

Review

Analysing, completing, and generating influent data for WWTP modelling: A critical review

Cristina Martin ^{a, b, *}, Peter A. Vanrolleghem ^a^a modelEAU, Département de génie civil et de génie des eaux, Université Laval, 1065 Avenue de la Médecine, Québec, QC, G1V 0A6, Canada^b Deusto Institute of Technology – DeustoTech Energy, University of Deusto, Avda. Universidades 24, 48007, Bilbao, Spain

ARTICLE INFO

Article history:

Received 5 July 2013

Received in revised form

4 March 2014

Accepted 7 May 2014

Available online 5 July 2014

Keywords:

Activated sludge models

Influent generator

Influent disturbance model

Phenomenological model

Probabilistic design

Uncertainty assessment

ABSTRACT

This paper makes a critical review of the available techniques for analysing, completing and generating influent data for WWTP modelling. The solutions found in literature are classified according to three different situations from engineering practice: 1) completing an incomplete dataset about the quantity and quality of the influent wastewater; 2) translating the common quality measurements (COD, TSS, TKN, etc.) into the ASM family components (*fractionation problem*); 3) characterising the uncertainty in the quality and quantity of the influent wastewater. In the first case (Situation 1), generators based on Fourier models are very useful to describe the daily and weekly wastewater patterns. Another specially promising solution is related to the construction of phenomenological models that provide wastewater influent profiles in accordance with data about the catchment properties (number of inhabitant equivalents, sewer network, type of industries, rainfall and temperature profiles, etc.). This option has the advantage that using hypothetical catchment characteristics (other climate, sewer network, etc.) the modeller is able to extrapolate and generate influent data for WWTPs in other scenarios. With a much lower modelling effort, the generators based on the use of databases can provide realistic influent profiles based on the patterns observed. With regard to the influent characterisation (Situation 2), the WWTP modelling protocols summarise well established methodologies to translate the common measurements (COD, TSS, TKN, etc.) into ASM family components. Finally, some statistical models based on autoregressive functions are suitable to represent the uncertainty involved in influent data profiles (Situation 3). However, more fundamental research should be carried out to model the uncertainty involved in the underlying mechanisms related to the wastewater generation (rainfall profiles, household and industries pollutant discharges, assumed daily and weekly patterns, etc.).

© 2014 Published by Elsevier Ltd.

1. Introduction

Wastewater Treatment Plants (WWTPs) are urban infrastructures that reproduce the biodegradation processes that naturally occur in rivers in an intensified manner. The high capital and operational costs associated to these facilities (Vanrolleghem et al., 1996; Liu et al., 2011; Rodriguez-Garcia et al., 2011) have fostered the use of simulation models to optimise their performance, and in this sense, the ASM family models (Henze et al., 2000) have become a standard. In the last years, many publications have illustrated the usefulness of simulation models for WWTP design (Bixio et al., 2002; Benedetti et al., 2010; Rivas et al.,

2008); operation (Ostace et al., 2011) and control (Ayesa et al., 2006; Nopens et al., 2010; Flores-Alsina et al., 2008; Yong et al., 2006).

One of the main limitations for a more widespread utilization of WWTP models is generally related to the scarce data sets measured at the inlet of the WWTP (Rieger et al., 2010a). The high cost (both in terms of workload and financial resources) related to experimental collection of an extended dynamic influent dataset is one of the main reasons. The use of on-line sensors still remains complicated, since the sticky materials of raw wastewater and the heavy deposit of pollutants make their maintenance cost considerable. Moreover, in view of risk analysis, models are generally used to predict the behaviour of the system under some hypothetical conditions (population growth; strong rain events; uncontrolled spills, etc.) for which real data might not exist.

Influent data for municipal WWTP modelling consist of time series data of the flow and concentrations of the water quality

* Corresponding author. Energy Unit, DeustoTech, Universidad de Deusto, Facultad de Ingeniería, Av. Universidades 24, 48007, Bilbao, Spain. Tel.: +34 944 139 073x2045.

E-mail address: cristina.andonegui@deusto.es (C. Martin).

parameters (COD, TKN, TSS, BOD, NH₄, NO₃, TP, PO₄, etc.). These profiles depend on many factors: size of the catchment, type of the sewer system, number of person equivalents, industrial discharges, soil type, rainfall patterns, temperature, etc (Butler et al., 1995; Bott and Parker, 2010; Schilperoort, 2011). The complexity of the wastewater generation is so big that there is still no clear relationship between the generating mechanisms and the expected water quality profiles. The problem becomes even more difficult if the objective is a modelling exercise because the characterisation of the aforementioned lab measurements in terms of the components of the ASM family models is not straightforward. This paper gathers the solutions found in literature to analyse, complete, and generate influent data for WWTP modelling.

2. Engineering practice necessities and classification followed

This paper makes a critical review of the available techniques for WWTP influent data analysis and generation. It is intended to guide engineers in their particular situation to get the most of their experimental data (if any) in view of performing simulation studies of real or hypothetical WWTPs. The solutions found in the literature are classified according to three situations from engineering practice as shown in Box 1.

The solutions in Situation 1 help to understand the information in the available data and to complete it on the basis of the patterns expected or of characteristics of the catchment area (temperature, population activity, etc.). The methods in Situation 2 include several protocols and mathematical techniques able to translate the generally available experimental measurements (COD, TSS, TKN, etc.) into ASM family components (S_s, S_i, X_s, NH₄, etc.). In Situation 3, we cover the scientific effort made to characterise the uncertainty of influent data profiles (Friedler and Butler, 1996) and the available techniques to generate other similar data series in a stochastic manner (Benchmann et al., 1999; Rousseau et al., 2001). Situation 1 and Situation 2 appear in any simulation study of a WWTP. Situation 3 follows the other two for simulation-based uncertainty analysis (usually performed by Monte Carlo techniques).

Although most of these methods are intended for simulation analysis, some of the protocols for wastewater fractionation (Situation 2) are encouraged to be followed even if no simulation analysis is scheduled. For example, the STOWA protocol (Hulsbeek

Box 1

Classification of solutions found in the literature to analyse, complete, and generate influent data for WWTP modelling: Three situations from engineering practice.

Situation 1

An engineer has an incomplete dataset (or no data) about the quantity and quality of the influent wastewater. She has some ideas about how the influent of the plant should be. She might also have some data about the catchment area.

Situation 2

An engineer has information enough about the quantity and quality of the influent wastewater. She would like to translate that into some ASM family model state variables. This is the so-called *characterisation* problem.

Situation 3

The engineer has information enough about the quantity and quality of the influent wastewater. However, that is only a single realisation of the problem. She would like to feature the uncertainty around these data and to generate other similar influent profiles.

et al., 2002) presents a wastewater characterisation method that is very helpful for the daily WWTP operation (Roeleveld and van Loosdrecht, 2002) i.e., the authors explain that the soluble COD available will be related to the capacity of the plant for high nitrogen and biological phosphorus removal.

Sections 3–5 of the manuscript describe the solutions found in the literature related to Situation 1, Situation 2 and Situation 3 respectively. At the end of each section, Box 2–4 provide a list of the most frequently asked questions about the solutions. These boxes, although helpful and illustrative should not draw attention from the text since the manuscript includes more solutions and much more detailed explanations about the available methodologies and their advantages. Finally, Section 6 presents the methods presented in literature for calibration, validation and model performance evaluation of influent generation models; Section 7 discusses the models and methods presented, and Section 8 summarizes the main conclusions.

3. Solutions found for situation 1: engineer has an incomplete dataset about the quantity and quality of the influent wastewater

Under Situation 1, we gather the techniques found in literature to analyse, complete, and generate influent data for WWTP modelling. Three groups of solutions have been identified. A first group of solutions consists of methods that are based on the construction of databases (Section 3.1). From these datasets general patterns are identified and used to complete or generate influent profiles of similar characteristics. A second group of solutions is based on the use of harmonic functions (Section 3.2). Under dry weather conditions, the influent wastewater concentrations follow stable trajectories related to daily and weekly patterns. In a third group, we describe those methods that provide a sufficient basis to advance towards the modelling of the wastewater generation mechanisms (phenomenological models) (Section 3.3). This classification follows an increasing modelling/conceptualisation effort from the data driven methods (solutions in Section 3.1) passing by very simple models based on harmonic functions (solutions in Section 3.2) to end up with the phenomenological models (Section 3.3).

3.1. Solutions based on the construction and interpretation of databases

The most intuitive manner of generating influent data for WWTP modelling consists of gathering experimental data, understanding its main patterns and completing or generating new data on that basis. With regard to wastewater quality and quantity characterization many authors have gathered and analysed data at the very source (Butler et al., 1995; Almeida et al., 1999; Friedler, 2004); in the sewer system (Benchmann et al., 1999; Ort et al., 2005); and at the entrance of WWTPs (Günther, 2000; Siegrist et al., 1999; Petersen et al., 2002; Bott and Parker, 2010; Schilperoort, 2011). In this section we only cover the datasets compiled in view of creating influent generators.

From a very practical perspective, Devisscher et al. (2006) proposed an influent generator as part of a methodology to assess the benefits of using advanced control in wastewater treatment plants. The generator is able to provide synthetic data by interpolating values within the same time window of the available data. With that in mind, they define a procedure in which correction factors are applied to represent the weekend and first flush events; available quality data are aggregated in the form of seasonal averages and redistributed on a day-to-day basis using a normal distribution; a daily pattern is applied to the flow rate; and (if

Box 2

Frequently Asked Questions in Situation 1, i.e., an engineer wants to complete a dataset about the quantity and quality of the influent wastewater.

a) Is there any method to interpolate values so as to increase the frequency of some given WWTP influent data profiles?

Yes, the method of [Devisscher et al. \(2006\)](#) based on the use of databases, or the ones of [Langergraber et al. \(2008\)](#) or [Manina et al. \(2011\)](#) based on the use of harmonic functions can help to complete some scarce datasets. Also, the phenomenological model of [Gernaey et al. \(2011\)](#) has been used with this propose ([Flores-Alsina et al., 2013a; b](#)).

b) Is there any tool to derive influent WWTP data given some properties about the catchment area or the distribution of emission sources (households, industries, restaurants, dentists ...)?

Yes, the generator of [Gernaey et al. \(2011\)](#) provides influent data for urban WWTPs given the population and catchment area characteristics (PE, industrial load, sewer length, rain and temperature profiles, etc.). The generator of [De Keyser et al. \(2010\)](#) provides emission profiles expected from a given urban population (PE for households, road transport, restaurants, etc.). The former describes the effect of soil, first flush, sewer system, etc.; while the later is only concerned about the very source emission patterns providing profiles that can be used as inputs for sewer, WWTP, or surface water models.

c) Are there some benchmark influent data profiles or simple models to generate influent data for urban WWTPs?

Influent profiles for a WWTP of 100,000 PE are available at <http://www.benchmarkwwtp.org/> including dry, rain and storm weather conditions.

The models of [Langergraber et al. \(2008\)](#) or [Mannina et al. \(2011\)](#) (based on harmonic functions) provide influent profiles for dry weather conditions. The former ([Langergraber et al., 2008](#)) is accompanied by a set of parameters among which the modeller can select those corresponding to the desired size of tin WWTP under study.

d) Which would be the main references or tools to improve the understanding of the generating mechanisms in a given catchment area and their relationship with the WWTP influent data?

Interesting results can be found in the research work conducted by [Butler et al. \(1995\)](#), [Friedler et al. \(1996a,b\)](#) or [Almeida et al. \(1999\)](#) about domestic wastewater generation. This research has been followed by the EU project ScorePP (www.scorepp.eu) ([De Keyser et al., 2010](#)) and the work of [Ort et al. \(2005\)](#) analysing the most general patterns of daily, weekly, and yearly emissions. Another good manner of proceed is to use the phenomenological model of [Gernaey et al. \(2011\)](#) and calibrate it for some given experimental data (as shown by [Flores-Alsina et al. \(2013\)](#)); or to study the very recent modelling proposal of [Talebizadeh et al. \(2014\)](#).

exhaustive literature review ([Gevaert et al., 2008](#)) and typical release patterns are defined based on expert knowledge. The generator is based on the development of a database with a structure and quantitative description of emission sources and their patterns on a typical daily, weekly, and yearly basis. The generator is able to provide profiles of Priority Pollutants (PP), generic pollutants (nutrients, suspended solids and COD), and wastewater flow rates by using: a phenomenological description of the PP release based on individual events which are modelled as random peaks that become bell-shaped by being transported in the sewer network ([Ort et al., 2005](#)); typical patterns identified in literature ([Gevaert et al., 2008](#)); a dry weather flow generator (based on [Gernaey et al., 2006](#)); and adding two different types of stochasticity (Section 5). [Fig. 2](#) shows an example of synthetically generated daily profiles for generic pollutant fluxes.

The generator of [Devisscher et al. \(2006\)](#) interpolates influent data by increasing the data frequency for given dataset (and given properties of the catchment area). It does not, therefore, have any capacity to adapt the output to different catchment properties (for example, different person equivalents, soil properties, etc.). On the contrary, the generator of [De Keyser et al. \(2010\)](#) generates dynamic influent data profiles by aggregating punctual emissions. The characteristics of these emissions are coherent with the person equivalents in the catchment area, weekly profiles, daily maximum peaks, etc. as specified by the user. The model is based on emission patterns and does not consider (it was never its objective) properties related to the climate, soil or sewer system.

Although both approaches propose the use of datasets to generate influent wastewater data, the contexts and scopes are very different. The approach of [Devisscher et al. \(2006\)](#) was motivated by the effort of the Aquafin wastewater utility for managing, monitoring and controlling more than 100 WWTPs in Flanders (Belgium). In this case, the generating approach acts as an ad hoc interpolator for wastewater quality parameters. The generator of [De Keyser et al. \(2010\)](#) is motivated by a European research project that looks for international consensus for controlling the release of priority and common pollutants. In this case, the dataset is based on expert knowledge and bibliographic data. The generating procedure is much more descriptive and can be used even for simulation studies where no experimental data is available.

3.2. Solutions based on the use of harmonic functions

A pioneering approach of using harmonic functions in the context of WWTP modelling was proposed by [Carstensen et al. \(1998\)](#). They used a second order Fourier approximation to feature the daily profiles of the influent flow and to forecast the hydraulic load one hour ahead. Based on a similar idea, other authors ([Langergraber et al., 2008](#); [Mannina et al., 2011](#)) have proposed the use of Fourier equations to feature the wastewater quantity and quality at the inlet of a WWTP.

The Fourier-based models are generally used to describe the patterns of the wastewater under dry weather conditions. These models are very useful to interpolate hourly values given average daily data ([Mannina et al., 2011](#)) or simply to complete the available data ([Langergraber et al., 2008](#)). The influent model proposed by [Mannina et al. \(2011\)](#) (already outlined by [Mannina and Viviani, 2009](#)) introduces the moving average concept to account for inter-daily variation. The generator provides hourly values for 43 days (based on 43-day average daily data in their case study) by introducing a moving average value that draws a straight line between two given daily averages on top of which the third order Fourier-based dynamics are added. In a simpler Fourier modelling framework, [Langergraber et al. \(2008\)](#) assumed that the overall wastewater results are the sum of different streams (infiltration

necessary) daily flow rate values are generated from Poisson distributions. This pragmatic method results in profiles comparable with the available data and able to provide reliable simulation responses. As an example, [Fig. 1](#) shows simulated and real data of ammonium during a 365-day period. The same approach was adopted by [Benedetti et al. \(2008\)](#) in a set of tools to assess WWTP performance in a probabilistic way. The method was also employed as an influent generator ([Gevaert et al., 2009](#)) in a dynamic pollutant fate model.

Another influent generator was proposed by [De Keyser et al. \(2010\)](#) within a framework of a pollutants release study in urban areas. In this case, the experimental data are gathered from an

water, urine with flush water and domestic wastewater without urine) which can be modelled as second order Fourier models. The authors investigated the dependence of the Fourier model parameters on the WWTP size by analysing data from 19 urban WWTPs of different sizes, and they provided approximate values of these parameters for any other plant under study. The model parameters were estimated by using the minimum and maximum values of flow and concentration and their timings as boundary conditions. Fig. 3 shows a comparison between the modelled and measured (a) daily flow profiles and (b) daily concentrations of COD and TKN.

Spering et al. (2008) used the same approach as Langergraber et al. (2008) to perform dynamic simulation studies during the design of new WWTPs where no dynamic influent data is available. Alex et al. (2007) demonstrate the validity of these designs since they are in agreement with the design solutions proposed by stationary design rules (ATV-DVWK-A (ATV, 2000)). The same tool (Langergraber et al., 2008) is also recommended to perform simulation studies for control and operation of WWTPs (Alex et al., 2009). The authors claim that the method is especially useful when limited budget is available because meaningful results can be obtained with very little measurement effort.

3.3. Solutions based on the analysis of the wastewater generation mechanisms

The characterisation of the quantity and quality of the wastewater entering a WWTP involves phenomena of very different nature: rain, soil type (permeable or impermeable), invert level, temperature, population activity, industrial discharges, type and length of the sewer system, retention tanks, etc (Butler et al., 1995; Bott and Parker, 2010; Schilperoord, 2011). Trying to define a mechanistic model able to describe all these elements is nearly utopian. A more feasible (although still ambitious) objective consists of developing a phenomenological model able to represent the patterns observed without aiming for a detailed characterisation of the underlying generating mechanisms. Such effort should continuously move forward step wisely, incorporating basic knowledge about the causes and dynamics of the driving forces and behaviours.

The household activity constitutes an important contribution to the wastewater received by WWTPs. In this respect, an important analysis of the domestic wastewater was carried out in England in the nineties (Butler, 1991, 1993; Butler et al., 1995; Friedler et al., 1996a,b; Almeida, 1999). The effort culminated in the presentation of a general methodology (Almeida et al., 1999) (previously outlined by Butler et al., 1995) to derive patterns associated with the quantity and quality of domestic wastewater and to assess their uncertainty. The analysis is based on the results of three surveys carried out in the south of England (Butler, 1991, 1993; Butler et al., 1995; Friedler et al., 1996a,b) and a fourth one in Malta (Butler and Gatt, 1996). The main conclusions are:

- Daily flow patterns become apparent with a high morning peak and two lower evening peaks corresponding to dinner and bed times. Slight changes might be expected depending on the dwelling occupancy (Butler, 1993) and the geographical region (in Malta later than in England according to Butler and Gatt (1996)). Note that another study carried out in a large Brazilian city demonstrated that social strata too may influence the patterns (Campos and von Sperling, 1996). For instance, the rich areas of cities in developing countries seem to produce the wastewater load later during the weekends (two hours in the Brazilian case study) while the poorer ones maintain the same peak hours as the week-days.

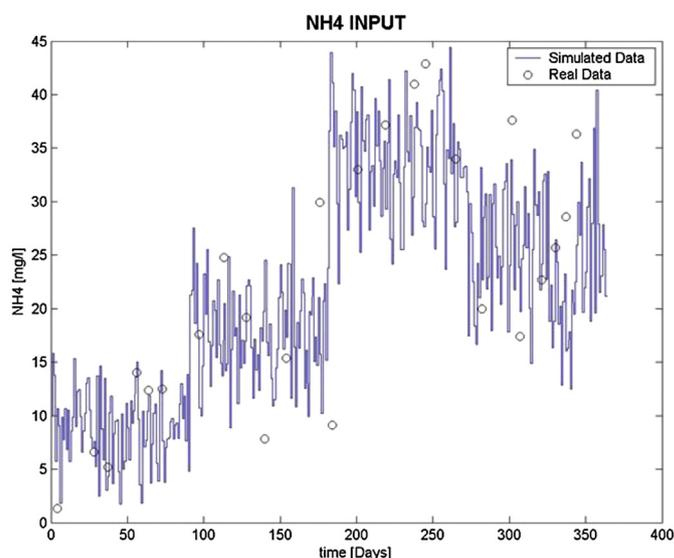


Fig. 1. (from Devisscher et al., 2006): Real and simulated ammonium data for a 365-day period.

- The daily patterns of the flow and loads (COD, TSS, NO_3 and PO_4) follow similar patterns while those of the concentrations present much lower variability (Almeida et al., 1999). The flow and concentration data seem to be quite independent (Friedler and Butler, 1996)
- The WC generates nearly 40% of the domestic wastewater. It is the major contributor to all pollutant concentrations and especially to ammonia (see Friedler et al., 1996a,b)
- The resulting values of the pollutant concentrations at the domestic source are higher than those reported at the entrance of WWTPs (Henze et al., 1995). This result outlines the importance of the dilution and biodegradation processes in the sewer systems (Butler et al., 1995)
- The uncertainties of the pollutant concentrations are featured by skewed distributions that represent not only the uncertainty of their magnitude but also the uncertainty about the timing of their peaks. Given the non-gaussian nature of these uncertainties, the evaluation of their combined effect should be assessed by numerical techniques, i.e. Monte Carlo simulation (Friedler and Butler, 1996)

Fig. 4 shows the relatively standard pattern observed in domestic daily load (COD, PO_4 , TSS, NH_3 and NO_3) and the proportions produced per household appliance.

The effort of such an exhaustive analysis (where different appliances and modes of use are distinguished) was intended to gain knowledge about the wastewater generation; and therefore to improve the operation and control of WWTPs. However, the connection of this methodology with the existing WWTP models was not investigated. Note that, the daily patterns (of flow and load) observed at the very source of the domestic generation may no longer be observable at the entrance of the WWTP due to intermediate sewer retention tank storage, different pumping regimes, etc (Ort et al., 2010).

The use of simulation models to find improved operation and control of WWTPs has been largely discussed (Flores-Alsina et al., 2008; Benedetti et al., 2010; Guerrero et al., 2011). Good examples of that work are the successful Benchmark Simulation Model No.1 (BSM1) (Copp, 2002) and Benchmark Simulation Model No.2 (BSM2) (Jeppsson et al., 2007) and their numerous applications (over 300 scientific publications) (Jeppsson et al., 2013). While the

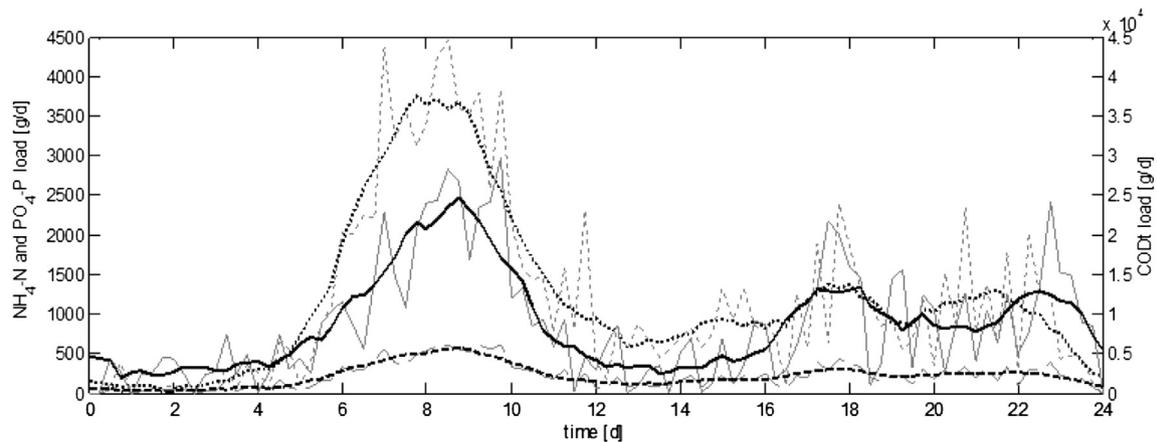


Fig. 2. (from De Keyser et al., 2010): Twenty-four hours of aggregated time series with temporal resolution of 15 min (jagged lines) and 2 h moving averages (smooth lines) for the generic pollutant fluxes of CODt (dotted lines, right axis), $\text{NH}_4\text{-N}$ (solid lines, left axis) and $\text{PO}_4\text{-P}$ (dashed lines, left axis) produced in a catchment with 100 PE household effluents.

BSM1 represents the water line (biochemical reactors and secondary clarifier), the BSM2 also includes the sludge line (thickener, anaerobic digester, dewatering unit and storage tank). The protocols define influent characteristics featuring a load of around 100,000 person equivalents or PE (80,000 from households and 20,000 from industrial origin) and fix performance evaluation protocols. The BSM1 simulation platform includes three influent files (Fig. 5), each of them representing 14 days of dynamic influent flow and load featuring: dry weather conditions, one rain event, and two storm events respectively. The BSM2 proposes to evaluate the model performance over one year of dynamic simulation, so that the effect of the temperature and the slow dynamics of the anaerobic digester can be investigated.

The phenomenological model of Gernaey et al. (2011) was born out of the BSM1_LT (Rosen et al., 2004) and BSM2 initiatives (Jeppsson et al., 2007). It is a flexible tool that describes the wastewater generation under typical urban conditions. The main objective of the phenomenological generator (Gernaey et al., 2011) is to create model-based influent data for the BSM2 platform in view of process control studies. Nevertheless, any modeller is invited to modify the original structure and include the specific characteristics of the particular case (catchment area, industrial discharges, etc.). The tool is implemented as 'open source' freely distributed on request. The first release of the phenomenological model (Gernaey et al., 2005, 2006) already provided influent data with the main characteristics of the most recent version (Gernaey et al., 2011). The main differences between the two of them are: a) the nitrogen load to the plant was reduced (Nopens et al., 2010) since it was discovered during BSM2 simulations that the BSM2

plant was chronically overloaded; b) model verification was conducted by a second, totally independent implementation (Gernaey et al., 2011) leading to full confidence in the results of the generator.

The phenomenological model (Gernaey et al., 2005, 2006, 2011) features the influent of a WWTP located in the Northern hemisphere and provides annual dynamic influent flow rate, pollutant concentrations and temperature profiles starting from the 1st of July. The pollutant concentrations can be expressed in terms of the state variables of ASM1, ASM2d or ASM3 models (Henze et al., 2000), as desired by the user. In accordance with the BSM2 protocol the dynamic profiles feature 609 days from which only the last 365 would be used for the evaluation of the simulation results. The generator is implemented in Matlab-Simulink by defining different interconnected modules. As an example, Fig. 6 shows the architecture of the flow generator (figure from Flores-Alsina et al., 2012) and the model block describing the wastewater flow coming from households (figure from Gernaey et al., 2005):

The main characteristics of the modules implemented in the phenomenological model (Gernaey et al., 2011) are:

- **Rain generation module:** It is an ingenious tool that transforms a random number (stochastic variable) into a rain event, the peak value of which is smoothed down by an exponential function until the next rain event appears. The module allows generating different rainfall patterns in relation with different climate conditions. The number of rain events and their maximum intensity can be manipulated by the user by adjusting different parameters (Flores-Alsina et al., 2013a, b): Q_{permm} defines the flow rate due to rain, L_{rain} determines the number

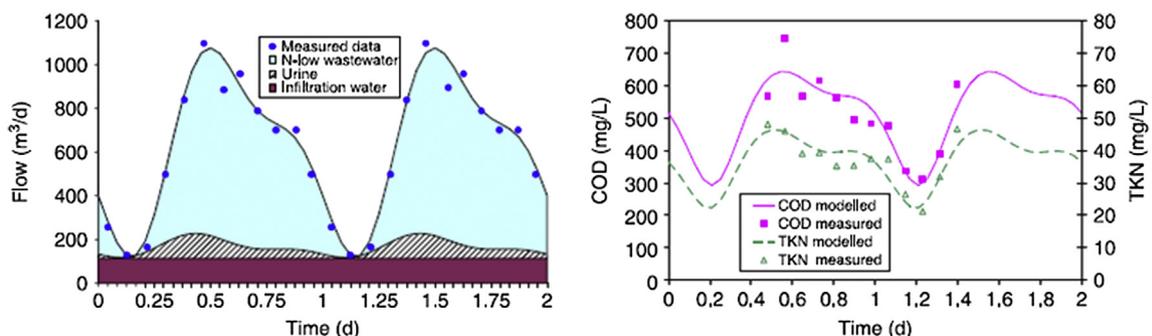


Fig. 3. (from Langergraber et al., 2008): Comparison of modelled and measured daily flow (left) and concentrations of COD and TKN (right). In the legend of the right figure, the COD measured (bottom) should be TKN measured.

of rain events, etc. Alternatively, some studies (Benedetti, 2006; Benedetti et al., 2006) have replaced the output of this module by actual rainfall time series data.

- **Household wastewater generation module:** It generates the pollutant load profiles (units of g/m^3) by multiplying the average load rates per PE (COD, NH_4 , TKN, etc.) with the number of PE, and assuming a daily pattern (1 h time step) with two peaks at 8 and 17 h (inspired by Almeida et al. (1999)). The flow generated by the household is also proportional to the PE and follows three patterns: daily (two peaks at 9 and 19 h), weekly (around 10% lower during the weekends) and yearly (25% lower during summer holidays). By modifying the default pollutant profiles, this module can represent the possible variability of the pollutant profiles with the geographical region (Butler and Gatt, 1996) or among neighbourhoods of different social strata (Campos and von Sperling, 1996).
- **Industrial discharges generation module:** The pollutant discharges from industries are represented by a single loading rate (COD, TKN, NH_4 , etc.) that follows a weekly pattern: a peak on Fridays represents the industrial cleaning and lower pollutant discharges represent the weekend. The wastewater flow generated by the industries presents weekly and yearly profiles: the first one is similar to the one of the pollutants and the yearly one represents two holiday periods (summer and Christmas) where the flow is lower (20% and 30% respectively). With the idea of proposing a simple model description that could be extensively used, the module assumes that the wastewater generated by industries and households can be fractionated according to the same principles. This default fractionation can easily be modified given the availability of the code.
- **Soil model module:** The soil is represented by a variable volume tank model to represent the storage of water in the soil. The flow entering the tank is composed of a fraction of the rainfall that falls in permeable areas, and the outflow represents the flow to the sewer system and the aquifers. The infiltration to the sewer system depends of the invert level (the maximum water level that will not cause infiltration in the sewer pipes) which at the same time depends of the temperature variation, as a surrogate for seasonal variation of the groundwater level.
- **Temperature module:** Daily and yearly temperature variations are defined by two sinusoidal functions. The yearly sinus has an average of 15 °C, amplitude of 10 °C and frequency of $2\pi/364$ rad/d. The daily sinus moves around the seasonal one with amplitude of 1 °C and frequency of $2\pi/1$ rad/d. In Flores-Alsina et al. (2013a, b), the temperature and rain generation are linked by a model that relates temperature with snow melting episodes. The snow melting water is generated by a new module similar to the rainfall module of Gerney et al. (2011).

- **Sewer model module:** The sewer system is defined by connecting a different number of subcatchments (one to eight) to represent more or less complex sewer systems. Each subcatchment is modelled by a tank-in-series approach using variable volume reactors where the outflow of each tank follows an exponential function on the water height. The first flush effect is included by a simple equation with two adjustable parameters to tune its strength. It also assumes an even complexity of all subcatchments by always using three tanks-in-series. Nevertheless, after calibration, the complexity of the sewer model has demonstrated to be sufficient to correctly describe influent profiles with one hour frequency (Flores-Alsina et al., 2013a, b).

As result, the influent generator of Gerney et al. (2011) achieves a simple phenomenological representation of the typical dynamics that are observed in a full scale WWTP influent, being able to feature diurnal phenomena; weekend effects (e.g. lower pollutant load); seasonal phenomena (e.g. decreased infiltration with temperature); holiday periods (e.g. with a lower average wastewater flow rate); rain events; etc. Currently, this generator seems to move forward to a more detailed description of the generating processes beyond the wastewater generation (see for example, the developments carried out in the EU-funded SANITAS project <http://www.sanitas-itn.eu/>). In this sense, and as a wish list for further developments, the following could be considered: a rain generator in accordance with state of the art models (Thauvin et al., 1998; Willems, 2001; Palynchuk and Guo, 2011); a more descriptive relationship between temperature and rain profiles for different climates (Birt et al., 2010); the building up and wash off of urban pollutants (Egodawatta et al., 2007; Wang et al., 2011); the description of the vertical infiltration in terms of the soil properties, such as specific moisture capacity, capillarity head, effective hydraulic conductivity, etc. (Richards, 1931; Corradini et al., 2011); and so on. In this context, the description of the wet weather flow and composition along the sewer system is still an important challenge but current progress is benefitting from (recently published) high frequency data sets (Schilperoord, 2011; Langeveld et al., 2013). New modelling proposals for influent generation are also better handling wet weather conditions or the sewershed and include a weather generator that can be calibrated on local climate data being able to describe inter-daily variations of flow and TSS (Talebizadeh et al., 2014).

Early applications of the phenomenological model (Gerney et al., 2005, 2006) have shown the usefulness of the tool for the modelling of integrated urban water systems. In some cases, the model is presented as a part of general methodologies: to check the WWTP performance (Ráduly et al., 2007) or to design correction measures in a probabilistic framework (Benedetti et al., 2008). The

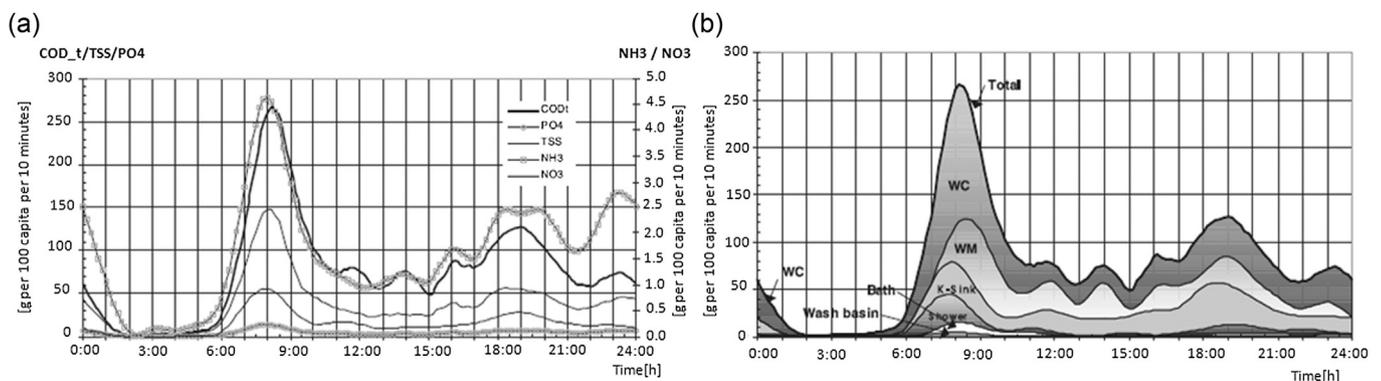


Fig. 4. (from Almeida et al., 1999): a) Diurnal pattern of COD, PO₄, TSS, NH₃ and NO₃ in wastewater; b) COD load contributions per appliance: WC, washing machine, kitchen sink, shower, bath and wash basin.

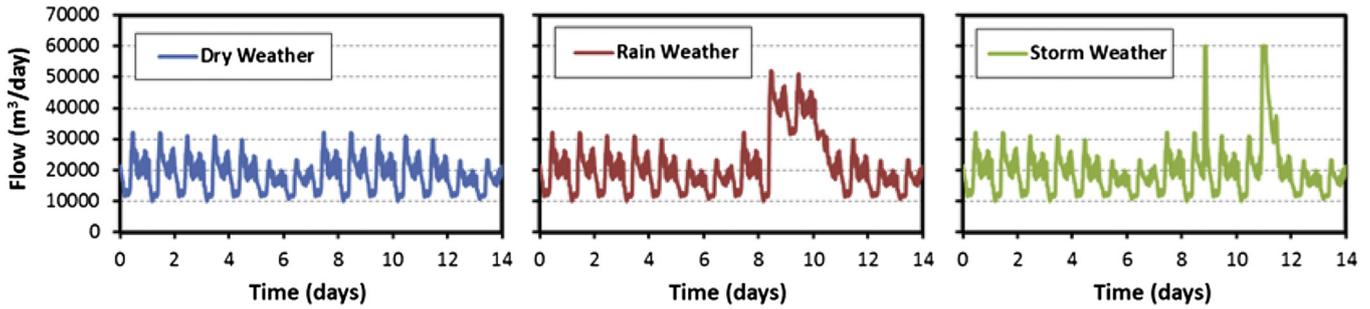


Fig. 5. BSM1 flow profiles featuring dry, rain and storm weather conditions (Copp, 2002).

phenomenological model has also been able to fulfil specific objectives as generating realistic influent data profiles for WWTP of different characteristics (Benedetti et al., 2008); adding new components to the ones in ASM family models (Lindblom et al., 2006); or generating training data with a wide range of influent conditions, such as: rain events, holiday effects, industrial discharges, etc (Ráduly et al., 2007).

A Global Sensitivity Analysis of the phenomenological model (Gernaey et al., 2011) has recently been published (Flores-Alsina et al., 2012). The analysis focuses on the influent flow rate and shows that while the main parameters affecting dry weather flow rate profiles are the catchment size and the wastewater flow per person, the ones determining the wet weather profile are the probability of occurrence of a rain event, the catchment size, and the quantity of rain falling on permeable areas. This analysis has been completed with a calibration of the BSM2 influent generator on full-scale plant data (Flores-Alsina et al., 2013a, b) (see Section 6 for further information).

One of the clearest advantages of the model is that the user can simply create influent wastewater characteristics according the features of the catchment area: number of PE, percentage of permeable surfaces, flow rate generated per unit of rain, complexity of the sewer network, etc.; or extrapolate to other scenarios by using hypothetical data. In this respect, Flores-Alsina et al. (2013a, b) evaluates three representative cases: the variation of influent flow rate for different rainfall patterns, the effect of parameter uncertainty on the predicted influent biodegradability, and the increase of frequency in a given dataset.

4. Solutions found for situation 2: the fractionation problem

In this section we cover the solutions found in literature with respect to the so-called fractionation problem. The translation of the experimental measurements (BOD, COD, TSS, TKN, TN, etc.) into the model components of ASM family models (readily biodegradable substrate, soluble inert organic matter, slowly biodegradable substrate, heterotrophic biomass, etc.) is not a straightforward problem. Indeed, it is a non-identifiable problem since the experimental data is never informative enough to uniquely determine the model components. Different protocols and scientific research have tried to overcome this situation.

Henze (1992) presented a summary of methods to determine the fractions of COD, the origin of different nitrogen compounds, and biomass populations. After analysing the wastewater composition of several WWTPs over the world, it seemed that for a specific WWTP the fractions of COD remain quite constant. However, when comparing the fractionation among the different WWTPs it was concluded that the sewer system and temperature variations might influence the composition significantly. Roeleveld and van Loosdrecht (2002) came to a similar conclusion. Motivated by the experience of modelling more than 100 WWTPs in The Netherlands, the Dutch Foundation of Applied Water Research stimulated the development of the STOWA protocol (Hulsbeek et al., 2002; STOWA, 1996; STOWA, 1999). On that effort, special attention was paid to the characterisation of the influent wastewater (Roeleveld and van Loosdrecht, 2002) in order to determine the ASM1 model components. After carefully comparing the COD

Architecture of the Flow Generator

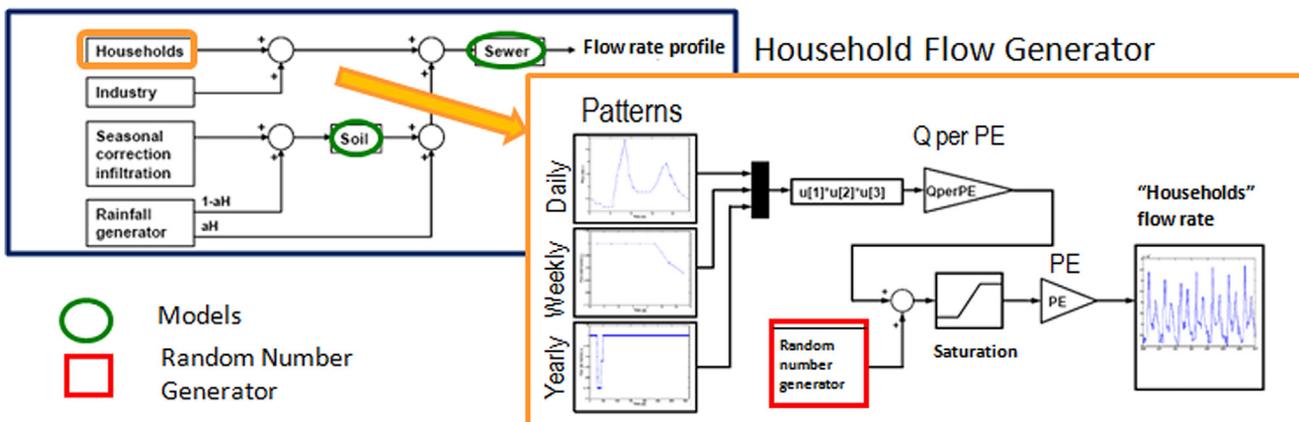


Fig. 6. Architecture of the wastewater flow generator in the influent disturbance model of Gernaey et al. (2011). Figure created from Flores-Alsina et al. (2012) and Gernaey et al. (2005).

Box 3

Frequently Asked Questions in Situation 2, i.e., the so called *fractionation problem*.

a) Is there any standard protocol for performing WWTP simulation studies?

STOWA (Hulsbeek et al., 2002), BIOMATH (Vanrolleghem et al., 2003), WERF (Melcer et al., 2003) and HSG (Langergraber et al., 2004) are well established protocols defining standard modelling steps such as: definition of objectives, general layout data collection and model selection, data quality control, simulation study, calibration/validation of models, assessment of the modelling application, etc. Improvements are described in Mannina et al. (2011) and Rieger et al. (2012).

b) Is there any analytical procedure to characterise the wastewater in terms of ASM model variables?

The STOWA, BIOMATH, WERF and HSG protocols include different methods to characterise COD and Nitrogen fractions. The main difference lies in the methods for the estimation of the COD fractions.

c) Is there any standard, experience-based, manner to calculate ASM model components from typical experimental measurements (COD, TKN, NH₄ etc.)?

Yes, the Benchmark Simulation Model No. 1 (Copp, 2002) includes a fractionation model in which the soluble and particulate components of the ASM1 model are estimated assuming fixed ratios. As an extension of that, Gernaey et al. (2011) proposes three different fractionation models to derive the state variables of ASM1, ASM2d and ASM3 models.

d) Is there any kind of generalisation of the above estimations so as to use all the experimental measurements available (Turbidity, COD, pH, etc.) to derive the state variables of any ASM family model?

Yes, Grau et al. (2007) proposed an optimisation algorithm that using as starting point a typical wastewater characterisation, leads to a consistent solution that minimises the difference with the provided experimental measurements. This solution is expressed in terms of an extended list of model components and can easily be translated into any set of ASM family state variables.

fractionation results of 21 WWTPs, they found that, except for the fraction of the soluble inert material that remains quite stable, a high variation is noticed in the other COD fractions (readily biodegradable substrate, slowly biodegradable substrate and particulate inert organic matter). This variation was attributed to the type of wastewater (contribution of industrial waters), the sewer system type and length (separate/combined) and the transformation processes in the sewer (pressure main/gravitational). However, the data showed insufficient evidence to draw firm conclusions with this regard. Another important conclusion of this modelling effort was that the wastewater characterisation is very helpful for the daily WWTP operation even if no modelling study is scheduled (Roeleveld and van Loosdrecht, 2002). For example, the authors explain that the determination of the soluble COD available is related to the capacity of the plant for high nitrogen and biological phosphorus removal.

Generally, the characterisation problem is related to the calibration/validation of mathematical models and therefore, it is generally presented as part of general protocols for activated sludge modelling, such as the STOWA protocol (Hulsbeek et al., 2002); the BIOMATH protocol (Vanrolleghem et al., 2003); the HSG guidelines (Langergraber et al., 2004); and the WERF protocol (Melcer et al., 2003). An SWOT analysis of the four protocols (Sin et al., 2005) identifies the detailed influent characterization of the BIOMATH and WERF protocols as strength, while it is identified as weakness

in the HSG guidelines. One of the main differences with regard to the influent characterization lies in the COD fractionation. The STOWA protocol (Roeleveld and van Loosdrecht, 2002) proposes the BOD test to determine the readily biodegradable soluble COD (S_5 in ASM models) while the BIOMATH protocol proposes respirometry tests (Sollfrank and Gujer, 1991; Spanjers and Vanrolleghem, 1995; Sperandio and Paul, 2000). Both protocols stress the importance of the wastewater characterization for a successful model calibration. With regard to the data collection, the STOWA protocol insists on the use of mass balances to check the data consistency and the BIOMATH protocol recommends the use of Optimal Experimental Design (Dochain and Vanrolleghem, 2001) to optimize the quality of the calibration results. The HSG guidelines are more flexible (than the STOWA or the BIOMATH protocols) and they recommend to use a COD fractionation based on respirometric tests (as in the BIOMATH protocol) or to combine physical-chemical methods with BOD measurements (as in the STOWA protocol). In the WERF protocol (Melcer et al., 2003) a detailed fractionation method is included in the last (and more sophisticated) type of the calibration procedure together with the estimation of the stoichiometric and kinetic model parameters.

Improvements to these protocols are described in Mannina et al. (2011) and Rieger et al. (2012). Mannina et al. (2011) proposed a novel calibration protocol for nutrient removal wastewater treatment models based on combining the use a global sensitivity analysis with the Generalized Likelihood Uncertainty Estimation (GLUE) approach. The most innovative feature consists of using a step-wise approach to define sub-groups of parameters to be calibrated on different model outputs. Rieger et al. (2012) summarizes the keys for good modelling practice in wastewater treatment. The IWA Task group has led important initiatives for standardizing: the qualitative criteria to evaluate simulation results (Hauduc et al., 2011), the notation of activated sludge models (Corominas et al., 2010), the uncertainty assessment (Belia et al., 2009), or the problem of data reconciliation (Rieger et al., 2010a). The effort is culminated with a critical review of wastewater treatment modelling and the presentation of future possibilities (Rieger et al., 2010b; Hauduc et al., 2013).

The Benchmark Simulation Model No. 1 (Copp, 2002) includes a fractionation model (inspired by Vanhooren and Nguyen (1996)) in which the soluble and particulate components of the ASM1 model are estimated incorporating influent wastewater characteristics of several studies (Butler et al., 1995; Campos and von Sperling, 1996; Londong, 1994; Verbanck, 1995) and assuming fixed ratios between the most common measurements: for example, 1.33 as the ratio between the particulate COD and the TSS. Following a similar approach, the influent generator of Gernaey et al. (2011), see Section 3.3, proposes three different fractionation models that transform the pollutant load of households and industries into the state variables of ASM1, ASM2d or ASM3 models. The pollutant load is characterized by the soluble COD, particulate COD, TKN and NH₄ concentrations for the ASM1 and ASM3 models; and also in terms of PO₄ in case of the ASM2d model. The module uses fixed relationships to derive the model state variables.

In a very pragmatic and engineering approach, based on the principle of elemental continuity, and without making any recommendation about the most suitable laboratory tests to use, Grau et al. (2007) proposed a mathematical tool for the characterization of the influent wastewater. The algorithm is based on the definition of an extended list of model components (ammonium, ammonia, dissolved carbon dioxide, bicarbonate, monosaccharide, aminoacids, LCFAs, etc.) and their relationship with commonly available analytical measurements (COD, BOD, TSS, TN, TKN, etc.). The algorithm is initialized by defining a typical urban wastewater in terms of an extended list of model components (default

solution); and it looks (by minimizing an error function) for another numerically *close* solution that reasonably fits the available experimental data. The optimal solution (expressed in terms of the extended list of model components) can be straightforwardly translated in terms of any other biochemical model (any ASM family model). By requiring a solution close to an initial estimation, the identifiability problem (caused by the fact that the experimental measurements are never informative enough to uniquely identify the model components) is overcome. De Gracia et al. (2011) presented an application of this methodology in which three types of WWTP sludge (sludge produced in primary clarifiers, sludge produced in secondary clarifiers; a mixture of the first two) are characterised in terms of the components of a generic digester model (de Gracia et al., 2009).

5. Solutions found for situation 3: engineer would like to characterize the uncertainty of the quantity and quality of the influent wastewater

The necessity of characterizing the uncertainty around the expected influent data profiles has been repeatedly outlined (Friedler

Box 4

Frequently Asked Questions in Situation 3, i.e., an engineer wants to characterize the uncertainty of the quantity and quality of tie influent wastewater.

a) How can the understanding of the uncertainty in the influent profiles be improved by analysing its sources?

The main references for they are those of Friedler and Butler (1996) that analyses the uncertainty in the generation of domestic wastewater, and of Ort et al. (2005) that studies the stochastic load variations in sewer systems. While the first one analyses the different wastewater compounds and their relationships, the second one is more focussed on the timing of the emissions.

b) In terms of model building, is there any manner to model the uncertainty in wastewater generation from its very source?

Yes, the phenomenological mode of Gernaey et al. (2011) includes some uncertainty in definition of the households' pollutant profiles, in the industrial pollutant loads, in the rain generation, the temperature profiles, etc. The generator of De Keyser et al. (2010) also includes uncertainty from the sources by including random variables in the daily, weekly and yearly profile definitions. The approach proposed by Talebizadeh et al. (2014) includes two statistical models for the synthetic generation of rainfall series and the water quality influent profiles during dry weather conditions.

c) Once a set of influent wastewater data is provided, is there any method to include some uncertainty over the time series that is coherent with its uncertainty?

Yes, Carstensen et al. (1998) propose an Autoregressive First Order model to represent the uncertainty over the response of a second order Fourier model. Similarly, Martin et al. (2007) proposed the use of ARMA models to introduce uncertainty (white and coloured noise) over a set of given time series data. If the influent data profiles are the result of a model, other realisations of the time series data can be obtained by using Bayesian calibration of the parameters of this mode (see Bechmann et al. (1999) or Talebizadeh et al. (2014)).

d) How can both approaches (in b and c) be combined?

The understanding of the wastewater generating mechanisms and the uncertainty of their sources lead to solutions as the ones presented in b). However, the complexity of the problem will unavoidably lead to synthetic data profiles that do not represent the whole reality of the catchment area. The solutions pointed out in c) allow describing the uncertainty due to the remaining lack of knowledge.

and Butler, 1996; Bixio et al., 2002; Benedetti et al., 2008; Gevaert et al., 2009). However, the difficulty of analysing the uncertainty in real data sets (Friedler and Butler, 1996), the short history of uncertainty analysis in water quality modelling (Belia et al., 2009), and the lack of stochastic formulations of biochemical models (Reichert and Mieleitner, 2009) have not yet given rise to standard methodologies with this in mind. Note that rigorous procedures for generating realistic time series profiles with similar statistical properties would be very useful for uncertainty based (i.e. Monte Carlo based) design and dimensioning of new WWTPs (Bixio et al., 2002; Martin et al., 2012), and also for validating robust control algorithms such as the QFT control strategy (García-Sanz et al., 2006).

Friedler and Butler (1996) tried to assess the uncertainty in the quantity and quality of domestic wastewater by analysing the data from two surveys carried out in the south of England (Butler, 1991, 1993). They concluded that the total volumes, pollutant loads and frequency of use are highly uncertain and that their distributions depend not only on the appliance and mode of use but also on the specific user. They explain that the uncertainty of these parameters is in general related to the magnitude of the pollutant profiles but also to their timing. In this sense, the uncertainty in the timing of the pollutant releases is highly skewed due the predominance of zero usage of the appliances. The authors conclude that Monte Carlo simulations would be needed to analyse the combined effect of the uncertainties in downstream elements of the wastewater infrastructure.

Rousseau et al. (2001) did also try to characterise the uncertainty found in real experimental data sets. In this case they measured the influent flow and influent load profiles at the inlet of 167 WWTPs in Flanders (Belgium) and certain relationships between the flow and wastewater pollutant concentrations (COD, TKN and NO_3) could be identified. These relationships (Fig. 7) together with the available flow data and a fixed fractionation model (of the COD into the ASM1 variables) were used to generate "shots of" dry weather influent profiles. The authors claim that the generated influent data is coherent with the variability observed among the 167 WWTPs but also with the inherent uncertainty of their patterns.

The phenomenological model of Gernaey et al. (2011) also follows a simple but effective approach to introduce randomness in the generated data. It uses random number generators (with zero mean and tuneable standard deviation) along the different blocks (households, rain generation, seasonal correction factor, fractionation model, etc.) so that different profiles featuring the uncertainty

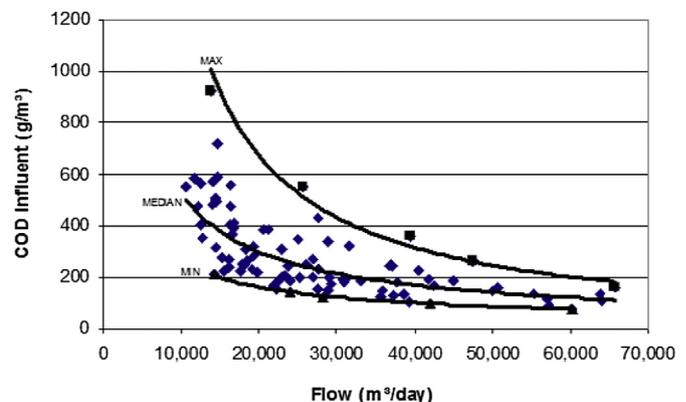


Fig. 7. (from Rousseau et al., 2001): Relation between influent COD concentrations and flow for medium strength wastewater and the indication of minimum and maximum concentration.

of the underlying mechanisms can be provided. More descriptive approaches can be implemented by representing some inference from uncertainty analysis of real data. For example, the inclusion of the uncertain correlation among different variables (as observed in Rousseau et al. (2001)) or the uncertainty in the patterns of domestic wastewater generation (Friedler et al., 1996) would be interesting.

The influent generator defined by De Keyser et al. (2010) is based on the use of 24 daily patterns, 11 weekly patterns and 14 yearly patterns resulting from an extensive literature review. The stochasticity of the resulting profiles is introduced in two different ways: (i) adding white noise to the obtained time series; and (ii) allowing some pattern parameters to be randomly sampled (so that the daily, weekly and yearly patterns become stochastic structures). The second strategy implies defining some pattern parameters to be random variables. For example, a uniform distribution can determine the time for an emission peak to occur (daily pattern), or the day in the week (weekly pattern), or the weekends in which it repeats itself (yearly pattern).

Another interesting manner of introducing uncertainty in the daily or weekly periodic profiles is to introduce the time warping approach (Vilchez et al., 2009; Gins et al., 2006) in which the profiles are disrupted by introducing the uncertainty of their timings. Fig. 8 shows an example in which the daily patterns of the ammonium and soluble organic nitrogen influent profiles of BSM1 model (Copp, 2002) have been manipulated by adding white noise not to their values but to the times at which they occur.

For cases in which the influent data of a WWTP is already the result of a mathematical model, different realisations of these data can be generated by using a stochastic formulation of the model. This can be achieved by defining the parameters of a deterministic model by random variables defined by a probability distribution function (Bayesian approach) (see Omlin and Reichert, 1999; Martin and Ayesa, 2010). This method has been used to estimate the COD and TSS mass load at the entrance of a WWTP by using measurements of UV absorption and turbidity along a sewer system (Bechmann et al., 1999). Talebizadeh et al. (2014) uses the same approach to incorporate an urban drainage model (CITYDRAIN, Achletner et al. (2007)) for the stochastic generation of water quantity and quality profiles under wet weather conditions. Another set of solutions to feature the uncertainty on a set of data profiles is based on the use of autoregressive models. This statistical approach does not try to understand the uncertainty in the generating mechanisms (domestic load, permeability of the soil, rainfall profiles, etc). On the contrary, it just tries to reproduce some uncertainty in coherence with the available experimental data or the assumed lack of knowledge. In this line, Carstensen et al. (1998) propose a grey box model to represent the uncertainty around the

hydraulic load expected at the inlet of a WWTP. They used an Autoregressive First Order model to generate some randomness over the response of a second order Fourier model (Section 3.2). They found that the inclusion of a first order autoregressive model considerably improves the description of the hydraulic data profile. In a similar way, Martin et al. (2007) proposed the use of Autoregressive Moving Average Models (ARMA) models to feature the uncertainty around given time series data (Bras and Rodríguez-Iturbe, 1985). The main advantage of the methodology is that it allows characterising the white and coloured noise and tuning the magnitude of the uncertainty to be generated by adjusting some of the ARMA model parameters.

6. Calibration, validation and model performance evaluation

In general, influent generators, as any other model, need to be calibrated and validated on the evidence of experimental data (Dochain and Vanrolleghem, 2001). However, the wide range of studies and methods presented in this paper do not necessarily follow these general model building steps. The influent generators based on databases (Section 3.1) for example, do not require any calibration. They can directly be used in coherence with the requirements of the user or the characteristics of the catchment area. On the contrary, the influent generators based on the use of harmonic functions (Section 3.2) or the phenomenological model (Section 3.3), do need some calibration of their parameters. The phenomenological model of Germaey et al. (2011), for example, has been calibrated using experimental data from two Scandinavian WWTPs (Flores-Alsina et al., 2013a, b). The authors follow a stepwise procedure based on adjusting the most influential parameters after a global sensitivity analysis (Flores-Alsina et al., 2012) at different time scales (from yearly profiles to hourly values). They demonstrate the validity of the model results by using a set of data not used during the calibration stage. The performance of the generator and the guidance provided by this calibration experience will facilitate the usage of the influent generator in other case studies or modelling scenarios.

In general, the model calibration follows either the Frequentist or Bayesian approach (Omlin and Reichert, 1999). Under the frequentist philosophy the modeller looks for an optimum parameter set that minimises the differences between the experimental and simulated results according to some statistics-based objective function. On the contrary, the Bayesian framework assumes that there is not a single optimum parameter set but an optimal region of parameter values in which the simulation result is reasonably good. An example of Frequentist estimation is the one presented by Flores-Alsina et al. (2013a,b). Examples of Bayesian estimation are the ones presented by Lindblom et al. (2011), Rieckermann et al.

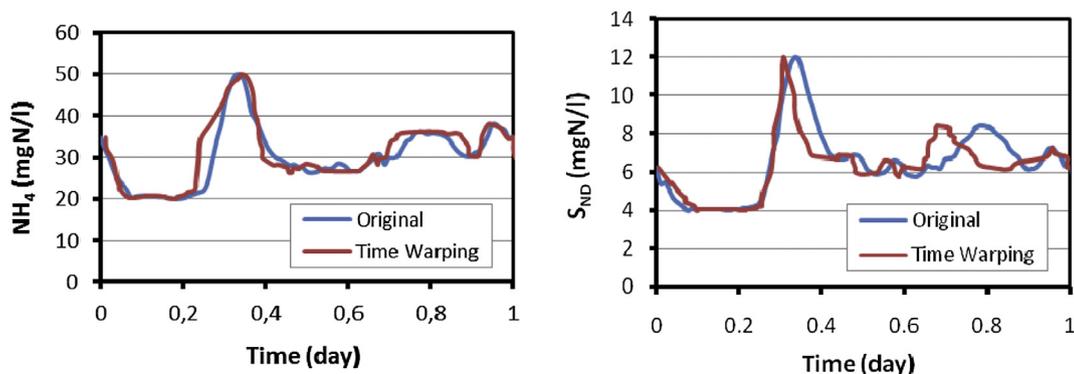


Fig. 8. Time warping approach used to introduce some uncertainty in the ammonium and soluble organic nitrogen influent profiles of BSM1 model (Copp, 2002).

(2011) or Talebizadeh et al. (2014). Lindblom et al. (2011) use the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992) to calibrate a stormwater surface accumulation-wash-off model based on an extensive measurement campaign of Zn, Cu, Pb and Cd field data. GLUE adopts the Bayesian framework but substitutes the likelihood function by a subjectively chosen objective function or 'less formal likelihood' (Beven and Freer, 2001). The reason to use the GLUE approach instead of traditional Bayesian inference is that the latter requires of statistically derived likelihood functions which involve strong assumptions on the model residuals' distribution (generally to be independent and normally distributed). In this case, Lindblom et al. (2011) adapted the classical likelihood function with an extra scaling parameter that makes the likelihood expression more or less sensitive to the residual errors. This scaling parameter is manually adjusted for the different datasets so that the forward propagation of parameter uncertainty covers 95% of the experimental data. The second example of using Bayesian inference is the one presented by Rieckermann et al. (2011) that calibrates a stochastic model describing wastewater micropollutant loads by using three Approximate Bayesian Computation (ABC) methods (Beaumont et al., 2002; Marjoram et al., 2003). ABC methods derive the posterior distribution of parameters without using the likelihood function of model residuals. The article compares three different numerical algorithms to approximate the posterior distribution of a single parameter: the number of wastewater pulses contained in three high-resolution data series of benzotriazole and total nitrogen loads in sewers. The main contribution of the article consists of divulging the use of ABC methods (not usual in biochemical modelling) since they are simple and with sound statistical properties when compared to traditional Bayesian inference or with the GLUE approach (Mantovan and Todini, 2006). Finally, Talebizadeh et al. (2014) used Bayesian calibration to adjust the output of the open-source CITYDRAIN model (Achleitner et al., 2007), with which it is possible to generate influent wastewater under wet weather conditions incorporating the effect of climate and sewershed characteristics.

The validation of any modelling approach should be accompanied by an analysis of the model performance by comparison of the simulated and experimental results. Different authors have suggested standard methodologies and presented appropriate use of metrics in this respect (Beck, 2006; Moriasi et al., 2007; Mattot et al., 2009). This effort has recently moved forward with the presentation of a position paper on water quality models (Bennet et al., 2013) that analyses quantitative and qualitative methods in literature for characterising model performance, and a critical review paper on the use of different objective functions for wastewater models (Hauduc et al., 2011). The former (Bennet et al., 2013) stresses that the adoption or rejection of a model is not only related to the goodness-of-fit measurements but also to other factors such as cost, simplicity, applicability, intelligibility, etc (Giordano et al., 2012; Krueger et al., 2012; Rowan et al., 2012). Based on the use of the objective criteria, Bennet et al. (2013) propose a stepwise performance evaluation method that includes the definition of the modelling objective, the data check, a visual performance analysis, etc. The latter paper (Hauduc et al., 2011) reviews 31 goodness-of-fit criteria generally used in wastewater treatment modelling. The study concludes that most of them are intrinsically correlated, and proposes a two-step cluster analysis using Kendall correlation and dendrograms to select criteria that provide non-redundant information.

Other approaches covered in this paper (for Situation 2 and Situation 3) do not need model calibration/validation nor performance evaluation, but they use these methods in different manners. For example, the standard protocols for using activated sludge

models (Situation 2, Section 4) define systematic calibration protocols. They recommend step-wise procedures where steady-state and dynamic simulations are compared with experimental data at different stages. Usually, subsequent validation and retrofit model improvement steps are included. On the other hand, some of the methods described to include uncertainty in influent generators (Situation 3, Section 5) are based on the use of Bayesian statistics to calibrate their parameters and assess their uncertainty (Bechmann et al., 1999; Talebizadeh et al., 2014). This approach enables the realisation of different influent profiles coherent with the mechanisms defined in the model structure and the uncertainty of their parameters.

7. Discussion

This paper makes a critical review of the techniques available in the state of the art to analyse, complete and generate influent data for WWTP modelling. This section summarizes the main points of the analysed studies and provides some tips for further development:

- The knowledge about the typical flow, concentration and load patterns helps to understand what the domestic wastewater generation at the household level resembles. In dry weather conditions, the daily patterns of flow and loads present very clear profiles with one morning and two evening peaks (dinner and bedtime preparation). In terms of concentrations the variability within the day is lower than the variability of the loads, and seems not to be correlated with the flow. The uncertainty of concentrations and loads is featured by skewed distributions that represent not only the uncertainty of the magnitude of the pollutant profiles but also of their timings. Note that these household level patterns may no longer be observable at the entrance of the WWTP due to different pumping regimes, sewer retention tanks, etc.
- Understanding the influent flow and concentration patterns in dry weather conditions has given rise to very simple and effective influent generators for WWTP modelling based on the use of harmonic functions (second or third order Fourier models). The parameters of these models seem to be related to the size of the WWTP under study.
- The construction of databases from real wastewater data profiles has provided influent generators able to provide synthetic data for long term WWTP modelling (dry and wet weather conditions). On the one hand, the methodology behind these approaches can be applied to complete or increase the frequency of real influent data. On the other hand, if no real data are available, the generators themselves are good tools to generate synthetic profiles which at least are coherent with the patterns identified in bibliography or accepted by the scientific community. The main limitation of these tools is that they do not allow extrapolating to new situations such as population growth, climate change, transformation of combined to separate sewer systems, rehabilitations of sewers to reduce infiltrations, etc.
- The phenomenological modelling of the wastewater generation mechanisms is a very young and promising research area. An influent generator for long term WWTP modelling has been proposed together with the Benchmark Simulation Model No. 2 initiative. The generator provides influent data in coherence with the catchment characteristics, such as: person equivalents, industrial discharges, rainfall profiles, percentage of permeable surface, etc. Using hypothetical data about the catchment area the modeller is able to a certain extent to extrapolate to other scenarios or other similar cities. Some of the remaining

challenges for the research community are related to: improve the rain and temperature generators (and their relationship); include some representation of the build-up and wash off of urban pollutants; or upgrade the soil model in terms of soil properties as the specific moisture capacity, the capillarity head, the effective hydraulic conductivity, etc. The interest already shown in the research community (for example in the aforementioned SANITAS project), the flexibility of the tool, and the fact that it is available as open-source code, will accelerate the development of new, more comprehensive versions of the generator.

- The STOWA (Hulsbeek et al., 2002) and BIOMATH protocols (Vanrolleghem et al., 2003) include detailed wastewater characterisation methods to translate the generally available measurements (COD, BOD, TKN, etc.) into the components of the ASM family models. These protocols include specific recommendations about the most adequate laboratory experiments and batch conditions, as well as data collection techniques. In simulation studies where there is no information about the origin of the available measurements the fractionation problem can be solved by numerical optimisation techniques.
- The stochastic nature of the wastewater influent profiles has been represented in the influent generators following a simple but effective approach. Both in databases built generators and in the phenomenological model, the stochasticity is achieved by including random values at some points of the generating process. More fundamental research would make possible to model the uncertainty involved in the underlying mechanisms (rainfall profiles, household and industry pollutant discharges, assumed daily and weekly patterns, etc.).
- Statistical models based on autoregressive functions are suitable to represent the uncertainty involved in influent data profiles. These statistical models are able to generate white and coloured noise over a given data series in representation of the uncertainty observed in the experimental data or assumed as modeller lack of knowledge. This statistical approach should move forward into the formulation of stochastic models able to capture the correlation among the typical wastewater measurable parameters (COD, TSS, NH₄, etc.).

8. Conclusions

The state of the art already provides a wide range of solutions for characterising, generating and completing WWTP influent data series. The main challenge for engineering practice consists of being able to select the most appropriate solution for each situation. Some questions that should facilitate that choice are: Do we have enough data to characterize the influent quantity and quality? Do we need to generate influent data by only knowing some general characteristics about the catchment area? Would we prefer some simple approach to generate dry weather influent profiles or do we need to characterise as well wet weather influents? Do we need to translate typical experimental measurements (COD, TSS, TKN, etc.) into the ASM family model state variables? Do we need to perform an uncertainty analysis of the WWTP performance?

The influent generators for WWTP modelling and simulation are going to be important in the following years. For instance, quantitative risk analysis of WWTPs or probabilistic design will only be possible if we can generate influent profiles under hypothetical conditions (population growth, strong rain events, uncontrolled spills, etc.) and characterise their uncertainty. That process will increase the awareness about modelling robustness (good behaviour under uncertain conditions) during future projects for WWTP design or improved operation. In that respect, phenomenological

modelling for influent wastewater generation should deserve special attention because it can integrate important knowledge about the generating mechanisms. The state of the art should move forward step by step by introducing more detailed descriptions of the different components in the catchment area (build-up and wash off of urban pollutants, soil infiltration, sewer system, retention tanks, etc.) and the stochastic nature of the inputs (rainfall and temperature profiles, run-off generation, industrial and household loads, etc.). The result will be a stochastic model where the mean values are coherent with the catchment properties (hypothetical or real) and the uncertainty of the generated time series coherent with the stochastic nature of the problem. While pursuing that effort, the lack of details in the description of the catchment elements or stochastic inputs (modeller uncertainty) can be represented by using a statistical model (e.g. based on autoregressive functions) on top of the phenomenological core model.

Acknowledgements

This work was made possible through the financial support of the Hampton Roads Sanitation District, Virginia, USA. The authors wish to acknowledge the support provided by the IWA Task Group on Design and Operations Uncertainty (DOUT Specialist Task Group) and the engineering experience of Evangelia Belia from Primodal Inc, Québec, QC, Canada. The authors wish also to thank the financial support of the Natural Science and Engineering Research Council of Canada (NSERC). Peter A. Vanrolleghem holds the Canada Research Chair in Water Quality Modeling.

References

- Achleitner, S., Möderl, M., Rauch, W., 2007. CITY DRAIN[®] – an open source approach for simulation of integrated urban drainage systems. *Environ. Model. Softw.* 22 (8), 1184–1195.
- Alex, J., Hetschel, M., Ogurek, M., 2009. Simulation study with minimised additional data requirements to analyse control and operation of WWTP Dorsten, Germany. *Wat. Sci. Tech.* 60 (6), 1371–1377.
- Alex, J., Wichern, M., Halfit, N., Sperring, V., Ahnert, M., Frehmann, T., Hobus, I., Langergraber, G., Plattes, M., Winkler, S., Woerner, D., 2007. A method to use dynamic simulation in compliance to stationary design rules to refine WWTP planning. In: Poster Presentation at the 10th IWA Specialised Conference on Design, Operation and Economics of Large Wastewater Treatment Plants, 10–13 September 2007, Vienna, Austria.
- Almeida, M.C., Butler, D., Friedler, E., 1999. At-source domestic wastewater quality. *Urban Water* 1, 49–55.
- ATV-DVWK, 2000. A 131E Dimensioning of Single-Stage Activated Sludge Plants. German Association for Water, Wastewater and Waste, Hennef, Germany.
- Ayesa, E., De la Sota, A., Grau, P., Sagarna, J.M., Salterain, A., Suescun, J., 2006. Supervisory control strategies for the new WWTP of Galindo-Bilbao: the long run from the conceptual design to the full-scale experimental validation. *Wat. Sci. Tech.* 53 (4–5), 193–201.
- Beaumont, M.A., Zhang, W.Y., Balding, D.J., 2002. Approximate Bayesian computation in population genetics. *Genetics* 162, 2025–2035.
- Bechmann, H., Nielsen, M.K., Madsen, H., Poulsen, N.K., 1999. Grey-box modelling of pollutant loads from a sewer system. *Urban Water* 1, 71–78.
- Beck, B., 2006. Model evaluation and performance. In: *Encyclopedia of Environmental Metrics*. John Wiley & Sons, Ltd.
- Belia, E., Amerlinck, Y., Benedetti, L., Johnson, B., Sin, G., Vanrolleghem, P.A., Gernaey, K.V., Gillot, S., Neumann, M.B., Rieger, L., Shaw, A., Villez, K., 2009. Wastewater treatment modelling: dealing with uncertainties. *Wat. Sci. Tech.* 60, 1929–1941.
- Benedetti, L., 2006. Estimating Uncertainty in Discharge and Water Quality Data for Monitoring and Modeling Application. PhD thesis. Ghent University, Belgium.
- Benedetti, L., De Baets, B., Nopens, I., Vanrolleghem, P.A., 2010. Multi-criteria analysis of wastewater treatment plant design and control scenarios under uncertainty. *Environ. Modell. Softw.* 25, 616–621.
- 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 systems that considers uncertainty. *Environ. Modell. Softw.* 23, 1082–1091.
- Benedetti, L., Bixio, D., Vanrolleghem, P.A., 2006. Assessment of WWTP design and upgrade options: balancing costs and risks of standards' exceedance. *Wat. Sci. Tech.* 54 (6–7), 371–378.

- Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. *Environ. Modell. Softw.* 40, 1–20.
- Beven, K., Binley, A., 1992. Future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* 6, 279–298.
- Beven, K.J., Freer, J.E., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* 249, 11–29.
- Birt, A.G., Valdez-Vivas, M.R., Feldman, R.M., Lafon, C.W., Cairns, D., Coulson, R.N., Tchakerian, M., Xi, W., Guldin, J.M., 2010. A simple stochastic weather generator for ecological modeling. *Environ. Modell. Softw.* 25, 1252–1255.
- Bixio, D., Parmentier, G., Rousseau, D., Verdonck, F., Meirlaen, J., Vanrolleghem, P.A., Thoeye, C., 2002. A quantitative risk analysis tool for design/simulation of wastewater treatment plants. *Wat. Sci. Tech.* 46 (4–5), 301–307.
- Bott, C., Parker, D.S., 2010. WEF/WERF Study Quantifying Nutrient Removal Technology Performance. Water Environment Research Foundation (WERF), Alexandria, VA, USA.
- Bras, R.L., Rodríguez-Iturbe, I., 1985. *Random Functions and Hydrology*. Addison-Wesley, U.S.A.
- Butler, D., 1993. The influence of dwelling occupancy and day of the week on domestic appliance wastewater discharge. *Build. Environ.* 28 (1), 73–79.
- Butler, D., 1991. A small-scale study of wastewater discharges from domestic appliances. *J. Inst. Water Environ. Manag.* 5, 178–185.
- Butler, D., Gatt, K., 1996. Synthesising dry weather flow input hydrographs: a Maltese case study. *Wat. Sci. Tech.* 34 (3–4), 55–62.
- Butler, D., Friedler, E., Gatt, K., 1995. Characterising the quantity and quality of domestic wastewater inflows. *Wat. Sci. Tech.* 31 (7), 13–24.
- Campos, H.M., von Sperling, M., 1996. Estimation of domestic wastewater characteristics in a developing country based on socio-economical variables. *Wat. Sci. Tech.* 34 (3–4), 71–77.
- Carstensen, J., Nielsen, M.K., Strandbaek, H., 1998. Prediction of hydraulic load for urban storm control of a municipal WWTP plant. *Wat. Sci. Tech.* 37 (12), 363–370.
- Corominas, L., Rieger, L., Takács, I., Ekama, G., Hauduc, H., Vanrolleghem, P.A., Oehmen, A., Gernaey, K.V., van Loosdrecht, M.C.M., Comeau, Y., 2010. New framework for standardized notation in wastewater treatment modelling. *Wat. Sci. Tech.* 61 (4), 841–857.
- Copp, J.B., 2002. The COST Simulation Benchmark – Description and Simulator Manual. Office for Official Publications of the European Communities, Luxembourg, ISBN 1-84339-031-0.
- Corradini, C., Flammini, A., Morbidelli, R., Govindaraju, R.S., 2011. A conceptual model for infiltration in two-layered soils with a more permeable upper layer: from local to field scale. *J. Hydrol.* 410, 62–72.
- de Gracia, M., Huete, E., Beltrán, S., Grau, P., Ayesa, E., 2011. Automatic characterisation of primary, secondary and mixed sludge inflow in terms of the mathematical generalised sludge digester model. *Wat. Sci. Tech.* 64 (3), 557–567.
- de Gracia, M., Grau, P., Huete, E., Gómez, J., García-Heras, J.L., Ayesa, E., 2009. New generic mathematical model for WWTP sludge digesters operating under aerobic and anaerobic conditions: model building and experimental verification. *Wat. Res.* 43, 4626–4642.
- 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. Modell. Softw.* 25, 554–561.
- Devisscher, M., Ciacci, G., Fé, L., Benedetti, L., Bixio, D., Thoeye, C., De Gueldre, G., Marsili-Libelli, S., Vanrolleghem, P.A., 2006. Estimating costs and benefits of advanced control for wastewater treatment plants – the MAgIC methodology. *Wat. Sci. Tech.* 53 (4–5), 215–223.
- Dochain, D., Vanrolleghem, P.A., 2001. *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, London, UK.
- Egodawatta, P., Thomas, E., Goonetilleke, A., 2007. Mathematical interpretation of pollutant wash-off from urban road surfaces using simulated rainfall. *Wat. Res.* 41, 3025–3031.
- Flores-Alsina, X., Saagi, R., Lindblom, E., Thirsing, C., Thornberg, D., Gernaey, K.V., Jeppsson, U., 2013a. Calibration and validation of a phenomenological influent pollutant disturbance scenario generator using full-scale data. *Wat. Res.* 47, 172–185.
- Flores-Alsina, X., Saagi, R., Lindblom, E., Thirsing, C., Thornberg, D., Gernaey, K.V., Jeppsson, U., 2013b. Calibration and validation of a phenomenological dynamic influent pollutant disturbance scenario generator using fullscale data. In: Proceedings 11th IWA Conference on Instrumentation, Control and Automation (ICA 2013), 18–20 Sept 2013, Narbonne, France.
- Flores-Alsina, X., Gernaey, K., Jeppsson, U., 2012. Global sensitivity analysis of the BSM2 dynamic influent disturbance scenario generator. *Wat. Sci. Tech.* 65 (11), 1912–1922.
- Flores-Alsina, X., Rodríguez-Roda, I., Sin, G., Gernaey, K., 2008. Multi-criteria evaluation of wastewater treatment plant control strategies under uncertainty. *Wat. Res.* 42, 4485–4497.
- Friedler, E., 2004. Quality of individual domestic greywater streams and its implication for on-site treatment and reuse possibilities. *Environ. Technol.* 25, 997–1008.
- Friedler, E., Brown, D.M., Butler, D., 1996a. A study of WC derived sewer solids. *Wat. Sci. Tech.* 33 (9), 17–24.
- Friedler, E., Butler, D.M., Brown, D., 1996b. Domestic WC usage patterns. *Build. Environ.* 31 (4), 385–392.
- Friedler, E., Butler, D., 1996. Quantifying the inherent uncertainty in the quantity and quality of domestic wastewater. *Wat. Sci. Tech.* 33 (2), 13–24.
- García-Sanz, M., Eguinoa, I., Ayesa, E., Martín, C., 2006. Non-diagonal multivariable robust QFT control of a wastewater treatment plant for simultaneous nitrogen and phosphorus removal. In: Proceedings of the 5th IFAC Symposium on Robust Control Design ROCOND 2006. Toulouse, France, 5–7 July 2006.
- Gernaey, K.V., Flores-Alsina, X., Rosen, C., Benedetti, L., Jeppsson, U., 2011. Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach. *Environ. Modell. Softw.* 26, 1255–1267.
- 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., Jeppsson, U., 2005. Phenomenological modeling of wastewater treatment plant influent disturbance scenarios. In: Proceedings of 10th International Conference on Urban Drainage (10ICUD), 21–26 August, 2005, Copenhagen, Denmark.
- Gevaert, V., Verdonck, F., Benedetti, L., De Keyser, W., De Baets, B., 2009. Evaluating the usefulness of dynamic pollutant fate models for implementing the EU Water Framework Directive. *Chemosphere* 76, 27–35.
- Gevaert, V., De Keyser, W., Benedetti, L., Nilsson, M.-L., Jonsson, A., Lützhøft, H.-C.H., Donner, E., Vanrolleghem, P.A., De Baets, B., Verdonck, F., 2008. Development of a Classification System for Dynamic, Generic Release Patterns and Evaluation of the Pollutant's Exposure Potential. Internal ScorePP Project Deliverable. p. 32. European Commission, Luxembourg.
- Gins, G., Espinosa, J., Smets, I.Y., Van Brempt, W., Van Impe, J.F.M., 2006. Data alignment via dynamic time warping as a prerequisite for batch-end quality prediction. *Lect. Notes Artif. Intell.* 4065, 506–510.
- Giordano, R., Liersch, S., 2012. A fuzzy GIS-based system to integrate local and technical knowledge in soil salinity monitoring. *Environ. Model. Soft.* 36, 49–63.
- Grau, P., Beltrán, S., de Gracia, M., Ayesa, E., 2007. New mathematical procedure for the automatic estimation of influent characteristics in WWTPs. *Wat. Sci. Tech.* 56 (8), 95–106.
- Guerrero, J., Guisasaola, A., Vilanova, R., Baeza, J.A., 2011. Improving the performance of a WWTP control system by model-based setpoint optimization. *Environ. Modell. Softw.* 26, 492–497.
- Günther, F., 2000. Wastewater treatment by greywater separation: outline for a biologically based greywater purification plant in Sweden. *Ecol. Eng.* 15, 139–146.
- Hauduc, H., Rieger, L., Oehmen, A., Comeau, Y., van Loosdrecht, M.C.M., Héduit, A., Vanrolleghem, P.A., Gillot, S., 2013. Critical review of activated sludge modelling: state of process knowledge, modelling concepts and limitations. *Bio-technol. Bioeng.* 110 (1), 24–46.
- Hauduc, H., Neumann, M.B., Muschalla, D., Gamerith, V., Gillot, S., Vanrolleghem, P.A., 2011. Towards quantitative quality criteria to evaluate simulation results in wastewater treatment – a critical review. In: Proceedings 8th International IWA Symposium on Systems Analysis and Integrated Assessment in Water Management (WATERMATEX2011). San. Sebastian, Spain, June 20–22 2011, pp. 36–46.
- Henze, M., 1992. Characterization of wastewater for modeling of activated sludge processes. *Wat. Sci. Tech.* 25 (6), 1–15.
- Henze, M., Gujer, W., Mino, T., van Loosdrecht, M.C.M., 2000. *Activated Sludge Models: ASM1, ASM2, ASM2d and ASM3*. Scientific and Technical Report. IWA Publishing, London, UK.
- Henze, M., Harremoës, P., Jansen, J.C., Arvin, E., 1995. *Wastewater Treatment - Biological and Chemical Processes*. Springer, Berlin, Germany.
- Hulsbeek, J.J.W., Kruij, J., Roelvelde, 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.
- Jeppsson, U., Alex, J., Batstone, D.J., Benedetti, L., Comas, J., Copp, J.B., Corominas, L., Flores-Alsina, X., Gernaey, K.V., Nopens, I., Pons, M.-N., Rodríguez-Roda, I., Rosen, C., Steyer, J.-P., Vanrolleghem, P.A., Volcke, E.I.P., Vrecko, D., 2013. Benchmark simulation models, quo vadis? *Wat. Sci. Tech.* 68 (1), 1–15.
- 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.
- Krueger, T., Page, T., Hubacek, K., Smith, L., Hiscock, K., 2012. The role of expert opinion in environmental modelling. *Environ. Modell. Softw.* 36, 4–18.
- Langergraber, G., Alex, J., Weissenbacher, N., Woerner, D., Ahnert, M., Frehmann, T., Half, N., Hobus, I., Plattes, M., Spering, V., Winkler, S., 2008. Generation of diurnal variation for influent data for dynamic simulation. *Wat. Sci. Tech.* 57 (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.
- Langeveld, J.G., Schilperoord, R.P.S., Weijers, S.R., 2013. Climate change and urban wastewater infrastructure: there is more to explore. *J. Hydrol.* 476, 112–119.
- Lindblom, E., Ahlman, S., Mikkelsen, P.S., 2011. Uncertainty-based calibration and prediction with a stormwater surface accumulation-washoff model based on coverage of sampled Zn, Cu, Pb and Cd field data. *Water Res.* 45 (13), 3823–3835.
- Lindblom, E., Gernaey, K.V., Henze, M., Mikkelsen, P.S., 2006. Integrated modelling of two xenobiotic organic compounds. *Wat. Sci. Tech.* 54 (6–7), 213–221.
- Liu, C., Li, S., Zhang, F., 2011. The oxygen transfer efficiency and economic cost analysis of aeration system in municipal wastewater treatment plant. *Energy Proc.* 5, 2437–2443.

- Londong, J., 1994. Consequences of the behaviour of activated sludge plants with combined sewage inflows. *Wat. Sci. Tech.* 30 (1), 139–146.
- Mannina, G., Cosenza, A., Vanrolleghem, P.A., Viviani, G., 2011. A practical protocol for calibration of nutrient removal wastewater treatment models. *J. Hydroinform.* 13, 575–595.
- Mannina, G., Viviani, G., 2009. Hybrid moving bed biofilm reactors: an effective solution for upgrading a large wastewater treatment plant. *Wat. Sci. Tech.* 60 (5), 1103–1116.
- Mantovan, P., Todini, E., 2006. Hydrological forecasting uncertainty assessment: incoherence of the GLUE methodology. *J. Hydrol.* 330, 368–381.
- Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S., 2003. Markov chain Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. U.S.A.* 100, 15324–15328.
- Martin, C., Ayesa, E., 2010. An integrated Monte Carlo methodology for the calibration of water quality models. *Ecol. Model.* 221, 2656–2667.
- Martin, C., Neumann, M.B., Vanrolleghem, P.A., 2012. A tool for optimum design of WWTPs under uncertainty: estimating the Probability of Compliance. In: *Proceedings of: 6th International Congress on Environmental Modelling and Software (iEMSS). Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig (Germany)*, 1–5 July 2012, ISBN 978-88-9035-742-8. In: <http://www.iemss.org/society/index.php/iemss-2012-proceedings>.
- Martin, C.M., Eguinoa, I., McIntyre, N.R., Garcia-Sanz, M., Ayesa, E., 2007. ARMA models for uncertainty assessment of time series data: application to Galindo-Bilbao WWTP. In: *Proceedings of Watermatex2007*, 7–9 May, 2007, Washington, USA.
- Matott, L.S., Babendreier, J.E., Purucker, S.T., 2009. Evaluating uncertainty in integrated environmental models: a review of concepts and tools. *Water Res. Res.* 45 (6), 1–14.
- 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 Modelling*. Water Environment Research Foundation (WERF), Alexandria, VA, USA.
- Moriassi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* 50 (3), 885–900.
- Nopens, I., Benedetti, L., Jeppsson, U., Pons, M.N., Alex, J., Copp, J.B., Gernaey, K.V., Rosen, C., Steyer, J.P., Vanrolleghem, P.A., 2010. Benchmark simulation model No 2: finalisation of plant layout and default control strategy. *Wat. Sci. Tech.* 62 (9), 1967–1974.
- Omlin, M., Reichert, P., 1999. A comparison of techniques for the estimation of model prediction uncertainty. *Ecol. Model.* 115, 45–59.
- Ort, C., Lawrence, M.G., Reungoat, J., Mueller, J.F., 2010. Sampling for PPCPs in sewers: comparison of different sampling modes and optimization strategies. *Environ. Sci. Tech.* 44 (16), 6289–6296.
- Ort, C., Schaffner, C., Giger, W., Gujer, W., 2005. Modeling stochastic load variations in sewer systems. *Wat. Sci. Tech.* 52 (5), 113–122.
- Ostace, G.S., Cristea, V.M., Agachi, P.S., 2011. Cost reduction of the wastewater treatment plant operation by MPC based on modified ASM1 with two-step nitrification/denitrification model. *Comput. Chem. Eng.* 35, 2469–2479.
- Palynchuk, B.A., Guo, Y., 2011. A probabilistic description of rain storms incorporating peak intensities. *J. Hydrol.* 409, 71–80.
- 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. Hydroinform.* 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. Modell. Softw.* 22, 1208–1216.
- Reichert, P., Mieleitner, J., 2009. Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters. *Water Resour. Res.* 45, W10402.
- Richards, L.A., 1931. Capillary conduction of liquids through porous mediums. *Physics* 1, 318–333.
- Rieckermann, J., Anta, J., Scheidegger, A., Ort, C., 2011. Assessing wastewater micropollutant loads with approximate Bayesian computations. *Environ. Sci. Technol.* 45 (10), 4399–4406.
- Rieger, L., Gillot, S., Langergraber, G., Ohtsuki, T., Shaw, A., Takács, I., Winkler, S., 2012. *Guidelines for Using Activated Sludge Models*. Scientific and Technical Report No.22. IWA Publishing, London, UK.
- Rieger, L., Takács, I., Villez, K., Siegrist, H., Lessard, P., Vanrolleghem, P.A., Comeau, Y., 2010a. Data reconciliation for WWTP simulation studies – planning for high quality data and typical sources of errors. *Water Environ. Res.* 82, 426–433.
- Rieger, L., Takács, I., Shaw, A., Winkler, S., Ohtsuki, T., Langergraber, G., Gillot, S., 2010b. Editorial: status and future of wastewater treatment modelling. *Wat. Sci. Tech.* 61 (4), 821–823.
- Rivas, A., Irizar, I., Ayesa, E., 2008. Model-based optimisation of wastewater treatment plants design. *Environ. Modell. Softw.* 23, 435–450.
- Rodriguez-García, G., Molinos-Senante, M., Hospido, A., Hernández-Sancho, F., Moreira, M.T., Feijoo, G., 2011. Environmental and economic profile of six typologies of wastewater treatment plants. *Wat. Res.* 45, 5997–6010.
- Roeleveld, P., van Loosdrecht, M.C.M., 2002. Experiences with guidelines for wastewater characterisation in the Netherlands. *Wat. Sci. Tech.* 45 (6), 77–87.
- 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.
- Rousseau, D., Verdonck, F., Moerman, O., Carrette, R., Thoeye, C., Meirlaen, J., Vanrolleghem, P.A., 2001. Development of a risk assessment based technique for design/retrofitting of WWTPs. *Wat. Sci. Tech.* 43 (7), 287–294.
- Rowan, J.S., Greig, S.J., Armstrong, C.T., Smith, D.C., Tierney, D., 2012. Development of a classification and decision-support tool for assessing lake hydromorphology. *Environ. Modell. Softw.* 36, 86–98.
- Schilperoord, R.P.S., 2011. *Monitoring as a Tool for the Assessment of Wastewater Quality Dynamics*. PhD thesis. Water Management Academic Press, Delft, the Netherlands. ISBN/EAN 978-90-8957-021-5.
- Siegrist, H., Koch, G., Kühni, M., 1999. Weitergehende Massnahmen auf der Kläranlage Werdhölzli. Untersuchung der Denitrifikation und der biologischen Phosphorelimination auf den Kläranlagen Werdhölzli und Neugut. EAWAG-Report, Zurich, Switzerland (in German).
- Sin, G., Van Hulle, S.W.H., De Pauw, D.J.W., van Griensven, A., Vanrolleghem, P.A., 2005. A critical comparison of systematic calibration protocols for activated sludge models: a SWOT analysis. *Wat. Res.* 39, 2459–2474.
- Sollfrank, U., Gujer, W., 1991. Characterization of domestic wastewater for mathematical modelling of the activated sludge process. *Wat. Sci. Tech.* 23 (4–6), 1057–1066.
- Spanjers, H., Vanrolleghem, P.A., 1995. Respirometry as a tool for rapid characterization wastewater and activated sludge. *Wat. Sci. Tech.* 31 (2), 105–114.
- Sperandio, M., Paul, E., 2000. Estimation of wastewater biodegradable COD fractions by combining respirometric experiments in various So/Xo ratios. *Wat. Res.* 34, 1233–1244.
- Spring, V., Alex, J., Langergraber, G., Ahnert, M., Half, N., Hobus, I., Weissenbacher, N., Winker, S., Yücesoy, E., 2008. Using dynamic simulation for design of activated sludge plants. In: *Proceedings of the International Symposium on Sanitary and Environmental Engineering-SIDISA08*, 24–27 June 2008, Florence, Italy.
- STOWA, 1999. *Influent Characterisation of Raw and Pre-treated Wastewater*. Report No. 99–13. STOWA-report 97–23, Hageman Fulfilment, Zwijndrecht, The Netherlands (in Dutch).
- STOWA, 1996. *Methods for Wastewater Characterization, Inventory and Guidelines*. STOWA-report 97–23, Hageman Fulfilment, Zwijndrecht, The Netherlands (in Dutch).
- Talebizadeh, M., Belia, E., Vanrolleghem, P.A., 2014. Influent generator for probabilistic design of nutrient removal wastewater treatment plants. In: *Proceedings of the 4th International Seminar on Wastewater Treatment (WWT) Modelling (WWTmod2014)*, 30 March–2 April, 2014, Spa (Belgium).
- Thauvin, V., Gaume, E., Roux, C., 1998. A short time-step point rainfall stochastic model. *Wat. Sci. Tech.* 37 (11), 37–45.
- Vanhooren, H., Nguyen, K., 1996. *Development of a simulation Protocol for Evaluation of Respirometry-based Control Strategies*. Technical Report. Ghent University, Belgium and University of Ottawa, Canada, p. 33.
- Vanrolleghem, P.A., Insel, G., Petersen, B., Sin, G., De Pauw, D., Nopens, I., Weijers, S., Gernaey, K., 2003. A comprehensive model calibration procedure for activated sludge models. In: *Proceedings: WEFTEC 2003, 76th Annual Technical Exhibition and Conference*, 11–15 October, 2003, Los Angeles, CA, USA.
- Vanrolleghem, P.A., Jeppsson, U., Carstensen, J., Carlsson, B., Olsson, G., 1996. Integration of wastewater treatment plant design and operation – a systematic approach using cost functions. *Wat. Sci. Tech.* 34 (3–4), 159–171.
- Verbanck, M., 1995. Variabilité des charges solides en suspension à l'exutoire des réseaux de collecte. *Journée d'étude CB-IAWQ 31 (May 1995)*. Liège, Belgium (in French).
- Villez, K., Steppe, K., De Pauw, D.J.W., 2009. Use of Unfold PCA for on-line plant stress monitoring and sensor failure detection. *Biosyst. Eng.* 103, 23–34.
- Wang, L., Wei, J., Huang, Y., Wang, G., Maqsood, I., 2011. Urban nonpoint source pollution buildup and washoff models for simulating storm runoff quality in the Los Angeles County. *Environ. Pollut.* 159, 1932–1940.
- Willems, P., 2001. A spatial rainfall generator for small spatial scales. *J. Hydrol.* 252, 126–144.
- Yong, M., Yongzhen, P., Jeppsson, U., 2006. Dynamic evaluation of integrated control strategies for enhanced nitrogen removal in activated sludge processes. *Control Eng. Pract.* 14, 1269–1278.