CODEN:LUTEDX/(TEIE-7231)/1-73/(2005)

Industrial Electrical Engineering and Automation

BSM2: A Model for Dynamic Influent Data Generation

Krist V. Gernaey, Christian Rosén, Ulf Jeppsson

Dept. of Industrial Electrical Engineering and Automation
Lund University

BSM2: A model for dynamic influent data generation

Technical report

Krist V. Gernaey*, Christian Rosen and Ulf Jeppsson

Dept. of Industrial Electrical Engineering and Automation
Lund University, Lund, Sweden

15 March 2005

* For questions about the model described in this document: krist.gernaey@iea.lth.se

Contents

Contents 2

1. Introduction 4

2. Influent characteristics 5

3. Model for influent flow rate 6

3.1. Introduction: design principles 6

3.2. 'Households' model block 7

3.3. 'Industry' model block 10

3.4. 'Seasonal correction infiltration' model block 12

3.5. 'Soil' model block 14

3.6. 'Sewer' model block 16

3.7. Dry weather influent flow rate generation: an example 19

3.8. 'Rain Generator' model block 23

3.9. Comparison of influent model rain generator with rainfall data 25

3.9.1. Rain distribution for data generated with rain generator 25

3.9.2. Rain distribution for Helsingør rain meter data 27

3.9.3. General discussion of results 29

3.10. Conclusions on the first design step of the influent flow rate model 30

4. Model for influent component concentrations 33

4.1. Influent component concentration design principles 33

4.2. 'Households pollutants' model block 34

4.2.1. Model block 34

4.2.2. Design of 'Household pollutants' model block input files 35

4.2.3. Design of 'Household pollutants' model block noise generators 37

4.3. 'Industry pollutants' model block 38

4.3.1. Model block 38

4.3.2. Design of 'Industry pollutants' model block input files 38

4.3.3. Design of 'Industry pollutants' model block noise generators 40

4.4. 'BSM1_fractionator' model block 41

4.5. 'Noise generator' model block 42

4.6. 'Sewer' model block 43

4.7. Dry weather simulation results 43

4.7.1. Simulation without noise generation 44

4.7.2. Simulation with noise generation 48

4.8. 'First flush effect generation' model block 51

4.8.1. Selection of an appropriate switch function 51

4.8.2. A simple model to represent first flush effects 52

4.9. Simulation results full influent model 53

4.10. Final tuning of input files 58

5. Model for temperature variations 62

6. A final BSM2 influent design proposal 64

6.1. Parameter changes for generating the BSM2 influent 64
6.2. Final influent model proposal: simulation 65
7. Acknowledgements *67*
8. Symbols *68*
9. References *70*
Appendix 1. List of influent model files *71*

1. Introduction

The IWA/COST benchmark system (www.benchmarkwwtp.org) is a protocol that allows objective comparison of the effectiveness of control strategies in biological nitrogen removal activated sludge plants. The first benchmark implementation (Copp, 2002), Benchmark Simulation Model no. 1 (BSM1), is a success. This is illustrated by the large number of scientific papers – more than 100 according to Jeppsson and Pons (2004) – using the benchmark or part of the benchmark (e.g. influent files, or plant performance evaluation criteria, etc.).

The evaluation of control strategies in BSM1 is done based on 3 different 'weather files', corresponding to dry, storm and rain weather disturbance scenarios. Basically, for each of these influent disturbance scenarios 1 week of data is used to evaluate the impact of a proposed control strategy on the simulated plant performance. There is a general consensus that 1 week of data is not sufficient to evaluate wastewater treatment plant (WWTP) controller performance, especially not when 'slow' actuators such as the waste sludge flow rate are manipulated. At the December 2003 meeting of the active benchmarkers (Lund, Sweden), it was decided to increase the control strategy evaluation period from 1 week to 1 year of data, with an extra 6 months of data needed to simulate the plant to a dynamic 'quasi' steady state before the start of the control strategy evaluation period. This extension of the control strategy evaluation period will be part of the Long-Term Benchmark Simulation Model no. 1 which is developed as an extension to BSM1, BSM1_LT (Rosen *et al.*, 2004), and of the new BSM1 based plant-wide definition of the simulation benchmark, Benchmark Simulation Model no. 2 (BSM2) (Jeppsson *et al.*, 2004). With respect to influent characteristics, the main difference between BSM1_LT and BSM2 is that BSM2 includes a primary clarifier, whereas BSM1_LT does not. It will be assumed here that the same influent model can be used for both BSM1_LT and BSM2, where the BSM1_LT influent corresponds to the BSM2 influent after its passage through the primary clarifier.

This document summarises the work that has been done at IEA in the period April 2004 to March 2005 with respect to the development of models that can generate 1.5 years of influent data for BSM1_LT and BSM2 (half a year for the start up of the simulations, one year for the control strategy evaluation period). The document can be used as a basis for the further development of long term benchmarking efforts. This technical report will focus on the generation of influent data for both BSM1_LT and BSM2, and should be accompanied by a set of Matlab and Simulink files. The report is quite extended, since it aims (1) at explaining the motivation for the structure of specific influent model components, (2) at illustrating the different features built into the model, and (3) at comparing a number of specific influent characteristics of the model influent with data obtained from full-scale systems.

At this moment, parts of the model described in this document are submitted as conference papers. When using the influent model in own research work, you are asked to refer to the ICA2005 conference paper (Gernaey *et al.*, 2005b), where general model principles are outlined. A second paper with focus on the different model blocks implemented to model the influent flow rate dynamics is currently under review (Gernaey *et al.*, 2005a). In case of questions about the influent model files and how to use them: krist.gernaey@iea.lth.se.

2. Influent characteristics

Most of the influent characteristics that appeared to be desirable for creating a realistic control strategy evaluation framework in BSM1_LT and BSM2 are described in Jeppsson *et al.* (2004). These characteristics are repeated here, except for the generation of truckloads of waste for the anaerobic digester.

Contrary to BSM1, the influent for BSM1_LT and BSM2 will be based on a model rather than defined by data files. The influent model should include the typical phenomena that are observed in a year of full-scale WWTP influent data. For dry weather the influent model should include the following phenomena (Jepsson *et al.*, 2004):

(1) Diurnal phenomena, which can for example be modelled using a second-order harmonic function
(2) A weekend effect, consisting of a lower average flow rate and pollutant concentration during weekends compared to normal week days, in an attempt to simulate a WWTP that receives a mixed municipal-industrial wastewater
(3) Seasonal phenomena reflecting typical effects from the sewer system and urban drainage, i.e. increased infiltration in the rainy season compared to the dry season, related to higher ground water levels;
(4) Holiday periods during which time a lower average wastewater flow rate is maintained over a period of several weeks.

The dry weather model should be combined with rain and storm weather generation, to account for 'first flush' effects from the sewer network and dilution phenomena that are typically observed at full-scale WWTPs. In addition to flow rate and ASM1 model components, the influent model will include wastewater temperature as an additional variable for describing seasonal temperature variations.

Besides the authors of this report and the participants in the last WG1 COST meeting in Aix-en-Provence, a number of discussions with Jens Alex (IFAK, Germany), Lorenzo Benedetti (BIOMATH, Belgium), Martijn Devisscher (Aquafin NV, Belgium) and Peter Vanrolleghem (BIOMATH, Belgium) have substantially contributed to the development of the proposed benchmark influent model. Based on these inputs, the results of the work at IEA will be introduced and discussed stepwise: (1) model for influent flow rate; (2) model for influent component concentrations; (3) model for seasonal temperature variations. This detailed description includes a summary of the main features of the proposed influent. Finally, the final design of the BSM1_LT/BSM2 influent files is made.

All models were implemented and tested in Matlab/Simulink, respecting 3 basic rules: model parsimony, model transparency, and model flexibility. Limiting the number of parameters was evaluated to be of primary importance for the user to accept this modelling tool. It was attempted to reach model transparency by using model parameters that have a physical meaning, such that the future user of the model can easily reach a sufficient level of understanding to apply the model on other case studies. Model flexibility was needed, because this tool for the generation of influent data can also be used or extended for other applications that are not related to the simulation benchmark, but where long influent time series are also needed. Moreover, it is also possible to replace part of the model with data, for example the use of rainfall time series data instead of the rain generator that will be described in this report.

3. Model for influent flow rate

A first influent flow rate design step is based on the BSM1 influent, mainly aiming at evaluating the dynamics that can be generated with the different model blocks. In a second influent design step (see 6) the influent loads are adjusted to take into account the effect of the primary clarifier, and a final BSM1_LT/BSM2 influent model/file is proposed.

3.1. Introduction: design principles

The model blocks that were implemented to achieve the generation of influent flow rate profiles are shown in Figure 1, and will be described in more detail below. However, before going too much into details, a few design considerations will be provided for clarity.

Using the flow-weighted average dry weather influent composition of BSM1 as a reference (Q_{av} = 18 446 m^3/d; Q_{av} on weekdays = 19 346.3 m^3/d; Q_{av} on weekend days = 16 196.3 m^3/d), the following initial influent design assumptions were made:
On normal weekdays, 25% of the BSM1 Q_{av} (19 346.3 m^3/d) results from infiltration, corresponding to 4 836.6 m^3/d. This average infiltration influent flow rate was rounded off to 4 800 m^3/d. The remaining 75% of the BSM1 Q_{av} corresponding to normal week days is assumed to be distributed over household wastewater (62%, or 11 994.7 m^3/d) and industrial wastewater (13%, or 2 515.0 m^3/d). Assuming a wastewater production of 0.15 m^3/d for each person equivalent in the household wastewater, the household wastewater flow rate of 11 994.7 m^3/d corresponds to 79 965 person equivalents (p.e.). To make the numbers a bit easier to remember, this value was rounded off to 80 000 p.e., corresponding to an average flow rate of 12 000 m^3/d resulting from household wastewater production. Similarly, the average industrial wastewater production on week days was rounded off to 2 500 m^3/d. With respect to flow rate dynamics, it was attempted to approach the influent flow rate dynamics of the BSM1 dry weather file, meaning for example that it was attempted to have two daily flow rate peaks (see 3.7).

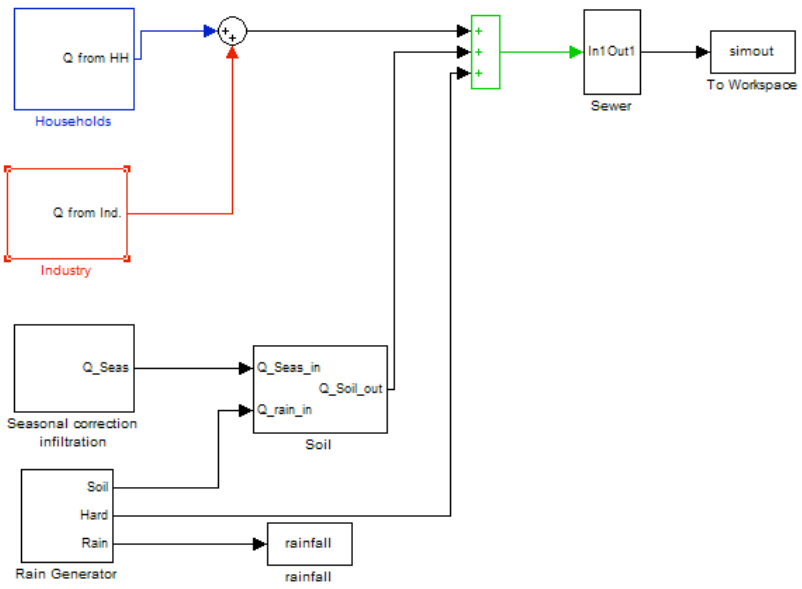

Figure 1. General overview of the influent flow rate model in Simulink

As illustrated in Figure 1, the generation of influent flow rate profiles was achieved in Simulink by combining contributions from households, industry, infiltration and rain. Rainfall

will contribute to the total flow rate in two ways: the largest part of the rainfall contribution to the flow rate is assumed to originate from run off from impervious surfaces, and is thus transported directly to the sewer. Rainfall on pervious surfaces will influence the groundwater level, and thus the contribution of infiltration to the influent flow rate. Assuming that there is a dry and a rainy season, the 'Seasonal correction infiltration' model block will create this seasonal effect. This seasonal correction is combined with the rainfall assumed to fall on pervious surfaces, and the sum of both flows is passed through the soil model block. Afterwards, the net contribution of infiltration, an output from the soil model block, is combined with the overall flow rate resulting from industry and households and the flow rate contribution from rainfall on impervious surfaces. The resulting flow rate is finally passed through a simple sewer system model. Each model block will be commented in detail below, and will be illustrated using simulation results that are based on the assumptions made in this first influent design step. In a second phase of the influent flow rate design, while implementing models for the BSM2 influent concentration profiles, a modification to the initial influent flow rate design will be needed (see 6).

3.2. 'Households' model block

Details of the 'Households' model block are shown in Figure 2. When it comes to dynamics in the generated influent data, this block will to a large extent contribute to the generation of diurnal influent flow rate variations. This is achieved by calling a user-defined data file from the workspace (day_HS), containing a profile corresponding to the diurnal variation of the wastewater production of one p.e. (see Figure 3A, one value per hour). The values in day_HS are normalised (average = 1). Multiplication with a gain QperPE, (150 liter/d/p.e., see 3.1) results in diurnal influent flow rate dynamics expressed in l/d (see Figure 3B). A second data file in the workspace (week_HS) contains information on the weekly household flow rate pattern (see Figure 3C, one value per day, dimensionless), and is used to create a weekend effect on the household flow rate data. In this case, the weekend effect is a slight reduction of the household wastewater production (8% reduction on Saturday, 12% on Sunday) compared to normal week days. A manual switch block allows deselecting the weekend effect (see Figure 2), which is helpful for example in testing the model. If selected, week_HS values will be multiplied with the diurnal influent flow rate profile, thus superimposing the weekend effect on the diurnal profiles. The effect of including this weekend effect is illustrated in Figure 4, for the case without and with noise added to the 'Households' model block respectively. Similar to the weekend effect, a holiday effect can also be added on to the household flow rate data via a data file in the workspace (year_HS). It is assumed that the holiday effect results in a decrease of the influent flow rate during 3 weeks, starting at the end of week 4 (see Figure 3D). The holiday effect results in a 25% decrease of the flow rate during the first two weeks, and a 12% decrease during the third holiday week.

It is left up to the model user whether noise is added or not. If noise is added, a random number is multiplied with a gain (this all takes place in the 'Random Number Flow Rate' block in Figure 2) and added to the flow rate values that were obtained by combining the diurnal, weekly and yearly variations of the wastewater production of one p.e. The amplitude of the noise signal can be selected by the model user. Note that this type of noise generation will be used in several other model blocks. Therefore, all random number generators that are used in the influent model generally start up with a different seed number, to avoid that two random number sequences are identical. A saturation block is inserted to avoid negative values, and to avoid unrealistically high flow rates. Finally, the wastewater production is multiplied with a gain PE, corresponding to the number of p.e. in the catchment area that is considered (see also Table 2). The parameter 'pop' in the 'Gain2' model block in Figure 2

acts as a switch, i.e. the complete influent model can be simulated with (pop = 100%) or without (pop = 0%) the contribution of the household wastewater, which is helpful in model testing, and furthermore leaves the flexibility to omit household wastewater contributions from a user-defined influent generation scenario.

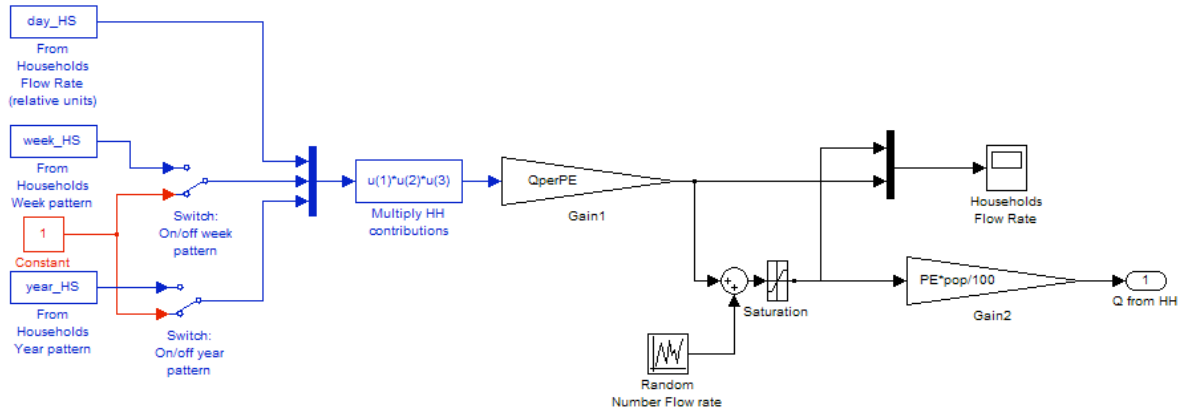

Figure 2. 'Households' model block, with diurnal flow rate variations (day_HS block), weekend effects (week_HS block) and holiday effects (year_HS block)

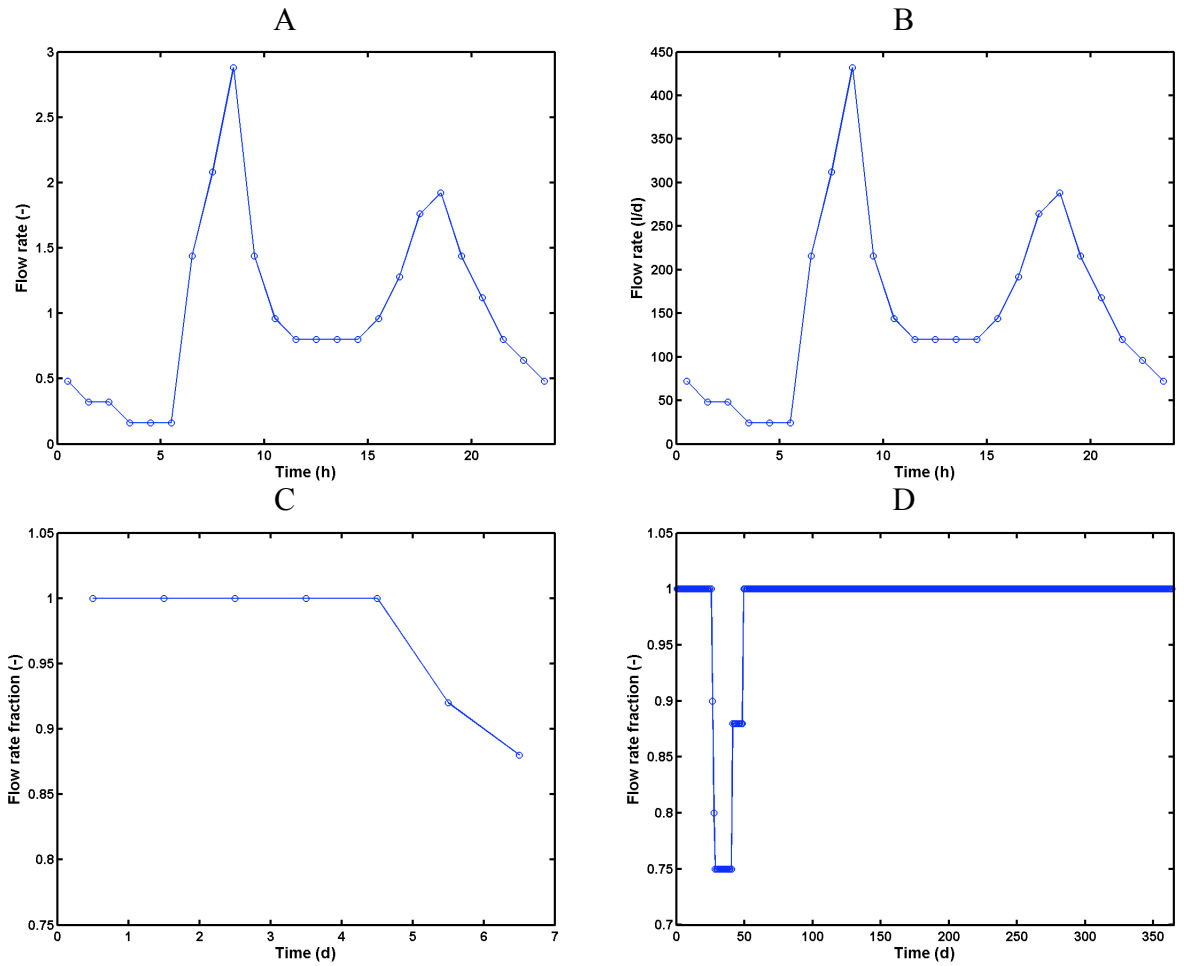

Figure 3. A: Diurnal variation of the household wastewater production (day_HS in Figure 2); B: Diurnal flow rate variation in l/d units; C: Weekly variation of the household water consumption (weekend effect; week_HS in Figure 2, one value per day); C: Yearly variation of the household water consumption (holiday effect; year_HS in Figure 2, one value per day). Note that the first day in figure C corresponds to July 1st

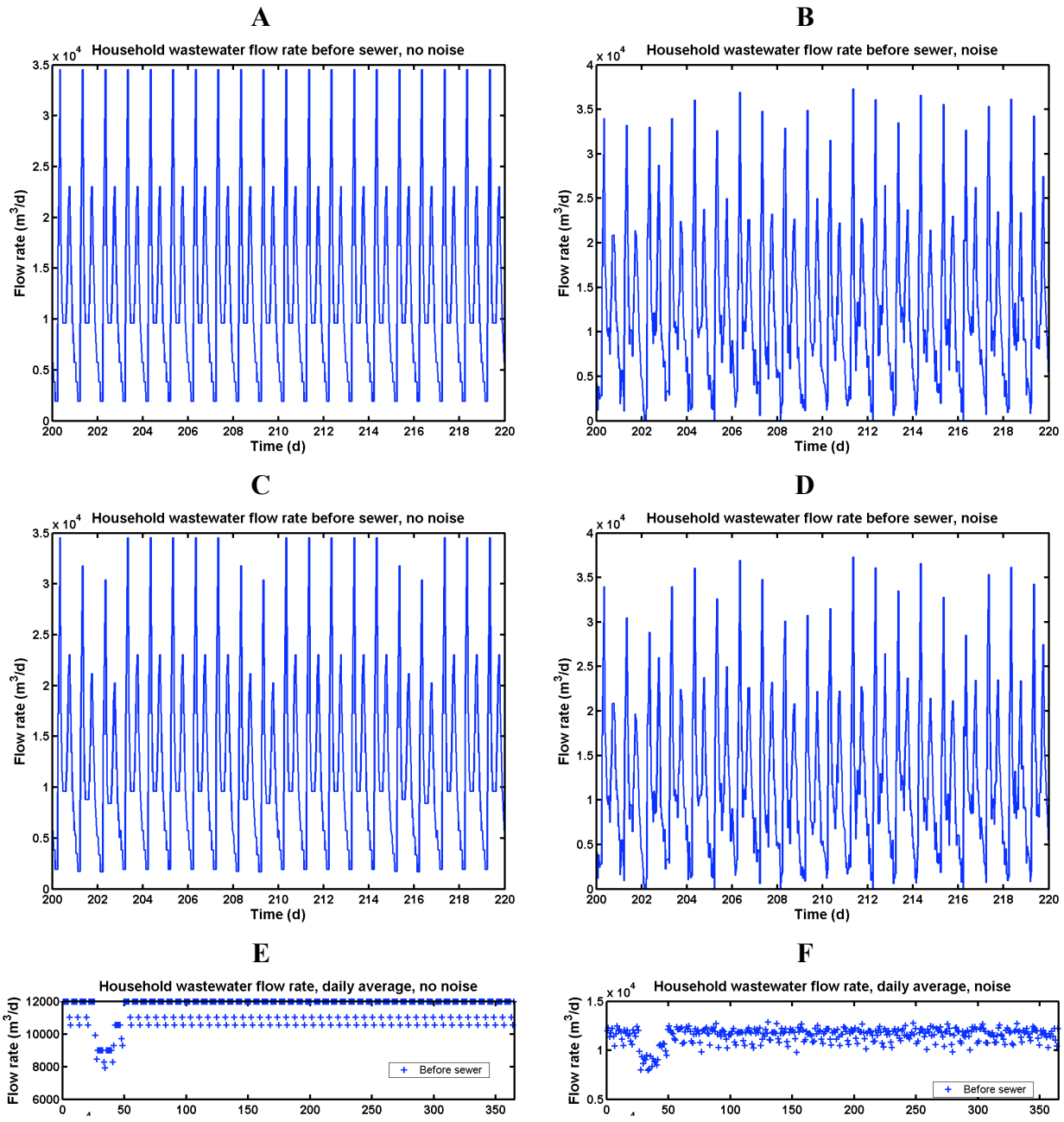

Figure 4. Dynamic flow rate profiles (15 min. sampling interval) resulting from the 'Households' model block: A: Only diurnal effect, no noise added; B: As A, but with noise added in the 'Households' model block; C: Diurnal effect combined with weekend effect, no noise added; D: As C, but with noise added; E: Daily average flow rates (each average daily flow rate value represents the mean of 96 samples), combining diurnal, weekend and holiday effect, no noise added, F: As E, but with noise added. Note that the first day in E and F corresponds to July 1st

For clarity, it is tried to illustrate the operation of the 'Households' model block with a number of figures that were made based on outputs of that model block. In other words, the graphs presented here have not been passed through the 'Sewer' model block (see Figure 1). Figure 4 provides an example of the diurnal flow rate profiles generated by the 'Households' model block. In Figure 4A and B only diurnal effects were considered, i.e. the weekend effect and the holiday effect were switched off. The effect of including noise is quite clear, when comparing the diurnal influent flow rate profiles without noise (Figure 4A) and with noise added to the 'Households' model block (Figure 4B). All dynamic simulation results that will

be presented in this report have a sampling interval of 15 minutes, similar to BSM1, except when explicitly mentioned otherwise. Clearly, adding noise will result in a data set that is more challenging when it comes to developing appropriate control strategies for the benchmark plant. When applying noise generation on different influent time series it will also allow to some extent to reduce the correlation between these influent time series.

In Figure 4C the diurnal influent flow rate profiles were combined with the weekend effect (no noise added). The lower wastewater flow rate from households during the weekends (see also Figure 3B) has a rather pronounced effect, for example on the value of the daily maximum influent flow rate. Again, the effect of adding noise to the 'Households' model block appears clearly when comparing Figure 4D with Figure 4C.

The holiday effect (see also Figure 3C) is illustrated in Figure 4E and F. In this case, average daily flow rates were used for clarity. Each average daily flow rate value represents the mean of 96 samples. The holiday effect results in a significant reduction of the influent flow rate. The data presented in Figure 4F correspond to the final contribution of the 'Households' model block to the influent flow rate before the influent is passed through the 'Sewer' model block. Without the noise generators, the yearly average wastewater flow rate from households is 11 513 m^3/d, i.e. a decrease of 487 m^3/d due to weekend and holiday effects compared to the design flow rate of 12 000 m^3/d.

3.3. 'Industry' model block

The industrial contribution to the influent flow rate is generated similarly to the 'Households' model block. The 'Industry' model block (Figure 5) will also call two user-defined data files from the workspace (week_IndS, and year_IndS). The file week_IndS contains a profile corresponding to the weekly variation of the industrial wastewater production (one value for each 4 hour period), which includes diurnal variations, a weekend effect and a flow rate peak on Friday afternoon. Contrary to the 'Households' model block the diurnal effect and the weekend effect were not split up in different data files for the 'Industry' model block, precisely because it is assumed that not all weekdays are similar due to the Friday afternoon flow rate peak. It was indeed agreed at the last WG1 COST meeting in Aix-en-Provence (May 2004) that the influent model should contain a 'Friday afternoon peak' corresponding to industrial cleaning at the end of the working week, with an increased wastewater flow rate and pollutant flux. This effect has been included, as will be illustrated in more detail below. Furthermore, as agreed at the WG1 COST meeting in Aix-en-Provence the effect of the variation of the industrial activity during a working day (not all industries have day and night shifts) is also included in the influent file. The dynamic flow rate pattern of week_IndS is multiplied with a gain QInd_weekday (2 500 m^3/d; see 3.1), similar to the generation of diurnal flow rate patterns in the 'Households' model block.

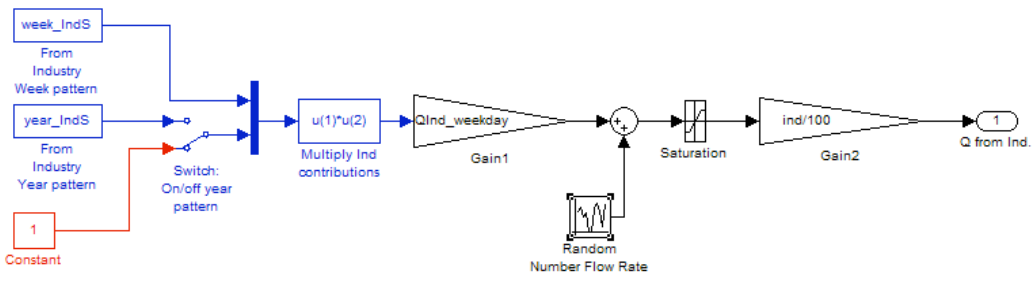

Figure 5. 'Industry' model block

The weekly industrial wastewater production flow rate pattern, obtained by multiplying the week_IndS pattern with QInd_weekday, is illustrated in Figure 6. On normal weekdays (in this case Monday to Thursday), the average industrial wastewater production resulting from the model is 2 500.0 m^3/d, corresponding to the influent design assumption (see 3.1). During weekend days, the industrial wastewater production is reduced to an average flow rate of 583.3 m^3/d. This creates a weekend effect in the flow rate data. During the Friday afternoon peak the maximum industrial wastewater production is 4 000.0 m^3/d, and as a consequence the average flow rate on Fridays is slightly higher compared to normal weekdays (2 645.8 m^3/d). The model user has of course the possibility here to define a different profile, when that is desirable. The industrial wastewater production profile can also be defined in more detail, e.g. one data point every hour, if needed.

The holiday effect on the industrial wastewater production is added on similarly to the 'Households' model block, and can be switched off if desirable. The holiday effect is illustrated in Figure 6: the industrial wastewater production is reduced with 80% during 3 weeks starting at the end of week 4, and corresponding to the 3 weeks with holiday effect in the 'Households' model block. In addition, the industrial wastewater production is reduced with 70% during week 26, to simulate shutdown of industrial activity in the Christmas period.

Noise is added similar to the 'Households' model block, by adding a random number to the industrial wastewater flow rate, and subsequently passing the result through a saturation model block. The parameter 'ind' in the 'gain' model block in Figure 5 is a switch, similar to the 'pop' parameter in the 'Households' model block.

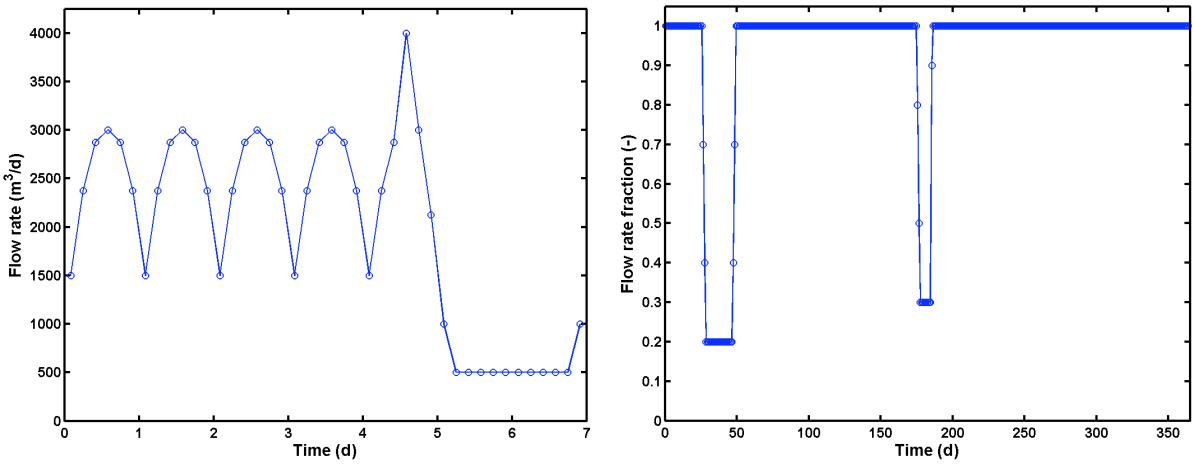

Figure 6. Left: Weekly variation of the industrial wastewater production (week_IndS in Figure 5; one value every 4 hours), including a diurnal effect on weekdays, a weekend effect and a Friday afternoon peak effect. Right: Holiday effect on industrial wastewater production (year_IndS in Figure 5; one value per day). Note that the first day corresponds to July 1st

Figure 7 provides an example of the weekly flow rate profiles generated by the 'Industry' model block. As for the 'Household' model block, the graphs represent time series data that have not been passed through the 'Sewer' model block. The weekend effect is quite pronounced, but will of course be less visible when all the influent sources are combined into the total flow rate entering the sewer system. Again, the effect of adding noise is quite clear (e.g. compare Figure 7A and Figure 7B), and results in a data set where there is a variation of the contribution of industry to the instantaneous influent flow rate.

The holiday effect on the industrial wastewater production is illustrated in Figure 7C. The holiday periods can be recognised directly as longer periods with a low average industrial wastewater flow rate. The weekends (low daily average influent flow rate) can also be distinguished clearly from the week days (high daily average influent flow rate). The data presented in Figure 7C correspond to the final contribution of the 'Industry' model block to the influent flow rate before the influent is passed through the 'Sewer' model block. Without the noise generators, the yearly average wastewater flow rate from industry is 1 845.0 m^3/d, which is significantly lower than the design flow rate of 2 500 m^3/d due to weekend and holiday effects.

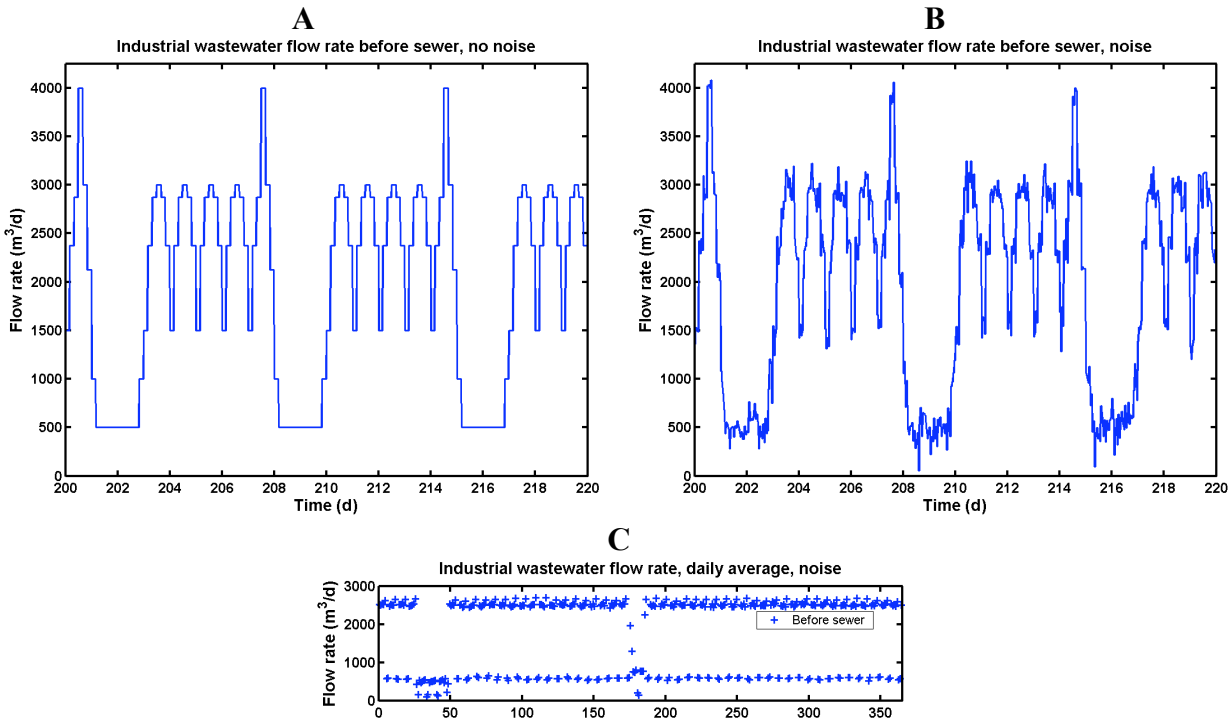

Figure 7. Flow rate profiles resulting from the 'Industry' model block. A. Weekly patterns, no noise added to 'Industry' model block; B: Same as A, but with noise added; C. Daily average flow rates (each average daily flow rate value represents the mean of 96 samples), combining diurnal, weekend and holiday effect, with noise added. Note that the first day corresponds to July 1st

3.4. 'Seasonal correction infiltration' model block

Details of the 'Seasonal correction infiltration' model block are shown in Figure 8. With respect to the dynamics in the generated influent data, this model block will be responsible for the generation of seasonal influent flow rate variations, and will consequently also result in seasonal variations in the influent pollutant concentrations. The seasonal changes of the amount of infiltration are attributed to changes in the groundwater levels over the year: during the rainy season the groundwater level is high, resulting in high infiltration. During the dry season the groundwater level is low, resulting in a lower infiltration. These seasonal effects are assumed to be the result of seasonal, temperature-related changes of the amount of evaporation. As mentioned earlier (see 3.1), it was assumed that the infiltration represents 25% of the average BSM1 dry weather influent flow rate on weekdays, corresponding to a flow rate of 4 800 m^3/day. This effect will be obtained by combining the 'Seasonal correction infiltration' model block with the 'Soil' model block.

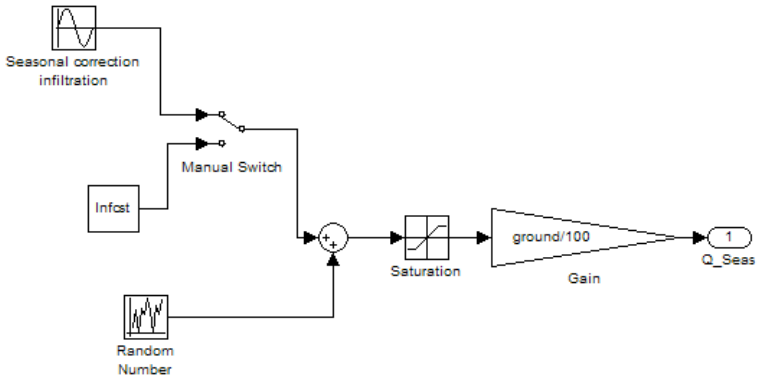

Figure 8. 'Seasonal correction infiltration' model block

The 'Manual Switch' model block in Figure 8 allows either to select a constant value for the 'seasonal correction infiltration' or a seasonal variation. This will then result in a constant infiltration or a seasonal variation of the infiltration in the 'Soil' model block (see 3.5), respectively. The seasonal variation is implemented as a sine function, with an average level of 7 100 m^3/d, an amplitude of 1 200 m^3/d, and a frequency of (2π/364) rad/day. Note indeed that one year of data is assumed to consist of 52 weeks of data, or 364 days, in the influent model. Modelling a year of data as 364 days does not substantially modify the length of the training and evaluation period, whereas handling of seed files via 'cyclic repetition' in Simulink is facilitated considerably because a year now corresponds to an exact number of weeks.

The flow rate profile resulting from the 'Seasonal correction infiltration' model block is illustrated in Figure 9. Note that the flow rate in that figure is an output of the 'Soil' model block, since it is the combination of the 'Seasonal correction infiltration' and the 'Soil' model blocks that generates the infiltration water flow rate entering the sewer system (see also Figure 1 for details on the data flow between the model blocks). An average water flow rate from infiltration of 4 858.0 m^3/d results from the parameters selected for the 'Seasonal correction infiltration' and the 'Soil model blocks (Figure 9 left). This is close to the design value of 4 800 m^3/d. The minimum infiltration (3 759.8 m^3/d), corresponding to the dry season, is reached at day 26, whereas maximum infiltration (5 944.1 m^3/d) is reached at day 209. Note again that day 1 corresponds to July 1st. When modifying the amplitude of the sine wave to 2 400 m^3/d (see Figure 9 right), the average water flow rate from infiltration (4 841.7 m^3/d) is again close to the design value, with a minimum flow rate of 2 631.9 and a maximum of 7 002.7 m^3/d.

It is again up to the model user to define other parameters, in case that is desirable. Noise generation is implemented similarly to the 'Households' and 'Industry' model blocks. Thus, the model user has the possibility to add on noise in this model block. However, it was preferred not to add any noise here for the BSM1_LT/BSM2 influent flow rate generation, since the changes in the ground water level, and thus the changes of the infiltration flow rate, are assumed to proceed slowly and smoothly. In fact, the rain generator will create noise on the output of the 'Seasonal correction infiltration' model block (see 3.8), and that was deemed to be sufficient. The parameter 'ground' functions as a switch in the 'Seasonal correction infiltration' model block, similar to the parameters 'pop' in the 'Households' model block and 'ind' in the 'Industry' model block.

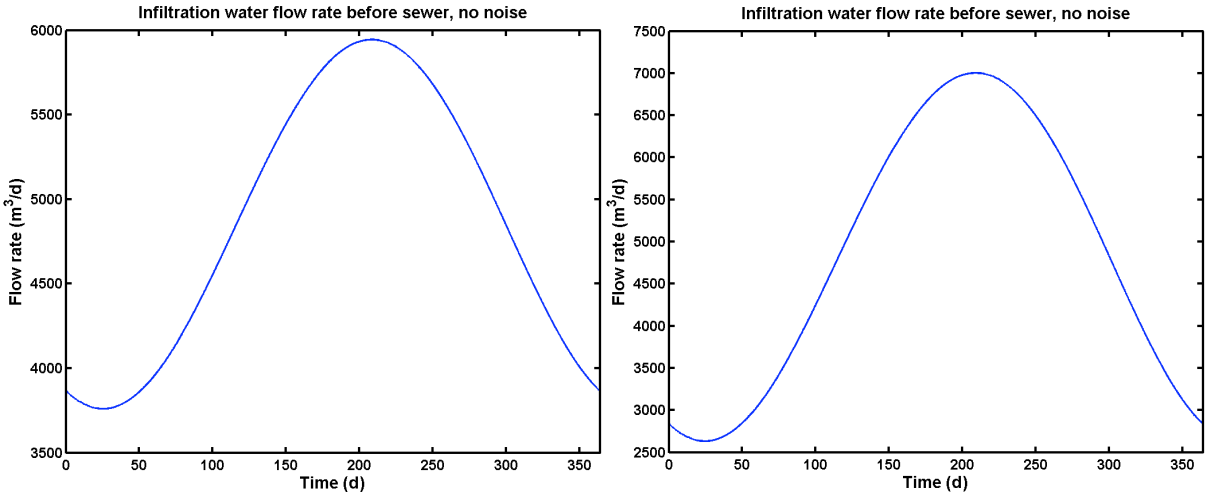

Figure 9. Seasonal influent flow rate variation generated by the 'Seasonal correction infiltration' model block (flow rate as output of the 'Soil' model block with rain generator switched off). Left: Amplitude = 1 200 m^3/d, Average = 7 100 m^3/d; Right: Amplitude = 2 400 m^3/d, Average = 7 100 m^3/d. Day 1 corresponds to July 1st

3.5. 'Soil' model block

One could assume that the joint contributions of the 'Households', 'Industry' and 'Seasonal correction infiltration' model blocks are sufficient for generating dry weather influent flow rate profiles for BSM1_LT/BSM2. However, a closer look at Figure 1 reveals that two other model blocks were deemed to be useful for the influent flow rate generation: the 'Soil' model block and the 'Sewer' model block. The 'Soil' model block (Figure 10) is inserted between the 'Seasonal correction infiltration' and 'Rainfall' model blocks on the one hand, and the 'Sewer' model block on the other hand (see Figure 1). The two inputs to the 'Soil' model block are contributions from the 'Seasonal correction infiltration' model block (see 3.4), and the rainfall (see 3.8) on pervious areas. The former is modelled as a sine wave, while the latter is assumed to be a fraction of the total amount of rainfall water resulting from the 'Rain generator' model block.

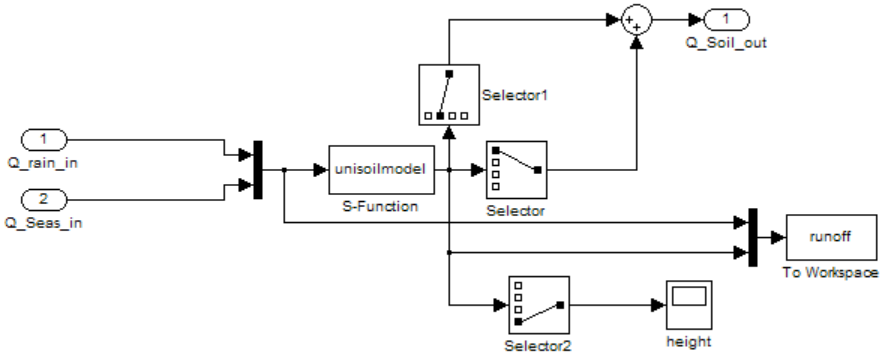

Figure 10. 'Soil' model block lay-out

The core of the 'Soil' model block is an S-function (unisoilmodel.c), which is an implementation of a variable volume tank model. This virtual tank (Figure 11) is used to describe the storage of water in the soil. Parameters for that S-function related to the tank dimensions are A_1 (the surface area of the groundwater storage tank in the soil model; 36 000 m^2), H_{MAX} (the maximum level in the tank; 2.8 m) and H_{INV} (the invert level, i.e. the

maximum water level in the groundwater storage tank that will not cause infiltration, corresponding with the bottom level of the sewer pipes; 2.0 m). Other S-function parameters are K (a measure for the permeability of soil for rainwater penetration; 1.0 $m^3/m^2.d$), K_{inf} (infiltration gain, a measure for the quality of the sewer system pipes, 10 000 $m^{2.5}/d$), and K_{down} (gain to adjust the flow rate to the downstream aquifers; 1 000 m^2/d).

The influent flow rate to the (virtual) groundwater storage tank (Figure 11), Q_{intot}, is considered to be the sum of the groundwater flow received from upstream aquifers (u2), corresponding to the flow rate generated with the 'Seasonal correction infiltration' model block, and the contribution of rain falling on pervious areas and leaking into the soil (u1), corresponding to a fraction (1 - aH) of the rainfall generated in the 'Rain generator' model block (see 3.8). The tank has two outputs, the infiltration flow rate into the sewer system, and the flow rate to downstream aquifers, which is not further considered in the influent model.

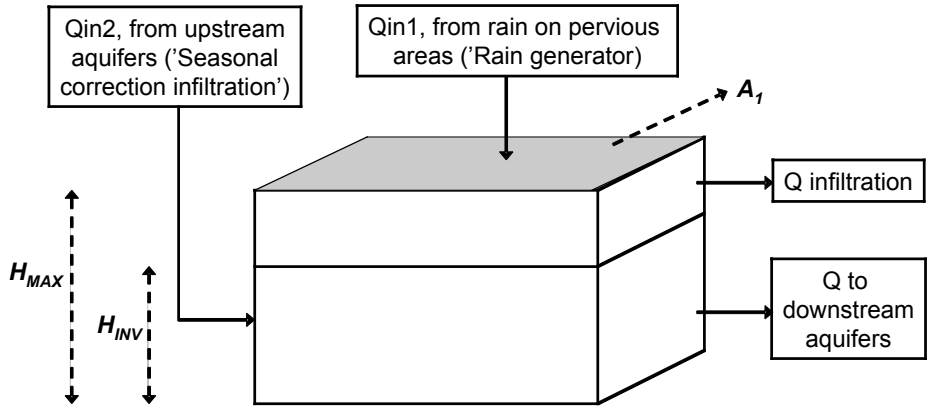

Figure 11. Scheme of the tank model implemented in the unisoilmodel.c S-function. Model parameters are indicated in bold, and using dashed arrows. Model inputs and outputs are indicated with full arrows.

There is only one simple mass balance to describe the evolution of the water level in the groundwater storage tank (Eq. 1).

$$\frac{dh_1}{dt} = \frac{1}{A_1} \cdot (Qin1 + Qin2 - K_{inf} \cdot \sqrt{H_{inf}} - K_{down} \cdot h_1) \quad (1)$$

In Equation 1, h_1 is the water level in the storage tank. The first input, Qin1, corresponds to the contribution of the rain water and is restricted by the permeability of the soil for water (Eq. 2). If $Qin1 > K \cdot A_1$, the remaining flow ($Qin1 - K \cdot A_1$) is assumed to be transported directly to the sewer system.

$$\begin{aligned} &\text{if } (Qin1 > K \cdot A_1) \\ &\quad Qin1 = K \cdot A_1 \\ &\text{else} \\ &\quad Qin1 = Qin1 \end{aligned} \quad (2)$$

The second input, Qin2, is assumed to be zero when the groundwater storage tank is completely filled with water (Eq. 3).

$$\begin{aligned} &\text{if } (H \geq H_{MAX}) \\ &\quad Qin2 = 0 \\ &\text{else} \\ &\quad Qin2 = Qin2 \end{aligned} \quad (3)$$

The first output, corresponds to the infiltration flow rate, and is proportional to the square root of the difference between h (the actual water level in the tank) and the invert level (H_{INV}), as long as h_1 is above H_{INV}. This is implemented according to Equation 4.

$$\begin{aligned} &\text{if } (h_1 > H_{INV}) \\ &\quad H_{inf} = h_1 - H_{INV} \\ &\text{else} \\ &\quad H_{inf} = 0 \end{aligned} \tag{4}$$

The second output, $K_{down} \cdot h_1$, is the flow rate to downstream aquifers.

This simple model attempts to represent a number of mechanisms: (1) the permeability of the top soil layer to receive rain water is limited. The excess will run off via the surface, and is then assumed to reach the sewer system directly. (2) the model allows including periods where the infiltration is zero, corresponding to $h_1 < H_{INV}$. If needed, the model can be extended easily to also include exfiltration. (3) the infiltration flow rate is not constant, but will depend on the difference between the actual water level and the invert level. Parameters used in the BSM2/BSM1_LT influent generation simulations are provided in Table 2. During dry weather conditions, the water level in the tank (h_1) will follow the variations of the infiltration water flow rate, as illustrated in Figure 12. However, the 'Soil' model will also have an impact on the behaviour of the flow rate following a rain event. This will be illustrated below, when introducing the rain generator.

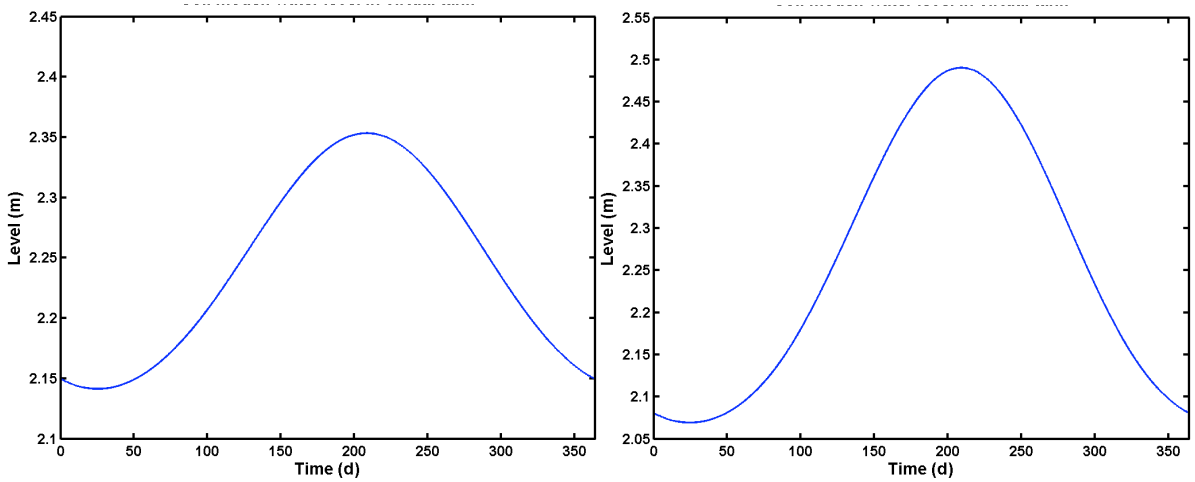

Figure 12. Water level in the virtual tank ('Soil' model block) for the dry weather situation, corresponding to the infiltration water flow rate profiles depicted in Figure 9. Left: Amplitude = 1 200 m^3/d, Average = 7 100 m^3/d for 'Seasonal correction infiltration' model block parameter values; Right: Amplitude = 2 400 m^3/d, Average = 7 100 m^3/d for 'Seasonal correction infiltration' model block parameter values. Note that day 1 corresponds to July 1st

3.6. 'Sewer' model block

The 'Sewer' model block (Figure 13) is an important model block because of its large contribution to the dynamics of the influent flow rate values generated with the influent model. Again, it was attempted to achieve a high level of flexibility for the model user. This was realised by allowing the user to select the size of the sewer system, assuming that a relatively small sewer system will result in sharp diurnal concentration peaks, whereas a large sewer system will result in smooth diurnal concentration variations.

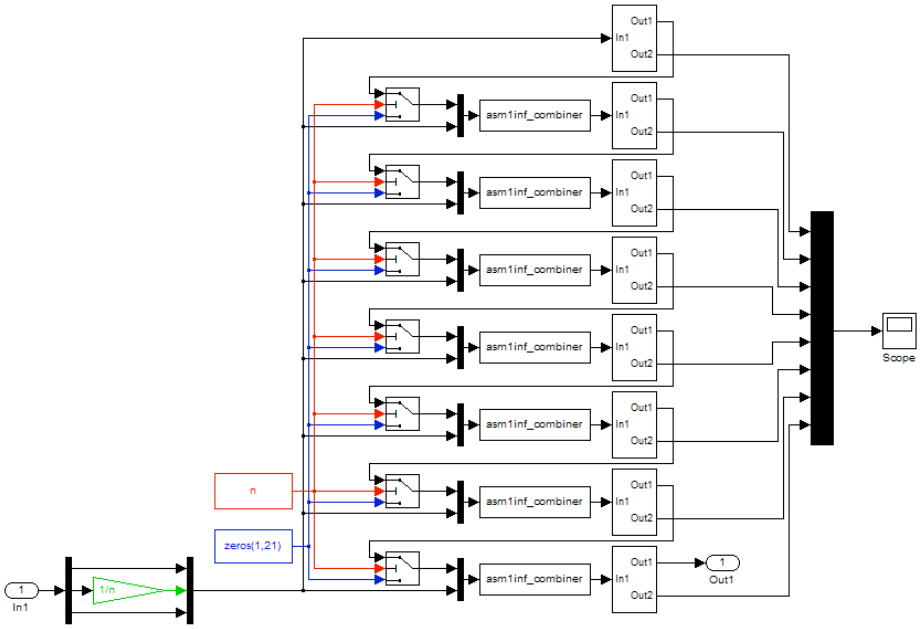

Figure 13. 'Sewer' model block, implementation with deterministic variable volume tank models in series

In the 'Sewer' model block implementation of Figure 13, the influent flow rate is passed through a number of deterministic variable volume tank models, which are grouped in subsystems (Figure 14). Each subsystem consists of 3 variable volume tank models in series. The parameter 'subareas', an integer that can vary between 1 and 8, determines how many subsystems are actively used in the influent flow rate generation.

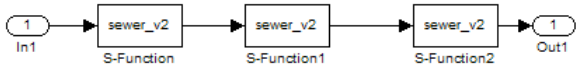

Figure 14. Detail of the 'Sewer' model block, one subsystem. The S-function sewer_v2.c contains a model of a variable volume tank system

The principle of the deterministic variable volume tank model is illustrated in Figure 15. For the deterministic sewer model implementation, the mass balance over the tank is simply written as:

$$\frac{dh_2}{dt} = \frac{1}{A_2} \cdot (Qin - Qout) \tag{5}$$

Where Qout is given by:

$$Qout = C \cdot h_2^{1.5} \tag{6}$$

The parameters A_2, corresponding to the tank surface, and C, corresponding to a gain, need to be provided by the model user. Their values used for BSM1_LT/BSM2 generation are provided in Table 2. It is inherently assumed that the flow rate entering the 'Sewer' model block will never be zero for a long time, meaning that there is always some liquid in the tank. Alternatively, the model also allows using the following expression for the effluent flow rate

$$Qout = C \cdot (h_2 - H_{min})^{1.5} \tag{7}$$

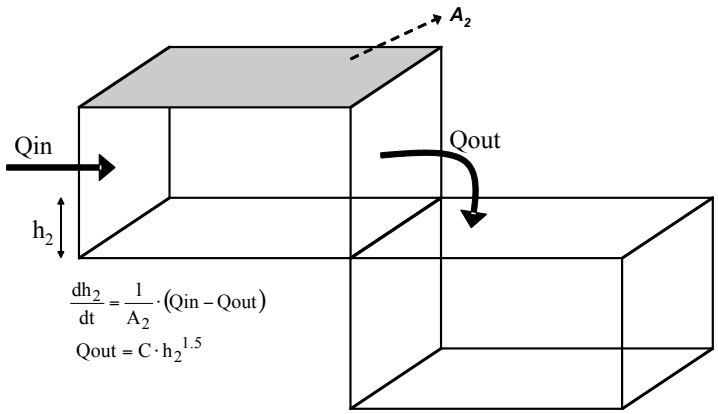

Figure 15. Principle of the deterministic variable volume tank model

When using the model in this way, i.e. with $H_{min} > 0$, the liquid volume in the tank will always be equal to or higher than a threshold $A_2 \cdot H_{min}$. The parameter H_{min} represents the minimum liquid level in the tank that has to be exceeded to obtain a flow rate out of the tank (weir).

Note that it is inherently assumed in the concept of the 'Sewer' model block that the pollution is uniformously discharged along the sewer system. Thus, the dynamics of the output of this model block will try to mimic the fact that part of the pollutants pass through the complete sewer system, whereas another part might only pass through half of the sewer system. As mentioned before, the size of the catchment area in this 'Sewer' model block implementation is adjustable through the use of the parameters 'subareas'. Depending on the value of the subareas parameter (represented by 'n' in Figure 13), an integer which can vary between 1 and 8 in this implementation, the influent pollutant flux is distributed into n equal parts by dividing the flow rate of the original input to the 'Sewer' model block by 1/n. For example, for subareas = 2 the influent flow rate entering the 'Sewer' model is divided by 2. The first half is only passed through the last subsystem, containing 3 variable volume tank models in series (Figure 14), whereas the second half will be passed through the last two subsystems. This division is realised by inserting switch functions before each subsystem (Figure 13), which are triggered by the parameter subareas. The switch function preceding the one but the last subsystem will only pass on the output of the 6^{th} subsystem to the next subsystem if $n \geq 3$. If this condition is not fulfilled, a vector of zeros will be passed on instead, meaning that subsystems 1 to 6 will not contribute to the output dynamics of the 'Sewer' model block. The latter will for example be the case in the example, where n = 2. Thus, in the example the 7^{th} (= the one but the last) subsystem will only receive 50% of the influent flow rate. The output of the 7^{th} subsystem is passed on to the last subsystem, since the switch function preceding the last subsystem will pass on the output of the 7^{th} subsystem if $n \geq 2$. An additional S-function was implemented in C (asm1inf_combiner.c). This model block acts as an ideal mixing tank with no volume, and will combine two inflow vectors, in this example the output of the 7^{th} subsystem and 50% of the influent flow to the sewer system. The output flow rate generated by this model block is the sum of the two input flow rates; whereas output concentrations and output temperature are the flow weighted mean of the two concentration/temperature inputs.

Similar to the other model blocks, the possibility to add noise to the output of the 'Sewer' model block was considered. However, at this point it is important to reflect a little on the purpose of the influent model. The important question to be answered is whether the purpose of the influent model is to generate influent data or influent measurements, where influent

measurements should be understood as a signal that also contains the noise of the flow rate sensor. Clearly, the purpose of this BSM1_LT/BSM2 influent model is to generate influent data, not influent measurements, and therefore a noise generator was not added to the output of the 'sewer' model block for generation of BSM1_LT/BSM2 influent files. Noisy influent flow rate measurements that can be used as measured variable for a control loop in the BSM1_LT/BSM2 simulation models can be obtained by combining the BSM1_LT/BSM2 influent flow rate data generated with this model with one of the sensor models developed earlier for BSM1.

The effect of the number of subsystems (parameter 'subareas') on the output of the 'Sewer' model block is illustrated in Figure 16. It is up to the model user to select the appropriate value of this parameter, providing the flow rate dynamics that are desirable. The size of the first and the second flow rate peak can of course also be influenced by modifying the data files that form the basis for the 'Households' and the 'Industry' flow rate profiles. For the BSM2 influent model, it seemed appropriate to select 4 subareas, since that allowed to reach a more or less similar maximum level for the two daily flow peaks as for the peaks in the BSM1 dry weather data file.

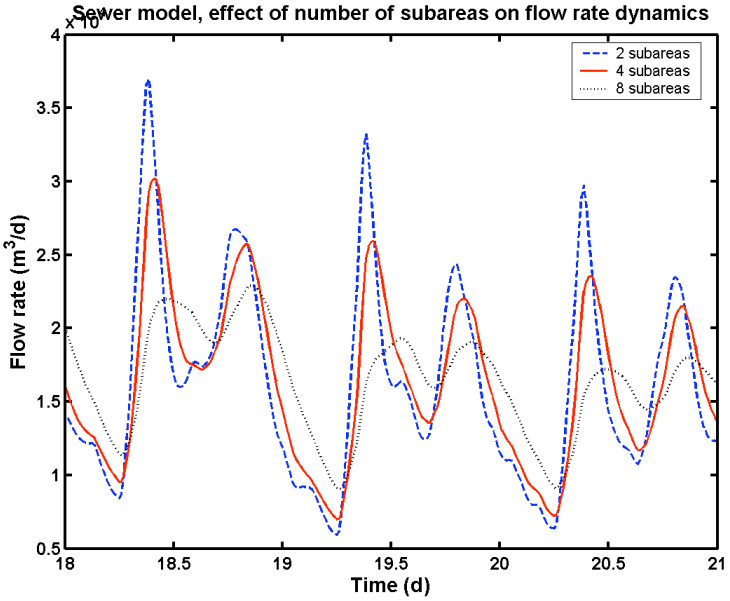

Figure 16. Effect of the number of subareas in the 'Sewer' model block on the dynamics of the generated influent flow rate profile (dry weather). Simulation results obtained for a sine wave amplitude of 1 200 m^3/d in the 'Seasonal correction infiltration' model block

3.7. Dry weather influent flow rate generation: an example

The combination of the 3 contributions to the influent flow rate that were described thus far ('Households', 'Industry', 'Seasonal correction infiltration'), together with the 'Soil' and the 'Sewer' model blocks, is sufficient to generate dry weather influent flow rate variations. This influent flow rate model was used to generate 1.5 years of data with a 15 minute sampling interval, starting on January 1st, and ending on June 30th the next year, or a total of 546 days of data. The simulation was repeated for two different values of the amplitude of the sine wave in the 'Seasonal correction infiltration' model block: 1 200 and 2 400 m^3/d, corresponding to 25 and 50% of the design value of the average infiltration flow rate respectively. Both simulations were performed with and without adding noise to the input

files. According to BSM2 conventions, the first half a year of data (26 weeks) is assumed to be training data, allowing the model to reach steady state using the generated influent data as inputs, whereas the final 364 days of data (52 weeks) are assumed to be used for control strategy evaluation purposes. At this point, it was decided to check for the degree of realism in the dry weather influent data, by comparing these data with influent flow rate data measured on full-scale plants. Of course, one should immediately remark here that such a comparison with full-scale plant data does not say that much, since one can undoubtedly find all kinds of influent flow rate profiles on full-scale plants considering the diversity of household and mixed household + industrial activated sludge plants that exist. However, it is still important to check whether the influent flow rate dynamics generated with the model are comparable to treatment plant data.

For the dry weather influent flow rate data, a histogram was made using the last year of influent flow rate data, i.e. the data that correspond to the BSM1_LT/BSM2 evaluation period (364 days with 15 minute sampling interval = 34 944 data points). The influent flow rate data were divided into 25 bins, and the results (see Figure 17) were expressed as the % of the total number of samples that are classified in each bin. Some statistical properties of the data are also summarised in Table 1 for all simulations.

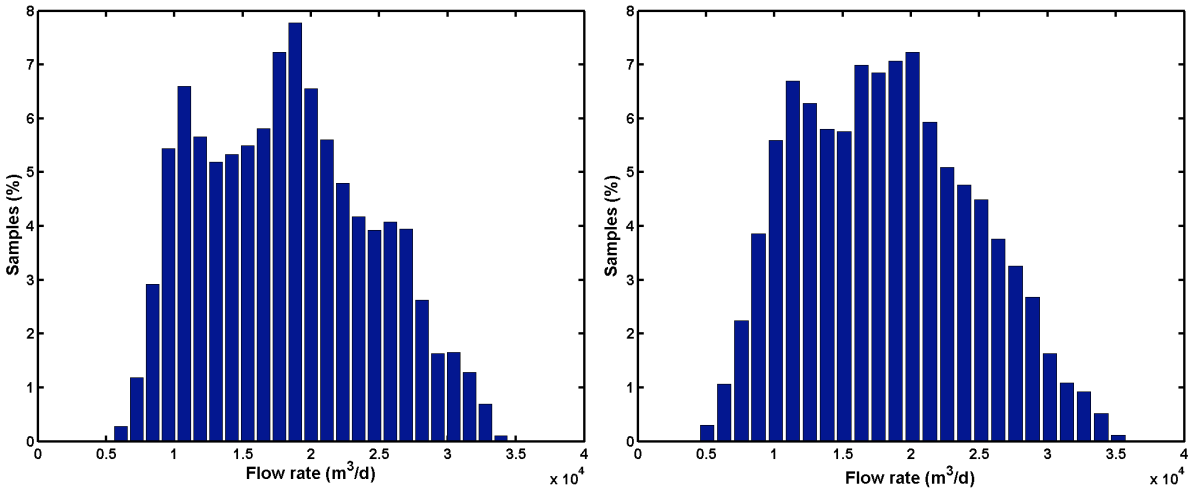

Figure 17. Histogram (25 bins) of the dry weather influent flow rate data generated with the influent model based on the design rules presented in the previous paragraphs (1 year of data, 34 944 samples; n = 4 subareas). Left: Sine wave amplitude = 1 200 m^3/d ('Seasonal correction infiltration' model block); Right: Sine wave amplitude = 2 400 m^3/d

Influent flow rate data were also obtained for a full-scale plant, in this case the Helsingør wastewater treatment plant (Denmark), via www.wwcontrol.dk. Data with a sampling interval of 6 minutes were downloaded for the period January 1st 2002 until August 8th 2004, or a total of 228 243 data points. Of course, the full-scale influent flow rate data contained dry weather as well as rain weather days. Therefore, the raw data were processed through the following data pre-treatment steps:

- Removal of missing values (in this case, time stamps where no measurement value was available). These values were just removed from the data set, since it was assumed that this would not influence the final shape of the histogram.
- Removal of data corresponding to rain weather days. To avoid that rain events would influence the dry weather flow rate histogram, only influent flow rate data collected after a period of at least 48 hours without rain were considered as dry

weather influent flow rate data. In this way, increased infiltration which typically lasted a number of days after large rain events, did not influence the selected data too much. Information on occurrence of rain events was obtained from on-line rain gauge data collected at the WWTP.

- Removal of remaining outliers. Following visual evaluation of typical influent flow rate profiles for this plant, it was assumed that all influent flow rate values above a limit of 650 m^3/h did not correspond to normal dry weather conditions.

As a result of the pre-treatment steps, only 108 972 samples were left, and were used to make the histogram shown in Figure 18.

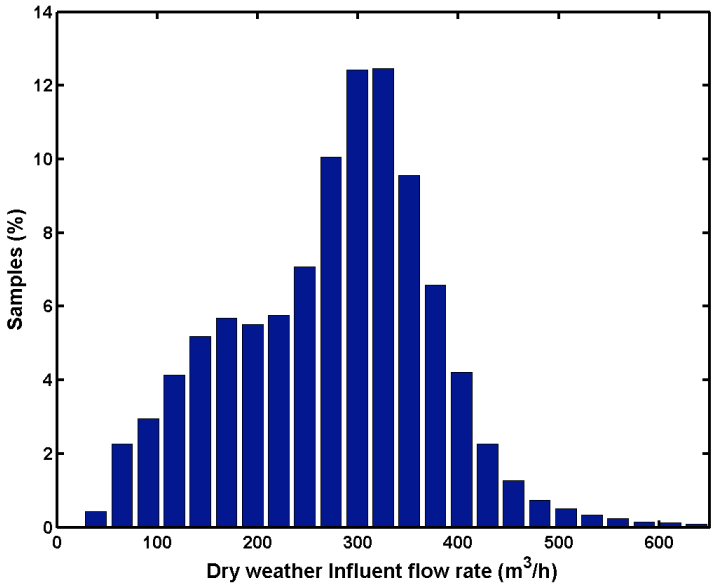

Figure 18. Histogram (25 bins) of the dry weather influent flow rate data of the Helsingør WWTP (period January 1st 2002 to August 8th 2004)

As can be seen from Figure 17 and Figure 18, the histograms calculated from simulated data are quite different from the histogram obtained from the Helsingør WWTP data. The histogram for the dry weather influent flow rate data generated with the model is almost symmetric for a sine wave amplitude of 2 400 m^3/d in the 'Seasonal correction infiltration' model block (Figure 17, right), i.e. it almost looks like a Gaussian distribution. The histogram obtained from simulated data for a sine wave amplitude of 1 200 m^3/d (Figure 17, left) is less symmetric, and it contains two clear peaks: one around 11 000 m^3/d and one around 18 000 m^3/d. The histogram resulting from the full-scale plant data is completely asymmetric (Figure 18), and shows a sharp peak between 300 and 350 m^3/h, meaning that almost 25% of the measured influent flow rate values lay in that range. Moreover, the histogram has a long tail, which is probably due to the fact that the effect of smaller local rain events that were not detected by the rain gauge at the treatment plant cannot be removed completely from the WWTP influent flow rate data using the simple criteria mentioned above. Indeed, one can often observe in full-scale plant data that there is an increase in the influent flow rate, probably due to a local rain event somewhere in the catchment area, whereas the rain meter at the plant does not measure any rainfall.

However, the size of the catchment area also has an influence on the shape of the histogram. This is best illustrated by comparing the histograms in Figure 17 with the histograms in Figure 19. As explained earlier (see 3.6), the 'Sewer' model block has a parameter 'subareas',

which will influence the shape of the influent flow rate profiles (see Figure 16). The higher the value of this parameter, the larger the catchment area is assumed to be. Dry weather influent flow rate data were also generated for 2 and 8 subareas respectively, keeping all other influent model parameters used for the simulation that resulted in the histograms of Figure 17 constant. Histograms were also generated for these two influent flow rate time series (Figure 19). Some statistical properties of the data are summarised in Table 1. Clearly, the shape of the histogram for 8 subareas for a sine wave amplitude of 2 400 m^3/d is much more similar to the histogram obtained on the dry weather data of the Helsingør treatment plant. This indicates that the parameter subareas of the sewer model can be used to tune influent flow rate profiles to resemble data obtained from a specific plant, which can be useful if influent scenarios are to be generated for that plant. It also indicates that the flow rate dynamics generated with the influent model are similar to the flow rate dynamics that are observed on full-scale plants.

Table 1. Summary of results obtained for simulations with the dry weather influent flow rate model, for subareas = 2, 4 and 8 respectively. The effect of adding noise on the simulation output was evaluated by repeating the simulation for subareas = 4 without noise. The simulation for subareas = 4 was repeated for two values of the sine wave amplitude in the 'Seasonal correction infiltration' model block: 1 200 and 2 400 m^3/d The 5% percentile value corresponds to that flow rate value that is greater than or equal to 5% of the flow rate samples. The 95% percentile value corresponds to that flow rate value that is greater than or equal to 95% of the flow rate sample population

Amplitude = 1 200 m^3/d	No noise added, 4 subareas	Noise added, 4 subareas	Noise added, 2 subareas	Noise added, 8 subareas
Mean (m^3/d)	18 215.9	18 248.5	18 248.2	18 248.8
Standard deviation (m^3/d)	6 144.2	6 176.4	7 283.8	4 243.7
Minimum (m^3/d)	5 889.9	5 508.9	4 741.3	7 197.9
Maximum (m^3/d)	32 677.8	34 487.0	41 850.6	27 080.8
5% percentile (m^3/d)	9 167.0	9 135.8	8 374.2	11 007.4
95% percentile (m^3/d)	28 714.3	28 888.9	31 648.0	24 183.2
Median (m^3/d)	18 121.6	18 099.6	17 504.6	18 916.1
Amplitude = 2 400 m^3/d	No noise added, 4 subareas	Noise added, 4 subareas	Noise added, 2 subareas	Noise added, 8 subareas
Mean (m^3/d)	18 199.6	18 231.6	18 231.3	18 231.9
Standard deviation (m^3/d)	6 320.5	6 356.3	7 437.5	4 503.7
Minimum (m^3/d)	4 859.9	4 443.1	3 902.9	6 177.0
Maximum (m^3/d)	34 038.8	35 798.3	42 558.4	28 062.0
5% percentile (m^3/d)	8 714.5	8 703.8	7 964.5	10 555.3
95% percentile (m^3/d)	29 055.5	29 136.8	31 970.3	24 926.3
Median (m^3/d)	17 947.5	17 975.5	17 405.6	18 793.3

The results in Table 1 indicate that adding noise to the data will lead to an increased standard deviation of the resulting flow rate data, a lower minimum flow rate value and a higher maximum flow rate value. Accordingly, the distance between the 5% percentile and the 95% percentile is larger, indicating that adding noise increases the spread of the data. Similar phenomena are observed for both values of the sine wave amplitude in the 'Seasonal correction infiltration' model block. Interestingly, adding noise seems to lead to a slightly higher mean flow rate value, although the noise generators used add on zero mean noise. This contradiction can be explained by the use of the 'saturation' model blocks, which are inserted to avoid that adding noise to a flow rate leads to negative flow rate values.

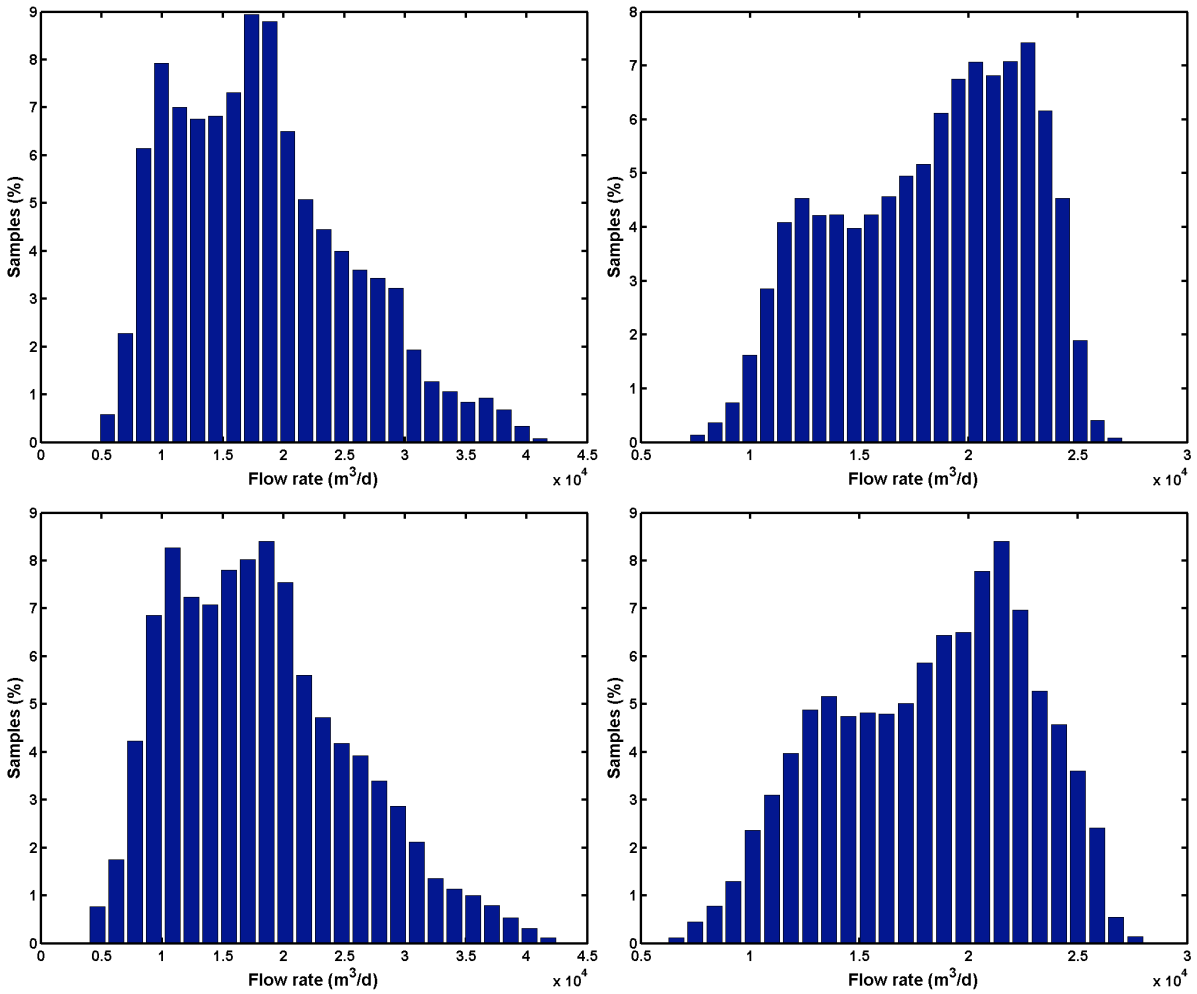

Figure 19. Histogram (25 bins) of the dry weather influent flow rate data generated with the model (34 944 samples) for 2 subareas (left) and 8 subareas (right). Top: Sine wave amplitude = 1 200 m^3/d ('Seasonal correction infiltration' model block); Bottom: Sine wave amplitude = 2 400 m^3/d

3.8. 'Rain Generator' model block

All the model elements mentioned before are needed to generate the dry weather influent flow rate profile, including diurnal variations, weekend effects and seasonal variations. Since it is assumed that the benchmark WWTP is treating wastewater from a combined sewer system, rainfall has to be included as a disturbance. To this purpose, a simple 'Rain Generator' model block was implemented (Figure 20). In the 'Rain Generator' model block, a random number is first generated (mean = 1, variance = 400), and is subsequently passed through a first order transfer function model block. A constant (LLrain) is subtracted, and the resulting number is passed through a saturation model block to avoid negative numbers. The resulting signal is subsequently passed through two 'Gain' model blocks: the first one converts the rainfall values generated by the random generator to a value representing rainfall intensities in mm/day, whereas the second 'Gain' model block includes a constant (Qpermm = 1 500), whose units can be interpreted as m^3/mm rain.

The parameter aH, varying from 0 to 100 %, multiplied with the constant Qpermm corresponds to the contribution of rainfall falling on impervious surfaces in the catchment area. It is indeed assumed in the model that the rainfall that is collected on impervious

surfaces will immediately be collected in the sewers. In other words, during a rain it is assumed that the size of the catchment connected to the treatment plant corresponds to an impervious area of 150 hectares (Qpermm = 1500) multiplied by aH. A part of the rain collected on pervious areas will first penetrate into the soil, as far as the capacity of the soil to further adsorb water is not exceeded (see 3.5 for details on the 'Soil' model block). After passing the 'Soil' model block, this fraction (1 - aH) of the rainfall will thus reach the sewer system via infiltration. The parameter 'rain' functions as a switch and allows to completely switch off the rain generator, e.g. during model testing.

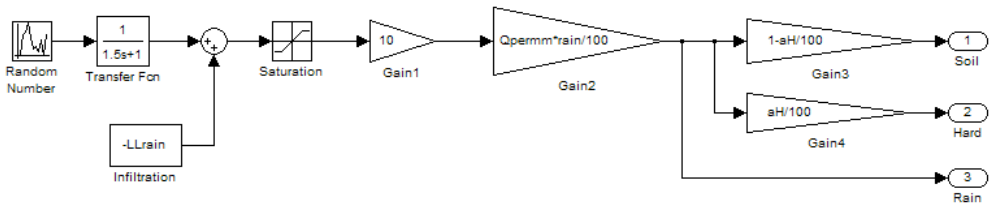

Figure 20. 'Rain Generator' model block

Figure 21 illustrates the extra influent rain flow rates resulting from applying the 'Rain Generator' model block. The height of the peaks gives an indication on the rain intensity. Note that each major rain event, corresponding to the highest peaks in Figure 21, is followed by a tail. This tail illustrates the effect of passing a fraction (1 – aH) of the rainfall through the soil model block, and corresponds with observations made on full-scale plants where it often takes a few days after a major rain event before the flow rate has completely returned to the dry weather situation.

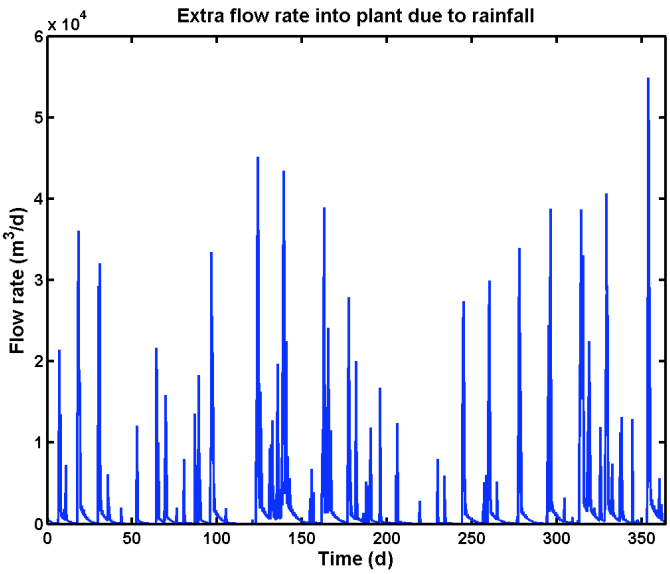

Figure 21. One year of data: extra influent flow rate due to rain resulting from the 'Rain Generator' model block

The effect of including the 'Soil' model block is further illustrated in Figure 22. The results were obtained with the parameter values that are provided in Table 2. In the overflow, corresponding to periods where $Q_{intot} > K \cdot A_1$ (see 3.5), major rain events should result in runoff to the sewer system. However, for the benchmark plant the parameter K was chosen sufficiently high, such that this contribution will not play any role in the sewer output

dynamics. At the same time, major rain events result in a significant increase of the level in the virtual groundwater storage tank, and thus a significant increase of the flow rate due to infiltration. The seasonal variation of the amount of water due to infiltration is very clear in Figure 22. As mentioned before, an interesting feature of this simple soil model is that it allows including a 'memory effect', following a rainfall (see also Figure 21).

As for the other model blocks, there is again a great flexibility included, since the user can modify the parameters of the soil model to reach the desired influent profile.

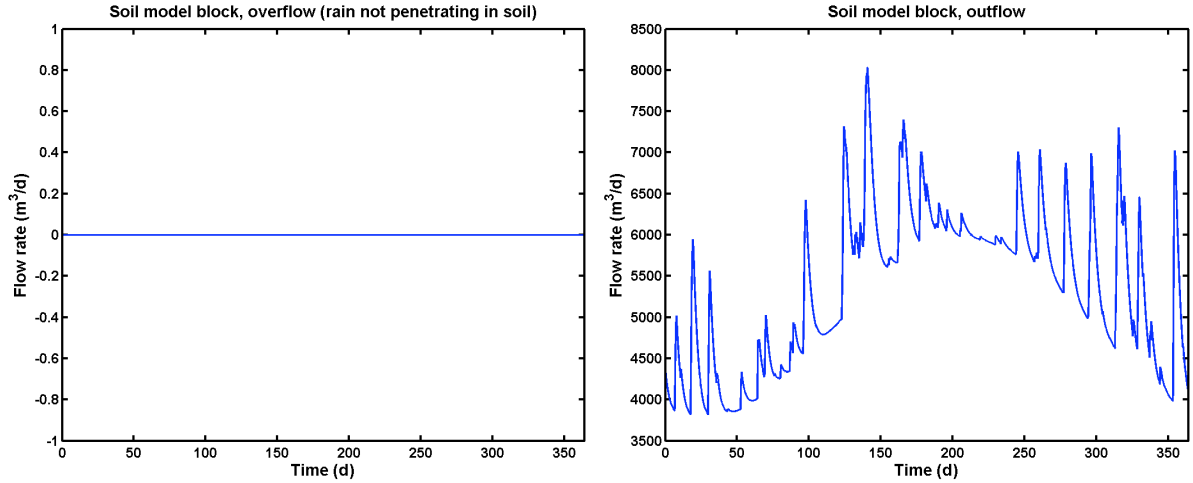

Figure 22. Overflow (left), i.e. flow in excess of $K \cdot A_1$ (see 3.5), and outflow (right) leaving the groundwater storage tank to the sewer system. The total flow rate leaving the 'Soil' model block to the sewer system, i.e. the output of the 'Soil' model block, corresponds to the sum of both contributions

3.9. Comparison of influent model rain generator with rainfall data

An attempt was made to compare the rainfall data generated with the rain generator (see 3.8) with rainfall data obtained from rain meters installed at different locations in Europe. Rain meter data were available from Helsingør (Denmark).

3.9.1. Rain distribution for data generated with rain generator

The output of the 'Rain generator' model block for a 1.5 year period (dynamic data with 15 minutes sampling interval) for the parameter values reported in Table 2, which is expressed as a rainfall intensity (in mm rain/d units), was processed further in a number of steps:

- All samples for which the output of the 'Rain generator' model block was equal to zero (= no rain) were discarded.
- The remaining samples were grouped into rain events, assuming that 2 subsequent samples with rainfall belonged to the same rain event as long as there was less than two hours between the time stamps of both samples.
- The rain depth (total mm of rain) for each rain event was calculated by integrating over all rainfall intensity values belonging to the same rain event (see Figure 23).
- The rain depth values were sorted in ascending order, and both a cumulative rain depth distribution plot and a rain depth histogram were made.

During the period considered, a total of 213 rain events were registered. The cumulative rain depth distribution plot and the corresponding rain depth histogram are provided in Figure 24.

About 85% of the registered rain events have a rain depth that is lower than 5 mm. The maximum rain depth registered for an event is 43 mm (Figure 23).

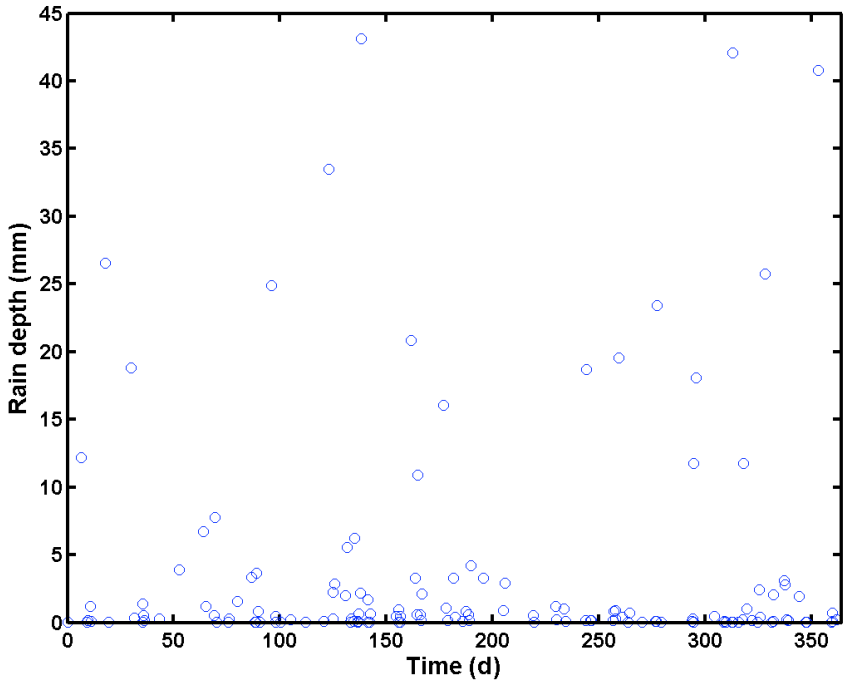

Figure 23. Rain depth calculated for the rain events for 1 year of data resulting from applying the 'Rain generator' model block with the influent model parameters of Table 2

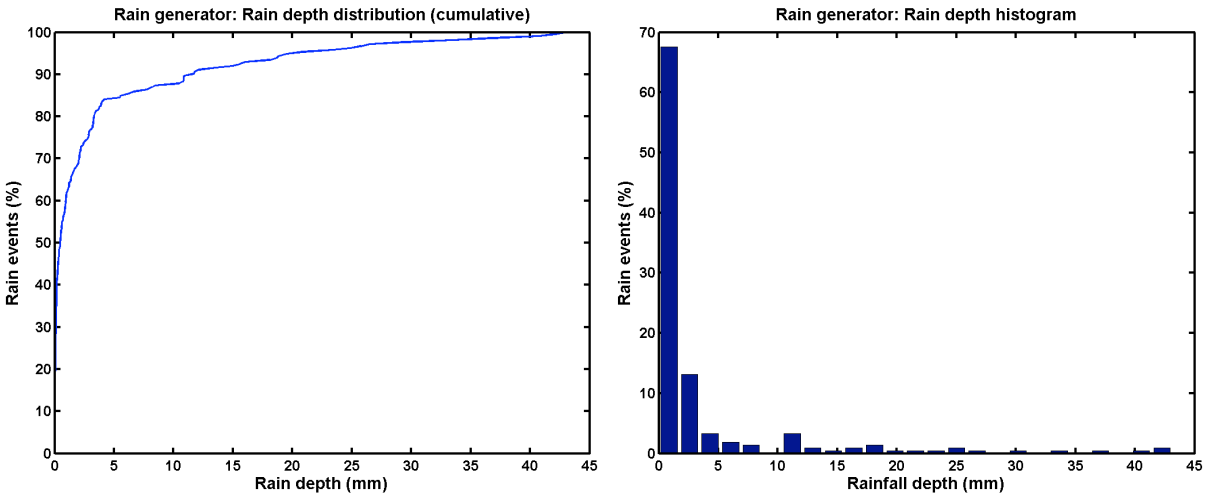

Figure 24. Cumulative rain depth distribution (left) and rain depth histogram (right) for 1.5 years of rain data (213 events) generated with the 'Rain generator' model block for the influent model parameters of Table 2

In view of the results of Figure 24, all rain events with a rain depth below 1 mm were considered insignificant. Therefore, these events were discarded from the data set, and a new cumulative rain depth distribution and rain depth histogram were made. For this particular data set, the number of rain events below 1 mm equalled 131 (61.5% of the total number of rain events), i.e. 82 rain events were retained as significant and were subsequently used for generating the plots in Figure 25. There are several reasons for discarding these rain events. First of all, such small rain events have very little impact on the influent flow rate and influent composition, and will thus not significantly disturb normal plant operation. Second, the rain

meters installed on full-scale systems do not allow measuring all rain events, that is a rain event needs to have a minimum rain depth before the rain meter will register it. In addition to that, several subsequent small rain events can only be registered as one rain event by the rain meter, when the intensity of the individual events is below the sensitivity of the rain meter. Typically, on-line rain gauges such as tipping bucket measuring systems will only register a rain when its rain depth has reached 0.1 to 0.2 mm.

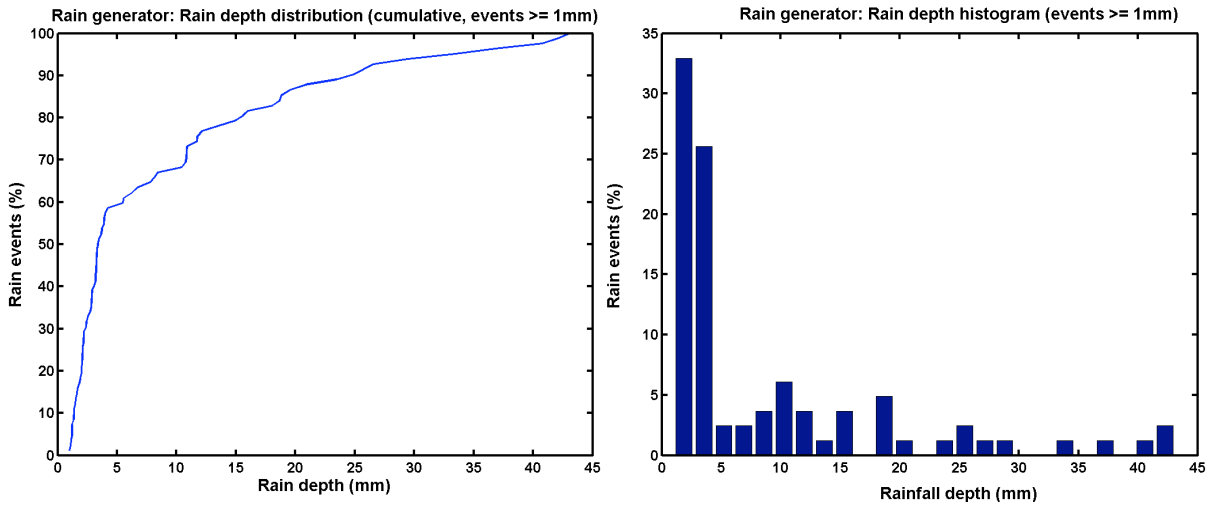

Figure 25. Cumulative rain depth distribution (left) and rain depth histogram (right) for 1.5 years of rain data (82 events) generated with the 'Rain generator' model block. Rain events with a cumulative rain depth below 1 mm were considered insignificant

3.9.2. Rain distribution for Helsingør rain meter data

The Helsingør WWTP data mentioned in 3.7 contained on-line rain gauge data. These data, lasting from January 1st 2002 until August 8th 2004 were analysed in different steps, to allow comparison with the data generated with the 'Rain generator' model block. The original data set, with a 6 minute sampling interval, contains 228 243 data points. Data were treated in a number of subsequent steps:

- Removal of missing values (in this case, time stamps where no measurement value was available), which was rather easy since they occurred as NaN values in the data set. In total, 7 764 samples were discarded, and 220 479 were retained for further processing. An example of missing values can be observed around t = 125 d in Figure 26.

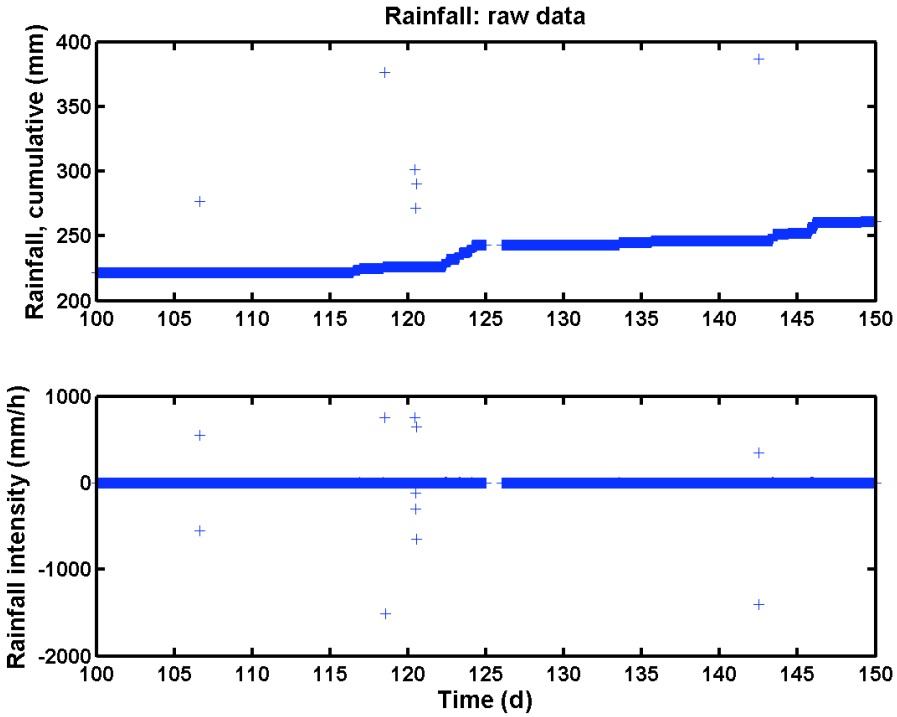

Figure 26. Illustration of the transformation of the original rainfall data (cumulative values, top) to rainfall intensity data (bottom)

- Time stamps occurring twice in the data set were removed (2 samples)
- The raw data were available as the cumulative amount of rainfall registered by the rain gauge (in mm rain units). Rain intensity values (in mm rain/24 h units) were obtained by taking the derivative of the raw data, which was approximated by calculating the difference between the actual and the previous measurement value for each sample. The result of this transformation is illustrated in Figure 26.
- Outliers were removed. As can be seen from Figure 26, the raw data contained quite some outliers. Removal of these outliers was easier on the rain intensity data. Negative values should not occur in the rainfall intensity data, and 92 samples were therefore removed from the data set. Positive outliers were also removed, assuming that rain intensity values higher than 25 mm/h were outliers. 105 samples were removed as positive outliers.
- A total number of 220280 samples were left, and these were treated similarly to the rainfall data generated with the model: rain data were grouped in rain events, and the cumulative rain depth of each rain event was calculated. The rain depth values were subsequently sorted in ascending order, and both a cumulative rain depth distribution plot and a rain depth histogram were made.

A total of 467 rain events were identified in the data set. The corresponding distribution is provided in Figure 27. Quite comparable to the data generated with the 'Rain generator' model block, about 80 % of the rain events have a rain depth that is below 5 mm. Similar to the rainfall data generated with the model, events with a rain depth below 1 mm were considered insignificant, and a new cumulative rain depth distribution and rain depth histogram were made (Figure 28). For the Helsingør data set, 229 rain events had a rain depth below 1 mm (49.0 % of the total number of rain events).

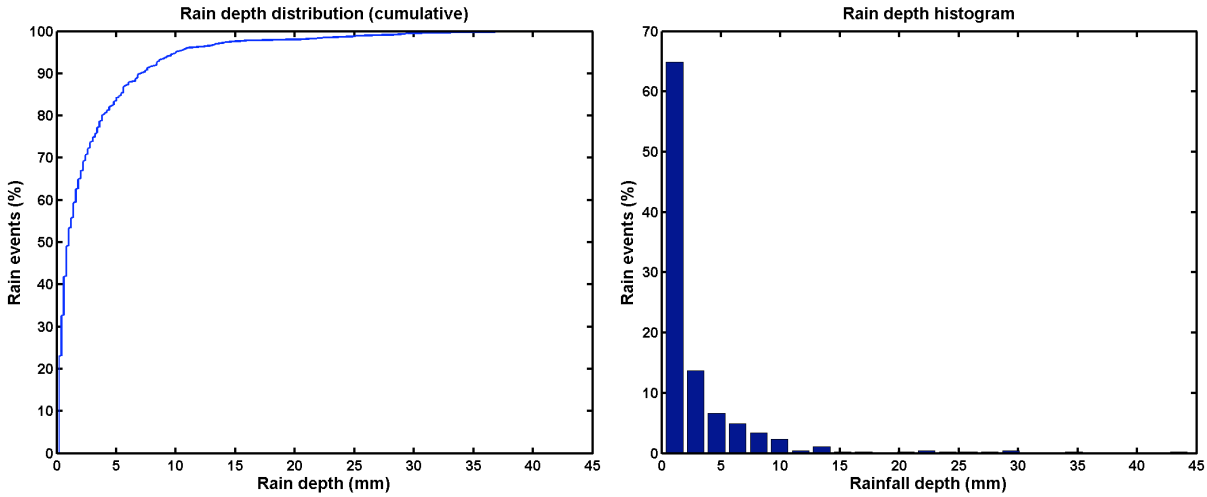

Figure 27. Cumulative rain depth distribution (left) and rain depth histogram (right) for rainfall data of the Helsingør WWTP (467 events)

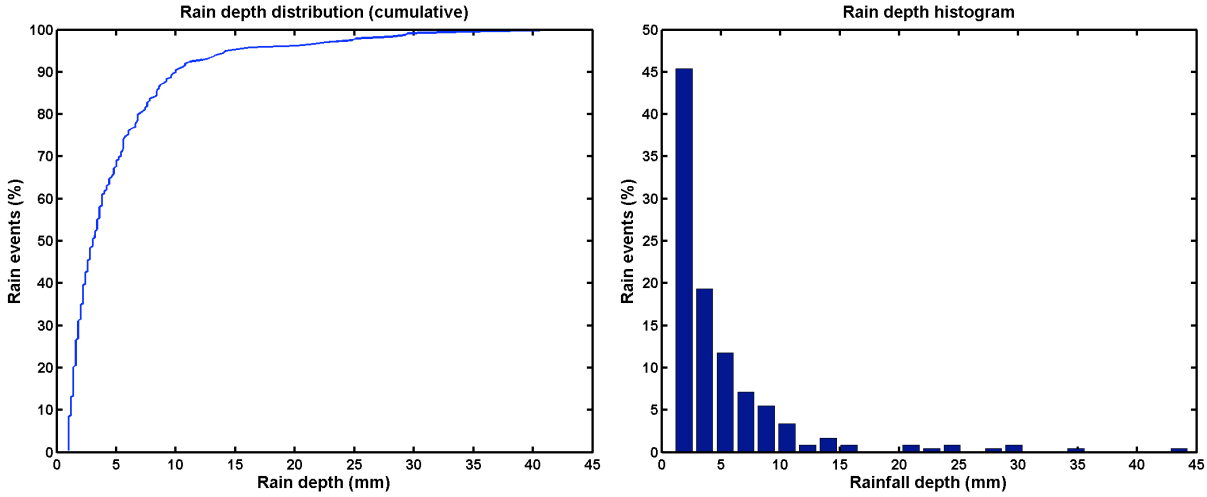

Figure 28. Cumulative rain depth distribution (left) and rain depth histogram (right) for rainfall data of the Helsingør WWTP (238 events). Rain events with a cumulative rain depth below 1 mm were considered insignificant

3.9.3. General discussion of results

The shape of the rain depth histograms obtained with the 'Rain generator' model block (Figure 25) is of course not exactly the same as the histogram for the Helsingør rainfall data (Figure 28). The 'Rain generator' model block produces a higher % of rain events below 5 mm, compared to the Helsingør data. Also, the frequency of occurrence of certain rain depths at the Helsingør plant is decreasing steadily with increasing rain depth. For the data obtained with the 'Rain generator' model block, this steady decrease of the frequency of occurrence of certain rain depths with increasing rain depths is less pronounced. It is probably also caused by the lower number of rain events (82) that is considered in this plot, compared to the Helsingør data set (238 rain events). In general, however, the results obtained with the 'Rain generator' model block seem to be acceptable for its intended purpose, i.e. generation of rain events for the BSM1_LT/BSM2 influent.

3.10. Conclusions on the first design step of the influent flow rate model

Based on the simulations performed during the development of the influent flow rate model blocks, and with the BSM1 influent flow rate profiles in mind, the parameters summarised in Table 2 were selected for the BSM2 influent flow rate profile generation. From the two influent design alternatives tested (see Figure 29 for average daily flow rate results), with InfAmp = 1 200 m^3/d and InfAmp = 2 400 m^3/d respectively, the simulations with InfAmp = 1 200 m^3/d were preferred. The seasonal effect was considered to be too strong for InfAmp = 2 400 m^3/d.

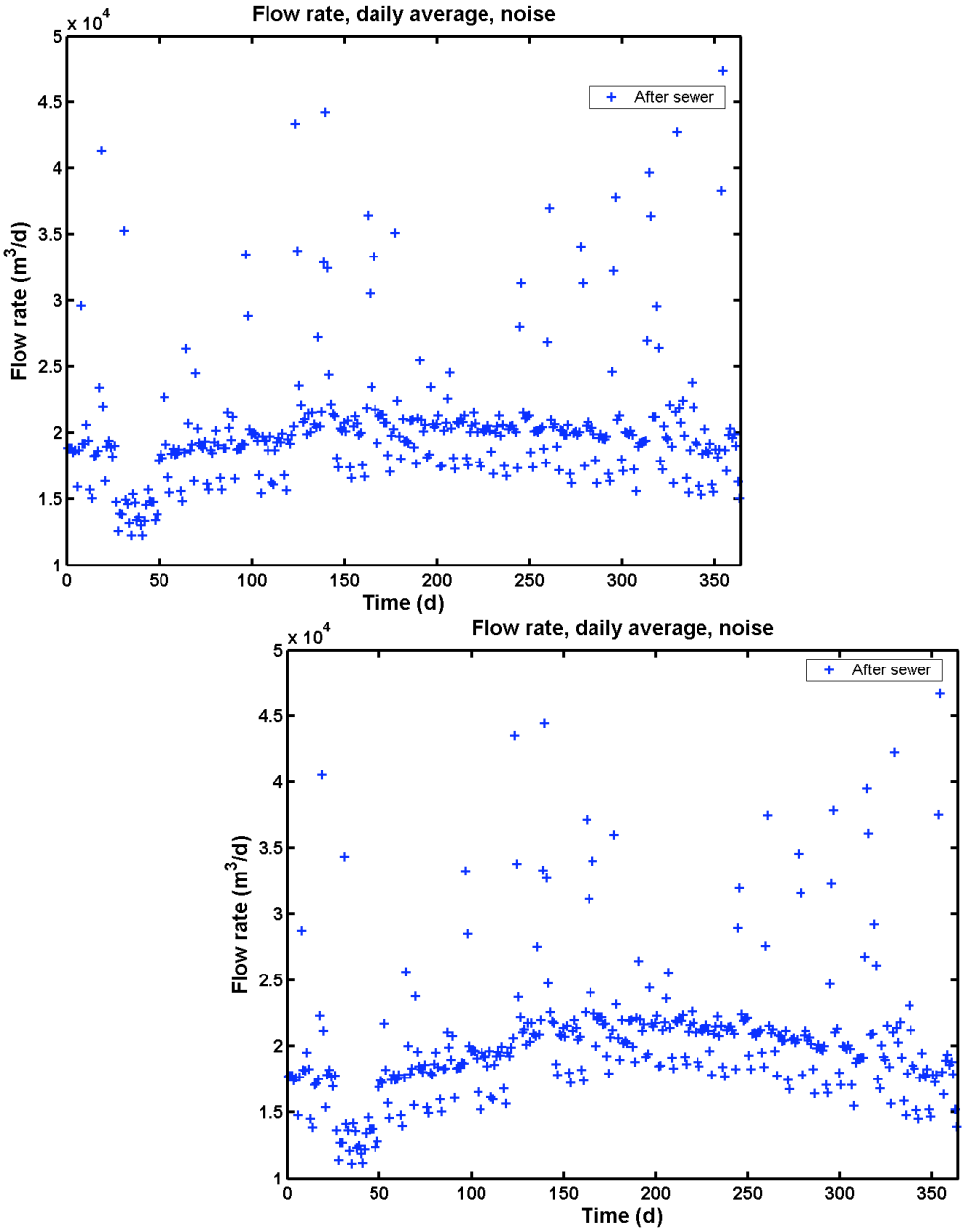

Figure 29. Average daily flow rates generated with the influent flow rate model (each data point in this figure corresponds to the average of 96 samples, i.e. one day of dynamic data with 15 minutes sampling interval). Top: InfAmp = 1 200 m^3/d; Bottom: InfAmp = 2 400 m^3/d

The simulation examples have illustrated that the simple phenomenological models allow generating diurnal variations, weekend effects, seasonal effects and rainfall effects for the

influent flow rate profile. When summing up the effects of the different model blocks, average daily flow rate results as in Figure 29 can be obtained. In that figure, the seasonal effect and severe rain events appear clearly. When looking more in detail at the generated data (Figure 30), the weekend effects also appear clearly in the average daily flow rate profiles. A detail of a full dynamic influent flow rate profile is also provided in the same figure.

Table 2. Main parameters for the influent flow rate model

Model block	Parameter	Value	Units	Remarks
'Households' (See 3.1 and 3.2)	QperPE	150	(l/d)/p.e.	Wastewater flow rate per p.e. for municipal wastewater
	PE	80	1000 person equivalent	Number of household p.e. connected to the WWTP
'Industry' (See 3.1 and 3.3)	QInd_weekday	2 500	m^3/d	Average wastewater flow rate from industry on normal week days (Monday to Thursday)
'Seasonal correction infiltration' (See 3.4.)	InfBias	7 100	m^3/d	Mean value of the sine wave signal for generating seasonal effects due to infiltration
	InfAmp	1 200	m^3/d	Amplitude of the sine wave for generating seasonal effects due to infiltration
	InfFreq	$2\cdot\pi/364$	rad/d	Frequency of the sine wave for generating seasonal effects due to infiltration
	InfPhase	$\pi\cdot15/24$	rad	Phase shift of the sine wave for generating seasonal effects due to infiltration
'Soil' (See 3.5)	A_1	36 000	m^2	Surface area of the groundwater storage in the soil model
	H_{MAX}	2.8	m	Total height of the groundwater storage in the soil model
	H_{INV}	2.0	m	Height of the invert level
	K	1.0	$(m^3/d)/m^2$	Soil permeability constant
	K_{inf}	10 000	$m^{2.5}/d$	Infiltration gain
	K_{down}	1 000	m^2/d	Gain for adjusting the flow rate to downstream aquifers
'Sewer' (See 3.6)	subareas	4	-	Number of sub-areas in the sewer system, a measure for the size of the sewer system/catchment area
	A_2	1 100	m^2	Surface area of a variable volume tank in the sewer model
	C	150 000	$m^{1.5}/d$	Gain for adjusting the flow rate out of the variable volume tank model
	H_{min}	0.0	m	Minimum liquid level in the variable volume tank
'Rain Generator'	LLrain	3.5	10 mm rain/d	Cut-off value for rain generator
	Qpermm	1500	m^3/mm rain	A measure for the size of the catchment area connected to the sewer system
	aH	75	%	Percentage of impervious surfaces in the catchment area (the remainder of the rainwater ((1-aH) fraction) is assumed to first percolate in the soil from pervious areas, and to reach the sewer system afterwards via infiltration

Summarising, it can be concluded that all the dynamic flow rate phenomena that were thought to be important for the BSM1_LT/BSM2 influent generation can be produced with the proposed models.

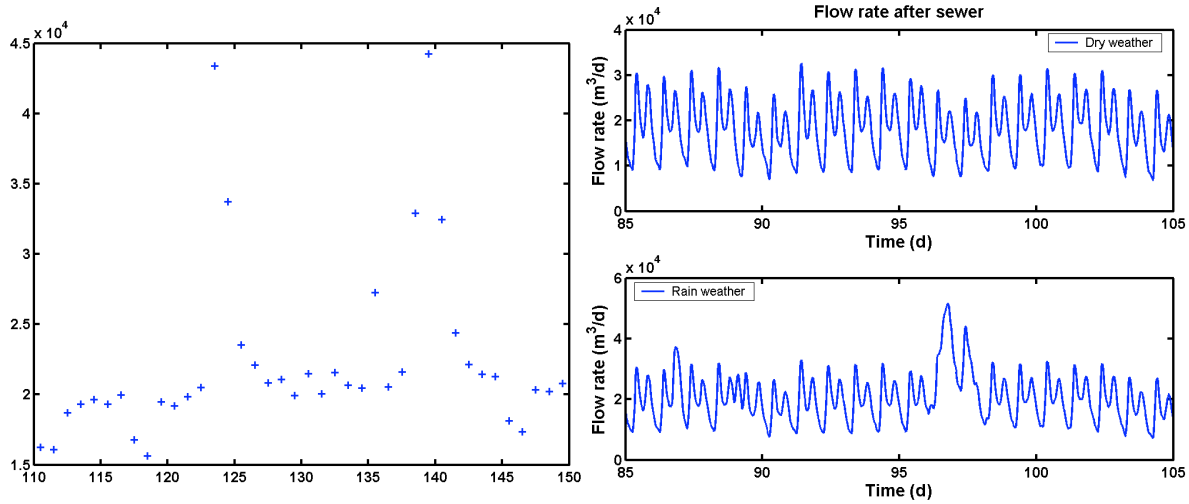

Figure 30. Details of influent flow rate profiles generated with the influent flow rate model (influent model parameters of Table 2). Average daily flow rates showing weekend effects and rain events (left); Dynamic influent flow rate profiles (right) with a 15 minutes sampling interval between data points. Simulation of dry weather conditions (top right) is compared with simulation of rain events superimposed on the dry weather flow rate (bottom right)

4. Model for influent component concentrations

In a first version of this influent model, the different ASM1 influent variables were generated separately (see Gernaey *et al.*, 2004). However, it was decided to adopt a more general approach, based on particulate and soluble COD fluxes and on ammonium and TKN fluxes, which could be used for example for generating influent files for other ASM models too. The model for generating influent component concentrations is quite similar to the influent flow rate model, and details will be explained below. Similar to the influent flow rate model, a first design of the model for generating influent component concentrations will be presented. The modifications made to achieve the final proposal for a BSM2 influent model will be provided later (see 6)

4.1. Influent component concentration design principles

The design principles of the influent component concentrations generation will be introduced briefly, before providing more details on each model block. An overview of all the model blocks in the influent model is provided in Figure 31.

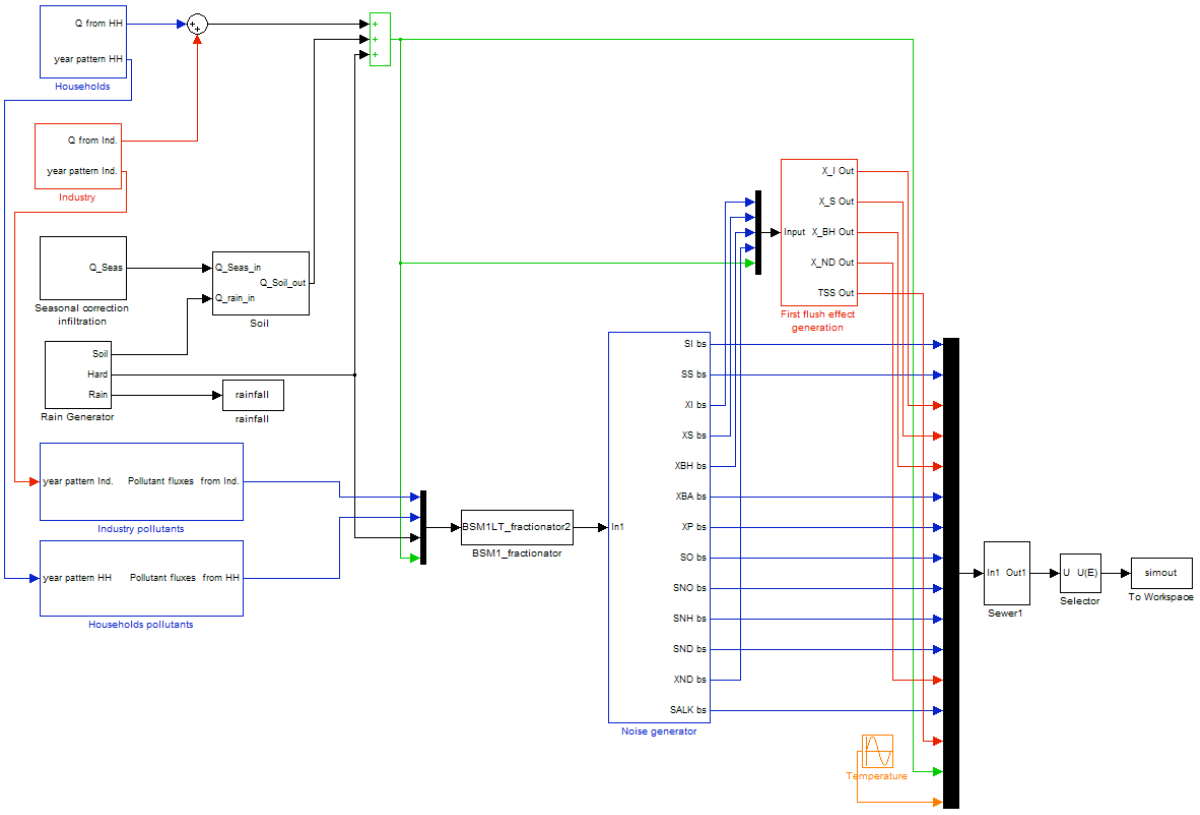

Figure 31. Overview of the modules in the proposed model for influent data generation

The model blocks related to the influent flow rate generation will not be commented any further (see 3). It is assumed in Figure 31 that there are two pollutant sources: households ('Households pollutants' model block) and industry ('Industry pollutants' model block), which is an acceptable simplification. The pollutant loads from both sources are defined as fluxes (particulate and soluble COD, ammonium and TKN respectively). The complexity of

the model is thus reduced significantly by neglecting contributions of the other influent sources (infiltration and rain). Pollutant fluxes from both sources are subsequently combined and converted to ASM1 compatible influent concentrations in the 'BSM1_fractionator' model block. Noise is added to the different influent ASM1 component concentrations in the 'Noise generator' model block. The X_I, X_S, X_{BH} and X_{ND} concentrations are subsequently passed through the 'First flush effect generation' model block, where first flush effects will be generated related to unusually high flow rates, for example during rain events. The resulting ASM1 component concentrations are finally combined with the flow rate (generated with the influent flow rate model, see 3) and the temperature (generated in the 'Temperature' model block, see 5) via a Mux model block, and are passed through the 'Sewer' model block. The resulting influent vector, consisting of the 13 ASM1 components, the TSS concentration, the influent flow rate and temperature, is saved in the Matlab workspace at the end of a simulation, and can be used directly as input to an ASM1 compatible simulation model.

4.2. 'Households pollutants' model block

4.2.1. Model block

The lay-out of the 'Households pollutants' model block is illustrated in Figure 32. Diurnal profiles for soluble and particulate COD, S_{NH} and TKN form the basis for the dynamic profiles generated with this model block. These profiles are user-defined and are called from the workspace (CODsol_day_HS, CODpart_day_HS, SNH_day_HS, TKN_day_HS). Each diurnal profile consists of 24 values (1h sampling interval). Similar to the influent flow rate profile in day_HS (see 3.2) the values in the 4 input files are normalised (average = 1). Note that the option of defining an influent S_{NO} contribution is also foreseen. However, for the benchmark plant influent this contribution is assumed to be zero.

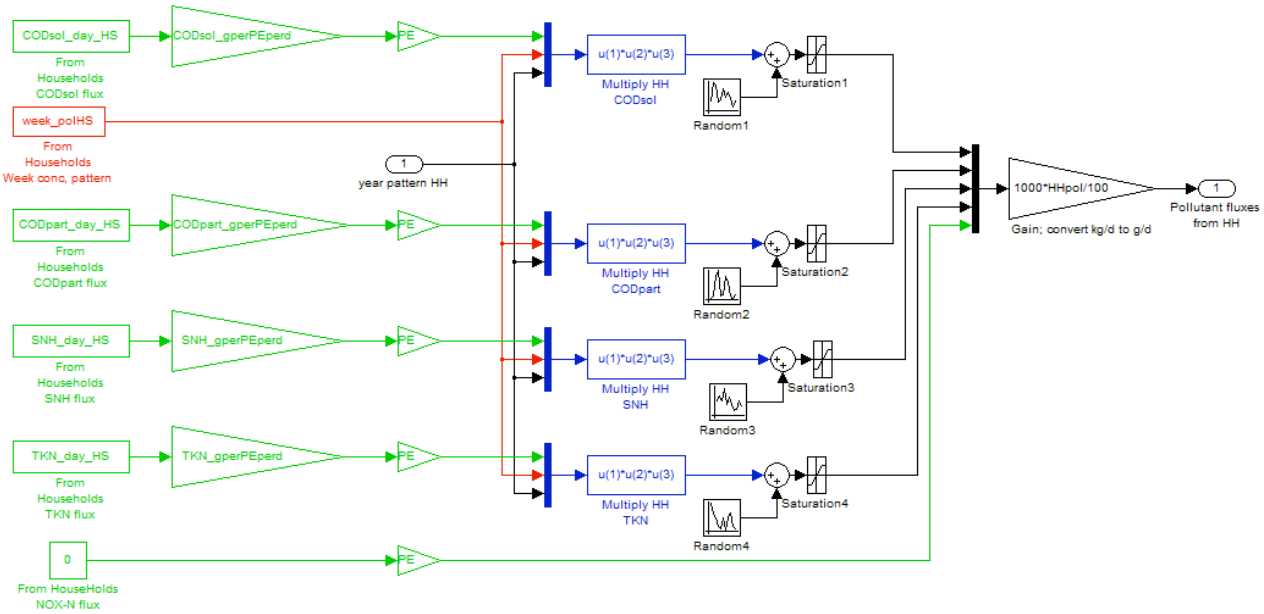

Figure 32. Lay-out of the 'Households pollutants' model block

The pollutant fluxes are transformed into g COD/p.e. per day and g N/p.e. per day units by multiplying the values in the input files with a gain (e.g. CODsol_gperPEperd for soluble COD). They are subsequently multiplied with a gain PE corresponding to the number of p.e.

in the catchment area, similar to the household contribution to the influent model (see 3.2). This gain will convert the fluxes into kg COD/d and kg N/d units. Similar to the household influent flow rate generation contribution (see 3.2), the diurnal pollutant flux profiles are multiplied with the contribution of two other data files to allow simulating a weekend effect (week_polHS) and a holiday effect. For the holiday effect, the same file (year_HS) is used as for the household influent flow rate generation. The weekend effect corresponds to a reduction of the household pollutant fluxes (12% reduction on Saturday, 16% on Sunday) which is 4% more than the reduction of the household wastewater production on the corresponding weekend days.

Noise is added to each pollutant flux, and a saturation model block is inserted to avoid negative values. Further details on the selection of the noise level follows below (see 4.7). Finally, the resulting fluxes are multiplied with a gain to convert units from kg/d to g/d. The parameter HHpol can be used during testing of the model to switch off the contribution of the 'Households pollutants' model block.

4.2.2. Design of 'Household pollutants' model block input files

The input files are designed in terms of fluxes (kg COD/p.e. per day or kg N/p.e. per day). In a first step, the input files to the 'Households pollutants' model block were designed based on the BSM1 dry weather influent pollutant flux dynamics. In a second step, the input file dynamics were adjusted such that the influent concentration profile dynamics were closer to the BSM1 influent dynamics.

For the BSM1 dry weather influent, pollutant fluxes dynamics are quite similar to the influent flow rate dynamics. In the design of the BSM1_LT/BSM2 influent dynamics, an attempt will be made to reduce this correlation (see 4.7). Keeping the practical aspects in mind, it is quite difficult to subdivide influent N into soluble and particulate N. Therefore, influent N fluxes were defined for the directly measurable quantities S_{NH} and TKN, where TKN includes organic N and S_{NH}. The BSM1 influent S_{NH} and TKN fluxes are also strongly correlated to the influent flow rate.

In the first design step, the following calculations were performed to convert the BSM1 dry weather influent pollutant concentration profiles to the BSM1_LT/BSM2 household pollutant flux input files in Figure 33:

- Average hourly pollutant fluxes were calculated based on the BSM1 dry weather influent pollutant fluxes on the 5 week days, thus resulting in diurnal influent profiles with 24 data points each for soluble and particulate COD, S_{NH} and TKN.
- It was assumed that the Households produce 80% of the influent COD flux and 90% of the influent N flux. Thus, the BSM1 influent COD fluxes obtained in the previous calculation were multiplied with 0.8, whereas BSM1 influent N fluxes were multiplied with 0.9.
- The fluxes were converted from kg COD/d and kg N/d to g COD/p.e. d and g N/p.e. d units, where the value for the parameter PE (number of p.e. in the catchment area) was obtained from the influent flow rate model (see also 3.2, and Table 2). It was preferred to express the household input pollutant fluxes per p.e. since that is consistent with the philosophy underlying the construction of the influent flow rate model.
- The resulting diurnal influent pollutant flux profiles were then moved 3 hours ahead in time. The diurnal pollutant flux profiles were moved ahead in time to compensate for

the delays that will be introduced in the influent model via the 'Sewer' model block. As can be seen in Figure 33, N fluxes are a little bit ahead in time compared to COD fluxes.

- The information on pollutant flux dynamics on the one hand and the average daily pollutant flux on the other hand was separated, to increase the flexibility for the model user. To this purpose, the average daily pollutant fluxes were first calculated for soluble and particulate COD, S_{NH} and TKN based on the diurnal influent pollutant flux profiles resulting from the previous step. The values of the average daily pollutant fluxes (in g COD/p.e. per day or g N/p.e. per day) are considered parameters in the influent model (see Table 12). The diurnal influent pollutant flux profiles were subsequently normalised by dividing them with the average daily pollutant flux values. As a result, 4 input files with an average value of 1 are now available. Similar to the flow rate input file (day_HS, see 3.2), these files only contain information on the dynamics, not on the load.

The dynamics in the resulting input files are shown in Figure 33, and are compared with the households flow rate input file (day_HS, see 3.2). All input fluxes have two peaks. A large pollutant flux peak occurs in the morning, and a smaller peak occurs in the evening. The particulate COD peak lags a bit behind compared to the other pollutant fluxes.

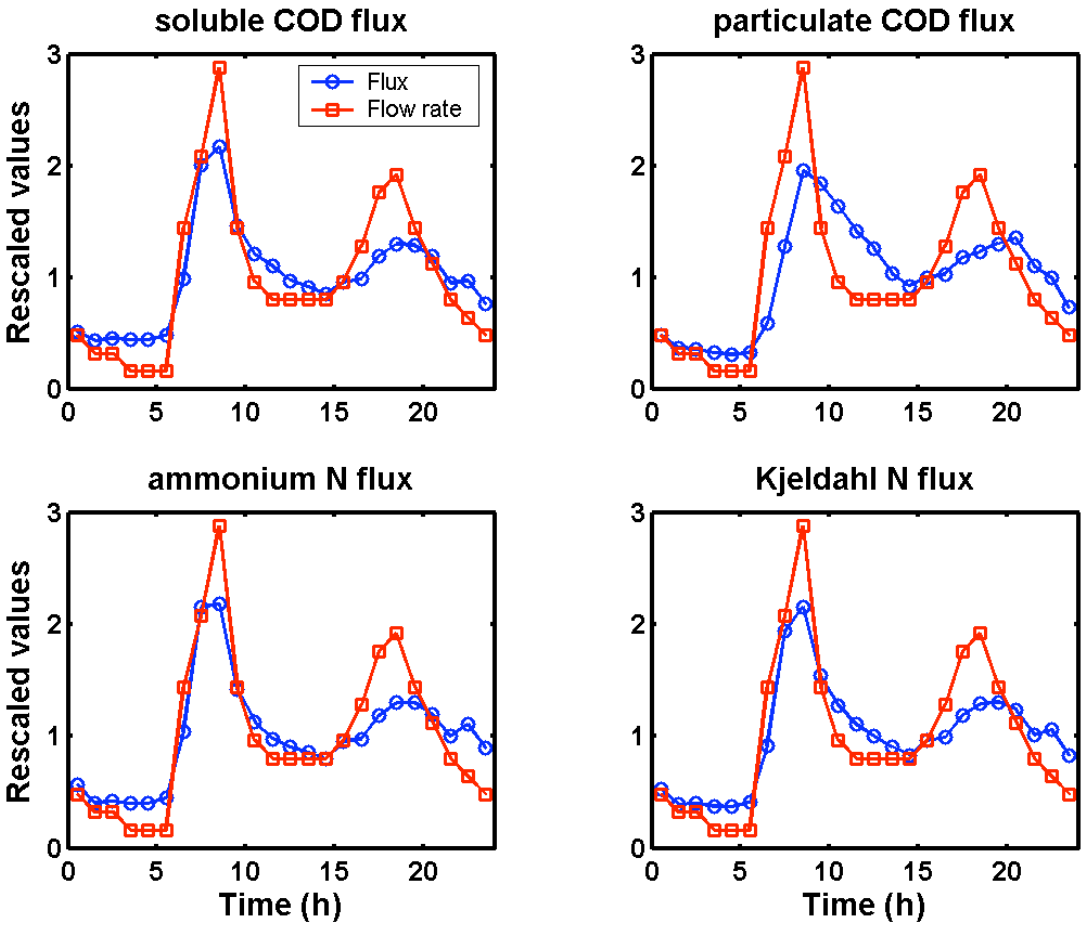

Figure 33. Input files for household pollutant fluxes generation resulting from influent pollutant load dynamics design step 1 (CODsol_day_HS, CODpart_day_HS, SNH_day_HS, TKN_day_HS)

In design step 2, simulations were performed with the influent model, and the pollutant flux dynamics profiles in Figure 33 were iteratively adjusted until the shape of the influent

pollutant profiles resulting from the simulations was satisfactory. The input pollutant flux dynamics for households resulting from design step 2 are illustrated in Figure 34. Compared to the result of design step 1, the pollutant flux peaks have become much sharper. In design step 3 (see 4.10), finally, further adjustments were made to the input files while tuning the noise levels to pollutant concentrations.

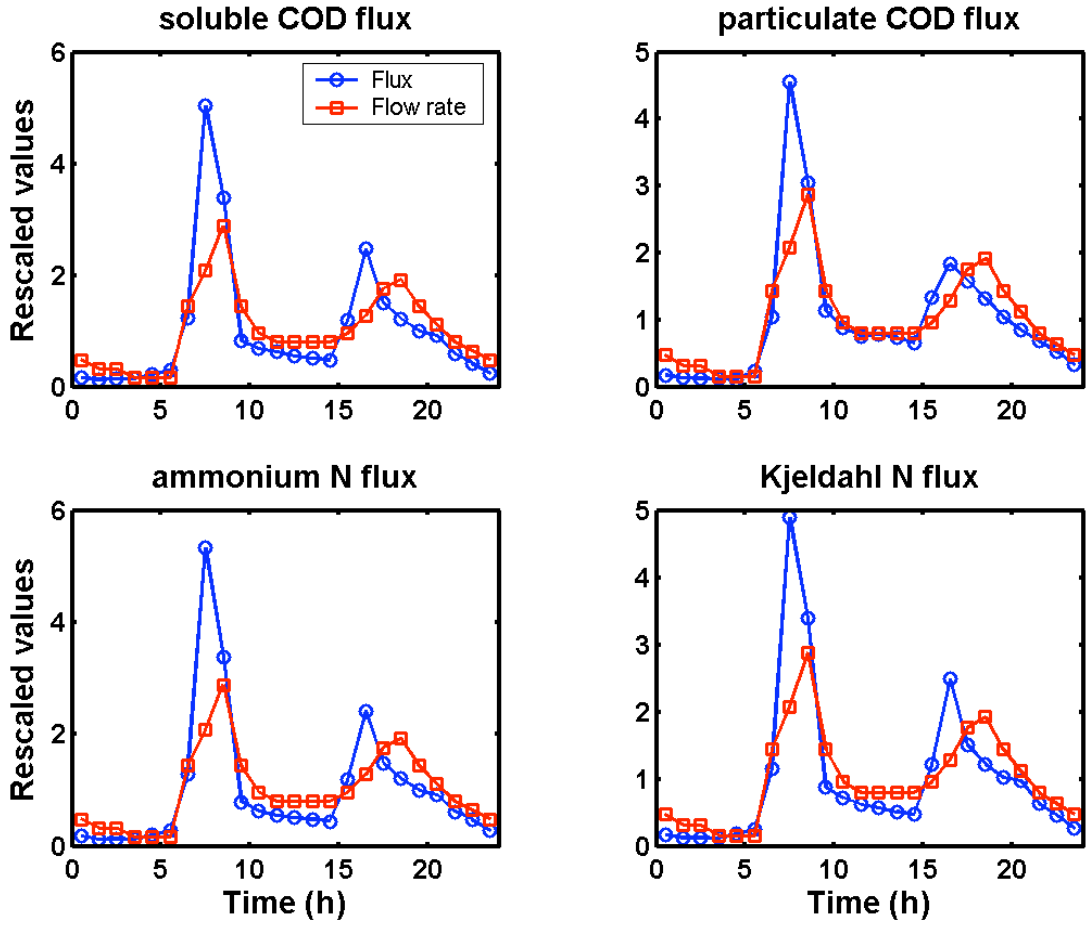

Figure 34. Input files for household pollutant fluxes generation resulting from influent pollutant load dynamics design step 2 (CODsol_day_HS, CODpart_day_HS, SNH_day_HS, TKN_day_HS)

4.2.3. Design of 'Household pollutants' model block noise generators

Returning to Figure 32, it can be seen that zero-mean white noise is added to the influent pollutant fluxes. As for the influent flow rate generation, care is taken such that different noise seeds are selected. The noise variances are parameters in the random number generators, and are set by the user in the influent model initialisation file. The noise variances are proportional to the pollutant fluxes, as illustrated in Table 3. The parameter f_I is a constant that can be tuned by the user (see also 4.7 for details on tuning this constant). Note that it is assumed that the noise variance on the soluble COD and S_{NH} fluxes is twice as high as the noise variance for particulate COD. Furthermore, the noise variance on TKN is 1.5 as high as the noise variance for particulate COD. There are no data to support this assumption. Different noise levels were mainly selected to create slightly different characteristics of the time series that are generated.

Table 3. Calculation of the noise variances for the household pollutant fluxes

Pollutant flux	Noise variance calculation
Soluble COD	$f_1 \cdot 2 \cdot CODsol\_gperPEperd \cdot PE$
Particulate COD	$f_1 \cdot 1 \cdot CODpart\_gperPEperd \cdot PE$
S_{NH}	$f_1 \cdot 2 \cdot SNH\_gperPEperd \cdot PE$
TKN	$f_1 \cdot 1.5 \cdot TKN\_gperPEperd \cdot PE$

4.3. 'Industry pollutants' model block

4.3.1. Model block

The 'Industry pollutants' model block (Figure 35) is quite similar to the 'Households pollutants' model block. Similar to the 'Industry' model block that is part of the influent flow rate model (see 3.3), the 'Industry pollutants' model block only uses 2 user-defined data files for generating each dynamic pollutant flux profile assumed to be produced due to industrial activity. Weekly profiles corresponding to the weekly variation of the industrial pollutant fluxes (one value for each 4 hour period) are made available for soluble COD (CODsol_week_IndS), for particulate COD (CODpart_week_IndS), for S_{NH} (SNH_week_IndS) and for TKN (TKN_week_IndS). Each of these files (Figure 37) contains diurnal pollutant flux variations, a Friday afternoon pollutant flux peak load and a weekend effect. The Friday afternoon pollutant flux peak corresponds to industrial cleaning at the end of the working week. A holiday effect is added to the industry pollutant fluxes completely similar to the 'Industry' model block in the influent flow rate model (see 3.3), and the same input file is used to achieve this (year_IndS). After multiplying the weekly flux pattern with the yearly pattern (year_IndS), zero mean white noise is added via random number generators.

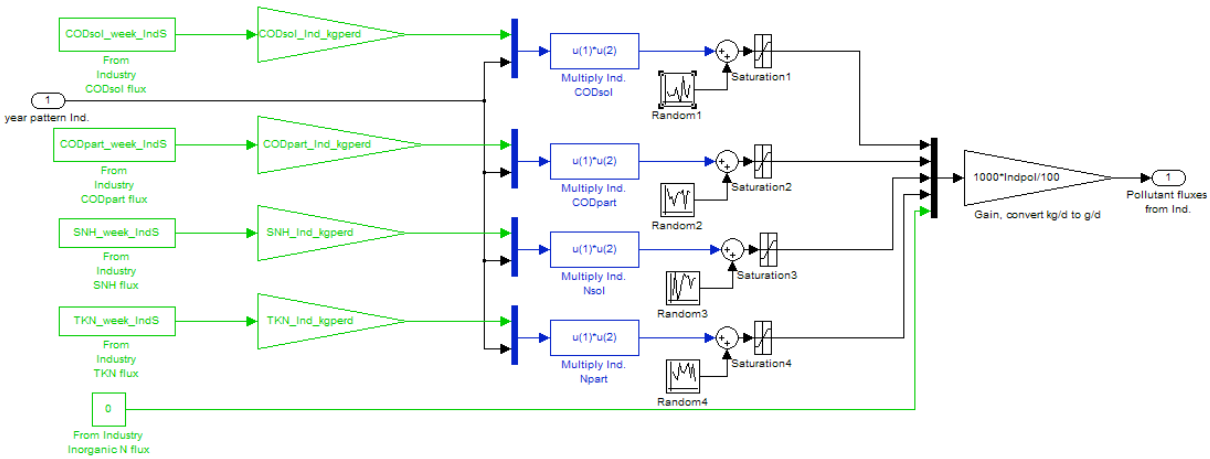

Figure 35. Lay-out of the 'Industry pollutants' model block

4.3.2. Design of 'Industry pollutants' model block input files

The pollutant loads are designed based on the BSM1 pollutant loads. It is assumed that 20% of the influent COD load and 10% of the influent N load is due to industrial activity. The following steps were needed to convert the BSM1 dry weather influent pollutant concentration profiles to the BSM2 industry pollutant flux input files in Figure 37:

- Average hourly pollutant fluxes were calculated based on the BSM1 dry weather influent pollutant fluxes on the 5 week days, thus resulting in diurnal influent profiles with 24 data points each for soluble and particulate COD, S_{NH} and TKN. These data were moved 2 hours ahead in time, to take time delays that will originate from the sewer system into account.
- The 24 hours diurnal profiles obtained in the previous step were averaged 4 by 4, to result in pollutant fluxes with a 4 hour sampling interval.
- It was then assumed that the industry produces 20% of the influent COD flux and 10% of the influent N flux that will reach the treatment plant. Thus, the COD fluxes obtained in the previous calculation were multiplied with 0.2, whereas N fluxes were multiplied with 0.1.
- It was furthermore assumed that the diurnal variations in the industry pollutant load are less extreme compared to the household pollutant fluxes. Therefore the values resulting from the previous calculation were filtered using an average filter, where each new pollutant flux data point corresponds to the weighted mean of three data points (one data point before and one data point after the actual data point were included and each get a weight of 0.5, whereas the actual data point gets a weight equal to 1). The effect of such filtering is illustrated in Figure 36.

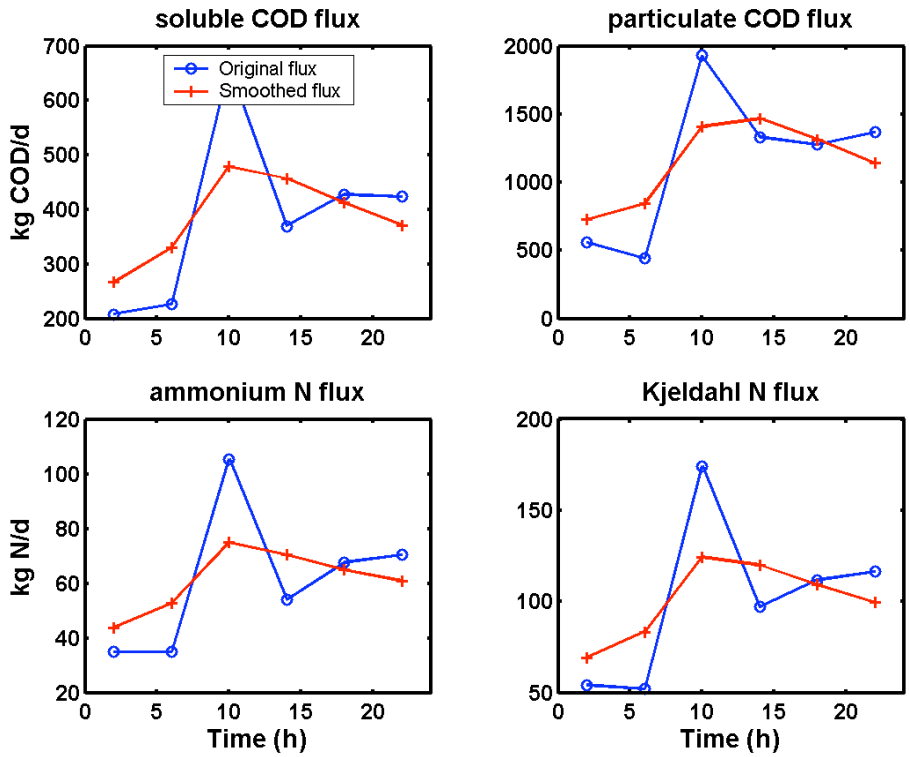

Figure 36. Illustration of the smoothing of the original BSM1 diurnal pollutant flux dynamics (4 hour sampling interval), resulting in the dynamics of the industry pollutant fluxes

- The resulting 6 data points were sampled 7 times, resulting in 1 week of data (42 data points with 4 hour sampling interval). The first data point corresponds to Monday morning, the last data point to Sunday evening.
- On Friday afternoon (data point 28 in the sequence), the industry pollutant fluxes are doubled, which is assumed to correspond to the pollutant load increase due to industrial Friday afternoon cleaning activity.

- It was assumed for the weekend days that the industry pollutant fluxes were considerably lower compared to week days. This will contribute to a weekend effect in the WWTP influent data. The original flux for sample 31 (Saturday morning) and 42 (Sunday evening) were reduced with 60 %, whereas the original pollutant fluxes for samples 32 to 41 were reduced with 80% to obtain weekend day pollutant fluxes.
- Similar to the input files for the household pollutants model block, the information on pollutant flux dynamics on the one hand and the average daily industry pollutant flux on the other hand was separated. The average daily pollutant fluxes were calculated using the first 24 samples of each input sequence, corresponding to normal week days (Monday to Thursday). The values of the average daily industry pollutant fluxes (in kg COD/d or g N/d) are considered parameters in the influent model (see Table 12). The weekly industry pollutant flux input profiles were subsequently normalised by dividing them with the average daily pollutant flux values. As a result, 4 input files with an average value of 1 for the first 4 week days are now available. These files only contain information on the dynamics, not on the load.

The resulting input files are shown in Figure 37, where each of the weekly pollutant flux dynamic input files is compared to the weekly industry flow rate input file (week_IndS).

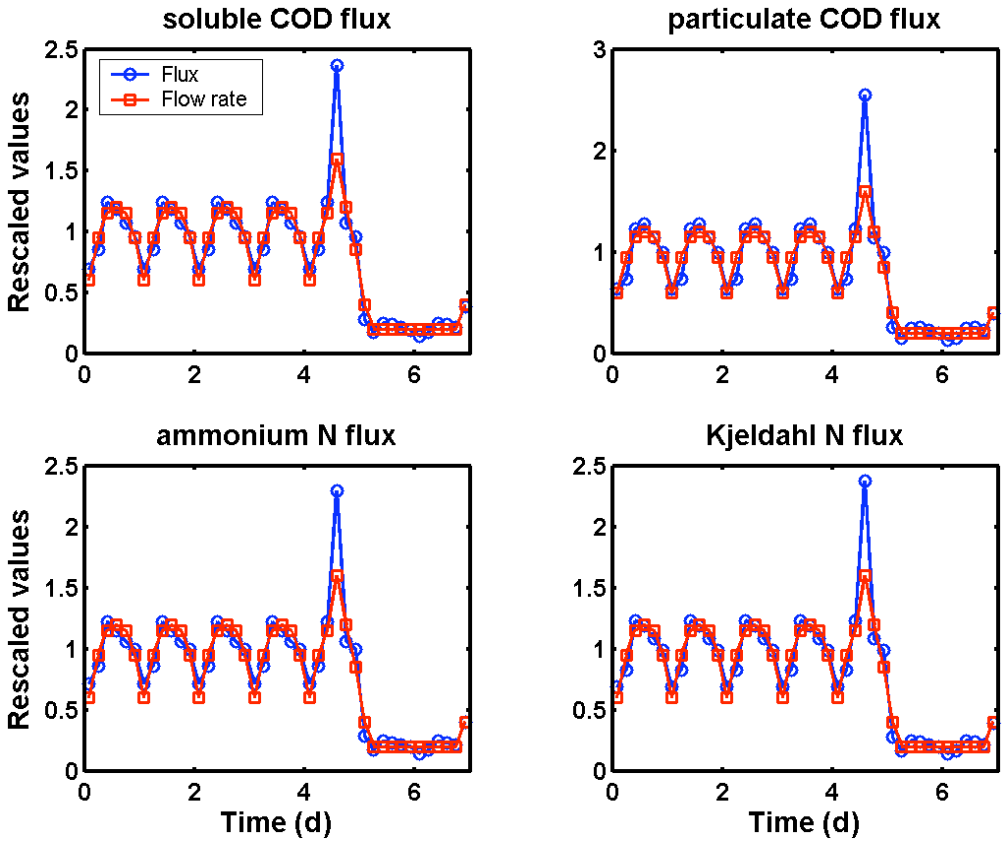

Figure 37. Comparison between the input data for industry pollutant flux generation (CODsol_week_IndS, CODpart_week_IndS, SNH_week_IndS and TKN_week_IndS), and the input data for industry flow rate generation (week_IndS)

4.3.3. Design of 'Industry pollutants' model block noise generators

The procedure for obtaining the noise variances for the random number generators is similar to the generation of noise variances for the 'Household pollutants' model block (see 4.2.3), that is the noise variances are proportional to the industry pollutant fluxes. The parameter f_2 is

a constant that can be tuned by the user (see also 4.7 for details on tuning this constant). It is assumed that the noise levels on the 4 industry pollutant fluxes are similar.

Table 4. Calculation of the noise variances for the industry pollutant fluxes

Pollutant flux	Noise variance calculation
Soluble COD	$f_2 \cdot CODsol\_Ind\_kgperd$
Particulate COD	$f_2 \cdot CODpart\_Ind\_kgperd$
S_{NH}	$f_2 \cdot SNH\_Ind\_kgperd$
TKN	$f_2 \cdot TKN\_Ind\_kgperd$

4.4. 'BSM1_fractionator' model block

The 'BSM1_fractionator' model block will convert the pollutant fluxes to pollutant concentrations that are compatible with the ASM1 model, the model that is used in BSM1_LT/BSM2 to describe the biological processes in the WWTP. The block is written as an S-function in c. It has 12 inputs: two times five pollutant fluxes (5 from households and 5 from industry), the total wastewater flow rate before the sewer system and the flow rate resulting from rainwater collected on impervious surfaces, corresponding to a fraction aH of the output of the 'Rain generator' model block, see 3.8). Note (see also Figure 31) that the influent fractionation, that is the conversion of global variables such as soluble and particulate COD to model-specific variables, in this case for the ASM1, takes place after the fluxes from households and industry have been combined. Although leading to a simpler model, this also means a practical limitation since it inherently introduces the assumption that the fractionation of the wastewater generated in industry and households can be fractionated according to similar principles. Applying a separate fractionation to households and industry wastewater before combining both fluxes would leave the possibility to apply for example different distributions of particulate COD over X_I, X_S and X_{BH}. The latter, combined with the pollutant flux variations of both pollutant sources would have contributed to a reduction of the correlation between these time series as generated at the output of the influent model.

The principles of the influent fractionation implemented in the 'BSM1_fractionator' model block are as follows:

- The S_I concentration is assumed to be constant, equal to 30 g COD/m^3. It is furthermore assumed that S_I is present in the dry weather wastewater flow, meaning that all the flow that has passed the soil model block (e.g. the flow originating in the 'Seasonal correction infiltration' model block) is assumed to contain S_I. Only rainfall runoff from impervious areas, corresponding to a fraction aH of the output of the 'Rain generator' model block (see 3.8), can dilute S_I to concentrations that are significantly below 30 g COD/m^3. Of course, in case the soluble COD flux is not sufficient to provide 30 g COD/m^3 as S_I for the dry weather flow rate, then S_I concentrations will also be diluted below 30 g COD/m^3. In other words, the 'BSM1_fractionator' model block is not a COD source. This situation with lower S_I concentrations due to low soluble COD fluxes will typically occur in low loaded periods such as weekends and holiday periods.
- The S_S concentration is obtained based on the difference between the soluble COD flux, which contains both a contribution from households and from industry, and the S_I pollutant flux.
- X_{BA} and X_P are assumed to be zero.

- The particulate COD is distributed over X_I, X_S and X_{BH}. The different COD fractions are determined based on the BSM1 flow weighted average dry weather influent composition (see Table 5). The X_I fraction of the particulate COD, for example, is equal to $X_{I,inf}/(X_{I,inf}+X_{S,inf}+X_{BH,inf})$. This results in the following particulate COD fractions: 18.2% X_I, 71.8% X_S and 10.0 % X_{BH}.
- The S_O concentration is assumed to be zero.
- The S_{NO} concentration is zero, since both the households and the industry S_{NO} flux are zero.
- The S_{NH} concentration is calculated based on the S_{NH} flux.
- The organic N flux is obtained by subtracting the S_{NH} flux and the N flux corresponding to $i_{XB} \cdot X_{BH}$ from the TKN flux. If the result of this calculation is negative, influent S_{ND} and X_{ND} are assumed to be zero. If the result of this calculation is positive, the organic N is distributed between S_{ND} and X_{ND} based on the ratio between S_{ND} and X_{ND} in the BSM1 flow weighted average dry weather influent composition (Table 5). The resulting S_{ND} and X_{ND} fractions are 39.6% and 60.4% respectively.

Table 5. BSM1 flow weighted average dry weather influent composition

	Concentration	Unit		Concentration	Unit
$S_{I,inf}$	30.00	g COD/m^3	$S_{O,inf}$	0.00	g - COD/m^3
$S_{S,inf}$	69.50	g COD/m^3	$S_{NO,inf}$	0.00	g N/m^3
$X_{I,inf}$	51.20	g COD/m^3	$S_{NH,inf}$	31.56	g N/m^3
$X_{S,inf}$	202.32	g COD/m^3	$S_{ND,inf}$	6.95	g N/m^3
$X_{BH,inf}$	28.17	g COD/m^3	$X_{ND,inf}$	10.59	g N/m^3
$X_{BA,inf}$	0.00	g COD/m^3	$S_{ALK,inf}$	7.00	g HCO_3^-/m^3
$X_{P,inf}$	0.00	g COD/m^3	Q_{inf}	18446.0	m^3/d

4.5. 'Noise generator' model block

The 'Noise generator' model block (Figure 38) is inserted between the 'BSM1_fractionator' model block and the 'Sewer' model block, and will add on zero-mean white noise to the concentrations of S_I, S_S, X_I, X_S, X_{BH}, S_{NH}, S_{ND} and X_{ND} resulting from the 'BSM1_fractionator'. This model block is assumed to be necessary to avoid for example that influent X_I, X_S and X_{BH} concentrations are perfectly correlated to each other. The structure of the noise generators is similar to the other noise generators that have been included in the influent model, i.e. noise is added via a random number generator (see Figure 38 for an example). The noise variances are based on the BSM1 flow weighted average dry weather influent pollutant concentrations (see Table 5) multiplied with a gain f_3.

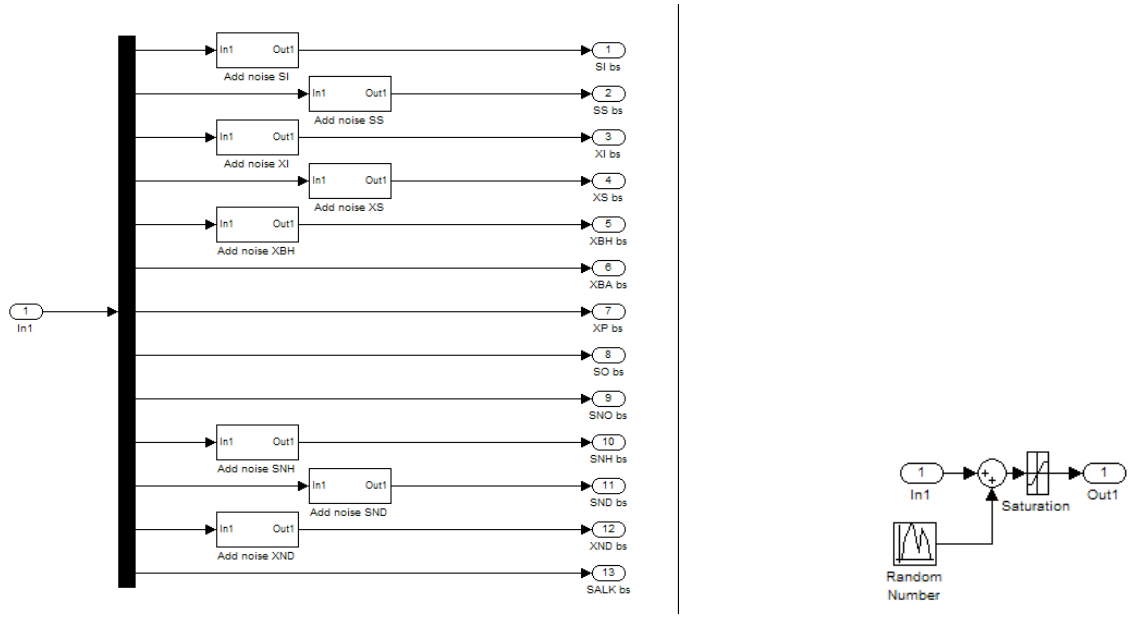

Figure 38. The 'Noise generator' model block. Left: The lay-out of the 'Noise generator' model block; Right: Detail of one subsystem in the 'Noise generator' model block

4.6. 'Sewer' model block

The pollutant concentrations are passed through the same 'Sewer' model block as the flow rate (see 3.6), and will therefore not be commented in detail. The same parameter 'subareas' is used in this model block as for the 'Sewer' model block (see also Table 2). No noise is added on to the output of the 'Sewer' model block. Note that the construction of the sewer model block (see Figure 13) implies the assumption that the pollutant discharges are distributed evenly along the sewer system as soon as the parameter 'subareas' is greater than 1.

As an example, the differential equation resulting from the mass balance for one of the pollutant concentrations in a tank in the 'Sewer' model block is provided in Eq. 8

$$\frac{d(S_S \cdot V)}{dt} = \left(Qin \cdot S_{S,in} - Qout \cdot S_S\right) \quad (8)$$

In Equation 8, $S_{S,in}$ represents the S_S concentration in the inlet to the tank, Qin is the flow rate into the tank, and Qout is the flow rate out of the tank, which is given by Eq. 6, and the volume change of the liquid in the tank is given by Eq. 5.

4.7. Dry weather simulation results

The different model blocks for pollutant concentration generation available at this point were validated through a number of simulations. The aim of these simulations was two-fold: 1) verify whether the mass balances hold, as a general check for the correct implementation of the model blocks implemented thus far; 2) investigate the effect of the pollutant concentration noise generators. The different simulations are summarised in Table 6.

Table 6. Overview of the simulations performed at this stage of the model development. The influent model parameters of Table 2 and Table 12 were used for all simulations

Simulation	Description
'no noise'	No random noise generated
'noise 1'	Noise generated, $f_1 = 0.5$; $f_2 = 0.5$;$f_3 = 0.5$; Q_HH_nv=400
'noise 2'	Noise generated, $f_1 = 1.0$; $f_2 = 1.0$;$f_3 = 1.0$; Q_HH_nv=400
'noise 3'	Noise generated, $f_1 = 2.0$; $f_2 = 2.0$;$f_3 = 2.0$; Q_HH_nv=400
'noise 4'	Noise generated, $f_1 = 2.0$; $f_2 = 2.0$;$f_3 = 2.0$; Q_HH_nv=4000
'noise 5'	Noise generated, $f_1 = 3.0$; $f_2 = 3.0$;$f_3 = 3.0$; Q_HH_nv=4000

4.7.1. Simulation without noise generation

It was first checked whether the mass balances hold. For the 'no noise' simulation this can be validated on the available simulation results since the fluxes at the input of the influent model should be identical to the fluxes at the output. The daily fluxes in the output of the influent model are illustrated in Figure 39. In that figure, the first day is a Monday, and the resulting pollutant fluxes are slightly lower compared to the fluxes on Tuesday. That can be explained because there is still a dilution effect of the lower weekend influent flow rates and pollutant loads that propagates to the Monday via the wastewater stored in the sewer system. The total daily pollutant fluxes are identical on Tuesday, Wednesday and Thursday. These fluxes can be used for validation, i.e. they can be compared to the pollutant fluxes applied at the input of the influent model. The comparison between input (= design) fluxes and simulation outputs is summarised in Table 7. The results in that table demonstrate that the design fluxes equal the output fluxes, and thereby indicate that there are no significant leaks in the different model blocks used for influent flow rate and concentration generation. It can thus be assumed that the model blocks described thus far work properly.

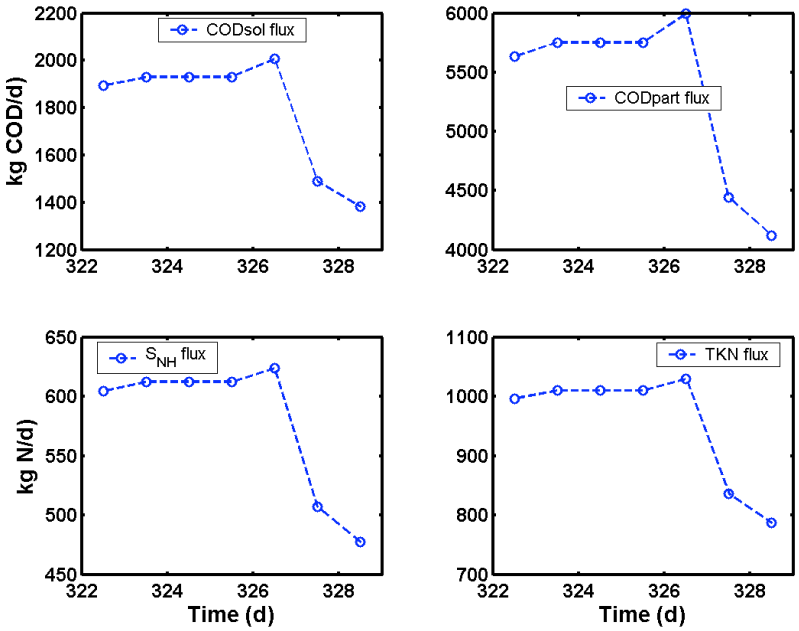

Figure 39. Simulation 'no noise': total daily fluxes for soluble and particulate COD, S_{NH} and TKN

Table 7. Simulation 'no noise': comparison between input and output pollutant fluxes. Output pollutant fluxes correspond to Tuesday-Wednesday-Thursday (day 324 - 326) in Figure 39

	Influent design (input)			Simulation
	Households	Industry	Total	Output
Soluble COD (kg COD/d)	1 545.0	386.2	1 931.2	1 931.0
Particulate COD (kg COD/d)	4 603.4	1 150.9	5 754.3	5 753.8
S_{NH} (kg N/d)	551.3	61.3	612.5	612.4
TKN (kg N/d)	909.0	101.0	1 010.0	1 009.8

The concentration dynamics resulting from the 'no noise' simulation are illustrated in Figure 40. Contrary to what could be expected, the S_I concentration is not constant for the 'no noise' simulation. The decrease of the influent S_I concentration generated with the model during day 329 (Figure 40) originates from the 'BSM1_fractionator' model block, and occurs during periods with low soluble COD input fluxes. During these periods, the total input soluble COD flux will not be sufficient to generate an influent with a concentration of 30 g COD/m^3 as S_I, since it is assumed in the 'BSM1_fractionator' model block that soluble COD is first used to generate S_I, whereas the remainder is assumed to be S_S. This dilution of S_I is further increased by the assumption that the influent flow rate resulting from the 'soil model block', corresponding to infiltration water, is also assumed to contain 30 g COD/m^3 S_I.

The concentrations of soluble and particulate COD fractions show a clear decrease in weekends, whereas the concentration profiles of S_{NH}, S_{ND} and X_{ND} are not significantly different on weekend days, when compared to normal week days.

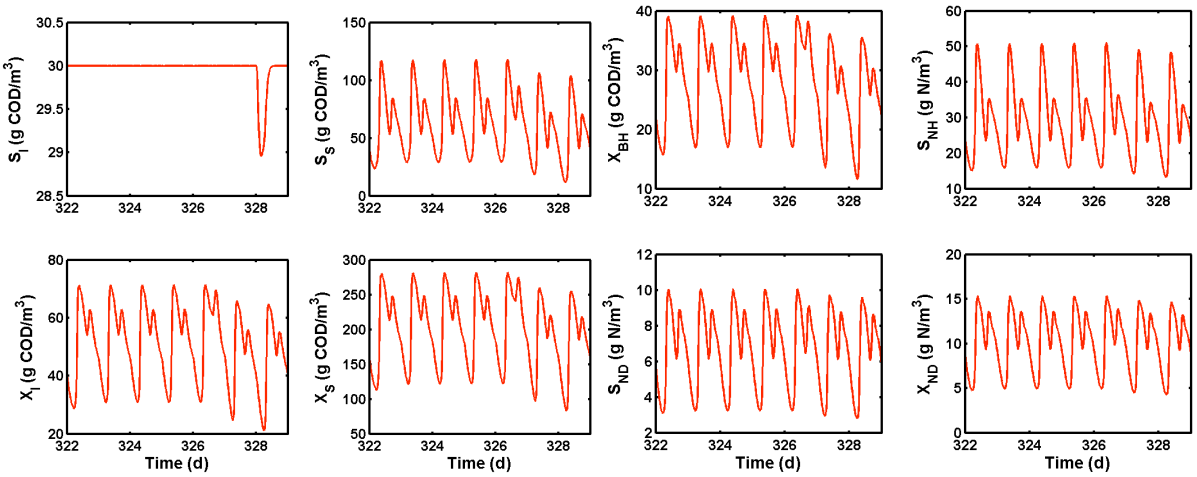

Figure 40. Simulation 'no noise': an example of influent concentration dynamics for the ASM1 influent variables S_I, S_S, X_I, X_S, X_{BH}, S_{NH}, S_{ND} and X_{ND}

The correlation coefficients between different influent time series were calculated (Table 8). In general, correlation coefficients between different influent concentration time series are high, except for S_I. The low correlation coefficient between S_I and the other time series is due to the fact that S_I is almost constant, apart from dilution effects in situations with low soluble COD input fluxes, whereas the dynamics of all other influent time series will include pronounced diurnal effects. As expected, the correlation coefficients between X_I, X_{BH}, X_S and X_{TSS} concentrations are 1. This is evident, since X_I, X_{BH} and X_S are obtained as constant fractions of the same quantity in the 'BSM1_fractionator' model block, whereas X_{TSS} is obtained by multiplying the sum of X_I, X_S and X_{BH} with a constant. The latter effect is

because the conversion factors for X_I, X_S and X_{BH} are identical (0.75 g TSS/g COD). The correlation coefficient between S_{ND} and X_{ND} is also 1 since S_{ND} and X_{ND} are obtained as constant fractions of the same quantity (remaining TKN after subtracting S_{NH} and N content of X_{BH}) in the 'BSM1_fractionator' model block. Reducing the correlation between X_I, X_S and X_{BH} on the one hand, and between S_{ND} and X_{ND} on the other hand, will be possible due to the noise generators implemented in the 'Noise generator' model block. Reducing the correlation between particulate COD and TSS is only possible by applying different COD to TSS conversion factors for X_I, X_S and X_{BH}.

Table 8. Simulation 'no noise': correlation coefficients between different influent time series corresponding to 1 week of data (see Figure 40) generated with the influent model

	S_I	S_S	X_I	X_S	X_{BH}	S_{NH}	S_{ND}	X_{ND}	X_{TSS}	Q
S_I	1.0000									
S_S	0.3126	1.0000								
X_I	0.3581	0.9523	1.0000							
X_S	0.3581	0.9523	1.0000	1.0000						
X_{BH}	0.3581	0.9523	1.0000	1.0000	1.0000					
S_{NH}	0.2439	0.9805	0.9084	0.9084	0.9084	1.0000				
S_{ND}	0.2905	0.9411	0.9589	0.9589	0.9589	0.9383	1.0000			
X_{ND}	0.2905	0.9411	0.9589	0.9589	0.9589	0.9383	1.0000	1.0000		
X_{TSS}	0.3581	0.9523	1.0000	1.0000	1.0000	0.9084	0.9589	0.9589	1.0000	
Q	0.2912	0.8912	0.8784	0.8784	0.8784	0.8750	0.9145	0.9145	0.8784	1.0000

Table 9. Simulation 'no noise': correlation coefficients between different global variables calculated from 1 week of influent time series data (see Figure 40) generated with the influent model

	CODsol	CODpart	S_{NH}	TKN	X_{TSS}	Q
CODsol	1.0000					
CODpart	0.9527	1.0000				
S_{NH}	0.9801	0.9084	1.0000			
TKN	0.9814	0.9437	0.9915	1.0000		
X_{TSS}	0.9527	1.0000	0.9084	0.9437	1.0000	
Q	0.8912	0.8784	0.8750	0.9027	0.8784	1.0000

The cross-correlations between relevant influent time series are plotted in Figure 41 and Figure 42 for 96 sample time lags (= 1 day). In general, the cross-correlation is rather high for the first 10 sample times. The fact that the cross-correlation increases again when approaching 96 samples is due to the diurnal influent flow rate and pollutant concentration variations.

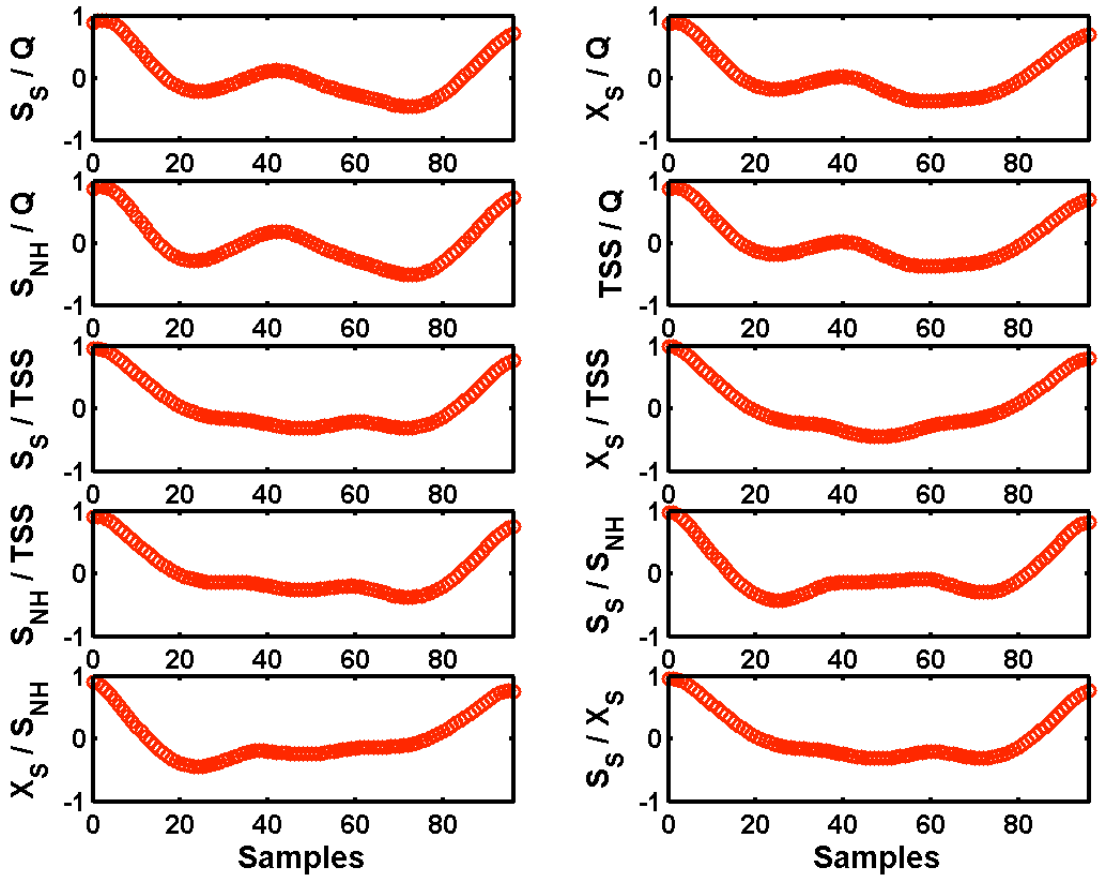

Figure 41. Simulation 'no noise': cross-correlation between different time series generated with the influent model, calculated for the data presented in Figure 40;. The sampling interval is 15 minutes

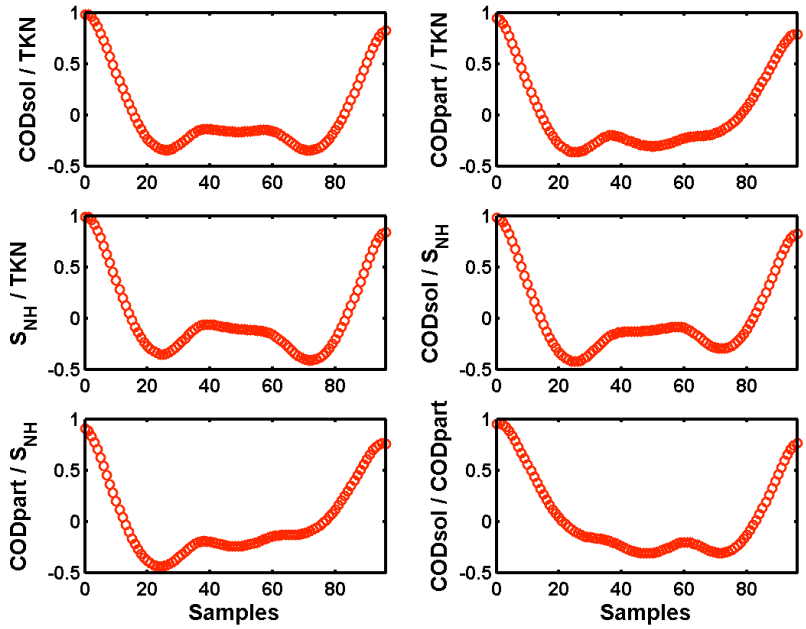

Figure 42. Simulation 'no noise': cross-correlation between different time series generated with the influent model, calculated for the data presented in Figure 40; The sampling interval is 15 minutes

4.7.2. Simulation with noise generation

Only results of simulation 'noise 4' (Table 6) will be discussed in detail. Details on the other simulations (with increasing noise levels) are not provided. The daily fluxes in the output of the model are illustrated in Figure 43, which should be compared with the results presented in Figure 39. The effect of adding on noise is quite clear. Where daily pollutant fluxes were identical for Tuesday, Wednesday and Thursdays in Figure 39, fluxes in Figure 43 are quite different on these days due to the effect of the random number generators in the 'noise 4' simulation. Still, noise levels are not that high that they would obscure the difference between week days and weekend days. Figure 44 provides an example of the dynamic influent profiles generated in the 'noise 4' simulation. Contrary to the results for the 'no noise' simulation (Figure 40), diurnal concentration profiles on subsequent dry weather days can look rather different.

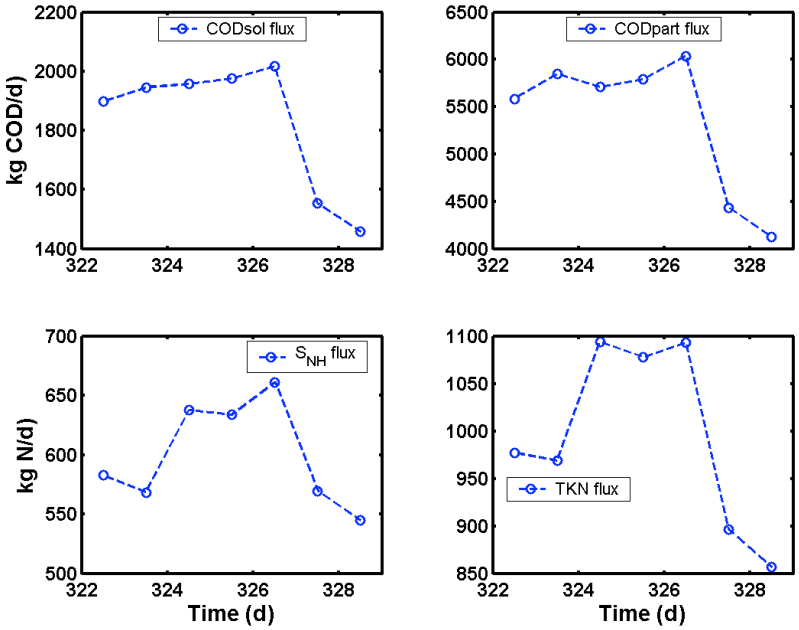

Figure 43. Total daily fluxes for soluble and particulate COD, S_{NH} and TKN; Simulation 'noise 4'

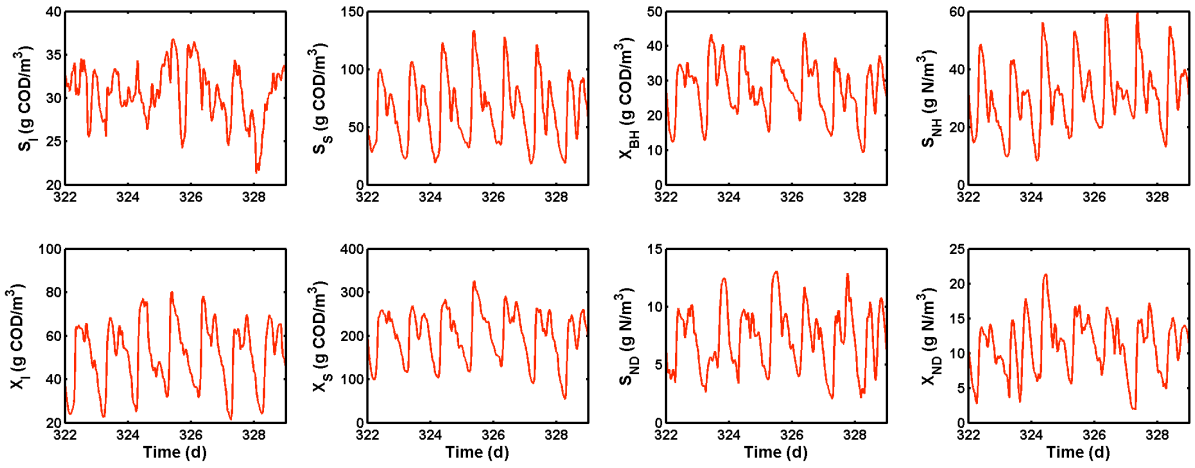

Figure 44. Influent concentration dynamics for S_I, S_S, X_I, X_S, X_{BH}, S_{NH}, S_{ND} and X_{ND}; Simulation 'noise 4'

The correlation coefficients in Table 10 should first be compared to the results presented in Table 8. As a consequence of the random noise generators, the correlation coefficients between different time series have in general decreased. Especially the correlations between pollutant concentration values and the flow rate Q has reached more acceptable levels. Other coefficients, for example the correlation coefficients between S_{NH} and S_S or between X_S and S_S, are still too high to be acceptable. This becomes clear when comparing for example the correlation coefficient between X_S and S_S for BSM1, which is only 0.51. This high correlation coefficient in the 'noise 4' simulation should be reduced further before we can consider the quality of the generated influent data resulting from the influent model as acceptable.

Table 10. Correlation coefficients between different influent time series generated with the influent model for the data presented in Figure 44; Simulation 'noise 4'

	S_I	S_S	X_I	X_S	X_{BH}	S_{NH}	S_{ND}	X_{ND}	TSS	Q
S_I	1.0000									
S_S	0.2190	1.0000								
X_I	0.1644	0.9012	1.0000							
X_S	0.2215	0.9199	0.9534	1.0000						
X_{BH}	0.2583	0.8696	0.8944	0.9203	1.0000					
S_{NH}	0.1589	0.9149	0.8028	0.7978	0.7456	1.0000				
S_{ND}	0.1323	0.7817	0.8238	0.8584	0.7709	0.6734	1.0000			
X_{ND}	0.0990	0.8048	0.8511	0.8601	0.7972	0.7302	0.8404	1.0000		
TSS	0.2174	0.9242	0.9696	0.9974	0.9359	0.8047	0.8553	0.8642	1.0000	
Q	0.0312	0.7400	0.6727	0.7274	0.7042	0.6685	0.6429	0.6384	0.7250	1.0000

Results in Table 11 should be compared to results in Table 9. Similar to the previous table, correlations between different influent variables have decreased by adding noise. In some cases, for example the TSS versus CODpart correlation coefficient, the correlation coefficient is not influenced by noise since there is an identical COD to TSS conversion coefficient for all particulate COD fractions, as mentioned before. Again, correlation coefficients between flow rate and other influent concentration variables are reaching more acceptable (low) levels, whereas other correlation coefficients can still be considered too high.

Table 11. Simulation 'noise 4': correlation coefficients between different global variables calculated from 1 week of influent time series data (see Figure 40) generated with the influent model

	CODsol	CODpart	S_{NH}	TKN	TSS	Q
CODsol	1.0000					
CODpart	0.9206	1.0000				
S_{NH}	0.9051	0.8047	1.0000			
TKN	0.9350	0.9012	0.9629	1.0000		
TSS	0.9206	1.0000	0.8047	0.9012	1.0000	
Q	0.7214	0.7250	0.6685	0.7167	0.7250	1.0000

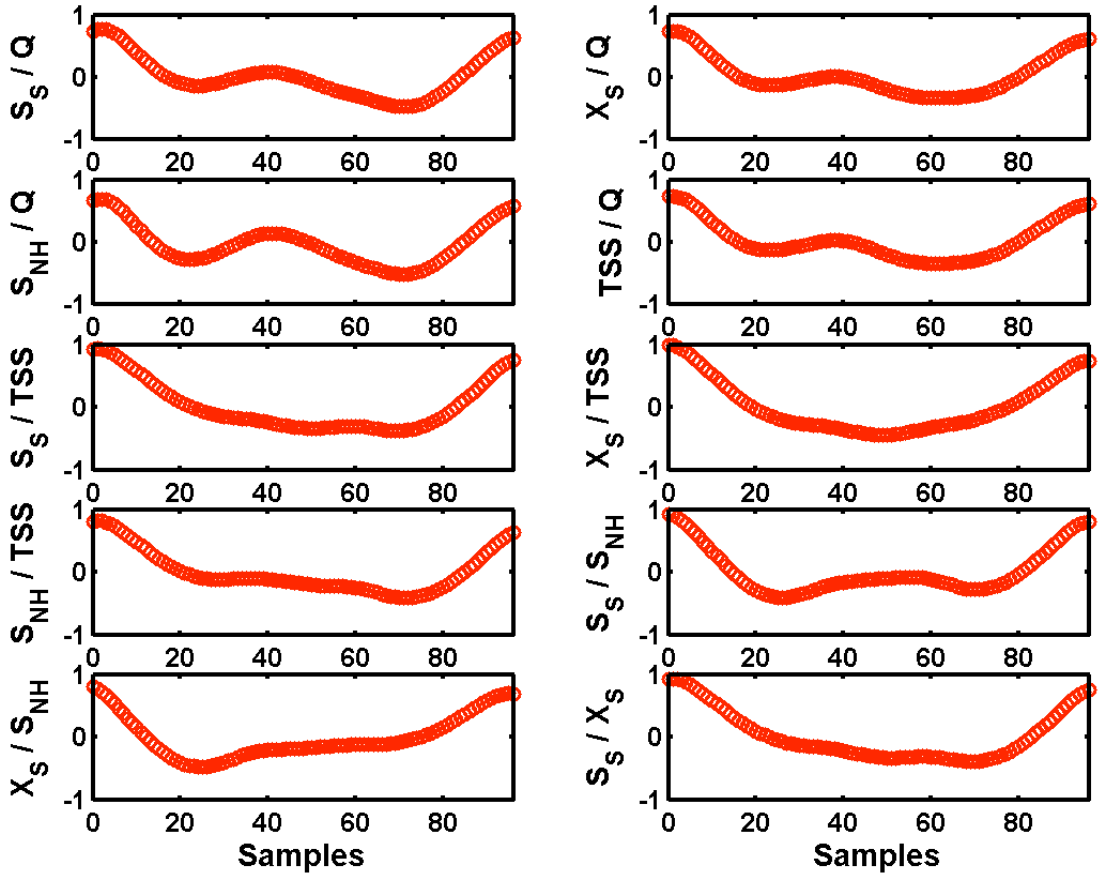

Figure 45. Cross-correlation between time series generated with the influent model for the data presented in Figure 44; Simulation 'noise 4'. The sampling interval is 15 minutes

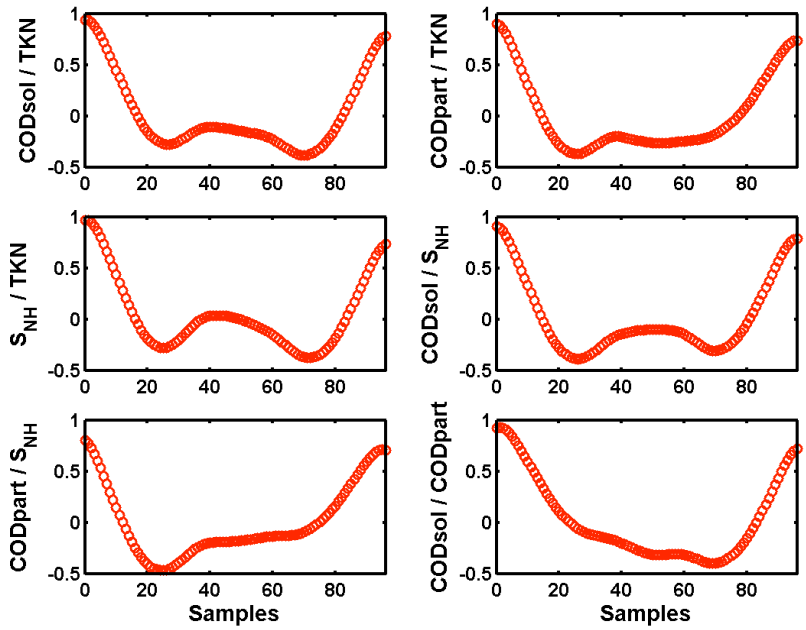

Figure 46. Cross-correlation between time series generated with the influent model for the data presented in Figure 44; Simulation 'noise 4'. The sampling interval is 15 minutes

4.8. 'First flush effect generation' model block

As mentioned before, it was attempted to include a 'first flush' effect in the TSS profile, meaning that severe rain events will result in a flushing of the sewer system. First of all, it was attempted to select a suitable mathematical function to realise the switching between two situations, i.e. in this case the switching between sedimentation in the sewer system and wash-out of all sediments from the sewer system. Furthermore, a simple way of modelling the first flush effect is proposed, and some practical problems are discussed. Finally, the lay-out of the 'First flush effect generation' model block is explained in more detail.

4.8.1. Selection of an appropriate switch function

This part of the influent model includes two effects: (1) sedimentation in the sewer system during dry weather conditions; (2) a quick re-suspension of the deposited solids in the sewer for increasing flow rates, typically corresponding to rain weather conditions, to create a first flush effect. The switching between both behaviours is obtained by a Hill function (Eq. 9).

$$V = V_{max} \frac{Q^n}{Q_{lim}^n + Q^n} \qquad (9)$$

In this equation, V is the re-suspension rate of sediments from the sewer, whereas V_{max} is the maximum re-suspension rate. Q corresponds to the flow rate and Q_{lim} is a flow rate limit triggering the switching. The Hill function is illustrated in Figure 47. For n = 1, the Hill function reduces to standard Monod kinetics typical for activated sludge models. However, increasing the value of the parameter n results in a sigmoidal function, that can serve as a switch between two behaviours. For the re-suspension of sediments in the sewer, low flow rates correspond to a situation where the switch is off (V = 0 in the example in Figure 47), whereas high flow rates will cause wash-out of sediments from the sewer (V = V_{max} in Figure 47). Q_{lim} is the parameter that determines for which flow rate the switch is activated.

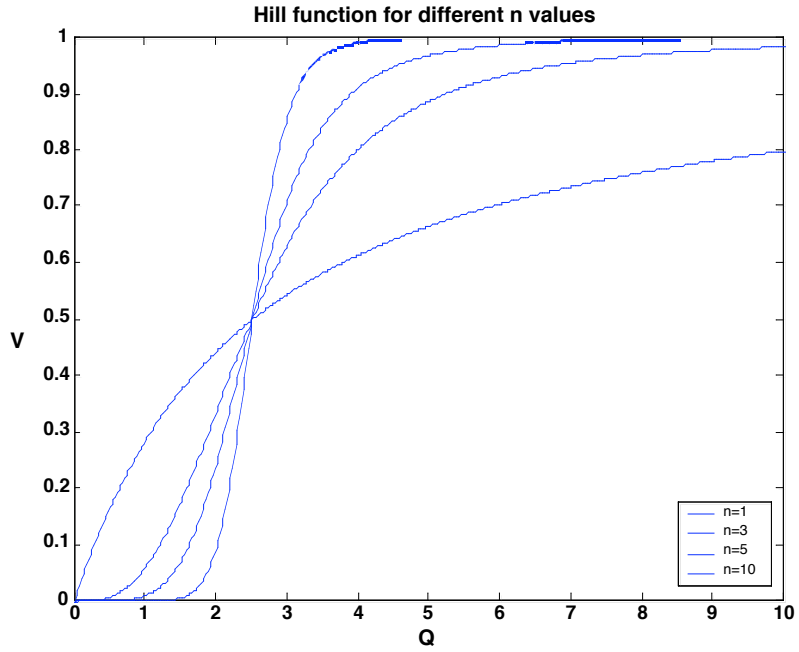

Figure 47. Hill function (Eq. 1) for different values of n. V_{max} = 1, Q_{lim} = 2.5 (arbitrary flow rate units)

4.8.2. A simple model to represent first flush effects

A model was implemented as an S function in c (see Figure 48) to provide a description of first flush effects in the sewer system. The total wastewater flow rate entering the sewer system and the concentrations of particulate ASM1 model components are the inputs. Standard ASM1 parameters for conversion of particulate COD to TSS are used to calculate the TSS concentration of the wastewater flow entering the 'First flush effect generation' model block. It is assumed that only part of the particulate material can settle in the sewer system. This settleable fraction is a parameter (FFfraction) that is to be chosen by the model user, and was here chosen to be 0.25. The settleable fraction of the wastewater TSS concentration is passed through the S-function, which forms the core of the 'First flush effect generation' model block (see Eq. 10)

$$\frac{dM}{dt} = Q_{in} \cdot TSS_{in} \cdot (1 - \frac{M}{M_{max}}) - \frac{(Q_{in})^n}{(Q_{lim})^n + (Q_{in})^n} \cdot M \cdot FF \qquad (10)$$

In this equation Q_{in} represents the influent flow rate (m^3/d), whereas TSS_{in} represents the SS concentration that forms the input to the S-function, i.e. corresponding to the parameter FFfraction multiplied with the TSS concentration (g/m^3) in the wastewater. M_{max} is the maximum amount of TSS that can be stored in the sewer system, and was selected to be equal to 1 000 kg. Q_{lim} (70 000 m^3/d) is the flow rate limit triggering the first flush effect, and FF (equal to 500) is a dimensionless adjustable parameter to tune the desired strength of the first flush effect. The dimensionless parameter n was set equal to 15.

There are 2 different phenomena included in Equation 10: (1) there is a dry weather phenomenon, where TSS can accumulate in the sewer system (settling, build up of sediments at the bottom of the sewer pipes). For dry weather conditions, the mass balance in Equation 10 reduces to $\frac{dM}{dt} = Q_{in} \cdot TSS_{in} \cdot (1 - \frac{M}{M_{\max}})$, meaning that the amount of sediments in the sewer will increase as long as the total amount of sediments stored in the sewer system is below M_{max}; (2) for rain events, the sudden increase of the flow rate should result in a washout of the sediments from the sewer (first flush). The switching function in the second part of the equation will induce the switching: its value will be zero for dry weather flow and 1 for rain weather flow conditions. It should be mentioned explicitly that no biological or chemical reactions are implemented in the 'First flush effect generation' model block.

Mass balances for particulate COD components and for X_{ND} are closed in the 'First flush effect generation' model block. To that purpose, a separate differential equation was implemented for each COD fraction and for X_{ND}. An example is presented in Equation 11.

$$\frac{dM_{X_I}}{dt} = Q_{in} \cdot X_{I,in} \cdot (1 - \frac{M}{M_{max}}) - \frac{(Q_{in})^n}{(Q_{lim})^n + (Q_{in})^n} \cdot M_{X_I} \cdot FF \qquad (11)$$

In Equation 11, M_{X_I} represents the mass of X_I that is stored in the sewer system, corresponding to a fraction of the TSS that is accumulated in the sewer system. $X_{I,in}$ corresponds to the X_I concentration in the inlet to the 'First flush effect generation' model block.

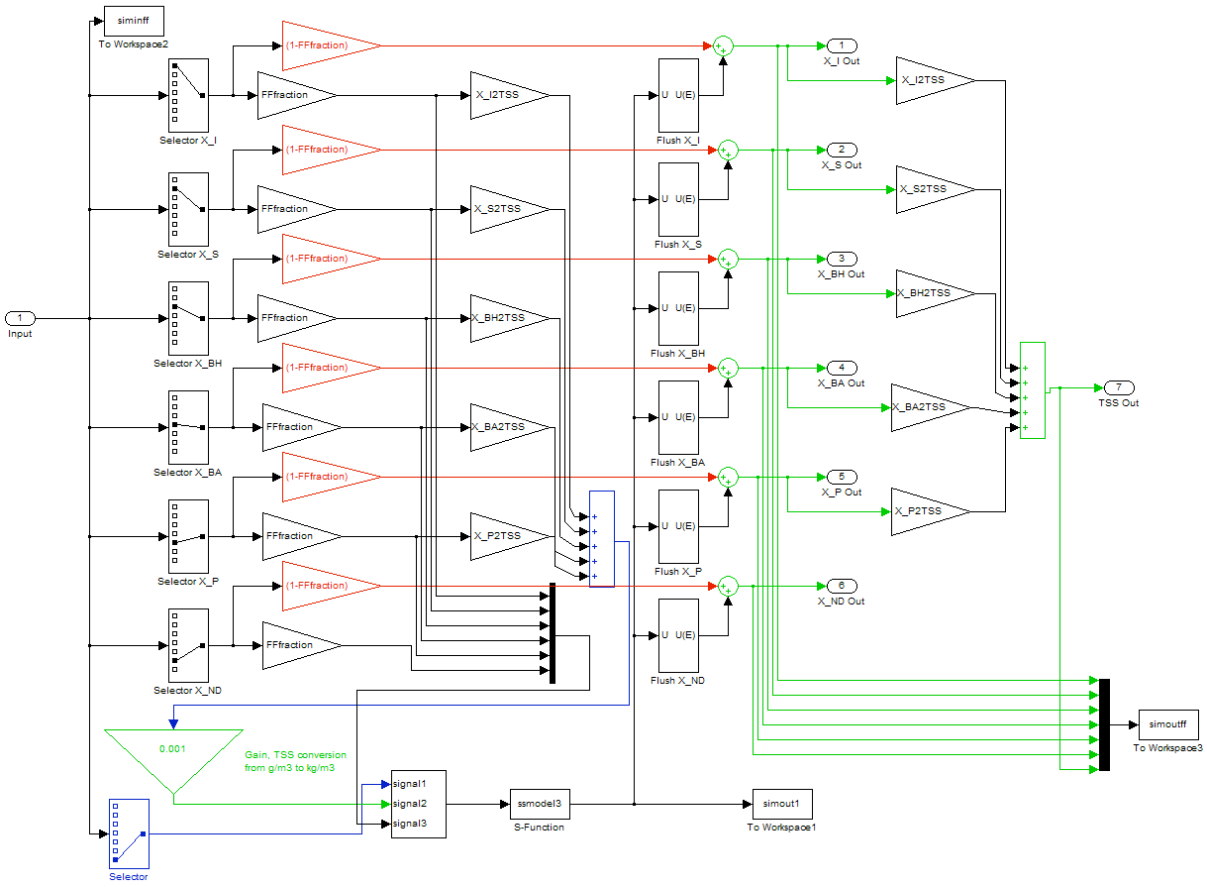

Figure 48. Lay-out of the 'First flush effect generation' model block

4.9. Simulation results full influent model

A few plots with results are provided to illustrate the dynamic influent concentration profiles that can be generated by the proposed model. The plots correspond to the 'noise 4' simulation in Table 6. A year of average influent S_S concentration and flow rate data generated with the model is shown in Figure 49. The simulation leading to the results in Figure 49 was done using the complete functionality of the influent model, meaning that the first flush effect and the generation of rain events were included. The seasonal dilution effect of the infiltration water should be responsible for slight seasonal concentration variations because the input pollutant loads over the year are constant. However, these seasonal concentration variations are to a large extent hidden as a consequence of the noise that was added to the influent model input files. It is a bit easier to notice the seasonal variation in the influent flow rate profile (Figure 49). The low S_S concentration values that sometimes appear represent the dilution effect of rainfall on the pollutant concentrations, and correspond to flow rate peaks.

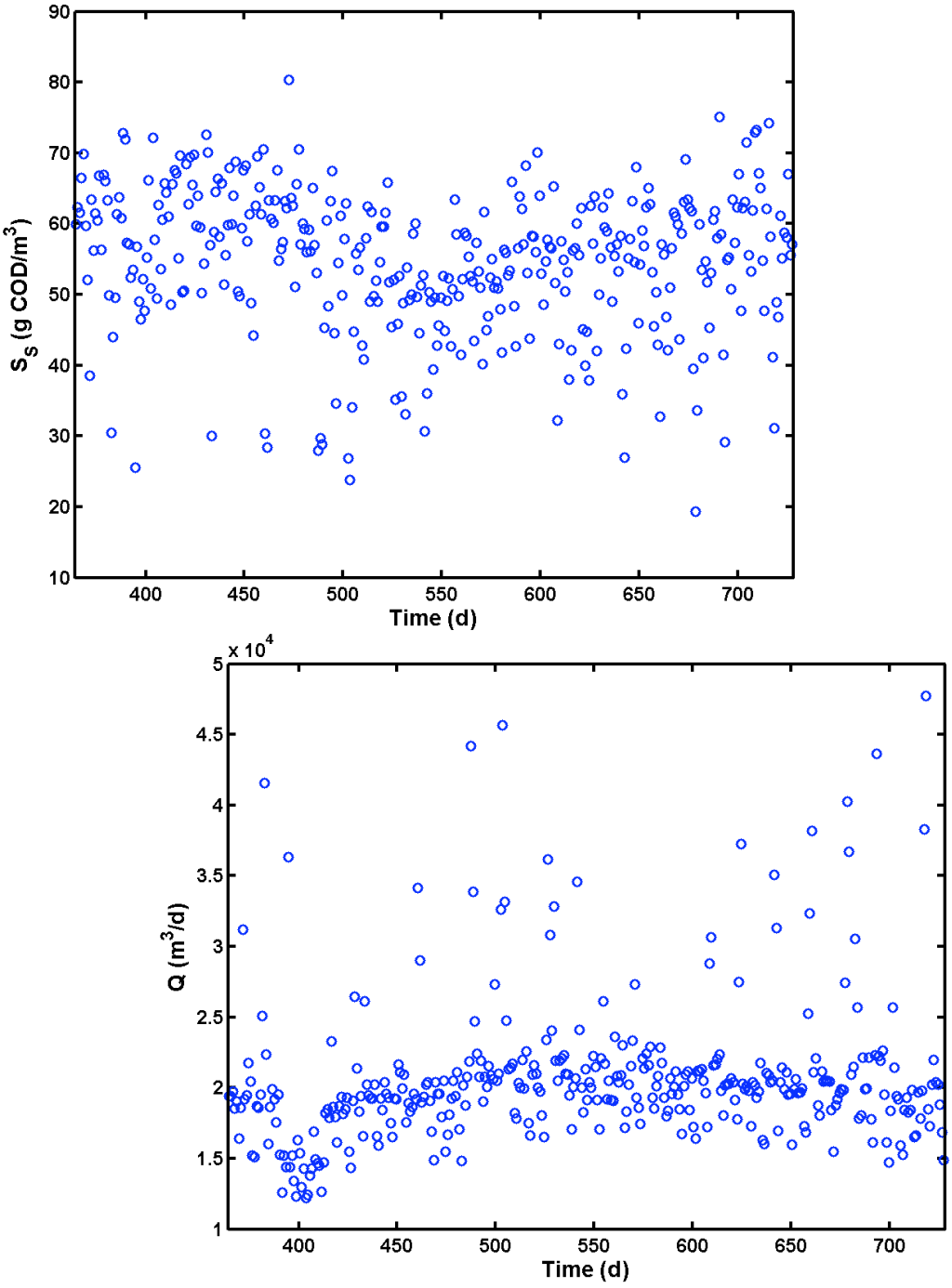

Figure 49. One year of average S_S concentration (top) and flow rate data generated with the influent model (each data point in this figure corresponds to the average of 96 samples, i.e. one day of dynamic data with 15 minutes sampling interval); Simulation 'noise 4'

A detail of the dynamic influent concentration data resulting from the model is provided in Figure 50. The corresponding flow rate profile is provided in Figure 51. The major rain event starting at around t = 487 d appears clearly in the flow rate data, and results in strong dilution effects for the influent pollutant concentrations. When looking in detail at the influent concentration profiles for particulate components such as for example X_S, a concentration peak is visible at around t = 487 d, at the start of the rain event. This concentration peak does not appear for the soluble components, and is the effect of including the 'first flush effect generation' model block.

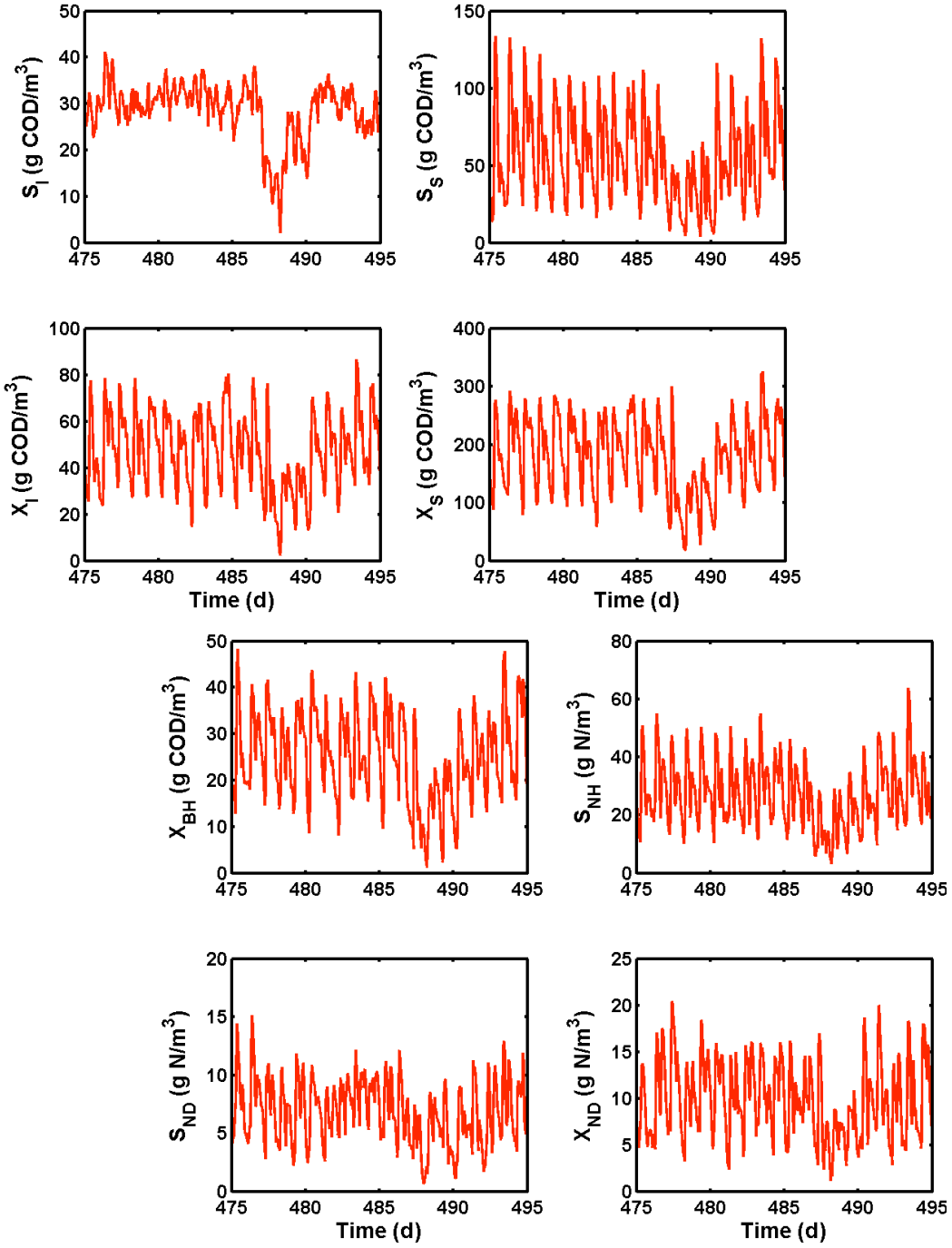

Figure 50. Detail of the concentration dynamics generated with the influent model; Simulation 'noise 4'

The influent model parameters used for generating the flow rate and pollutant concentration data provided in Figure 50 and Figure 51 are provided in Table 2 and Table 12. The cross-correlation between different ASM1 influent variables and between different global variables is illustrated in Figure 52 and Figure 53 respectively. Note that Figure 52 and Figure 53 were obtained by simulating the influent model without the rain generator, such that there is a fair basis to compare the cross-correlations for data generated with the influent model with the cross-correlations for BSM1 dry weather influent data.

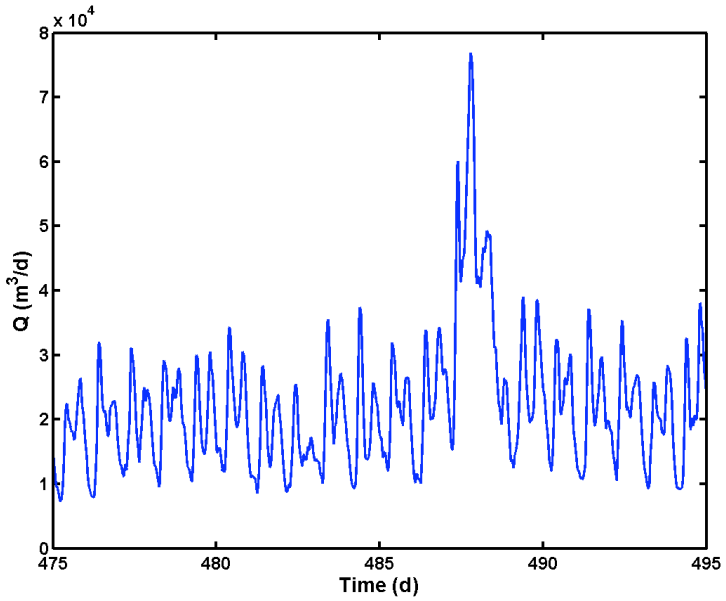

Figure 51. Detail of the flow rate dynamics generated with the influent model; Simulation 'noise 4'

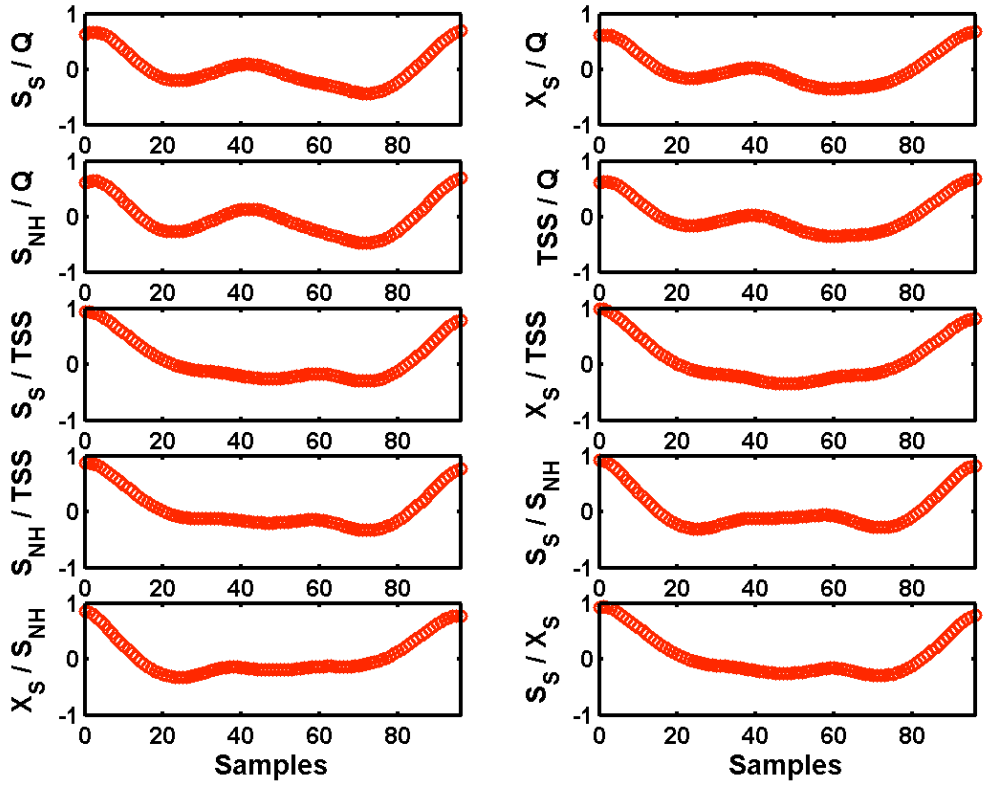

Figure 52. Cross-correlation between time series of ASM1 variables generated with the influent model for 1 year of dynamic data (see Figure 50 and Figure 51 for examples); Simulation 'noise 4'. The sampling interval is 15 minutes

Table 12. Main parameters for the influent pollutant concentration model

Model block	Parameter	Value	Units	Remarks
'Household pollutants'	CODsol_gperPEperd	19.31	g COD/(p.e. d)	
	CODpart_gperPEperd	57.54	g COD/(p.e. d)	
	SNH_gperPEperd	6.89	g N/(p.e. d)	
	TKN_gperPEperd	11.36	g N/(p.e. d)	
	f_1	2.0	-	Gain for tuning the noise levels on the household pollutant fluxes
'Industry pollutants'	CODsol_Ind_kgperd	386.2	kg COD/d	
	CODpart_Ind_kgperd	1 150.9	kg COD/d	
	SNH_Ind_kgperd	61.3	kg N/d	
	TKN_Ind_kgperd	101.0	kg N/d	
	f_2	2.0	-	Gain for tuning the noise levels on the industry pollutant fluxes
BSM1_fractionator	SI_cst	30	g COD/m^3	Constant influent S_I concentration
	XI_fr	0.182	-	X_I fraction of CODpart
	XS_fr	0.718	-	X_S fraction of CODpart
	XBH_fr	0.100	-	X_{BH} fraction of CODpart
	XBA_fr	0.0	-	X_{BA} fraction of CODpart
	XP_fr	0.0	-	X_P fraction of CODpart
	SND_fr	0.396	-	S_{ND} fraction of TKN – S_{NH} – N in biomass
	XND_fr	0.604	-	X_{ND} fraction of (TKN – S_{NH} – N in biomass)
'Noise generator'	f_3	2.0	-	Gain for tuning the noise levels on the pollutant concentrations
'First flush effect generation'	FFfraction	0.25	-	Fraction of TSS that can settle in the sewer system
	Q_{lim}	70 000	m^3/d	Limit flow rate triggering a first flush effect
	n	15	-	Exponent for Hill function
	M_{Max}	1 000	kg	Maximum sediment mass stored in sewer system
	FF	500	-	Gain for first flush effect

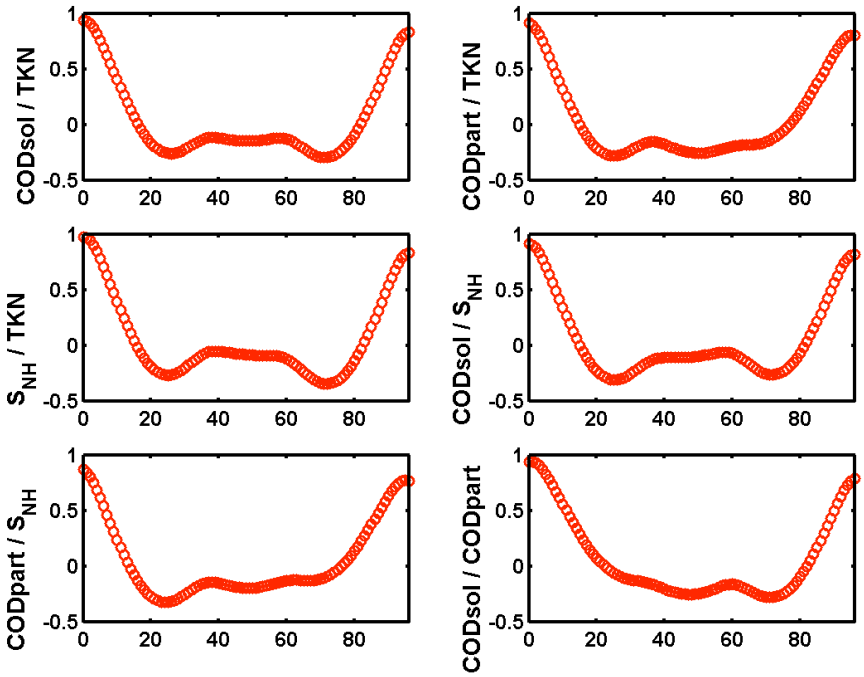

Figure 53. Cross-correlation between time series of global variables generated with the influent model for 1 year of dynamic data (see Figure 50 and Figure 51 for examples); Simulation 'noise 4'. The sampling interval is 15 minutes

4.10. Final tuning of input files

During the influent generation, noise is added to the inputs via random number generators. Several simulations were performed to decide on appropriate noise levels. While doing this simulation work, it appeared that one of the more annoying properties of the 'sewer' model block is that it will counteract the effect of any noise that is added to the inputs received. Thus, despite the rather high noise levels added to the inputs, the correlation between different influent flow rate variables is still quite high under dry weather conditions. As mentioned before, some correlations (e.g. between CODpart and TSS concentrations) cannot be avoided, in this case because all CODpart to TSS conversion factors are identical. Note that the BSM1 has a similar problem, and that it can be avoided, but only to a limited extent, by applying different COD to TSS conversion factors for all the particulate COD fractions. When comparing to the BSM1 dry weather influent files, the following problems can be mentioned when considering the global variables:

- The correlation coefficient between CODsol and CODpart influent concentrations is too high for the data generated with the influent model (0.9331 compared to 0.5874 for BSM1 dry weather influent). As a consequence, the correlation coefficient between CODsol and TSS is also too high. Of course, the correlation coefficient only provides information on the correlations between variables that were sampled at the same moment. The cross-correlation curves contain information on lagged variables too. When studying the cross-correlation curves in more detail, the BSM1 data (Figure A1.6) initially show an increase for the CODsol/CODpart cross-correlation curve. This peak is not there for the dynamic data generated with the influent model. The BSM1 CODsol/CODpart cross-correlation curve indicates that CODpart concentration peaks, and thus also influent TSS peaks, lag 5 to 6 samples behind on the CODsol, S_{NH} and TKN concentration peaks. Similar delays can be noticed in the S_S/TSS, S_S/X_S and S_{NH}/TSS cross-correlation curves (Figure A1.5)

- There is a rather high correlation between influent CODpart and S_{NH} concentrations for the data generated with the influent model (0.8665 compared to 0.4881 for BSM1 dry weather influent). The reasoning is the same as for the previous point. Similarly, there is a too high correlation between influent TSS and S_{NH} concentrations.
- There is a rather high correlation between influent CODpart and TKN concentrations for the data generated with the influent model (0.9112 compared to 0.7303 for BSM1 dry weather influent). The reasoning is the same as for the first point. Similarly, there is a too high correlation between influent TSS and TKN concentrations.
- There is a too high correlation between influent S_{NH} and TKN concentrations for the data generated with the influent model (0.9749 compared to 0.9514 for BSM1 dry weather influent). Note that a rather high correlation between S_{NH} and TKN concentrations is unavoidable, since a major part of the TKN actually corresponds to the S_{NH}.

The use of non-ideal sensors on the influent generated with the influent model will further decrease the correlations between influent time series. However, the same applies for the BSM1 influent data. There is a clear indication that the CODpart/TSS peaks in the BSM1 dry weather influent lag behind some samples on the CODsol/S_{NH} influent concentration peaks. This time lag is at this point not present in the influent dynamics resulting from the influent model. It was therefore decided as part of influent design step 3 to modify the household input file, shifting the households CODpart input profile one hour behind compared to households CODsol, S_{NH} and TKN input profiles. The new households pollutant flux input files are provided in Figure 54.

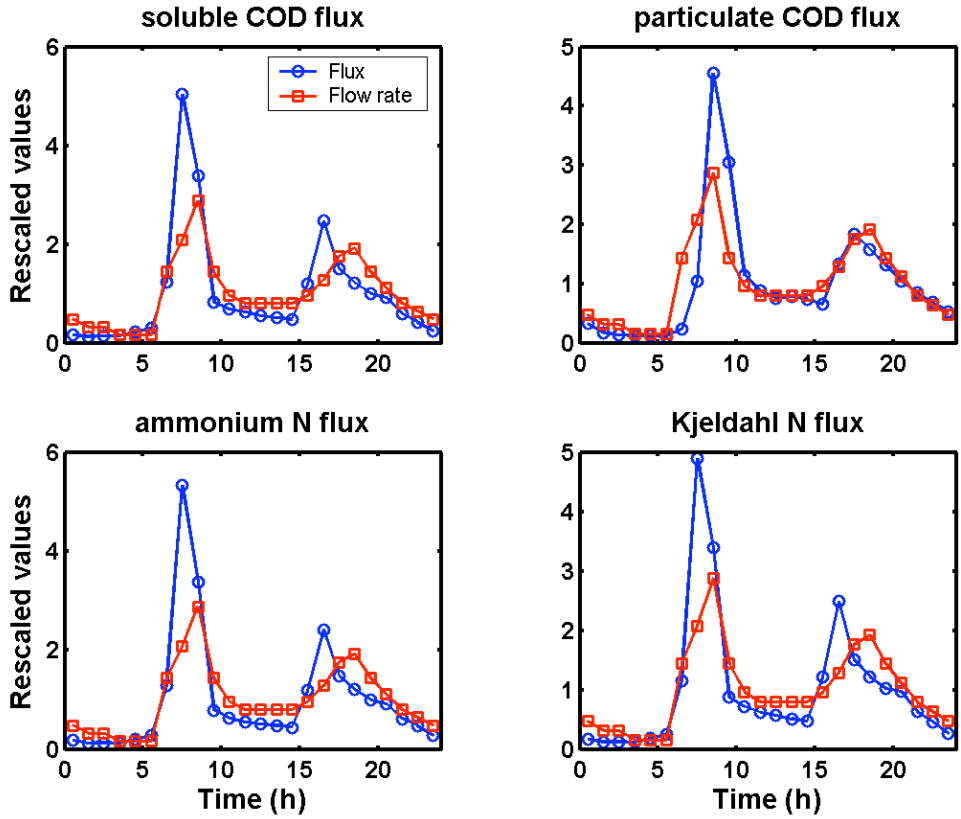

Figure 54. New input files for household pollutant fluxes generation resulting from influent pollutant load dynamics design step 3 (CODsol_day_HS, CODpart_day_HS, SNH_day_HS, TKN_day_HS)

Simulation results (dynamic input profiles) with the new input files are not shown here, since the influent dynamics are not significantly influenced by the change in the households CODpart input profile. The cross-correlation between different ASM1 variables and different global variables calculated from the influent model outputs is illustrated in Figure 55 and Figure 56. The effect of the modification to the input files is that some of the cross-correlation plots, for example the one for S_S/TSS, are shifted a number of samples when comparing with the plots in Figure 52 and Figure 53. As a consequence, several correlation coefficients have decreased considerably to levels that are close to or below the correlation coefficients for the BSM1 dry weather influent.

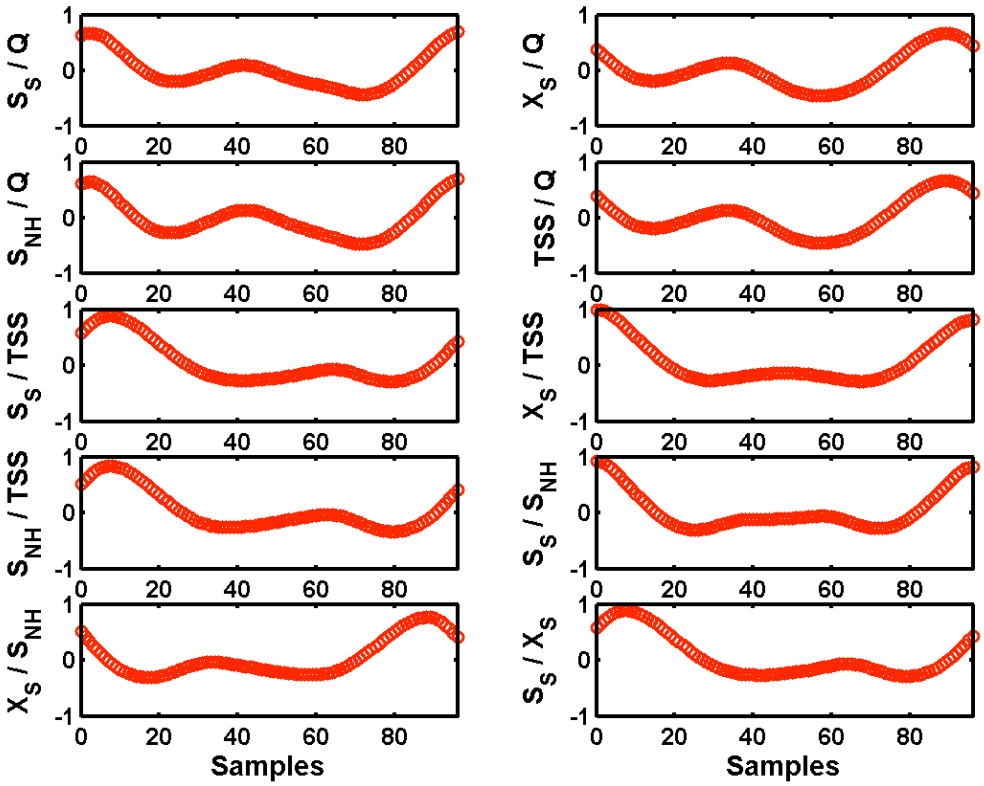

Figure 55. Influent design step 3: cross-correlation between time series of ASM1 variables generated with the influent model for 1 year of dynamic data; Simulation 'noise 4'. The sampling interval is 15 minutes

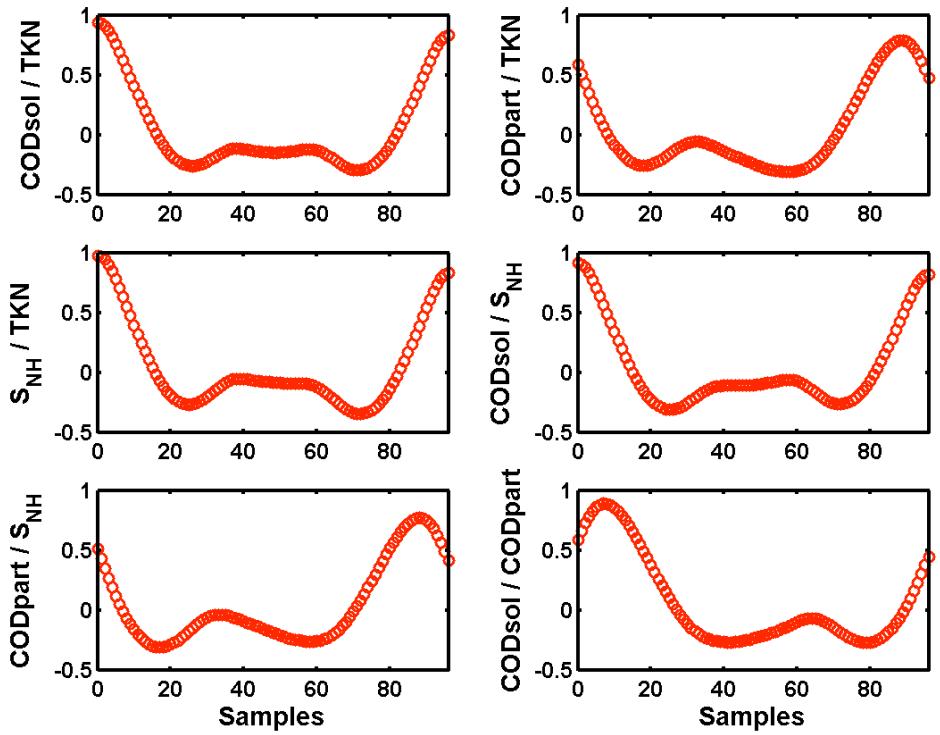

Figure 56. Influent design step 3: cross-correlation between time series of global variables generated with the influent model for 1 year of dynamic data; Simulation 'noise 4'. The sampling interval is 15 minutes

5. Model for temperature variations

Temperature was included as an additional state in the influent model. The temperature profile includes a seasonal effect, i.e. there is a warm and a cold season. The temperature profile is modelled with a sine function (see Figure 57), with an average temperature TBias of 15 deg C., an amplitude TAmp of 5 deg. C, and a frequency TFreq of (2·π/364) rad/day. A phase shift TPhase of (8.5·π/24) rad was applied, such that the maximum flow rate due to infiltration (rainy season, see 'Groundwater' model block) approximately corresponds to the lowest temperature, and vice versa. No noise is added. It is again up to the user to modify the temperature profile. The average temperature, the amplitude of the sine function and the phase shift can be modified.

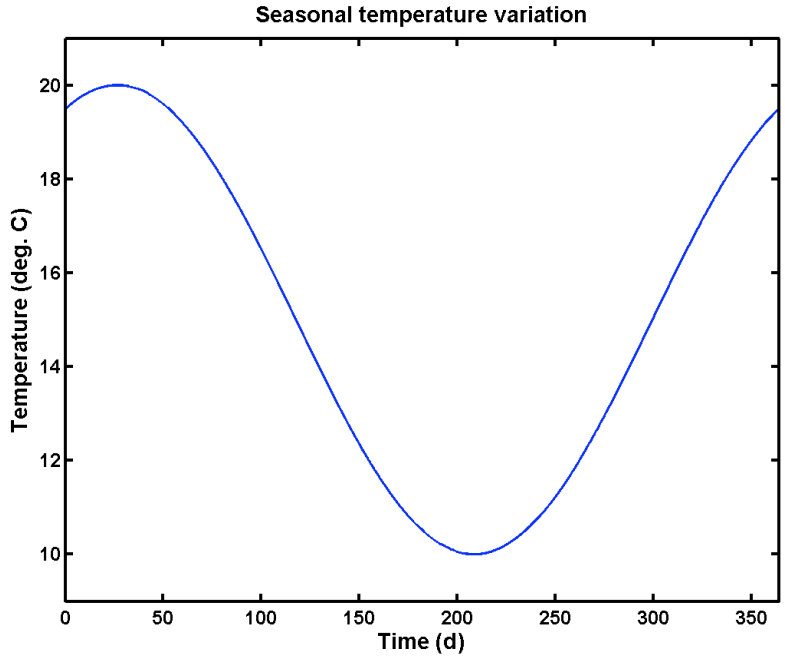

Figure 57. An example of an influent temperature profile resulting from the influent model. Note that the first day corresponds to July 1st

In addition to the seasonal variations, and following the results of preliminary simulation tests with the BSM1_LT by Darko Vrecko, it was decided to add on a daily temperature effect to the temperature model, assuming that the temperature in the wastewater treatment plant varies according to a sinusoidal wave with an amplitude TdAmp of 0.5 deg. C. The daily temperature variation obtained this way is illustrated in Figure 58. The parameters for the daily T effect were tuned such that the minimum temperature for each day occurs around 8:00 h in the morning. Note here that the temperature generated by the influent model is assumed to be the temperature in the treatment plant. The seasonal and daily temperature variations are summed up in the influent model. The parameters for the temperature model are summarised in Table 13.

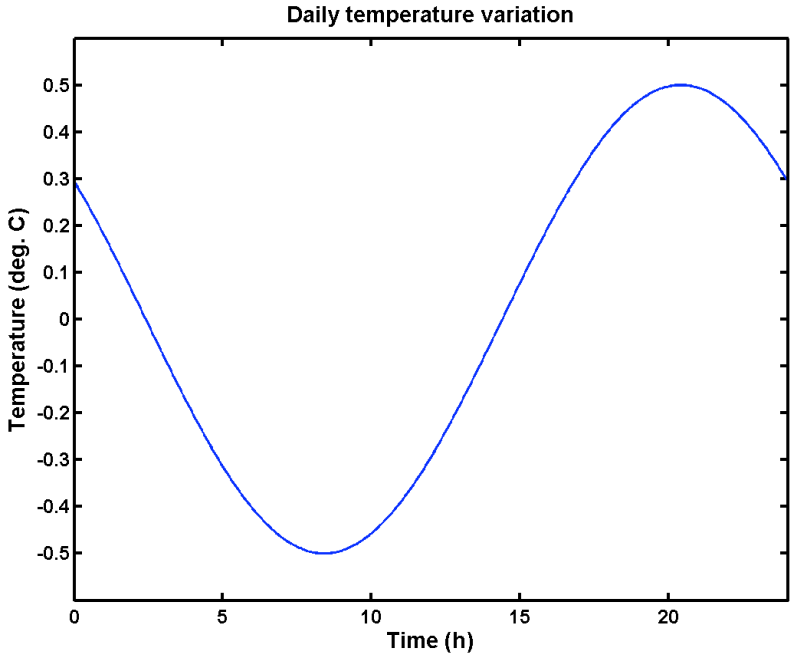

Figure 58. A daily influent temperature profile resulting from the influent model

Table 13. Parameters of the temperature model

Parameter	Value	Units	Remarks
TAmp	5	deg C	Seasonal temperature variation, amplitude
TBias	15	deg C	Seasonal temperature variation, average
TFreq	2·π/364	rad/d	Seasonal temperature variation, frequency
TPhase	8.5·π/24	rad	Seasonal temperature variation, phase shift
TdAmp	0.5	deg C	Daily temperature variation, amplitude
TdBias	0.0	deg C	Daily temperature variation, average
TdFreq	2·π	rad/d	Daily temperature variation, frequency
TdPhase	0.8·π	rad	Daily temperature variation, phase shift

6. A final BSM2 influent design proposal

All simulation work described previously was based on the BSM1 influent loads, corresponding to a pre-settled influent. However, the idea of the influent modelling work is to provide influent flow rate and concentration data for the BSM2, corresponding to raw influent that is not pre-settled. Therefore, the influent loads generated with the influent model need to be increased. The increase of the influent loads will be based on the guidelines agreed on with the group of active benchmarkers, where the effect of the primary settler in BSM2 was proposed to correspond to 50% removal of TSS, with a 3% TSS concentration in the underflow. The primary clarifier model parameters will be tuned to obtain the 50% TSS removal and the 3% TSS concentration in the underflow. Tuning of the primary clarifier parameters has already been done, and will be discussed in another technical report.

6.1. Parameter changes for generating the BSM2 influent

A few parameter changes were made to increase the influent particulate components pollutant load for the BSM2 (Table 14), since all the influent model development work reported until now has been done assuming BSM1 influent loads. Taking into account that the primary clarifier will remove about 50 % of the particulate material in the influent, and considering that the load in the effluent of the primary clarifier should correspond to the BSM1 loads, the CODpart fluxes were doubled for households and industry. Increasing the TKN load was a bit more complicated. The original TKN load was increased such that the X_{ND} concentration in the output of the influent model was doubled compared to the simulations with BSM1 loads. In addition to that, it was also considered that the extra influent biomass, which is a fraction of the CODpart that was already added when increasing the COD loads, would also contribute to an increase of the influent TKN load with a factor i_{XB} times the biomass concentration. The calculation for the new TKN_gperPEperd parameter value is provided in Equation 12.

$$(TKN\_gperPEperd)_{BSM2} = (TKN\_gperPEperd)_{BSM1} + [(TKN\_gperPEperd)_{BSM1} - SNH\_gperPEperd - i_{XB} \cdot XBH\_fr \cdot (CODpart\_gperPEperd)_{BSM1}] \cdot (XND\_fr)_{BSM1} + i_{XB} \cdot XBH\_fr \cdot (CODpart\_gperPEperd)_{BSM1} \quad (12)$$

In that equation the subscript BSM1 refers to BSM1 input load parameters, whereas the subscript BSM2 refers to BSM2 input load parameters (see also Table 14). Other parameters are explained in Table 12. A similar formula is applied for the calculation of the TKN_Ind_kgperd. Note that the parameter XBH_fr does not get a subscript, since the value of this parameter is identical in the BSM1 and the BSM2 load scenario.

As a consequence of the increase in the influent XND concentration for the BSM2 influent, compared to an influent with BSM1 loads, the parameters determining the fractionation of the (TKN – S_{NH} – N in biomass) part of the TKN need to be modified. The modified parameters can simply be calculated according to Equation 13 and 14.

$$(XND\_fr)_{BSM2} = \frac{2 \cdot (XND\_fr)_{BSM1}}{1 + (XND\_fr)_{BSM1}} \quad (13)$$

$$(SND\_fr)_{BSM2} = 1 - (XND\_fr)_{BSM2} \quad (14)$$

In Equation 13 and 14, the subscripts BSM1 and BSM2 refer to BSM1 and BSM2 fractionator parameters, respectively.

Finally, the noise levels for particulate components X_I, X_S, X_{BH} and X_{ND} was doubled in 'Noise generator' model block, as a consequence of the fact that average concentrations of these pollutants were also doubled when applying the BSM2 loads, compared to the situation with the BSM1 influent loads.

Table 14. Influent model parameter changes to influent model to convert the particulate COD loads from BSM1 to BSM2

Parameter	BSM1 value	BSM2 value	Units
CODsol_gperPEperd	19.31	19.31	g COD/(p.e. d)
CODpart_gperPEperd	57.54	115.08	g COD/(p.e. d)
SNH_gperPEperd	6.89	6.89	g N/(p.e. d)
TKN_gperPEperd	11.36	14.24	g N/(p.e. d)
CODsol_Ind_kgperd	386.2	386.2	kg COD/d
CODpart_Ind_kgperd	1 150.9	2 301.8	kg COD/d
SNH_Ind_kgperd	61.3	61.3	kg N/d
TKN_Ind_kgperd	101.0	128.62	kg N/d
SND_fr	0.396	0.247	-
XND_fr	0.604	0.753	-

6.2. Final influent model proposal: simulation

The parameters in Table 2 and Table 12, combined with the parameter updates summarised in Table 14 were used to generate the BSM2 influent file. The files needed to generate that BSM2 influent file are summarised in Appendix 1. The BSM2 influent file consists of 9 + 26 + 52 weeks of dynamic data. Since initialisation will take place at 15 deg. C, the dynamic influent time series starts also with a temperature that is about 15 deg. C. During the first 9 weeks of data, temperature decreases steadily. These weeks are intended to be used for simulating the plant to a dynamic 'pseudo steady state', i.e. model outputs during that period should not be used. The following 26 weeks, starting at January 1st, are the training data. During that period, control algorithms can be tuned, and monitoring algorithms can be trained. The final 52 weeks of dynamic data, starting on July 1st, corresponds to the performance validation period, and is used to compare the effect of control or monitoring strategies on the plant performance.

No simulation results are provided here. Instead, it is much better to run the model code included with this document. For the last year of data generated with the influent model, the average flow-weighted influent concentrations should correspond to the values provided in Table 15.

Table 15. Average flow-weighted influent concentrations calculated for 1 year of influent time series data generated with BSM2influentmodel.mdl, corresponding to the BSM2 control strategy evaluation period

	Concentration	Unit		Concentration	Unit
$S_{I,inf}$	27.21	g COD/m^3	$S_{NO,inf}$	0.00	g N/m^3
$S_{S,inf}$	58.15	g COD/m^3	$S_{NH,inf}$	27.91	g N/m^3
$X_{I,inf}$	92.46	g COD/m^3	$S_{ND,inf}$	6.66	g N/m^3
$X_{S,inf}$	363.77	g COD/m^3	$X_{ND,inf}$	19.35	g N/m^3
$X_{BH,inf}$	50.66	g COD/m^3	$S_{ALK,inf}$	7.00	g HCO_3^-/m^3
$X_{BA,inf}$	0.00	g COD/m^3	TSS_{inf}	380.17	g/m^3
$X_{P,inf}$	0.00	g COD/m^3	Q_{inf}	20 668.44	m^3/d
$S_{O,inf}$	0.00	g - COD/m^3	T_{inf}	15.0	deg. C

7. Acknowledgements

The authors wish to thank the participants in the influent model discussions during a COST Action 624 Working Group No. 1 meeting in Aix-en-Provence, France in May 2004, including (apart from the authors): Jens Alex, Eduardo Ayesa, Bengt Carlsson, Jeremy Dudley, Celal Gokcay, Nadja Hvala, Güclü Insel, Boris Kompare, Erik Lindblom, Marie-Noelle Pons, Antonio Salterain, Paul Schosseler, Henri Spanjers, Peter Vanrolleghem, Darko Vrecko, and Bernard Wett.

8. Symbols

A_1	Surface area of the tank in the unisoilmodel.c S-function (m^2)
A_2	Surface area of the tank in the sewer variable volume tank model S-function (m^2)
aH	A parameter determining the direct contribution of rainfall falling on impervious surfaces in the catchment area to the flow rate in the sewer (%).'Rain generator' model block
ASM1	Activated Sludge Model No. 1
BSM1	Benchmark Simulation Model No. 1
BSM1_LT	Long-Term Benchmark Simulation Model no. 1
BSM2	Benchmark Simulation Model No. 2
C	Gain relating Qout to the liquid level in the variable volume tank in the sewer model block ($m^{1.5}$/d)
COD	Chemical oxygen demand
CODpart	Particulate COD
CODsol	Soluble COD
CODtot	Total COD
f_1	Tuning factor for noise levels
f_2	Tuning factor for noise levels
f_3	Tuning factor for noise levels
f_{corr}	Primary clarifier model, correction factor removal efficiency
h_1	Actual height of the water level for the tank in the soilmodel S-function (m)
h_2	Actual height of the water level for the sewer variable volume tank model S-function (m)
H_{inf}	Height of the water level above the invert level in the unisoilmodel.c S-function (m)
H_{INV}	Height of the invert level in the unisoilmodel.c S-function (m)
H_{MAX}	Maximum height of the water level for the tank in the soilmodel S-function (h)
H_{min}	The minimum liquid level that needs to be exceeded before there is a flow rate out of the variable volume tank model (sewer model block)
InfBias	Mean value of the sine wave signal for generating seasonal effects due to infiltration ('Seasonal correction infiltration' model block)
InfAmp	Amplitude of the sine wave for generating seasonal effects due to infiltration ('Seasonal correction infiltration' model block)
InfFreq	Frequency of the sine wave for generating seasonal effects due to infiltration ('Seasonal correction infiltration' model block)
InfPhase	Phase shift of the sine wave for generating seasonal effects due to infiltration ('Seasonal correction infiltration' model block)
IWA	International Water Association
i_{XB}	Fraction of N in biomass (g N/g COD)
K	Soil permeability constant in the unisoilmodel.c S-function
K_{down}	Gain for adjusting the flow rate to downstream aquifers in the unisoilmodel.c S-function
K_{inf}	Infiltration gain in the unisoilmodel.c S-function
LLrain	A constant, used in the 'rain generator' model block
PE	Person equivalent (parameter in the influent model)
p.e.	Person equivalent (as a unit)

Q_{av}	Average flow rate (m^3/d)
QInd_weekday	Average wastewater flow rate from industry on normal week days (Monday to Thursday) (m^3/d)
Qin	Flow rate in to a variable volume tank in the sewer model block (m^3/d)
Q_{intot}	The sum of the infiltration water flow rate and the flow rate originating from the fraction of rainfall that is not collected on impervious surfaces
Qin1	Contribution of rain water to the inflow of the tank model in the soil model block
Qin2	Contribution of 'Seasonal correction infiltration' model block to the inflow of the tank model in the soil model block
Qout	Flow rate out of a variable volume tank in the 'sewer' model block (m^3/d)
Qpermm	Flow rate per mm rain (m^3/mm) in the 'rain generator' model block
QperPE	Wastewater flow rate per p.e. for municipal wastewater (l/d)
S_{ALK}	Alkalinity
S_I	Inert soluble COD
S_{NH}	Ammonium nitrogen
S_{NO}	Nitrite + nitrate nitrogen
S_O	Dissolved oxygen
S_S	Readily biodegradable COD
subareas	Parameter of the sewer model block. A measure of the size of the catchment area. It will determine the number of variable volume tanks in series that will be used for describing the sewer system.
TKN	Total Kjeldahl nitrogen
TSS	Total suspended solids concentration
V_{pc}	Volume of the primary clarifier
WG1	Working Group 1
WWTP	Wastewater treatment plant
X_{BA}	Autotrophic biomass
X_{BH}	Heterotrophic biomass
X_I	Inert particulate COD
X_{ND}	Particulate organic nitrogen
X_P	Inert particulate COD resulting from biomass decay
X_S	Slowly biodegradable particulate COD
X_{TSS}	Total suspended solids concentration

9. References

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

Gernaey K.V., Rosen C. and Jeppsson U. (2004). Benchmark long term: A model for influent data generation. Technical report, 19 May 2004.

Gernaey K.V., Rosen C., Benedetti L. and Jeppsson U. (2005a). Phenomenological modelling of wastewater treatment plant influent disturbance scenarios. Paper submitted to the 10th International Conference on Urban Drainage, Copenhagen/Denmark, 21-26 August 2005.

Gernaey K.V., Rosen C. and Jeppsson U. (2005b). WWTP dynamic disturbance modelling – an essential module for long-term benchmarking development. Proceedings 2nd IWA Conference on Instrumentation, Control and Automation for Water and Wastewater Treatment and Transport Systems (ICA2005), May 29-June 2, Busan, Korea.

Jeppsson U. and Pons M.-N. (2004) Editorial: The COST benchmark simulation model – current state and future perspective. Control Engineering Practice, **12**, 299 – 304.

Jeppsson U., Rosen C., Alex J., Copp J., Gernaey K.V., Pons M.-N. and Vanrolleghem P.A. (2004) Towards a benchmark simulation model for plant-wide control strategy performance evaluation of WWTPs. In: *Proc. 6th Int. Symposium on Systems Analysis and Integrated Assessment in Water Management (WATERMATEX 2004)*, Nov. 3 – 5, Beijing (China).

Rosen C., Jeppsson U. and Vanrolleghem P.A. (2004) A common benchmark for long-time control evaluation and monitoring methodology testing. *Proc. 2nd Int. IWA Conference on Automation in Water Quality Monitoring (AutMoNet 2004)*, April 19 – 20, Vienna (Austria), 205 - 212.

Appendix 1. List of influent model files

The table below contains an overview of the Matlab files that come with the influent model. File names are ordered alphabetically. All files were created using Matlab 6.5 on a Windows XP computer.

When running the influent model, the S-function c files first need to be compiled (using the 'mex' command, for example 'mex asm1inf_combiner.c'). The model parameters are initialised by running the BSM2influent_init.m script. Finally, one can run a simulation with BSM2influentmodel.m. A simulation with the influent model should take a few minutes, on condition that Real-time workshop is installed. Without Real-time workshop, simulation time might be considerably higher (factor 10 or more). Dynamic BSM2 influent data are then generated by executing the script Generate_BSM2_influent.m.

Filename	File description
asm1inf_combiner.c	S-function in c; model block combining the contributions of two BSM1_LT/BSM2 compatible flows (ideal mixing tank, no volume). The function includes handling of 5 dummy states. The two flows consist of 13 ASM1 states + TSS + flow rate + temperature + 5 dummy states.
BSM1LT_fractionator2.c	S-function in c; influent fractionator to convert pollutant fluxes to ASM1 compatible influent concentrations
BSM2influent_init.m	Initialisation file for BSM2 influent generation
BSM2influentmodel.mdl	Simulink file, containing the influent model
CODpart_day_HS.mat	Influent model input file; diurnal particulate COD flux variation at source (Households)
CODpart_week_IndS.mat	Influent model input file; weekly particulate COD flux variation at source (Industry)
CODsol_day_HS.mat	Influent model input file; diurnal soluble COD flux variation at source (Households)
CODsol_week_IndS.mat	Influent model input file; weekly soluble COD flux variation at source (Industry)
day_HS.mat	Influent model input file; diurnal flow rate variation at source (Households)
Generate_BSM2_influent.m	A script to generate a MATLAB workspace and a text file containing the BSM2 influent. Execute the script after finishing a simulation with BSM2influentmodel.mdl
sewer_asm1dum.c	S-function in c; deterministic variable volume tank model used to model one tank in the 'Sewer' model block (22 outputs: 13 ASM1 states + TSS + flow rate + temperature + 5 dummy states + tank level)
SNH_day_HS.mat	Influent model input file; diurnal S_{NH} flux variation at source (Households)
SNH_week_IndS.mat	Influent model input file; weekly S_{NH} flux variation at source (Industry)
ssmodel3.c	S-function in c; model describing storage of TSS and particulate ASM1 model components in the sewer system. Model used in the 'First flush effect generator' model block.

TKN_day_HS.mat	Influent model input file; diurnal TKN flux variation at source (Households)
TKN_week_IndS.mat	Influent model input file; weekly TKN flux variation at source (Industry)
unisoilmodel.c	S-function in c; variable volume tank model. Model used in the 'Soil' model block
week_HS.mat	Influent model input file; weekend effect on flow rate (Households)
week_IndS.mat	Influent model input file; weekly flow rate variation at source (Industry)
week_polHS.mat	Influent model input file; weekend effect on pollutant fluxes (Households)
year_HS.mat	Influent model input file; holiday effect on flow rate (Households)
year_IndS.mat	Influent model input file; holiday effect on flow rate (Industry)