-1}]
InfBias Mean value of the sine wave signal for generating seasonal effects due to infiltration ('Seasonal correction factor' model block) [m^3 d^{-1}]
InfFreq Frequency of the sine wave for generating seasonal effects due to infiltration ('Seasonal correction factor' model block) [rad d^{-1}]
InfPhase Phase shift of the sine wave for generating seasonal effects due to infiltration ('Seasonal correction factor' model block) [rad]
IWA International Water Association
K Soil permeability constant ('Soil' model block) [m^3 m^{-2} d^{-1}]
K_{down} Gain for adjusting the flow rate to downstream aquifers ('Soil' model block) [m^2 d^{-1}]
K_{inf} Infiltration gain ('Soil' model block) [$m^{2.5}$ d^{-1}]
LLrain A constant converting the random number generator output to a value representing rainfall intensities ('Rain generator' model block) [mm rain d^{-1}]
M_{in} Sediment flux entering the sewer system ('First flush effect' model block) [kg d^{-1}]
M_{out} Sediment flux leaving the sewer system ('First flush effect' model block) [kg d^{-1}]
M_{max} Maximum mass of stored sediment in the sewer system ('First flush effect' model block) [kg]
N Exponent for Hill function ('First flush effect' model block) [-]
N_{tot} Total N concentration [g N m^{-3}]
N_{org} Total organic N concentration [g N m^{-3}]
PAOs Polyphosphate accumulating organisms
PE Person equivalent ('Households' model block) [-]
Q_{in} Flow rate in to a variable volume tank ('Sewer model block) [m^3 d^{-1}]
QInd_weekday Average wastewater flow rate from industry on normal weekdays (Monday to Thursday) ('Industry' model block) [m^3 d^{-1}]
Q_{intot} The sum of the infiltration water flow rate and the flow rate originating from the fraction of rainfall that is not collected on impermeable surfaces ('Soil' model block) [m^3 d^{-1}]
Q_{in1} Contribution of rain water to the inflow of the variable volume tank ('Soil' model block) [m^3 d^{-1}]
Q_{in2} Contribution of 'Seasonal correction factor' model block to the inflow of the variable volume tank ('Soil' model block) [m^3 d^{-1}]
Q_{lim} Flow rate limit triggering a first flush effect ('First flush effect' model block) [m^3 d^{-1}]
Q_{out} Outlet flow rate of a variable volume tank ('Sewer' model block) [m^3 d^{-1}]
Q_{out1} Outlet flow rate to the sewer of the variable volume tank ('Soil' model block) [m^3 d^{-1}]
Q_{out2} Outlet flow rate to the downstream aquifers of the variable volume tank ('Soil' model block) [m^3 d^{-1}]
Q_{out3} Overflow flow rate of the variable volume tank ('Soil' model block) [m^3 d^{-1}]
Qpermm Flow rate per mm rain ('Rain generator' model block) [m^3 mm^{-1}]
QperPE Wastewater flow rate per person equivalent ('Households' model block) [m^3 d^{-1}]
Q_w Waste sludge flow rate [m^3 d^{-1}]
S_A Acetate [g COD m^{-3}]
S_{ALK} Alkalinity [mol HCO_3^- m^{-3}]
S_F Fermentable products [g COD m^{-3}]
S_I Inert soluble COD [g COD m^{-3}]
S_{ND} Soluble organic nitrogen [g N m^{-3}]
S_{NH} Ammonium nitrogen [g N m^{-3}]
SNH_gperPEperd S_{NH} load per person equivalent ('Households pollutants' model block) [(g N pe^{-1}) d^{-1}]
SNH_Ind_kgperd S_{NH} load from industry ('Industry pollutants' model block) [kg N d^{-1}]
S_{NO} Nitrite + nitrate nitrogen [g N m^{-3}]
S_O Dissolved oxygen [g (-COD) m^{-3}]
S_S Readily biodegradable COD [g COD m^{-3}]
Subareas A parameter that forms a measure of the size of the catchment area. It will determine the number of variable volume tanks in series that will be used for describing the sewer system ('Sewer' model block) [-]
S_{NKj} Total Kjeldahl nitrogen [g N m^{-3}]

S_{PO4} Soluble ortophosphate phosphorus [g P m^{-3}]
SPO4_gperPEperd S_{PO_4} load per person equivalent ('Households pollutants' model block) [(g P pe^{-1}) d^{-1}]
SPO4_Ind_kgperd S_{PO4} load from industry ('Industry pollutants' model block) [kg P d^{-1}]
TAmp Seasonal temperature variation, amplitude ('Temperature' model block) [°C]
TBias Seasonal temperature variation, average ('Temperature' model block) [°C]
TdAmp Daily temperature variation, amplitude ('Temperature' model block) [°C]
TdBias Daily temperature variation, average ('Temperature' model block) [°C]
TdFreq Daily temperature variation, frequency ('Temperature' model block) [rad d^{-1}]
TdPhase Daily temperature variation, phase shift ('Temperature' model block) [rad]
TFreq Seasonal temperature variation, frequency ('Temperature' model block) [rad d^{-1}]
TKN_gperPEperd S_{NKj} load per person equivalent ('Households pollutants' model block) [(g N pe^{-1}) d^{-1}]
TKN_Ind_kgperd S_{NKj} load from industry ('Industry pollutants' model block) [kg N d^{-1}]
TPhase Seasonal temperature variation, phase shift ('Temperature' model block) [rad]
TSS Total suspended solids concentration [g m^{-3}]
WWTP Wastewater treatment plant
X_{BA} Autotrophic biomass [g COD m^{-3}]
X_{BH} Heterotrophic biomass [g COD m^{-3}]
X_I Inert particulate COD [g COD m^{-3}]
X_{ND} Particulate organic nitrogen [g N m^{-3}]
X_P Inert particulate COD resulting from biomass decay [g COD m^{-3}]
X_{PAO} Phosphorus accumulating organisms [g COD m^{-3}]
X_{PP} Polyphosphate phosphorus [g P m^{-3}]
X_S Slowly biodegradable particulate COD [g COD m^{-3}]
X_{STO} Storage products [g COD m^{-3}]

1. Introduction

Activated Sludge Models (ASMs) are widely used for benchmarking (Copp, 2002; Jeppsson et al., 2007), diagnosis (Rodriguez-Roda et al., 2002), design (Flores et al., 2007; Benedetti et al., 2010), teaching (Hug et al., 2009) and optimization (Rivas et al., 2008) of wastewater treatment plants (WWTPs). In order to obtain proper simulation results for a specific WWTP, several studies have focused on developing standardized procedures for calibrating ASMs (Hulsbeek et al., 2002; Petersen et al., 2002; Melcer et al., 2003; Langergraber et al., 2004; Gillot et al., 2009), efforts that have also resulted in the formation of the IWA Task Group on Good Modelling Practice. Once a calibrated model is available and applied in the framework of WWTP simulation studies, the current practice is to simulate the plant for a relatively long period of time, preferably with dynamic influent inputs including a wide range of disturbance scenarios, i.e. daily, weekly, seasonal and annual variations amongst others (Gernaey et al., 2004). If no dynamic influent disturbances are applied to the WWTP in a simulation study, the system will not be sufficiently excited and thus the simulations will result in a too optimistic picture of the plant performance (Ráduly et al., 2007).

In essence, the success of a WWTP simulation study depends to a large extent on the availability of a sufficiently long set of influent data – the main disturbance of a typical WWTP – representing the inherent natural variability of the flow rate and pollutant concentrations at the plant inlet as well as possible. Dynamic influent flow rate data can be obtained easily from on-line sensors, every time more reliable and cheaper, but still require considerable efforts. In addition, there is often a severe mismatch between measured (assumed) flow rates and the real flow rate entering the treatment plant, for example as a consequence of drift of the sensor signal (Rieger et al., 2004). Contrary to influent flow rate, it is not that easy to obtain long dynamic influent pollutant concentration time series. A first bottleneck is the amount of work and the cost involved for analyzing samples taken from the influent for the most relevant influent pollutants (soluble and total chemical oxygen demand (COD), total suspended solids (TSS), ammonium nitrogen (NH_4^+–N), total Kjeldahl nitrogen (TKN), orthophosphate phosphorus (PO_4^{3-}-P), total phosphorus (TP), etc.). One should indeed consider a sampling interval of maximum 2 h if influent dynamics should be represented realistically in the measured influent concentration data. If a 1 h sampling interval is adopted, and if the aim is to obtain one year of dynamic influent data, then 8760 samples need to be analyzed to complement the influent flow rate data. A second bottleneck is that on-line collection of dynamic influent pollutant concentration data is expensive (e.g. cost of sensors, maintenance and calibration of sensors, consumption of chemicals), and does not guarantee that an influent pollutant concentration data set of high quality will be obtained. Indeed, several standard lab analyses, such as COD, can still not be performed reliably in on-line mode on the influent of a WWTP.

Model-based influent scenario generation is an alternative that recently has gained considerable interest. Dynamic influent flow rate data can be generated by means of a simple equation, such as a Fourier series (sum of sinusoids with varying frequencies and phase shifts) whose parameters are fitted to dynamic influent data (e.g. Carstensen et al., 1998; Langergraber et al., 2008; Alex et al., 2009), by means of a more sophisticated model (e.g. Gernaey et al., 2005; Béraud et al., 2007) or by a very complex and detailed deterministic model of the complete catchment area (e.g. Hernebring et al., 2002). The same methods used for the influent flow rate generation can be applied to describe the dynamics of the different influent pollutant loads (e.g. Bechmann et al., 1999). Note also that model-based influent disturbance generation for WWTPs has not been limited to flow rate and traditional ASM influent fractions. Lindblom et al. (2006) combined an influent model for ASM variables with the generation of influent pollutant profiles for two xenobiotic compounds, and used the resulting influent pollutant profiles to model the fate of pollutants in the WWTP. De Keyser et al. (2010) developed a model that generates time series of traditional and micro-pollutants from their emission sources in the urban catchment. Similarly, Ort et al. (2005) developed a stochastic model describing short-term variations of benzotriazole concentrations (a chemical contained in dishwasher detergents). Additionally, during the development of the BSM1_LT platform (BSM1 Long Term, Rosen et al., 2004), the influent model for ASM variables was combined with a Markov chain approach to describe the occasional occurrence of either toxic or inhibitory influent shock loads (Rosen et al., 2008).

In this paper, a phenomenological modelling approach is proposed to generate dynamic influent pollutant disturbance scenarios. The novelty of this approach relies on a set of generic model blocks that combine urban water system insights and knowledge on typical experimental influent data sets, but without pretending at any point to provide a basis for studying the complex drainage mechanisms in detail (Achleitner et al., 2007; Elliott and Trowsdale, 2007; Fu et al., 2008; Devesa et al., 2010). Three basic principles were applied during the model development: 1) model parsimony; i.e. limiting the number of parameters as much as

possible; 2) model transparency, by using model parameters that have a physical meaning where possible; and 3) model flexibility, such that the proposed influent model can be easily extended for other applications where long influent time series are needed. The main purpose of this paper is to present and demonstrate the capabilities of the new phenomenological modelling approach. The proposed influent model produces dynamic influent flow rate, pollutant loads and temperature profiles using four main model blocks. Each model block – flow rate generation model, pollutants generation model, temperature generation model, transport model – will be described in detail below, with focus on explaining the main underlying principles. The performance of the model is evaluated by generating several dynamic influent pollutant disturbance scenarios, where examples were taken from the influent file that was developed for the Benchmark Simulation Model no. 2 (BSM2, Jeppsson et al., 2007). Specifically for the selected case study a dynamic influent time series of 609 days is generated. Only the last 364 days of those influent data are used for plant performance evaluation in the BSM2. The influent data are assumed to correspond to the influent of a WWTP located in the Northern hemisphere, and are designed such that the evaluation period starts on July 1st. The first part of the influent data (245 days) has two purposes in the BSM2: the first 63 days are used to reach a pseudo-steady state, whereas the next 182 days are used for generating training data, e.g. for fine tuning control or monitoring algorithms before the start of the evaluation period. Each model block is commented in detail below, and is illustrated using simulation results. The *X*-axis labels (time labels) on the graphs and in the text refer to specific days in the BSM2 influent time series.

2. Flow rate generation model

The first model block contains the flow rate generation model. The generation of the influent flow rate is achieved by combining the contributions from households, industry, rainfall and infiltration (see Fig. 1). Rainfall contributes to the total flow rate in two ways: the largest part (*aH*) of the rainfall contribution to the flow rate originates from runoff from impermeable surfaces, and is thus transported directly to the sewer. Rainfall on permeable surfaces (fraction 100 − *aH* of the rainfall) will influence the groundwater level, and thus the contribution of infiltration to the influent flow rate. Assuming a dry and a rainy season, the 'Seasonal correction factor' model block creates the desired seasonal effect. This seasonal correction is combined with the rainfall falling on permeable surfaces, and the sum of both flows is passed through the 'Soil' model block. Afterwards, the net contribution of infiltration – an output of the 'Soil' model block – is combined with the overall flow rate resulting from industry and households and the flow rate contribution from rainfall on impermeable surfaces.

2.1. 'Households' model block

In the proposed approach, the 'Households' (HH) model block contributes to the final influent flow rate dynamics with diurnal variations, a weekend effect and a holiday effect. This is achieved by combining three normalized user-defined data files containing: 1) a diurnal profile; 2) a weekly household flow pattern including the weekend effect; and 3) a yearly pattern including the holiday effect. The relative contributions from the data files are combined by multiplication. The generated signal is then multiplied by two gains corresponding to the flow rate per person equivalent (*QperPE*) and the number of person equivalents in the catchment area (*PE*) respectively. The dynamic flow rate pattern is obtained by repeating the data files in a cyclic manner. See also the detailed flow diagram in the supplemental information for further details.

Fig. 2 provides an example of the effect of the different elements comprising the HH model block assuming *QperPE* = 0.150 $m^3\ d^{-1}$ and *PE* = 80,000 *person equivalents (pe)*. Thus, the resulting daily average flow rate from HH is 12,000 $m^3\ d^{-1}$. As can be seen in Fig. 2a, there is a daily flow rate profile that is repeated. The daily flow rate profile represents a general behaviour, namely one morning peak, one evening peak, and late night and mid-day minima. The morning and the evening peaks represent the residents going to or returning from work. The late night minimum flow rate corresponds to the sleeping hours with little water consumption. Finally, the daytime flow rate shows a small decline corresponding to the residents' working hours. The result of including the weekend and the holiday effect is quite clear (see Fig. 2c and e, respectively). When the weekly pattern is included there is an 8% flow rate reduction on Saturdays and a 12% reduction on Sundays. Finally, the holiday period (week 4–6) represents a 25% reduction of the flow rate during the first two weeks and a 12% decrease during the third week. For this specific case, the user-defined profiles correspond to the experimental observations reported in Butler et al. (1995). However, other input files could of course describe different influent characteristics.

Zero mean white noise can be added to this flow rate signal. White noise is a sequence of independent and identically distributed random numbers with zero mean and specified variance. Noise is added by adding a random signal to a time series, e.g. flow rate. Negative values of the resulting noisy signal, e.g. flow rate, or other unrealistic situations are avoided by also including a maximum and/or a minimum threshold for the noisy signal. The noise variance is a tunable model parameter and depends on 1) the nature of the

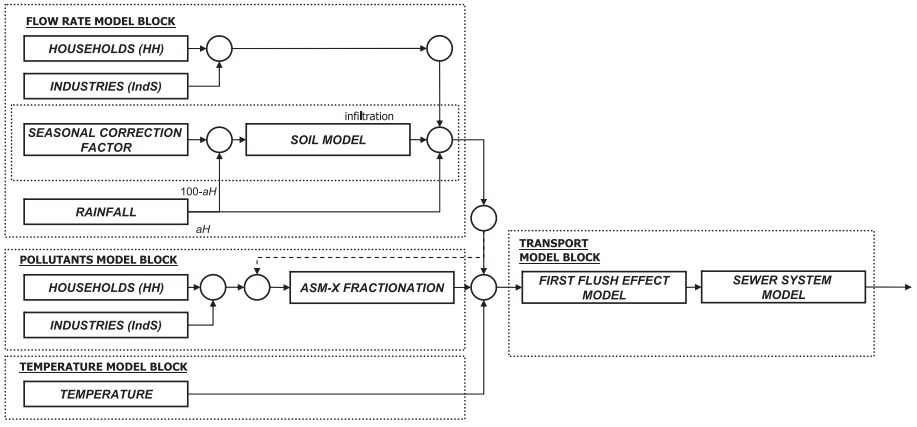

Fig. 1. Schematic representation of the phenomenological influent pollutant disturbance scenario modelling approach.

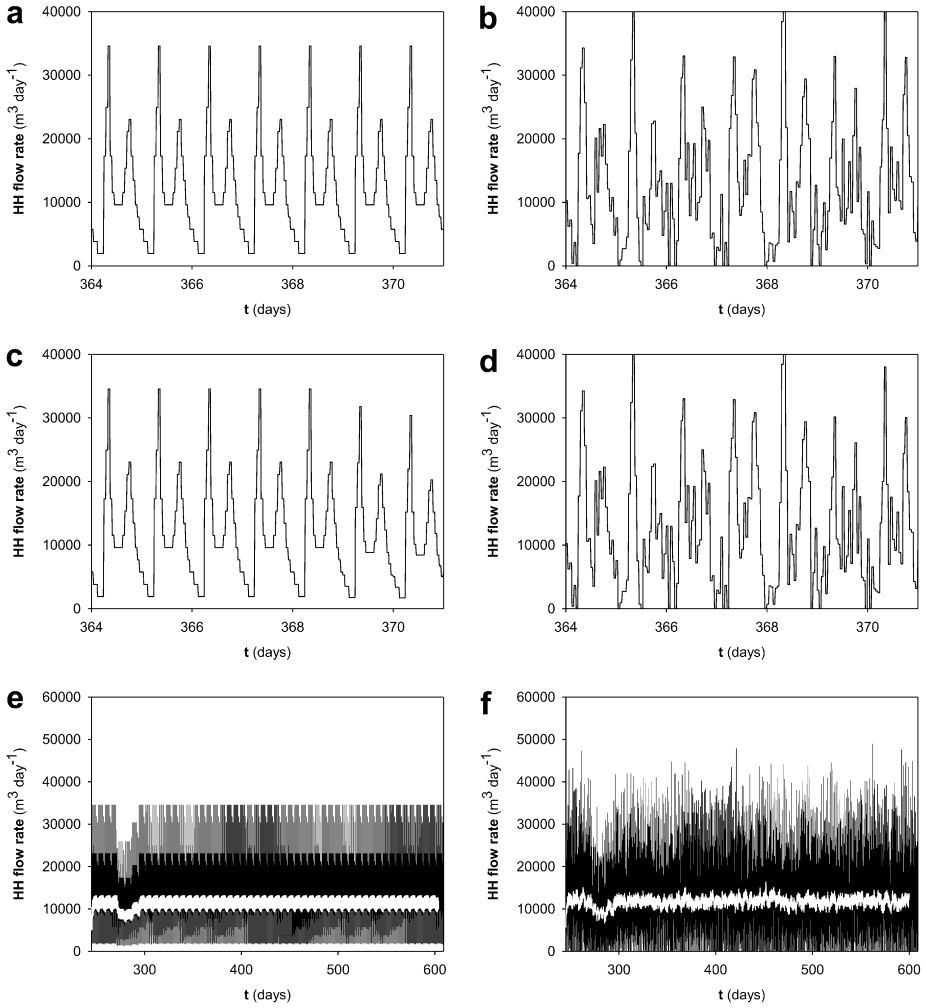

Fig. 2. Dynamic flow rate profiles (15 min temporal resolution) resulting from the HH model block: Only diurnal effect (a,b), diurnal and weekend effects (c,d) and diurnal, weekend and holiday effects (e,f). Zero mean white noise is added in b, d and f. One year of data is shown in figures e and f, starting on July 1st. An exponential smoothing filter (time constant = 1 day) is used in e and f to improve visualization of the data (in white).

temporal series (flow rate, loads, concentration) and 2) the temporal resolution (Gernaey et al., 2011a). Further details about the Simulink implementation of the noise generators can be found in the supplemental information section. It is up to the model user to decide whether or not to activate the random noise. Specifically for flow rate, adding noise is done with the purpose of avoiding that subsequent days have exactly the same flow rate profile (see Fig. 2,b, d and f). Applying noise to the different time series (influent flow rate, loads and concentrations) has the additional advantage that the correlation between the different variables is reduced, since each noise generator has a different seed value. Finally, it is important to mention that better approaches could be incorporated in the phenomenological models in order to correctly represent the noise structure of the data, for example by replacing white noise by an ARIMA model, where the parameters would be different depending on the temporal resolution. Nevertheless, the presented way of adding noise to the influent time series has been chosen since this is the way the IWA Task Group on Benchmarking of Control Strategies for Wastewater Treatment Plants decided to incorporate noise in the benchmark models (Gernaey et al., 2011a,b).

2.2. 'Industry' model block

The industrial contribution to the influent flow rate is generated similarly to the HH model block. The industry (IndS) model block is also based on user-defined files describing diurnal variations, and weekend and holiday effects. Again, the dynamic flow rate pattern sampled in a cyclic manner from source files is multiplied by a gain representing the use of water made by industry (*QInd_weekday*) similarly to the generation of the diurnal flow rate patterns in the HH model block (see flow diagram of this model block in the supplemental information for further details).

The weekly industrial wastewater production flow pattern, assuming that *QInd_weekday* = 2,500 $m^3\ d^{-1}$, is illustrated in Fig. 3. During the weekend, the industrial wastewater production is reduced to around 25% of the average flow rate generated on normal weekdays (see Fig. 3a). There is also a Friday afternoon peak included in the weekly profile, which is assumed to correspond to industrial cleaning at the end of the working week. The holiday effect on the industrial wastewater production corresponds to the three weeks with holiday effect in the HH model block, and is characterized by an 80% reduction of the flow rate during those 3 weeks. In addition, the industrial wastewater production is reduced by 70% during week 26, to simulate the shutdown of industrial activity during the winter holidays period (see Fig. 3b). Nevertheless, it should be emphasized here that this is just an example, and inputs can be modified by the user with the purpose of obtaining another type of dynamic influent profile, e.g. industries with daily cleaning. In both cases, adding zero mean white noise to the IndS flow rate adds realism to the industrial wastewater flow rate profiles.

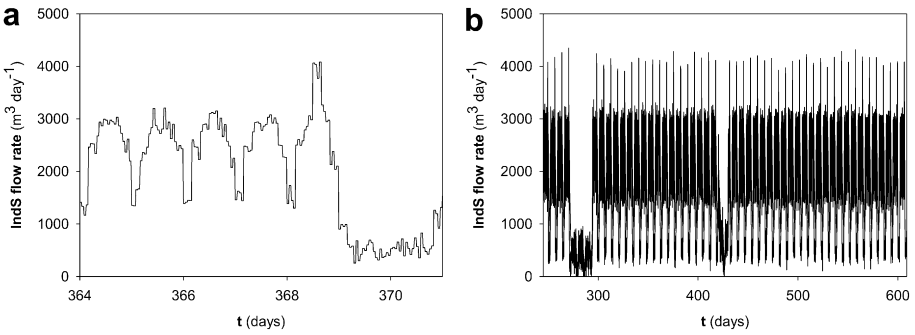

Fig. 3. Dynamic profiles resulting from the IndS model block: only diurnal and weekly effects (a) and holiday effect (b). Zero white mean noise is added to the flow rate data. One year of data is shown in figure b starting on July 1st.

2.3. 'Seasonal correction factor' model block

The seasonal changes of the amount of infiltration are attributed to changes in the groundwater levels over the year. During the cold season it is assumed that the groundwater level is high, resulting in a high infiltration of groundwater in the sewer system. On the other hand, during the warm season the groundwater level is low resulting in reduced infiltration. These effects are assumed to be the result of seasonal, temperature-related changes of the amount of evaporation.

Fig. 4 shows the flow rate resulting from the 'Seasonal Correction Factor' (SCF) model block. For the BSM2 example, the seasonal variation is implemented as a sine function, with an average level of 7,100 $m^3\ d^{-1}$ (*InfBias*), an amplitude of 1,200 $m^3\ d^{-1}$ (*InfAmp*), and a frequency of $2\pi/364$ rad d^{-1} (*InfFrq*). Note that the flow rate shown in this figure is the input to the soil model block (see also Fig. 1 for details on the data streams between the model blocks). Noise generation can be included as well, similarly to the other model blocks. Nevertheless, in the BSM2 case it was preferred not to add noise to this factor. Changes in the groundwater level, and thus changes in the infiltration flow rate, are assumed to proceed slowly and smoothly (the same would hold true for the other blocks depending on the temporal resolution of the simulated data)

2.4. 'Rain generator' model block

All the model elements mentioned before are needed to generate the dry weather influent flow rate profile, including diurnal variations, weekend and holiday effects and seasonal variations. Since it is assumed that the influent model generator has to be capable of generating an influent from combined sewer systems, rainfall has to be included. In the 'Rain Generator' model block, a random number is first generated and is subsequently multiplied by two factors. The first factor (*LLrain*) converts the numbers generated by the random generator to rainfall intensities in mm d^{-1}, whereas the second one (*Qpermm* [m^3 (mm rain)$^{-1}$]) converts the rainfall intensity into a flow rate. The parameter *Qpermm* thus represents the extra volume of water originating from rainfall that will enter the sewer system for each mm of rain that falls in the catchment. There is also a parameter *aH*, which varies from 0 to 100% and corresponds to the contribution of rainfall falling on impermeable surfaces in the catchment area (see also the detailed flow diagram in the supplemental information for further details).

Fig. 5 shows the output of the 'Rain generator' model block using a random number with mean = 1 and variance = 400. The parameters *LLrain* and *Qpermmm* are set at 3.5 mm rain d^{-1} and 1,500 m^3 (mm rain)$^{-1}$, respectively. The height of the peaks in Fig. 5 gives an indication of the rain intensity.

2.5. 'Soil' model block

The soil model is described as a variable volume tank model. It is used to represent the assumed storage of water in the soil. Parameters for that soil model related to the tank dimensions are: A_1 (the surface area of the groundwater storage tank in the soil model), H_{max} (the maximum level in the tank), H_{inv} (the invert level, i.e. the maximum water level in the groundwater storage tank that will not cause infiltration, corresponding to the bottom level of the

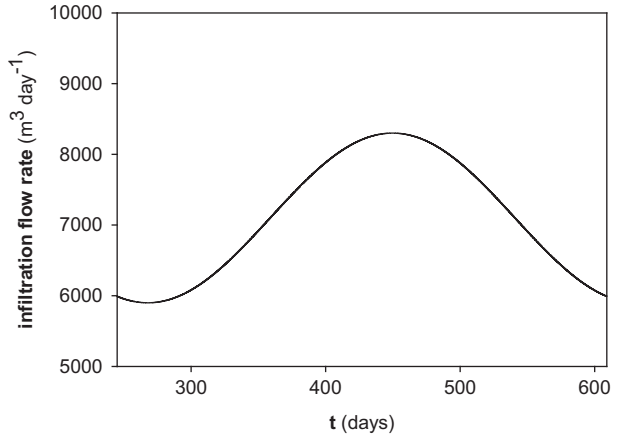

Fig. 4. Seasonal infiltration flow rate generated by the SCF model: *Amplitude* = 1,200 $m^3\ d^{-1}$, *Average flow rate* = 7,100 $m^3\ d^{-1}$. The start of the time series corresponds to July 1st (day 245).

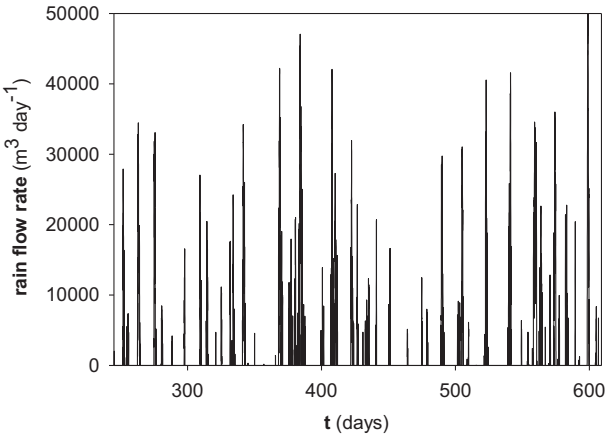

Fig. 5. One year of data: dynamics of the extra influent flow rate due to rain resulting from applying the 'Rain generator' model.

sewer pipes) while H_{inf} is defined as the difference between the actual water level h_1 and H_{inv}. Other model parameters are K (a measure of the permeability for rain water percolation), K_{inf} (infiltration gain, a measure of the quality of the sewer system pipes), and K_{down} (gain to adjust the flow rate to the downstream aquifers) (see also the scheme of this model block in the supplemental information section). There is only one simple differential equation to describe the evolution of h_1, the water level in the groundwater storage tank (Eq. (1))

$$\frac{dh_1}{dt} = \frac{1}{A_1}\cdot(Q_{in1} + Q_{in2} - Q_{out1} - Q_{out2}) \quad (1)$$

where Q_{in2} corresponds to the flow rate generated with the 'Seasonal Correction Factor' model block, i.e. the contribution of rain falling on permeable areas and leaking into the soil (see Section 2.3), and Q_{in1} is the percentage ($100 - aH$) of the rainfall generated in the 'Rain Generator' model block (see Section 2.4). The tank has two outputs, the infiltration flow rate into the sewer system (Q_{out1}) and the flow rate to downstream aquifers (Q_{out2}), which is not further considered in the influent model. Furthermore, the model includes a series of restrictions. Firstly, it is assumed that the permeability of the top soil layer to receive rain water is restricted (Q_{in1} can not be higher than $K\cdot A_1$). The excess of water is assumed to reach the sewer directly ($Q_{out3} = Q_{in1} - K\cdot A_1$). Secondly, the model allows periods where the infiltration to the sewer is zero ($Q_{out1} = 0$) when $h_1 < H_{inv}$. Eq. (2) is similar to Eq. (1), but the detailed calculation of the flow rates out of the tank is included.

$$\frac{dh_1}{dt} = \frac{1}{A_1}\cdot\left(Q_{in1} + Q_{in2} - K_{inf}\cdot\sqrt{H_{inf}} - K_{down}\cdot h_1\right) \quad (2)$$

Fig. 6 illustrates the behaviour of the soil model for the BSM2 case study assuming A = 36,000 m^2, H_{max} = 2 m, K = 0.4 m^3 m^{-2} d^{-1}, K_{inf} = 10,000 m$^{2.5}$ d^{-1}and K_{down} = 1 000 m^2 d^{-1}. Note that parameters of the 'Soil' model block were tuned to obtain the desired output dynamics of the infiltration flow rate to the sewer system generated by this model block. In the overflow, corresponding to periods where $Q_{in1} > K\cdot A_1$ (see Fig. 6), major rain events should result in runoff to the sewer system (Q_{out3}). However, for this case study the parameter K was chosen sufficiently high, such that this contribution will not play a substantial role in the sewer output dynamics. At the same time, major rain events result in a significant increase of the level of the groundwater (h_1), and thus a significant increase of the flow rate due to infiltration. The seasonal variation of the amount of water due to infiltration is clear in Fig. 6. An interesting feature of this simple soil model is that it allows including a 'memory effect', following a rainfall (see Fig. 6), i.e. each major rain event is followed by a tail. This tail illustrates the effect of passing a percentage ($100 - aH = 25\%$) of the rainfall through the soil model block (see Fig. 1), and corresponds with observations made on full-scale plants where it often takes a few days after a major rain event before the flow rate has completely returned to the dry weather situation.

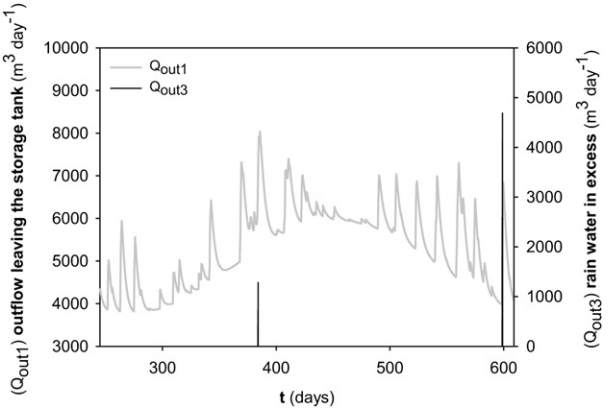

Fig. 6. Overflow, i.e. rain water in excess (Q_{out3}), and outflow leaving the groundwater storage tank, to the sewer system (Q_{out1}).

2.6. Total influent flow rate model

The simulation examples have illustrated that simple phenomenological models allow generating diurnal flow rate variations, with weekend and holiday effects for both households and industry. Moreover, rainfall and infiltration effects can be added to the influent flow rate profile as well. When summing up the effects of the previously described model blocks a total influent flow rate profile can be obtained as shown in Fig. 7. When studying the figure in detail, it is possible to observe the impact of the rainfall contribution on the flow rate profile, although the time series shown in the figure is smoothed using an exponential filter to facilitate visualization. Note that this exponential filter is only used for visualization, and is not included during the generation of the different time series. These results can be compared with the rainfall data of Fig. 5. Unfortunately, due to the large span of the Y-axis scale it is difficult to appreciate the seasonal variation as a consequence of reduced evaporation during the cold season, i.e. it only induces a flow rate variation of 2000 m^3 d^{-1}. Additional simulation results can be found in the supplemental information section, where a ten-day flow rate profile provides additional details on the contribution of each developed model block.

3. Model for influent pollutants

Similarly to the flow rate generation model, it is assumed (see Fig. 1) that there are two pollutant sources, households and industry, which is an acceptable simplification. Thus, the complexity of the model is reduced by neglecting other sources (infiltration and runoff). Including contributions from runoff can for example be important if micro-pollutant influent profiles are needed, but that is not the purpose of the work presented in this manuscript. Pollutants from both sources are combined and converted to ASM-X state variables, where X can be 1, 2, 2d or 3. Note that ASM2d is an extension of ASM2, adding denitrifying activity of the PAOs, but both models have the same state variables. Noise is added to the different influent ASM-X compound concentrations to reduce correlations between state variables (see below).

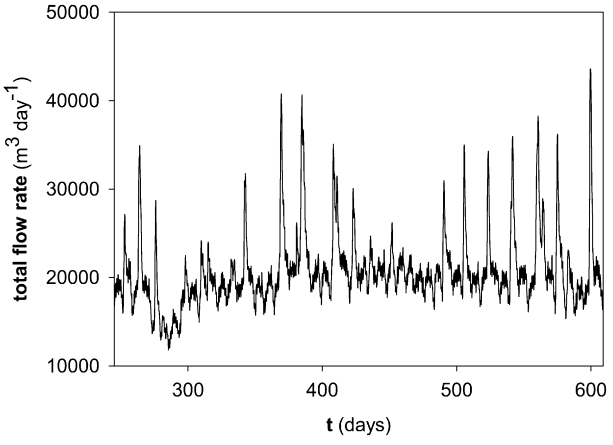

Fig. 7. One year influent flow rate profile. An exponential smoothing filter (time constant = 1 day) is used to improve visualization.

3.1. 'Households pollutants' model block

Similarly to the influent flow rate profiles, soluble and particulate COD (COD_{sol} and COD_{part}), Kjeldahl nitrogen (TKN), ammonium nitrogen (NH_4^+—N) and orthophosphate phosphorus (PO_4^{3-}-P) influent dynamics are generated on the basis of user-defined pollutant flux profiles. The pollutant fluxes are transformed into g COD pe^{-1}, g N pe^{-1} and g P pe^{-1} units, respectively, by multiplying the values in the input files with a gain (specified by the user) e.g. *CODsol_gperPEperd* for soluble COD. The pollutant fluxes are subsequently multiplied by another gain, *PE,* corresponding to the number of person equivalents in the catchment area, similarly to the HH flow rate model block (see Section 2.1). Multiplying the fluxes with this gain converts the fluxes into kg COD d^{-1}, kg N d^{-1} and kg P d^{-1}. Noise is added to each pollutant flux to give more realism (see flow diagram of this model block in the supplemental information for further details).

In the case study the values of *CODsol_gperPEperd*, *CODpart_gperPEperd, TKN_gperPEperd* and *SNH_gperPEperd* are 19.31, 115.08, 12.10 and 5.85 g d^{-1} pe^{-1} and the size of the catchment (PE) is 80,000 *pe.* The dynamics of the different pollutant loads have different characteristics. For the BSM2 influent example both COD and N profiles have two peaks. A pollutant peak occurs in the morning and a smaller one in the evening. On the other hand, the P profile (not included in the BSM2 example, only implemented for the ASM2d) has only one peak during the late afternoon. This is mainly due to the assumption that the main origin of P is detergents used in washing machines, whereas organic matter and nitrogen are primarily originating from kitchen sinks and WCs. The particulate peaks lag slightly behind compared to the soluble pollutant fluxes. During the weekend, there is a reduction of the pollution load of 12% on Saturday and 16% on Sunday. A similar reduction as for the influent flow rate generation is applied to describe the holiday effect on the pollutant fluxes (see Fig. 8, an example for particulate COD).

3.2. 'Industry pollutants' model block

Daily and weekly user-defined profiles describe the dynamics of soluble COD, particulate COD, NH_4^+-N, TKN and PO_4^{3-}-P. As in the IndS flow rate model block, there is a diurnal variation, a Friday afternoon pollutant peak and a weekend effect. The Friday afternoon pollutant peak is assumed to correspond to industrial cleaning at the end of the working week. The flux quantification is done similarly to the HH pollutants model block. The different profiles are each multiplied by a parameter depending on the pollutant, e.g. *CODsol_Indkg*. These parameter values are considered to correspond to average daily industrial pollutant fluxes (in kg COD d^{-1}, kg N d^{-1} and kg P d^{-1}). Finally, zero mean white noise is added (see flow diagram of this model block in the supplemental information for further details).

For the BSM2 influent generation, a daily dynamic pollutant flux for soluble COD (*CODsol_Ind_kgperd*), particulate COD (*CODpart_Ind_kgperd*), TKN (*TKN_Ind_kgperd*) and ammonium nitrogen (*SNH_Ind_kgperd)* is considered. The fluxes assume an industrial load of 386.24, 2,301.80, 109.32 and 52.06 kg d^{-1}, respectively. Fig. 9 represents a weekly TKN profile as an example. As can be seen, the variations in the industry pollutant fluxes are less extreme than the variations of the households' pollutant fluxes. Also, when these industry pollutant fluxes are compared to the industrial wastewater flow rate, the pollutant flux shows a 2-h time delay to account for the sewer dynamics. The Friday afternoon effect is also illustrated in Fig. 9, during which the pollutant fluxes are doubled, assuming to be the consequence of industrial cleaning. During the weekend, the industry pollutant fluxes are considerably lower compared to weekdays (60% decrease of the flux on Saturdays and 80% decrease on Sundays). The same holiday effect as described in the IndS model block is added to the pollutant fluxes (results not shown).

3.3. 'ASM-X fractionation' model block

The 'ASM-X fractionation' model block converts the pollutant fluxes to pollutant concentrations that are compatible with the family of ASMs (Henze et al., 2000). Note that the influent fractionation, that is the conversion of global variables, such as soluble and particulate COD to model specific variables, in this case for ASM1, 2d and 3, takes place after the fluxes from HH and IndS have been combined. Although leading to a simpler model, we are aware of the fact that this also forms a practical limitation since it inherently introduces the assumption that the wastewater generated in industry and households model blocks can be fractionated according to the same principles. Applying separate fractionations to household and industry wastewater before combining both fluxes would for example allow for both wastewater flows to apply different distributions of particulate COD for the different ASM particulate variables, e.g. X_I, X_S and X_{BH}. The latter, combined with the pollutant flux variations of both pollutant sources, would contribute to a reduction of the correlation between these time series in the output of the influent model. It is up to the user to extend the model with a separate fractionation for each wastewater source, if this is desirable.

The principles of the influent fractionation implemented in the 'ASM-X fractionation' model block are as follows:

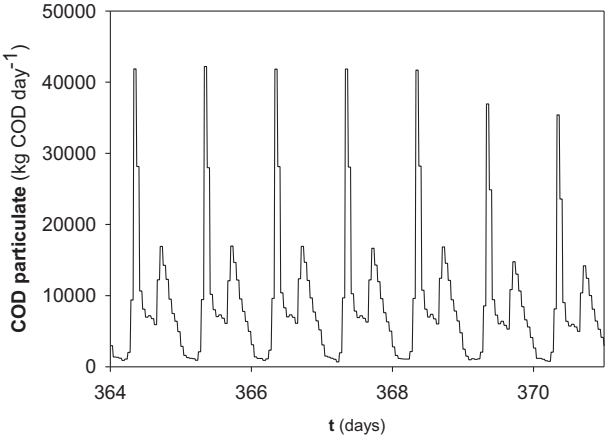

Fig. 8. Dynamic profiles resulting from the households for particulate COD.

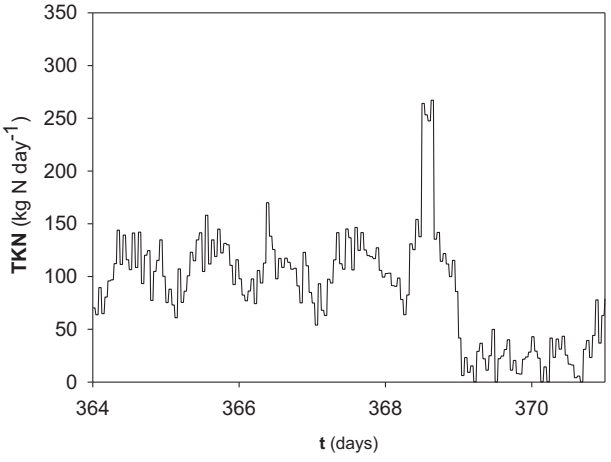

Fig. 9. Dynamic profiles resulting from the industry for TKN.

- The soluble inert material (S_I) concentration in all three models (ASM1, 2d and 3) is assumed to be constant, equal to 30 g COD m^{-3}. Only rainfall runoff from impermeable areas, corresponding to a fraction *aH* of the output of the 'Rain Generator' model block, can dilute S_I to concentrations that are significantly below 30 g COD m^{-3}, since this rain water flow is assumed not to contain any organic material. Of course, in case of insufficient soluble COD flux to provide 30 g COD m^{-3} as S_I for the dry weather flow rate, then the S_I concentration will also be below 30 g COD m^{-3}. In other words, the 'ASM-X fractionation' model block is not a COD source. The situation with reduced S_I concentrations due to low soluble COD fluxes would typically occur in low-loaded periods, such as weekends and holiday periods.
- The soluble organic substrate concentration (S_S) is obtained based on the difference between the soluble COD flux, which contains both contributions from households and from industry, and the S_I pollutant flux. If ASM2d is the selected model, it is assumed that S_S can be further sub-divided into 40% acetate (S_A) and 60% fermentable products (S_F), respectively.
- The concentration of autotrophic biomass (X_{BA}), polyphosphate accumulating organisms (X_{PAO}) and the different products resulting from cellular activity, e.g. X_P (in ASM1), X_{PP} and X_{PHA} (in ASM2d) and X_{STO} (in ASM3), are assumed to be zero in the influent.
- In all the three models the particulate COD is distributed over X_I, X_S and X_{BH}. The X_I fraction of the particulate COD flux, for example, is equal to $X_I/(X_I + X_S + X_{BH})$. For the BSM2 influent, this results in the following particulate COD fractions: 18.2% X_I, 71.8% X_S and 10. 0% X_{BH}.
- The S_O concentration is assumed to be zero (ASM1, 2d and 3).
- The S_{NO} concentration is zero (ASM1, 2d and 3).
- The S_{NH} concentration is calculated based on the NH_4^+-N flux (ASM1, 2d and 3).
- The S_{PO4} concentration is calculated based on the PO_4^{3-}-P flux (ASM2d).
- Just in ASM1, the organic N flux is obtained by subtracting the S_{NH} flux and the N flux corresponding to $i_{xB}\cdot X_{BH}$ and $i_{xI}\cdot X_I$ from the TKN flux. If the result of this calculation is positive, the organic N is distributed between S_{ND} and X_{ND}. The resulting S_{ND} and X_{ND} fractions for the BSM2 influent are 39.6% and 60.4%, respectively. If the result of this calculation of the organic N flux is negative, influent S_{ND} and X_{ND} are assumed to be zero.

As an example, Fig. 10 shows the soluble COD fractionation for ASM1. Thus, soluble COD is converted into ASM1 variables, in this case S_I (soluble inerts) and S_S (soluble organic substrate). It is also possible to see the dilution effect on S_I due to rainfall, for example on day 368. Additional simulation examples for different ASMs can be found in the supplemental information.

There is a possibility to add zero mean white noise to the concentrations of the state variables resulting from the 'ASM-X fractionation' model block (as in the previous sections). This is assumed to be necessary to avoid, for example, that X_I, X_S and X_{BH} concentrations in the generated influent are perfectly correlated to each other. The structure of the noise generators is similar to the other noise generators included in the influent model, i.e. noise is added using a random number generator (see supplemental information)

Table 1 summarizes the correlation coefficients among the different time series for the BSM2 influent example. The low correlation coefficient between S_I and the other time series is due to the fact that the S_I concentration is almost constant, apart from dilution effects in situations with rain events combined with low soluble COD input fluxes, whereas the dynamics of all other

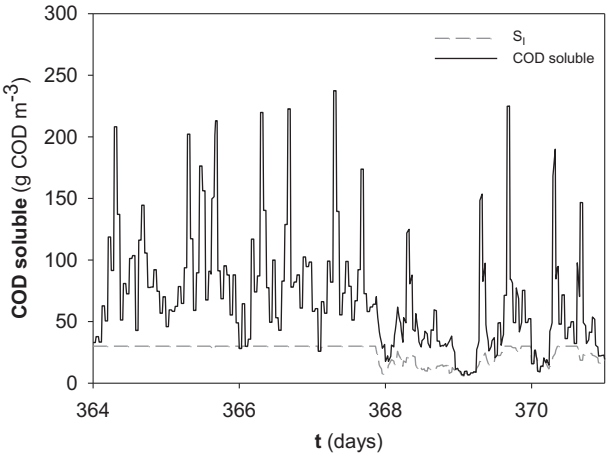

Fig. 10. Fractionation of the soluble COD for ASM1 and ASM3, i.e. COD soluble $- S_I = S_S$.

influent concentration time series will include pronounced diurnal effects. As expected, the correlation coefficients between X_I, X_{BH} and X_S concentrations are 1 if no random noise has been added. This is evident, since X_I, X_{BH} and X_S are obtained as constant fractions of the same quantity in the 'ASM-X fractionation' model block. The correlation coefficient between S_{ND} and X_{ND} is also 1 since S_{ND} and X_{ND} are obtained as constant fractions of the same quantity (remaining TKN after subtracting S_{NH} and N content of X_{BH} and X_I). Reducing the correlation between X_I, X_S and X_{BH} on the one hand, and between S_{ND} and X_{ND} on the other hand, is possible due to the activation of the implemented noise generators (see the results in bold summarized in Table 1).

4. Model for temperature variations

The temperature of the wastewater is added as an additional state variable in the model influent. The temperature profile includes a seasonal effect, i.e. there is a warm and a cold season. In addition, the model includes a daily temperature effect. Fig. 11 shows the dynamic profile of the wastewater temperature. Both daily and seasonal variations are modelled with a sinus function. The seasonal variation is modelled assuming an average temperature of 15 °C (*TBias*), an amplitude of 5 °C (*TAmp*) and a frequency of $2\pi/364$ rad d^{-1} (*TFreq*). A phase shift of $8.5\pi/364$ rad d^{-1} (*TPhase*) is applied, such that the maximum flow rate due to infiltration matches with the lowest temperature, and vice versa (see also Fig. 4). The parameters of the daily temperature profile are tuned assuming a temperature amplitude of 0.5 °C (*TdAmp*) and a frequency of 2π rad d^{-1} (*TdFreq*). As can be seen in Fig. 11 the minimum temperature of each day is situated around 8 o' clock in the morning. Again, it is up to the model user to either include noise or not.

Table 1
Correlation between the ASM1 state variables with (bold) and without including white noise.

	S_I	S_S	X_I	X_S	X_{BH}	S_{NH}	S_{ND}	X_{ND}	Q
S_I	*1.00*	**0.12**	**0.18**	**0.18**	**0.17**	**0.12**	**0.10**	**0.12**	**−0.18**
S_S	0.25	*1.00*	**0.34**	**0.34**	**0.33**	**0.87**	**0.71**	**0.83**	**0.10**
X_I	0.35	0.36	*1.00*	**0.96**	**0.92**	**0.27**	**0.21**	**0.25**	**−0.01**
X_S	0.35	0.36	1.00	*1.00*	**0.94**	**0.28**	**0.22**	**0.25**	**−0.01**
X_{BH}	0.35	0.36	1.00	1.00	*1.00*	**0.27**	**0.21**	**0.24**	**−0.02**
S_{NH}	0.28	0.96	0.32	0.32	0.32	*1.00*	**0.64**	**0.75**	**0.13**
S_{ND}	0.26	0.93	0.31	0.31	0.31	0.90	*1.00*	**0.69**	**0.10**
X_{ND}	0.26	0.93	0.31	0.31	0.31	0.90	1.00	*1.00*	**0.11**
Q	−0.33	0.11	−0.01	−0.01	−0.01	0.13	0.13	0.13	*1.00*

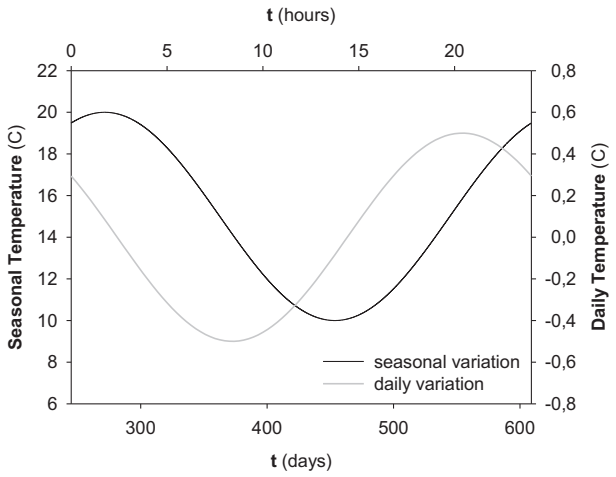

Fig. 11. Examples of the seasonal and daily temperature profiles of the influent model: note that the first day of the yearly profile corresponds to July 1st.

5. Transport model

The particulate pollutants of each model are subsequently passed through the 'First Flush Effect' model block, where first flush effects are mimicked as a function of flow rate, occurring for example during severe rain events following dry periods. Finally, the size of the sewer system can be incorporated in the influent dynamics as well: the larger the sewer system, the smoother the simulated diurnal flow rate and concentration variations.

5.1. 'First flush effect' (FFE) model block

The 'First flush effect' model block introduces additional influent dynamics corresponding to a flushing of the sewer system following a severe rain event. The model is based on the assumption that only part of the particulate material can settle in the sewer system, and is based on the following equations (Eqs. (3) and (4)):

$$\frac{dM}{dt} = M_{in} - M_{out} \tag{3}$$

$$\frac{dM}{dt} = Q_{in} \cdot TSS_{in} \cdot \left(1 - \frac{M}{M_{max}}\right) - \frac{(Q_{in})^n}{(Q_{lim})^n + (Q_{in})^n} \cdot M \cdot FF \tag{4}$$

This equation describes the accumulation of the total mass of solids in the sewer as a function of the flux of solids that is entering (M_{in}) and leaving (M_{out}) the system. Q_{in} represents the influent flow rate ($m^3\ d^{-1}$), TSS_{in} represents the SS concentration that forms the input to the model, M_{max} (kg) is the maximum amount of *TSS* that can be stored in the sewer system, Q_{lim} ($m^3\ d^{-1}$) is the flow rate limit triggering the first flush effect, and *FF* and *n* are dimensionless adjustable parameters to tune the desired strength of the first flush effect.

During dry weather conditions, *TSS* can accumulate in the sewer system (settling, build up of sediments at the bottom of the sewer pipes), i.e. the second term of Eq. (4) is equal to zero, meaning that the amount of sediments in the sewer will increase as long as the total amount of sediments stored in the sewer system is below M_{max}. For rain events, the function in the second part of the equation will induce the switching, i.e. from 0 to M·FF. Hence, the sudden increase of the flow rate should result in a washout of the sediments from the sewer system (first flush). No biological or chemical reactions are implemented in the 'First flush effect' model block.

Fig. 12 represents the dynamic behaviour induced by the first flush effect assuming that M_{max} is 1000 kg, Q_{lim} is 70,000 $m^3\ d^{-1}$ *FF* equals 500 and *n* is equal to 15. When the influent flow rate reaches the defined threshold rate (see influent flow profile in Fig. 7) and the sewer is full of sediments (Fig. 12a) the last term of the mass balance in Eq. (4) is active. As a result, the sewer is emptied of *TSS* and there is a sudden increase of the *TSS* load in the influent to the WWTP (Fig. 12b).

5.2. 'Sewer' model block

The conceptual principle of the 'Sewer' model block is based on a number of variable volume tank models, which are grouped in sub-systems. The principle of this variable volume tank can be described by a single differential equation, and the mass balance is written as:

$$\frac{dh_2}{dt} = \frac{1}{A_2}(Q_{in} - Q_{out}) \tag{5}$$

$$Q_{out} = C \cdot h^{1.5} \tag{6}$$

The parameter A_2 corresponds to the tank surface, *C* is a gain provided by the model user and h_2 is the water level in the tank. Note that the pollution model assumes that the pollution is discharged uniformly along the sewer. Thus, the dynamics of the output of this model will try to mimic the fact that part of the pollutants pass through the complete sewer system and some do not. The size of the catchment area is adjustable through the parameter *subareas*. It is up to the model user to select the appropriate model parameter value, providing the desirable dynamics (see also the detailed flow diagram in the supplemental information for further details).

When it comes to pollutant concentrations in the outlet of the sewer, the mass balance for one of the pollutant concentrations, e.g. X_I, in one of the variable volume tanks representing the sewer can be written:

$$\frac{d(X_I V)}{dt} = (X_{I,in} Q_{in} - X_{I,out} Q_{out}) \tag{7}$$

Fig. 13 shows an example of the smoothing effect of the sewer system on the flow rate dynamics, assuming that A_2 is 1100 m^2 and *C* is 150,000. When the value of the parameter *subareas* is increased, the sewer system is assumed to become larger and a more pronounced dispersion effect can be observed for the pollutant concentrations. The size of the flow rate peak in Fig. 13 can be influenced by modifying the data files that generate the households and industry profiles. Nevertheless it is important to highlight that there are special cases where even in large sewers, peak flow can be caused by pumping events and strongly determine the influent flow pattern (Ort et al., 2010). However, it is possible to also add such pump events in the proposed phenomenological models.

6. Discussion

For the BSM platform (see also www.benchmarkwwtp.org), the development of phenomenological models for the generation of dynamic influent flow rate and pollutant concentration time series was chosen on purpose for several reasons. First of all due to the relatively low complexity and the modularity of this approach (see Gernaey et al. (2006) for details). Moreover, another argument for generating influent data with a model was that the use of influent data collected from a full-scale system might result in a very

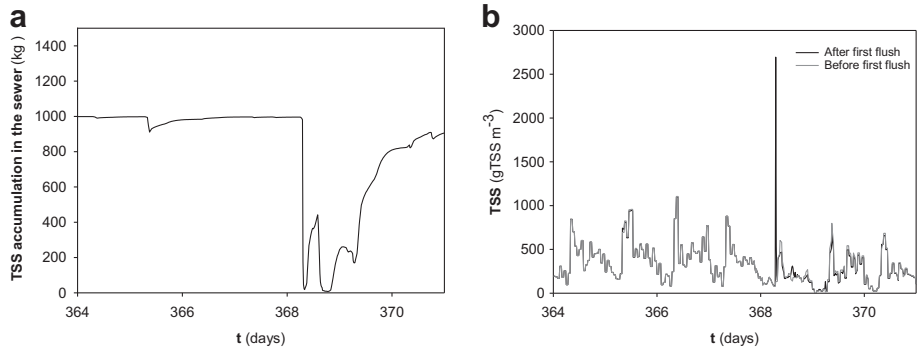

Fig. 12. Dynamics of the TSS in the sewer (a) and the consequences of the emptying of the sewer particulates storage on the TSS (b).

specific data set, where certain types of influent disturbances (for example seasonal flow rate variations) that were desired for the BSM2 influent data set would be lacking. Availability of an influent model allows the benchmark developers and other users to generate an influent file containing all the characteristics that are considered to be necessary for a thorough evaluation of the control systems in BSM2. Of course, this argument is equally valid for other plant models where simulation is applied for control strategy validation or scenario analysis.

The initial efforts of the WWTP benchmarking community in developing influent disturbance pollutant models has gradually found its way into several other modelling studies. Originally developed for the ASM1 model (Gernaey et al., 2011a), influent disturbance models have been reported that are compatible with ASM2d (Benedetti et al., 2008) and ASM3 (Ráduly et al., 2007) as well. Moreover, the flexibility of such influent models was demonstrated again by Benedetti et al. (2008), who combined a preliminary version of this approach with different rainfall time series to come up with a wide range of influent conditions including variation in weather scenarios (Alpine, Oceanic, Continental, Mediterranean), loading (ratio between households and industry) and holiday activities (tourism) to evaluate different plant designs.

Although the case study that illustrates this paper details the development of the well-known BSM2 influent, the presented set of model blocks is generic and therefore has a wide range of applications. Specifically, Flores-Alsina et al. (2011) show the preliminary results of a comprehensive global sensitivity analysis (GSA) examining the influence of the different influent model parameters on the model outputs. Also, the latter study gives an idea about the effort that the modeller has to invest in finding realistic values for a certain parameter if he/she does not have any

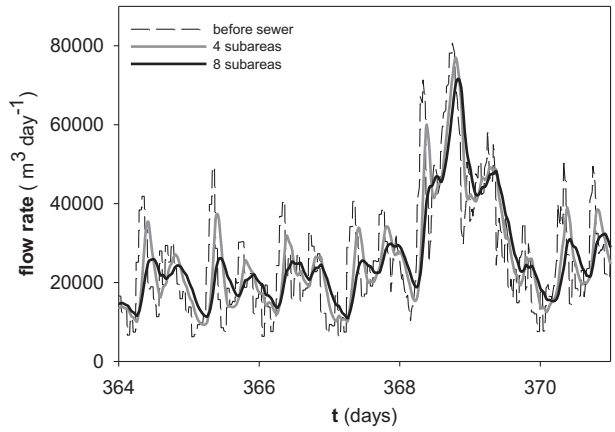

Fig. 13. Effect of the number of *subareas* used in the sewer model block on the flow rate dynamics.

measurements/expert knowledge available for the design study under investigation. Thus, in dry weather conditions the GSA indicated that the "Households" model block parameters (*QperPE*, *PE*) have the strongest influence on the dry weather influent flow rate. The "sewer" parameter *subareas* has a direct influence on the peaks. The seasonal correction factor (*InfBias*) and the soil (H_{inv}, K_{down}) related parameters affect (to a certain extent) average, minimum and maximum flow rate values. In wet-weather conditions, the "rain generator" model parameters (*LLrain*, *Qpermmm* and *aH*) have strong influence on wet-weather influent flow rate quantity and variability. In addition, rain related parameters (*LLrain*) clearly affect short-term evaluation criteria (hour, day peaks), but in the long run (month) households wastewater generation related parameters have a more important influence. Again, soil related parameters (H_{inv}, K_{down}) influence (like in dry weather conditions) minimum values at all time scales. These results match with the trial and error calibration procedure presented in Béraud et al. (2007), where *PE, QperPE* and *LLrain* were modified to generate continuous influent data from daily mean measurements available at industrial scale.

Flexibility is one of the major characteristics of the proposed approach. An obvious possibility not considered in the current study is the use of the proposed influent model for the modification of dynamic influent profiles and pollution loads, in order to allow simulations of a WWTP model to validate WWTP performance under a range of conditions. For example, in case it would be expected that the future load to a treatment plant might increase due to increased industrial activity, then this expected load increase can be incorporated in a dynamic influent scenario to investigate how the WWTP – and if needed also the receiving waters, the final destination of the WWTP effluent – would perform under such conditions.

The presented set of generic model blocks can also be extended to create other pollutant time series. Current research is, for example, focused on extending the presented phenomenological influent modelling approach with the generation of dynamic influent profiles of heavy metals and organic micro-pollutants, such as pharmaceuticals. An accumulating-releasing approach (like the one used to describe the first flush effect) is under construction for heavy metals, i.e. accumulation of metals on roads during dry weather periods and entering the sewer system after a rainy period. On the other hand, modelling the occurrence and variation of the influent load of hormones and antibiotics is expected to be done in a similar fashion as for the other pollutants currently included in the phenomenological influent model. The main difference in this case will be the user-defined pollutant generation profiles, which are in this case strongly correlated with the daily administration pattern of pharmaceuticals, their administration dosage and their expected human release. Finally it would also be feasible to describe catchments with significant pumping activities with the

proposed phenomenological influent modelling approach, overriding a"typical" diurnal flow pattern. A possible alternative would be to include an additional storage tank with minimum and maximum volumes in the sewer system model, where reaching the minimum and the maximum volumes would stop respectively start a pump to empty this storage tank.

Summarizing, the proposed collection of phenomenological models is presented as a useful tool to: a) fill gaps due to missing data; b) give higher frequency dynamics to data series consisting of daily average samples; c) create additional disturbance scenarios; or d) obtain validation of real influent data. The generic model blocks presented herein represent a valuable tool for scientists, process engineers and water professionals because they allow answering questions such as: What would be the effect of changing the rain regime or infiltration dynamics (due to for example climate change) on the wastewater influent profile? What is the effect of a population increase (changes in the number of population equivalents) or higher industrial presence (changing the kg pollutant d^{-1}) in the influent wastewater? Of course, when applying the proposed phenomenological models to a real plant, for example to generate dynamic influent profiles based on standard average influent data, the model user will normally have to invest some time to adjust the model parameters such that the generated influent dynamics fit the influent dynamics observed at the plant. In some cases, it might also be necessary to add additional model blocks in order to describe specific phenomena.

7. Conclusions

This paper introduces several generic model blocks to generate dynamic influent pollutant disturbance scenarios. The proposed phenomenological modelling approach allows to mimic influent flow rate and loading dynamics at the inlet of a WWTP (outlet of a sewer system), without the need of complex mechanistic drainage models. Based on the principles of parsimony, transparency and flexibility the influent model is constructed combining urban water system insights and experimental data/knowledge. Hence, it is now possible to reproduce some of the influent-related phenomena during the building of WWTP influent scenarios. The key findings of this research can be summarized in the following points:

- The influent flow rate model blocks allow reproducing flow rate dynamics including the daily, weekly, seasonal effects as well as irregular events, such as rainfall;
- The model blocks generate long time series describing the daily, weekly and yearly dynamics for each specific pollutant. In addition, it is possible to incorporate the effect of flow rate variations on the different pollution loads and concentrations, such as dilution by rainfall;
- The temperature model simulates seasonal and daily variations;
- The transport model modifies the flow rate and pollutants dynamics giving more or less pronounced peaks according to the size of the sewer system. In addition, the transport model has the possibility of simulating the first flush effect in the sewer.

The results reported in this paper show that the phenomenological modelling approach is an inexpensive and elegant way of generating dynamic WWTP influent data. Hence, it can be used to reduce the time consumption and the economic effort related to traditional sampling campaigns. Alternatively, based on available measurement campaign data the proposed phenomenological modelling approach can be used to extend those data, in order to obtain a more realistic and complete picture of how the WWTP performs under a wide range of disturbances.

Acknowledgements

Lorenzo Benedetti is a post-doctoral researcher of the Special Research Fund (BOF) of Ghent University.

Appendix. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.envsoft.2011.06.001.

References

Achleitner, S., Möderl, M., Rauch, W., 2007. CITY DRAIN © – An open source approach for simulation of integrated urban drainage systems. Environ. Model. Software 22 (8), 1184–1195.

Alex, J., Hetschel, M. and Ogurek, M., 2009. Simulation study with minimized additional data requirements to analyze control and operation of WWTP Dorsten. In: Proc. 10th IWA Conference on Instrumentation, Control and Automation, 14–17 June, Cairns, Australia.

Bechmann, H., Nielsen, M.K., Madsen, H., Poulsen, N.K., 1999. Grey-box modelling of pollutant loads from a sewer system. Urban Wat. 1, 71–78.

Benedetti, L., Bixio, D., Claeys, F., Vanrolleghem, P.A., 2008. Tools to support a model-based methodology for emission/immission and benefit/cost/risk analysis of wastewater treatment systems which considers uncertainties. Environ. Model. Software 23 (8), 1082–1091.

Benedetti, L., De Keyser, W., Nopens, I., Vanrolleghem, P.A., 2010. Probabilistic modelling and evaluation of waste water treatment plant upgrades in the EU water framework directive context. J. Hydroinformatics 12 (4), 380–395.

Butler, D., Friedler, E., Gatt, K., 1995. Characterising the quantity and quality of domestic wastewater inflows. Wat. Sci. Tech. 31 (7), 13–24.

Béraud, B., Steyer, J.P., Lemoine, C., Gernaey, K.V. and Latrille E., 2007. Model-based generation of continuous influent data from daily mean measurements available at industrial scale. In: Proc. 3rd International IWA Conference on Automation in Water Quality Monitoring (AutMoNet2007), 5–7 September, Gent, Belgium.

Carstensen, J., Nielsen, M.K., Strandbæk, H., 1998. Prediction of hydraulic load for urban storm control of a municipal WWT plant. Wat. Sci. Tech. 37 (12), 363–370.

Copp, J.B. (Ed.), 2002. The COST Simulation Benchmark: Description and Simulator Manual. Office for Official Publications of the European Community, Luxembourg, ISBN 92-894-1658-0.

De Keyser, W., Gevaert, V., Verdonck, F., De Baets, B., Benedetti, L., 2010. An emission time series generator for pollutant release modelling in urban areas. Environ. Model. Software 25 (4), 554–561.

Devesa, F., Comas, J., Turon, C., Freixó, A., Carrasco, F., Poch, M., 2010. Scenario analysis for the role of sanitation infrastructures in integrated urban wastewater management. Environ. Model. Software 24 (3), 371–380.

Elliott, A.H., Trowsdale, S.A., 2007. A review of models for low impact urban stormwater drainage. Environ. Model. Software 22 (3), 394–405.

Flores-Alsina, X., Gernaey, K.V and Jeppsson U, 2011. Global sensitivity analysis of a phenomenological wastewater treatment plant influent generator. Accepted for the IWA 8th International Symposium on Systems Analysis and Integrated Assessment (WaterMatex2011), San Sebastian, Spain, June 20-22, 2011.

Flores, X., Poch, M., Rodríguez-Roda, I., Bañares-Alcántara, R., Jiménez, L., 2007. Systematic procedure to handle critical decisions during the conceptual design of activated sludge systems. Ind. Eng. Chem. Res. 46 (17), 5600–5613.

Fu, G., Butler, D., Khu, S.T., 2008. Multiple objective optimal control of integrated urban wastewater systems. Environ. Model. Software 23 (2), 225–234.

Gernaey, K.V., Flores-Alsina, X., Rosen, C., Benedetti, L. and Jeppsson U, 2011a. BSM2: A model for dynamic influent data generation. Technical Report no. 8, IWA Task Group on Benchmarking of Control Strategies for Wastewater Treatment Plants.

Gernaey, K.V., Jeppsson, U., Vanrolleghem, P.A., Copp, J.B. and Steyer J.-P. (eds), 2011b. Benchmarking of control strategies for wastewater treatment plants. IWA Scientific and Technical Report, IWA Publishing, London, UK, in preparation.

Gernaey, K.V., Rosen, C., Jeppsson, U., 2006. WWTP dynamic disturbance modelling – an essential module for long-term benchmarking development. Wat. Sci. Tech. 53 (4–5), 225–234.

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

Gernaey, K.V., van Loosdrecht, M.C.M., Henze, M., Lind, M., Jørgensen, S.B., 2004. Activated sludge wastewater treatment plant modelling and simulation: state of the art. Environ. Model. Software 19 (9), 763–783.

Gillot, S., Ohtsuki, T., Rieger, L., Shaw, A., Takács, I., Winkler, S., 2009. Development of a unified protocol for good modeling practice in activated sludge modelling. Influents 4, 70–72.

Henze, M., Gujer, W., Mino, T. and van Loosdrecht, M.C.M, 2000. Activated sludge models ASM1, ASM2, ASM2d and ASM3. IWA Scientific and Technical Report No. 9 IWA, London.

Hernebring, C., Jonsson, L.-E., Thorećn, U.-B., Møller, A., 2002. Dynamic online sewer modelling in Helsingborg. Wat. Sci. Tech. 45 (4–5), 429–436.
Hulsbeek, J.J.W., Kruit, J., Roeleveld, P.J., van Loosdrecht, M.C.M., 2002. A practical protocol for dynamic modelling of activated sludge systems. Wat. Sci. Tech. 45 (6), 127–136.
Hug, T., Benedetti, L., Hall, E.R., Johnson, B.R., Morgenroth, E., Nopens, I., Rieger, L., Shaw, A., Vanrolleghem, P.A., 2009. Wastewater treatment models in teaching and training: the mismatch between education and requirements for jobs. Wat. Sci. Tech. 59 (4), 745–753.
Jeppsson, U., Pons, M.N., Nopens, I., Alex, J., Copp, J.B., Gernaey, K.V., Rosen, C., Steyer, J.P., Vanrolleghem, P.A., 2007. Benchmark simulation model No 2 – general protocol and exploratory case studies. Wat. Sci. Tech. 56 (8), 287–295.
Langergraber, G., Alex, J., Weissenbacher, N., Woener, D., Ahnert, M., Frehman, T., Halft, N., Hobus, I., Plattes, M., Spering, V., Winkler, S., 2008. Generation of diurnal variation for influent data for dynamic simulation. Wat. Sci. Tech. 50 (9), 1483–1486.
Langergraber, G., Rieger, L., Winkler, S., Alex, J., Wiese, J., Owerdieck, C., Ahnert, M., Simon, J., Maurer, M., 2004. A guideline for simulation studies of wastewater treatment plants. Wat. Sci. Tech. 50 (7), 131–138.
Lindblom, E., Gernaey, K.V., Henze, M., Mikkelsen, P.S., 2006. Integrated modeling of two xenobiotic organic compounds. Wat. Sci. Tech. 54 (6–7), 213–221.
Melcer, H., Dold, P.L., Jones, R.M., Bye, C.M., Takacs, I., Stensel, H.D., Wilson, A.W., Sun, P., Bury, S., 2003. Methods for Wastewater Characterization in Activated Sludge Modeling. Water Environment Research Foundation (WERF), Alexandria, VA, USA.
Ort, C., Lawrence, M.G., Rieckermann, J., Joss, A., 2010. Sampling for pharmaceuticals and personal care products (PPCPs) and illicit drugs in wastewater systems: are your conclusions valid? A critical review. Environ. Sci. Technol. 44 (16), 6024–6035.
Ort, C., Schaffner, C., Giger, W., Gujer, W., 2005. Modeling stochastic load variations in sewer systems. Wat. Sci. Tech. 52 (5), 113–122.
Petersen, B., Gernaey, K., Henze, M., Vanrolleghem, P.A., 2002. Evaluation of an ASM1 model calibration procedure on a municipal-industrial wastewater treatment plant. J. Hydroinformatics 4, 15–38.
Ráduly, B., Gernaey, K.V., Capodaglio, A.G., Mikkelsen, P.S., Henze, M., 2007. Artificial neural networks for rapid WWTP performance evaluation: methodology and case study. Environ. Model. Software 22 (8), 1208–1216.
Rieger, L., Thomann, M., Joss, A., Gujer, W., Siegrist, H., 2004. Computer-aided monitoring and operation of continuous measuring devices. Wat. Sci. Tech. 50 (11), 31–39.
Rivas, A., Irizar, I., Ayesa, E., 2008. Model-based optimisation of wastewater treatment plants design. Environ. Model. Software 23 (4), 435–450.
Rodriguez-Roda, I., Sanchez-Marre, M., Comas, J., Baeza, J., Colprim, J., Lafuente, J., Cortes, U., Poch, M., 2002. A hybrid supervisory system to support WWTP operation: implementation and validation. Wat. Sci. Tech. 45 (4–5), 289–297.
Rosen, C., Jeppsson, U., Rieger, L., Vanrolleghem, P.A., 2008. Adding realism to simulated sensors and actuators. Wat. Sci. Tech. 57 (3), 337–344.
Rosen, C., Jeppsson, U., Vanrolleghem, P.A., 2004. Towards a common benchmark for long-term process control and monitoring performance evaluation. Wat. Sci. Tech. 50 (11), 41–49.