Integrated dynamic urban scale sources-and-flux model for PPs

Deliverable No: D7.4, Date: 13 November 2009
Dissemination level: PU

Lorenzo Benedetti\textsuperscript{1}, Webbey De Keyser\textsuperscript{1}, Luca Vezzaro\textsuperscript{2}, Natasha Atanasova\textsuperscript{3}, Veerle Gevaert\textsuperscript{1}, Frederik Verdonck\textsuperscript{1}, Peter A. Vanrolleghem\textsuperscript{4} and Peter Steen Mikkelsen\textsuperscript{2}

\textsuperscript{1} Department of Applied Mathematics, Biometrics and Process Control, Ghent University
\textsuperscript{2} Department of Environmental Engineering, Technical University of Denmark
\textsuperscript{3} Institute of Sanitary Engineering, Faculty of Civil and Geodetic Engineering, University of Ljubljana
\textsuperscript{4} Département de génie civil, model\textsuperscript{EAU}

Source Control Options for Reducing Emissions of Priority Pollutants (ScorePP)
Sixth Framework Programme, Sub-Priority 1.1.6.3, Global Change and Ecosystems
Project no. 037036, www.scorepp.eu, Duration: 1 October 2006 – 30 March 2010
Abstract
Integrated urban water system (IUWS) modelling aims at assessing the quality of the surface water receiving the urban emissions through sewage treatment plants, combined sewer overflows (CSOs) and stormwater drainage systems. The unit process models developed in Task 7.2 were integrated in Task 7.4 in an urban water model. This allows describing the entire exposure pathway/route of PPs. The unit process models were implemented in the WEST® modelling and simulation platform. The unit process models are coupled using appropriate translations of the state variables. A case was used for testing. The tool developed in Task 7.1 was used to generate inputs to the integrated model. The multimedia fate and transport model (MFTM) developed in Task 7.3 was “wrapped around” the dynamic IUWS model for organic micropollutants to enable integrated environmental assessment. The combined model was tested on a hypothetical catchment using two scenarios: a reference scenario and a stormwater infiltration pond scenario, as an example of a sustainable urban drainage system (SUDS). A case for Bis(2-ethylhexyl) phthalate (DEHP) was simulated and the SUDS resulted in a modification of the fluxes to environmental compartments and the reduction of water concentration. The relationship between model inputs and outputs was explored in view of model complexity reduction in WP6.

Acknowledgement
The presented results have been obtained within the framework of the project ScorePP - “Source Control Options for Reducing Emissions of Priority Pollutants”, contract no. 037036, a project coordinated by Department of Environmental Engineer, Technical University of Denmark within the Energy, Environment and Sustainable Development section of the European Community’s Sixth Framework Programme for Research, Technological Development and Demonstration.

Cover picture by Lorenzo Benedetti: Rainwater sedimentation channel before discharge in storage tank, in the ancient fortress of Mandu, Madhya Pradesh, India.
1 Introduction

This deliverable presents the integration of models for the prediction of Priority Pollutants (PP) fate in the urban wastewater system. The final purpose of the integrated model is the estimation of the PP fluxes in the urban system and the assessment of the effects of different reduction options (Figure 1.1).

Integrated modelling of the urban water system has matured substantially over the last decade. In the past, sewers, urban drainage, wastewater treatment and surface waters were modelled separately. However, since all of these sub-systems are interconnected, and because the final aim of water management is a good ecological status of water bodies, these disciplines have evolved into integrated urban water system (IUWS) modelling (e.g. Schmitt and Huber, 2006).

As the presence of some micropollutants in surface water bodies can be of particular concern, the European Commission identified a list of “priority substances” (CEC, 2008). Monitoring programmes can be established for surveillance, operational or investigative purposes. However, monitoring is costly, often difficult to perform and, due to the limited sampling
frequency, restricted in the amount of collected data necessary to sufficiently reflect temporal variability. A possible solution is to use mathematical tools to complement water quality monitoring. Therefore, to model the origin, transport and transfer processes of micropollutants within the urban water system in full detail, it is necessary to expand existing urban water quality models with state variables representing these micropollutants (e.g. Lindblom et al., 2006). However, micropollutants have the tendency to occur in more than one environmental medium (i.e. air, water, sediment, soil, etc.). Harremoës (2002) therefore addressed the point of interrelating the environmental media water, air and soil in “integrated environmental assessment” as a scientific discipline going beyond integrated water system modelling.

The behaviour of pollutants in different interconnected environmental media is studied in multimedia fate and transport models (MFTMs) (e.g. Mackay, 2001). Compared to fully dynamic water quality models, the early “unit world” MFTMs show a limited complexity, lack spatial resolution and often assume steady-state or equilibrium distribution between the environmental media. Over the years, these MFTM evolved into more realistic and more dynamic models: connections with geographical information systems were established, time-variant parameter estimation was included and different sub-models were coupled together in order to represent geographical heterogeneity in the modelled system (e.g. Verdonck, 2003; Luo et al., 2007). Today MFTMs are widely accepted for evaluating the overall fate and transport of organic chemicals.

This deliverable introduces the implementation of the IUWS model for micropollutants and the integration with the MFTM to obtain fully integrated environmental assessment frameworks. This is illustrated by a hypothetical case study and should allow assessing the impact of applying source control strategies within the urban scale on surrounding environmental compartments. In addition, the multimedia model should provide the boundary conditions for the urban scale model, such as pollutant (anthropogenic) background concentrations in an upstream river part.

The complexity of the integrated model can be reduced for educational purposes or to improve the user friendliness of the model and its application for decision support, and to reduce the required computational resources. A model simplification based on the analysis of input/output relationships is also presented.
2 The Integrated Urban Water System

The elements and the boundaries of the urban system that are included in the IUWS model are sketched in Figure 2.1. The fate models used to represent the IUWS components can be subdivided into two families, according to the influence of human activities:

- "Technosphere" fate models: include anthropic systems, which are built and managed with specific objectives by the humans living in the urban area. For example, a sewer network is built and managed to route stormwater and wastewater out from the urban area.

- "Environmental" fate models: are used to simulate systems where natural processes are the driving force. Human activities can affect the behaviour of these systems, but cannot have a complete control of all the processes and reactions. An example is the river model.

Several components can be identified within the borders of the IUWS model

- Sources: water and pollutants flows are generated by point and diffuse sources across the catchment. This element is not included within the IUWS borders, but it is a *sine qua non* of the IUWS model. In fact, pollutant and water fluxes are estimated by the application developed in Task 7.1 and the resulting emission (or release) time series for the different sub-catchments are used as input to the IUWS model.

- Sewer network: water flows from combined and separate systems are collected and routed across the catchment. Also, water can be detained by storage units (e.g. detention basins, etc.) or discharged directly into receiving waters by overflows structures

- Stormwater treatment options: flows from stormwater separated systems are treated before the discharge to the receiving waters.

- Wastewater treatment options: several treatments and combinations thereof are used to treat wastewater (e.g. domestic, from combined systems, industrial, etc).

- Sludge treatment options: the residues from wastewater treatment are treated before the final disposal in the environment.

- Receiving water: natural waters (e.g. rivers, lakes) are the final recipient of the water flows from urban areas. The objective of the entire PP reduction strategy is the fulfilment of the quality requirements for this component of the IUWS model.
Figure 2.1 - Schematic representation of the IUWS elements and boundaries

For each component, state-of-the-art models have been used as starting point for the implementation of PP fate models. These models, usually dealing with “traditional” water quality (e.g. overall organic pollution, suspended solids, nutrients, etc.), have been extended with PP removal processes.

The urban system cannot be considered as a completely isolated system: interactions with the external environment are occurring and affecting some processes within the urban area. Therefore, the models developed in Task 7.2 can be connected to the multimedia model that is developed in Task 7.3. These connections include fluxes to atmosphere and groundwater.
3 Modelling of Priority Pollutants removal processes

The implementation of PP removal fate models require the identification of the various processes that should be coded in the models. The relevant PP removal processes have been identified by different tasks within WP5. This chapter introduces the information that is available in literature on the selected processes that have been included in the models. For a detailed description of the processes that are included in each category/unit model the reader is directed to Deliverable 7.2.

The selection of the equations used in the model implementation was based on simple principles:

- **process relevance:** the different PP removal processes are included in the unit models only if the process is relevant, i.e. it is likely to have a significant influence on the PP fate. Table 3.1 shows the processes that are included in the various model units. The choice of the relevant processes is the result of subjective choices, based on a qualitative assessment: an objective evaluation should confirm (or refute) with a quantitative estimation the structure of the matrix presented in Table 3.1;

- **common mathematical formulation:** common equations that can be applied in the different elements of the Integrated Urban Water System have been preferred, in order to provide a common structure across the integrated model. Different formulations and/or conceptual models have been utilized whenever the characteristics of the unit required peculiar features;

- **data availability:** the processes are modelled according to the parameters that can be retrieved from the ScorePP database (see Deliverable 3.1 - Lützøft et al., 2008). For example, (pseudo) first-order kinetics have been preferred whenever process half-lives are available in the database. As example, Table 3.2 shows the relevant parameters that are used in the mathematical formulation of the PP removal processes.
Table 3.1 – Relevant PP removal processes in the various units of the Integrated Urban Water System

<table>
<thead>
<tr>
<th>Processes</th>
<th>Unit model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sewer</td>
</tr>
<tr>
<td>Physical processes</td>
<td></td>
</tr>
<tr>
<td>Sedimentation</td>
<td>+</td>
</tr>
<tr>
<td>Resuspension</td>
<td>+</td>
</tr>
<tr>
<td>Volatilization</td>
<td>+</td>
</tr>
<tr>
<td>Sediment-water exchange</td>
<td>+</td>
</tr>
<tr>
<td>Physico-chemical</td>
<td></td>
</tr>
<tr>
<td>Sorption-desorption</td>
<td>+</td>
</tr>
<tr>
<td>Hydrolysis</td>
<td>+</td>
</tr>
<tr>
<td>Photolysis</td>
<td>+</td>
</tr>
<tr>
<td>Biological</td>
<td></td>
</tr>
<tr>
<td>Aerobic biodegradation</td>
<td>+</td>
</tr>
<tr>
<td>Anoxic biodegradation</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 3.2 – PP removal processes and relevant parameters

<table>
<thead>
<tr>
<th>Process</th>
<th>Parameter retrieved from PP database*</th>
<th>Other relevant parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical processes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedimentation and resuspension</td>
<td>-</td>
<td>Water depth, bottom shear stress, critical shear stress, erodibility constant</td>
</tr>
<tr>
<td>Volatilization</td>
<td>Henry’s law constant, molecular weight</td>
<td>Water depth, wind speed, water currents and temperature</td>
</tr>
<tr>
<td>Sediment-water exchange</td>
<td>Molecular weight</td>
<td>Sediment porosity</td>
</tr>
<tr>
<td>Physico-chemical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adsorption-desorption</td>
<td>Partition coefficient (k_d or k_OC)</td>
<td>TSS concentration, organic fraction</td>
</tr>
<tr>
<td>Hydrolysis</td>
<td>First-order degradation rate</td>
<td>pH, temperature</td>
</tr>
<tr>
<td>Photolysis</td>
<td>Half-life</td>
<td>Light intensity, water depth, water pollution</td>
</tr>
<tr>
<td>Biological</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aerobic biodegradation</td>
<td>Half-life</td>
<td>Oxygen concentration, temperature</td>
</tr>
<tr>
<td>Anoxic biodegradation</td>
<td>Half-life</td>
<td>Oxygen/nitrate concentration, temperature</td>
</tr>
</tbody>
</table>

* Lützhöft et al., 2008
4 MSL implementation

4.1 Connections between modelbases

To model the various components that are present in the Integrated Urban Water System (sketched in Figure 2.1) a new WEST modelbase has been developed by combining different modelbases already implemented in WEST. These separate modelbase represent the MSL implementation of widely applied models for prediction of water quality in the various parts of the urban water system. For example, the KOSIM hydrological catchment runoff and sewer transport model (as implemented in WEST by Solvi, 2007) is used as backbone for modelling the sewer network and the stormwater treatment units. The IWA standard activated sludge models (ASM - Henze et al., 2000), the Universal Stormwater Treatment Model (Wong et al., 2006), the Siegrist anaerobic sludge treatment model (Siegrist et al., 1993) and the river water quality model RWQM1 (Reichert et al., 2001), were chosen as basic water quality models

The resulting IUWS_MP modelbase employs different categories (i.e. components and reactions), developed from the original categories. The scheme of the various components, connections and interactions between the modelbase categories is showed in the Figure 4.1. Each different category is characterized by different number of components, reactions and definitions. Different names are also applied to similar components and reactions in order to avoid numerical problems during the model compiling (see Table 4.1 for an example).

The structure showed in Figure 4.1 requires the implementation of transformers units, which enable the integration between the different categories. These units connect the different components according to their definition. Three different component typologies can be identified (see Table 4.2 for some examples):

- **Absent**: these components are not included in/cannot be calculated from the category used by the upstream model. A fixed parameter (i.e. the component average concentration) is used to calculate the flux downstream the transformer

- **Equivalent**: components in the upstream category have an equivalent in the downstream category. In this case, no calculation is needed and the transformers just convert the component’s name.

- **Fraction**: the downstream category employs a group of components that are defined as fractions of a component used upstream. Each component’s flow is calculated by using fixed stoichiometric ratio (e.g. the classical example is the subdivision of the various COD fractions. 

---

1 Integrated Urban Water System modelbase for modelling of Micro-Pollutants
Table 4.1 - Examples of different names used for similar components and reactions.

<table>
<thead>
<tr>
<th>Components</th>
<th>KOSIM</th>
<th>ASM2_T_MP</th>
<th>RIVER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>H2O_sew</td>
<td>H2O</td>
<td>rH2O</td>
</tr>
<tr>
<td></td>
<td>S_MP_sew</td>
<td>S_MP</td>
<td>rS_MP</td>
</tr>
</tbody>
</table>

Reactions

- SewMPSorption
- SewMPVolatilization
- MPSorption
- MPVolatilization
- RivSorption
- RivVolatilization

Table 4.2 - Examples of component conversion from the KOSIM_MP to the ASM2_T_MP categories.

<table>
<thead>
<tr>
<th>KOSIM_MP component</th>
<th>Component typology</th>
<th>ASM2_T_MP component</th>
<th>Conversion equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>Absent</td>
<td>X_AUT</td>
<td>Fixed concentration</td>
</tr>
<tr>
<td>NH4_sew</td>
<td>Equivalence</td>
<td>S_NH</td>
<td>S_NH = NH4_sew</td>
</tr>
<tr>
<td>PO4_sew</td>
<td></td>
<td>S_PO</td>
<td>S_PO = PO4_sew</td>
</tr>
<tr>
<td>COD_sol</td>
<td>Fraction</td>
<td>S_F</td>
<td>( S_F = \varphi_{COD_{SS,F}} \cdot COD_{Sol} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S_A</td>
<td>( S_A = \varphi_{COD_{SS,A}} \cdot COD_{Sol} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S_I</td>
<td>( S_I = (1 - \varphi_{COD_{SS,F}} - \varphi_{COD_{SS,A}}) \cdot COD_{Sol} )</td>
</tr>
<tr>
<td>COD_part</td>
<td>Fraction</td>
<td>X_S</td>
<td>( X_S = \varphi_{COD_{R,X,S}} \cdot COD_{part} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X_H</td>
<td>( X_H = \varphi_{COD_{R,X,H}} \cdot COD_{part} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X_I</td>
<td>( X_I = COD_{part} - X_H - X_AUT - X_{PAO} - X_{PHA} - X_S )</td>
</tr>
</tbody>
</table>

Detailed information about the three transformers used to connect IUWS models (sewer to WWTP, WWTP to river and sewer to river) can be found in Appendix A.

The IUWS_MP modelbase should be connected with the multimedia model that is developed in Task 7.3. To enable the connection between the two models, new terminals have been added to all the units. All the units characterized by free water surfaces (i.e. with volatilization) are connected to the atmosphere compartment, while infiltration units have a terminal connected to the groundwater compartment.

An input terminal has been added to all the units to pass information about the system temperature, which is included in the input file to the model.
Figure 4.1 - Scheme of the connections between the different categories in the IUWS_MP modelbase.

4.2 Conversion models

The current version of the WEST ModelEditor does not enable the implementation of different matrices in the same modelbase. Indeed, the Petersen Matrix editor can be used to modify and write only the files employed in the WWTP models (e.g. ASM2_T_MP categories). The conversion models in the different modelbases are implemented by employing the similar file structure. Thus, any change made through the Petersen Matrix editor can be easily transferred to the other categories by simply copying and pasting the modified code into the corresponding files (with the caution of changing the name of reactions and components).
Table 4.3 - Examples of corresponding conversion models implemented in different categories.

<table>
<thead>
<tr>
<th>Models name</th>
<th>ASM2_T_MP*</th>
<th>KOSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td>File name</td>
<td>wwtp.VolumeASM2_T_MPConversionModel.parameters.msl</td>
<td>sewer.VolumeASM2_T_MPConversionModel.parameters.msl</td>
</tr>
</tbody>
</table>

* Modified by the Peterson Matrix Editor

4.3 Implementation of removal processes

All the removal processes are commonly described by the files connected to the Peterson Matrix editor, with the correspondence between modelbases that are described in the previous paragraph. However, as showed in Table 3.1, some removal processes are modelled only in certain units. To avoid the useless declaration of elements that would not be used, all the common processes are declared in main files. All the parameters, state variables and kinetics employed in the specific processes are declared by additional files, which extend the common conversion model.

All the units belonging to the same category share the same stoichiometric matrix, interface and sensors. As an example, the scheme of the conversion models that are used in the KOSIM_MP category is showed in the Table 4.4. The main conversion model (sewer) includes all the common processes. Volatilization is added to all the units characterized by free water surface by additional files (sewer.water or – only for circular pipes – sewer.pipe). Photodegradation, which is included only in the water compartment of the StormWater Treatment Unit, is declared by other specific files (storm).
### Table 4.4 - Example of extension of kinetic files for units in the KOSIM_MP category.

<table>
<thead>
<tr>
<th>Processes</th>
<th>Filename</th>
<th>KOSIM models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sediment</td>
</tr>
<tr>
<td>Adsorption-desorption</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Aerobic biodegradation</td>
<td><code>sewer.VolumeASM2_T_MPCversionModel.kinetics.msl</code></td>
<td>+</td>
</tr>
<tr>
<td>Anoxic biodegradation</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Hydrolysis</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Volatilization</td>
<td><code>sewer.water.VolumeASM2_T_MPCversionModel.kinetics.msl</code></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><code>sewer.pipe.VolumeASM2_T_MPCversionModel.kinetics.msl</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>storm.VolumeASM2_T_MPCversionModel.kinetics.msl</code></td>
<td></td>
</tr>
</tbody>
</table>

* Only for circular pipes
5 Example of application

5.1 The IUWS model

The IUWS model configuration was based on a setup described by Grum et al. (2000) and commonly used for educational purpose. It consists of a rural catchment, three urban sewer catchments connected to an intercepting combined sewer system, an activated sludge plant including primary settling, two aerated tanks and secondary settling. The treatment plant and the overflow structures at the three urban catchments discharge to a river modelled as a series of five completely mixed tanks, each of them in contact with river sediment. A scheme of the model setup is included in the overall IUWS/MFTM scheme shown in Figure 5.1.

5.2 The multimedia model

Most MFTMs rely on the fugacity concept (Mackay, 2001) to quantify the partitioning of a chemical between phases. A chemical’s fugacity $f$ is its tendency to escape from that phase to reach a concentration in equilibrium with the concentrations in other phases. By definition, a pollutant’s fugacity is proportional to its concentration $C$: $f = C/Z$, with $Z$ a constant for each pollutant in each environmental compartment. By assuming a certain volume $V$ for each environmental compartment, a relation between the fugacity and the mass $m$ of the chemical can be written: $m = ZfV$. For steady-state modelling (called level I and II), equilibrium distribution concentrations can be calculated from the fugacity, which is considered equal in all environmental media. Level III models assign different fugacities to each environmental medium, recognizing that the system can be in a non-equilibrium steady-state. By writing the fugacity equations as differential equations, dynamic conditions can be calculated in a level IV model.

The MFTM implemented in WEST is based on the ‘regional’ scale of the SimpleBox model (den Hollander et al., 2003). It is a level III / level IV Mackay-type model, but using the micropollutant’s mass or concentration in each environmental compartment as the main state variable in the differential equations, similar to the mass balances commonly used in IUWS sub-models. The main reasons not to use the fugacity are that (1) the constants $Z$ for each pollutant in each phase can be hard to determine, and (2) the fugacity is not an added value in dynamic models (in contrast to steady state modelling). As a result, the MTFM implemented in WEST can be used as a fully dynamic (level IV) model. Moreover, SimpleBox only requires a limited amount of data to be provided by the user, as a large number of parameter values can be estimated by the model based on the micropollutant’s physical-chemical properties. The environmental compartments considered in SimpleBox are: air, soil (natural,
agricultural and urban), water (fresh and sea), sediment and above ground vegetation (natural and agricultural).

For the purpose of this work, the WEST implementation of SimpleBox (De Keyser et al., 2008) was modified: the vegetation compartment was omitted and the soil and water compartments were implemented as generic soil and water models, which can be finetuned to different soil and water types by adjusting the parameter values. Next, the model was extended with a groundwater compartment, acting as a fixed volume completely stirred tank reactor (CSTR) where no biological degradation of the micropollutant occurs. A schematic overview of the model setup is part of the overall scheme in Figure 5.1.

![Schematic representation of the integrated environmental model (IUWS model (white blocks) and MFTM (grey blocks))](image)

**5.3 Integrated environmental model**

Using the two submodels discussed above, an integrated IUWS/MFTM model was set up. The characteristics of the two models can be found in De Keyser et al. (2008). In the reference scenario, the following links between both models were considered (Figure 5.1): an upstream MFTM water compartment provides the input to the IUWS river model, a downstream MFTM water compartment receives the IUWS river water and primary and secondary waste sewage sludge is conveyed to the MFTM soil compartment after treatment in a thickener.
Figure 5.2 – Addition of storm water infiltration ponds to the previously modelled setup

Furthermore, a combined sewer system was implemented with treatment in the WWTP before discharge into the surface water. In a second scenario, with separate sewer system, the installation of stormwater infiltration ponds as best management practice (BMP) was simulated, with volatilization and infiltration processes occurring (Figure 5.2 – see also Vezzaro et al., 2009).

Wet and dry deposition as well as diffusion are considered as exchange processes between the compartments air and soil in the multimedia model, but similar links between the air and the urban catchments were neglected because the surface area of the urban catchments causes the fluxes to be a factor 1000 smaller than the assumed emissions onto the urban surface.

5.4 Bis(2-ethylhexyl) phthalate (DEHP) as a case study

The parameter values were adapted to simulate the fate of bis(2-ethylhexyl) phthalate (DEHP) in the integrated system (Table 5.1). Because of the high production volume and widespread use of DEHP, the chemical’s presence in the environment is of growing concern. Cousins and Mackay (2003) assume EU production and consumption tonnages of 595,000 and 476,000 tons of DEHP per year respectively. They use emission factors of 0.0025, 0.00025 and 0.00005 ton/year to air, water and soil respectively, due to industrial production, industrial use and transport, and of 0.01, 0.00031 and 0.00065 due to product end use and disposal, based on Parkerton and Konkel (2001). Emission estimates were converted to a per capita basis and scaled to the size of the case study. Urban emissions to water were assumed to go to wastewater, whereas emissions to soil were supposed to accumulate on the urban surface and to be washed off with runoff.
Table 5.1 – Key parameters of the micropollutant DEHP used in the case study (European Commission, 2008; Lützhøft et al., 2008)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular weight</td>
<td>g·mole⁻¹</td>
<td>390.54</td>
</tr>
<tr>
<td>Vapour pressure</td>
<td>Pa</td>
<td>3.4E-5</td>
</tr>
<tr>
<td>Henry constant</td>
<td>Pa·m³·mole⁻¹</td>
<td>4.43</td>
</tr>
<tr>
<td>Water solubility</td>
<td>µg·l⁻¹</td>
<td>3.0</td>
</tr>
<tr>
<td>KOC</td>
<td>l·kg⁻¹</td>
<td>1E5</td>
</tr>
<tr>
<td>Kd</td>
<td>l·kg⁻¹</td>
<td>1.5E4</td>
</tr>
<tr>
<td>Half-life in air</td>
<td>d</td>
<td>1</td>
</tr>
<tr>
<td>Half-life in soil and sediment</td>
<td>d</td>
<td>300</td>
</tr>
<tr>
<td>Photolysis half-life in nearsurface water</td>
<td>d</td>
<td>0.5</td>
</tr>
<tr>
<td>Biodegradation half-life in aerobic water</td>
<td>d</td>
<td>25</td>
</tr>
<tr>
<td>Biodegradation half-life in anoxic water</td>
<td>d</td>
<td>3000</td>
</tr>
<tr>
<td>Adsorption rate constant</td>
<td>m³·d⁻¹·g solids⁻¹</td>
<td>0.1</td>
</tr>
<tr>
<td>Background concentration in air entering the system</td>
<td>ng·m⁻³</td>
<td>10</td>
</tr>
<tr>
<td>Background concentration in water entering the system</td>
<td>µg·l⁻³</td>
<td>0.10</td>
</tr>
</tbody>
</table>

5.5 Initialization

The IUWS model was initialized first by means of a steady-state simulation over a period of 120 days to obtain realistic starting conditions in all components of the IUWS (sewers, treatment plant, river). Then, the average DEHP mass fluxes from the IUWS to the multimedia compartments were calculated and used as steady-state input to initialize the stand-alone MFTM. When a steady-state was reached, both models were coupled and run with the obtained steady-state quantities as initial values, which were in the range of environmental DEHP concentrations reported in literature.

5.6 Environmental concentrations and fluxes in the different scenarios

The dynamic simulation results shown in Figure 5.3 indicate that the installation of stormwater infiltration ponds helps to avoid DEHP peak discharges into the surface water originating from the stormwater after treatment in the WWTP. On the other hand, the figure also shows that the stormwater infiltration ponds reallocate the DEHP flows to groundwater and air. The increased air concentrations are transient due to photochemical breakdown and advective transport out of the modelled system, whereas the accumulation of DEHP in the groundwater compartment could potentially cause problems in the long term.
The simulated river concentrations are relatively high compared with measured concentrations reported in the literature, which are for urban areas usually below or around one $\mu$g·L$^{-1}$ (Peijnenburg and Struijs, 2006; European Commission, 2008). Further research will reveal to what extent such simplified model setup can be tuned to yield more realistic environmental concentrations.

Figure 5.4 shows an overview of all mean DEHP mass fluxes, removal rates and concentrations in the modelled system, but with the IUWS represented as one compartment. The averages are calculated with data obtained from 120 days of dynamic simulation with a 15 minute output interval. Note that by summarizing the data as mean values, the dynamics are not apparent, but without a dynamic model (rainfall input time series) these results could not have been obtained. The main effect of the stormwater infiltration ponds on the outer-urban environment is that mass fluxes towards downstream water and soil have decreased, while DEHP fluxes to the groundwater and the air compartment have increased.
The latter may in this example not have a vast impact, but without the MFTM shell around the IUWS model, this effect could not have been assessed at all.

Two aspects to take into account in future studies to increase the confidence in the interpretation of simulation results are uncertainty and time scales. The uncertainty of the simulated environmental micropollutant concentrations in multimedia models can be several orders of magnitude (Hennes and Rapaport, 1989). For the more complex urban water models, the prediction performance is known to be better in terms of traditional pollutant concentrations (Mannina et al., 2006), although this is not known yet for micropollutants since to our knowledge no studies have been conducted so far on the whole IUWS. This topic deserves further attention in order to make integrated environmental assessment based on IUWS/MFTM coupled models more reliable.

For the more complex urban water models, the prediction performance is known to be better in terms of traditional pollutant concentrations (Mannina et al., 2006), although this is not known yet for micropollutants since to our knowledge no studies have been conducted so far on the whole IUWS. This topic deserves further attention in order to make integrated environmental assessment based on IUWS/MFTM coupled models more reliable.

Figure 5.4 – Mean DEHP mass fluxes (in g·d⁻¹) and concentrations in the modelled system: reference scenario (normal typeface) and after the implementation of stormwater infiltration ponds (italic typeface); values in the IUWS block are concentrations in the receiving water after mixing, i.e. at the outflow of river stretch 5
The second important factor in studies conducted with coupled IUWS/MFTM systems is the huge range of time constants. Processes in the IUWS model like adsorption, desorption, biodegradation, etc. can have small time constants and therefore make the system respond quite fast to dynamics acting on the system, like rainfall and time-varying emissions. In multimedia models, however, large time constants in the order of years are inherent to the system. It should be investigated whether the choice of the initial conditions for the dynamic simulations has an influence on the conclusions and, in case, whether there are better alternatives then a calculated steady-state.
6 Investigation of relationships between model inputs and outputs for model complexity reduction

6.1 Introduction

A dynamic integrated urban water system model for organic micropollutants was developed and used to simulate the fate of Bis (ethylhexyl) phthalate (DEHP) in hypothetical catchment. The IUWS model was extended with a multimedia fate and transport model so that all phases of the pollutant (gas, water, soil) can be modelled.

Although the usefulness of the model has been proven on hypothetical data, we need to point to the following issues:

- Putting such a complex model to real catchment can be quite challenging, mostly due to the difficulties with data acquisition. The data about micropollutants are typically scarce or not measured at all, but at the same time they are needed in a frequent time intervals for the model.
- The complexity of the model can be a bottle neck for its integration with other decision support systems (DSS).
- It is (too) expensive in terms of calculation time.

The goal of this research is to simplify the complex conceptual model by using machine learning tools. Simplifying means obtaining a model that requires less input information, contains simple rules or equations, and produces similar output results as the conceptual model. The computational time of such model is shorter and can therefore be integrated into a GIS based DB to perform different spatial analysis and obtain results in much shorter time.

Machine learning algorithms use the data-driven approach to modelling, i.e. they construct models from observations (see e.g. Michalski et al., 1998; Witten and Frank, 2000). Different ML algorithms include different methodologies for searching for patterns in data and thus many times successfully identify simple relations between the input and the output data. Successful applications in predicting runoff can be found in (Kompare et al., 1997, Atanasova and Kompare, 2002a, 2002b). In this task we use a ML algorithm for induction of regression tree (RT) model to analyze the simulated data by the IUWS model (input and output) and construct a simple RT model for the relations between the inputs and the outputs.

We should note here that the accuracy and the generality of the RT model will highly depend on the simulated data by the conceptual model, since they are used for inducing the RT model. Note also that such model is not transferable to other catchments and if the observed catchment changes significantly the RT model would have to be re-constructed on new data.
The main task of machine learning is to learn a concept from given examples. The entire procedure consists of a concept, examples (measurements), learning algorithm and learning scheme or model (Figure 6.1). Each example consists of attributes’ and class’ values. The attributes are descriptors of the class, i.e. independent variables, while the class (also called a target variable) represents the dependent variable. The learning algorithm then, from the examples and some background knowledge, generates the learning scheme (model), which is a presentation of what has been learned, i.e. the class values are presented in terms of the attributes values. Based on the class value type (e.g. numeric, nominal, discrete, continuous…) the learning scheme (model) can be a decision tree, regression tree, classification rules, decision tables, and so on (see also Witten and Frank, 2000; Michalski et al.,1998).

Typically the quality of data-driven models is dependent on the examples (data) quality and quantity. To successfully learn a concept sufficient number of representative examples is needed.

![Diagram](image)

**Figure 6.1 – Machine learning procedure**

### 6.2 Learning numerical (continuous) classes – regression and model trees

Linear regression is a method, which aims to express the dependent variable as a linear combination of the independent variables from the given measurements (examples). Each example is given by the values of the attributes and by the class values. If \( x \) is the class and \( a_i \) are the attributes of an example then:

\[
    x = a_1*w_1 + a_2*w_2 + \ldots + a_n*w_n = \sum a_i*w_i
\]

where \( w_i \) are weights, which are learned (calculated) from the training set.

While the simple linear regression calculates one equation (one weight vector) for the entire data set, piecewise or tree-structured regression divides the data set into several subsets on
which or linear equation (or a uniform class value) can be applied. The division to subsets is based on tests of the values of the input attributes, which are put as nodes in the regression tree. Thus, regression trees are hierarchical structures composed of nodes and branches, where the internal nodes contain tests on the input attributes. Each branch of an internal test corresponds to an outcome of the test, and the predictions for the values of the target attribute (class) are stored in the leaves, which are the terminal nodes in the tree.

Widely used algorithm for regression trees is the M5 algorithm (Quinlan, 1992). It works recursively, starting with the entire set of examples \((S)\) and selecting the best attribute and the best split of that attribute according to the splitting criterion, e.g. to get the most homogeneous subset regarding the class values or regression model. Algorithm M5 uses the maximization of the expected error reduction in predicted class values as a splitting criterion at a node.

Let \(S\) be the subset of examples that reaches a particular internal node of the tree. First test at that node is whether the set contains only a few examples or the class value of the examples vary only slightly. In this case the process is terminated and a leaf is constructed. Otherwise the set \(S\) is split to subsets \(S_i\) according to the outcomes of a test which is performed for each attribute. All outcomes are evaluated by calculating the expected error reduction in predicted class values (eq. 2):

\[
SDR = sd(S) - \sum \frac{|S_i|}{|S|} \cdot sd(S_i)
\]

where \(sd(S)\) is standard deviation of the class values in the set \(S\) and \(S_i\) are the sets that result from splitting the node according to the chosen attribute. The attribute that maximizes the expected error reduction is chosen for splitting that node.

6.3 Model accuracy and pruning

After the tree is constructed from the training (learning) set of data, it is necessary to assess the model quality, i.e., the accuracy of prediction. This can be done by providing an additional testing set of data. Another option is to employ cross-validation. The given data set is partitioned on a chosen number of folds \((n)\). In turn, each fold is used for testing, while the remainder \((n-1\) folds) is used for training. Finally the model that fits best to all folds is chosen.

The size of the error between the actual and the predicted values is calculated by several measures to evaluate the model accuracy: root mean-squared error, mean absolute error, root relative squared error, relative absolute error, and correlation coefficient. If \(x\) is actual value of the class, \(x'\) is the predicted value by the model and \(\bar{x}\) is average class value, then the size of the error between the actual and the predicted values is calculated by several measures to evaluate the model accuracy:
Pruning is a powerful technique to deal with noisy and incorrect data. Pruning works much the same way as the name says, i.e., cutting away the least accurate leaves and thus resulting in a smaller, more generalized tree (Figure 6.2). This method prevents the tree from overfitting the learning data and improves the quality (performance) of the tree on unseen data. Following we illustrate how pruning affects the model performance.

**Figure 6.2 – Pruning the tree.**
6.4 Building a regression tree model for DEHP

6.4.1 Data set for model construction

The data for model induction are obtained by simulating the IUWS model introduced in section Chapter 5. As input the model takes calculated DEHP loads from the catchment, presented in Table 6.1 and the catchment characteristics. In turn it calculates DEHP concentrations and loads in each simulated element in the catchment (river stretches, sewer sections, WWTP, and so on) at every time step in the simulated period. We selected one year simulation period with one hour time step resolution.

In order to have enough representative data, which indicate the system’s behaviour, we prepared and simulated 5 different scenarios by changing the input loads:

- basic load (21 emission strings) converted into dynamic loads and aggregated according to the compartment they contribute: DWF, ACC, RIVER, AIR, SOIL
- DWF increased by 50%
- ACC decreased by 20%
- basic load entirely decreased by 80%
- basic load entirely decreased by 50%

In Figure 6.3 we present simulated data of the DEHP loads in the river stretch 5 at the outflow of the catchment for the 5 scenarios. Expectedly, the highest peaks appear at the highest input loads (DWF increased by 50%) and the lowest at basic minus 80%.

Table 6.1 – Aggregated DEHP loads in the catchment

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREC</td>
<td>precipitation; summed within selected time interval (4 hours)</td>
<td>mm</td>
</tr>
<tr>
<td>DEHP_WW_A</td>
<td>Sum of all DEHP emissions connected to the sewer system in the urban catchment A</td>
<td>g/day</td>
</tr>
<tr>
<td>DEHP_WW_B</td>
<td>Catchment A, B, C respectively. This load contributes to dry weather flow.</td>
<td></td>
</tr>
<tr>
<td>DEHP_WW_C</td>
<td>Sum of all DEHP emissions connected to the sewer system in the urban catchment A, B, C respectively. This load contributes to dry weather flow.</td>
<td>g/day</td>
</tr>
<tr>
<td>DEHP_ACC_SP</td>
<td>Sum of all emissions in the urban catchment A that are accumulated on the surface. This load is washed from the surface during rain events</td>
<td>g/(day)</td>
</tr>
<tr>
<td>DEHP_ACC_DA</td>
<td>Sum of all emissions in the urban catchment A, B, and C that are accumulated on the surface. This load is washed from the surface during rain events</td>
<td>g/(day)</td>
</tr>
<tr>
<td>DEHP_ACC_SA</td>
<td>Sum of all emissions in the urban catchment A, B, and C that are accumulated on the surface. This load is washed from the surface during rain events</td>
<td>g/(day)</td>
</tr>
<tr>
<td>DEHP_AIR</td>
<td>Sum of all emissions going to the air</td>
<td>g/(day)</td>
</tr>
<tr>
<td>DEHP_SOIL</td>
<td>Sum of all emissions going to the soil</td>
<td>g/(day)</td>
</tr>
</tbody>
</table>
6.4.2 Data and model preparation

While the IUWS model takes substantial input data for simulations and returns output DEHP loads and concentrations at various points in the catchment, we set up the machine learning algorithm to induce a model that will take as input variables only the loads presented in Table 6.1, and simulate only the DEHP load at the river section after the WWTP (river stretch 5), see Figure 6.3. Thus, the task is to simplify the complex conceptual model in a way that the output load is calculated from the input loads only. The variables used for induction of regression tree model are schematically presented in Figure 6.4.

In line with the explanation of regression tree induction from section 2, the concept or the class to be learned is the DEHP load in the last river section and the attributes used for describing the concept are presented in Table 6.1. Additionally, we introduce attributes that include six hours history in the data, i.e., the present DEHP load in the river section will become dependent on the present load from the catchment as well as on the load that was released one to six hours ago. History attributes were given additional notation -X, where X represents the value of the attribute X hours ago. For example DEHP-SP-DWF-3, represents the load from Spangen that contributes to the dry weather flow, generated 3 hours ago. After introducing these attributes the data set comprised 70 attributes and a class DEHP-river.

The data used for training the model is taken from the 5 simulations, schematically presented in Figure 6.5. The reminder of the data was used for testing the model accuracy on new (unseen) data.
Figure 6.4 – Variables used by the ML algorithm for model induction: green – independent attributes, red – target or class

Figure 6.5 – Selection of training data for the ML algorithm
6.4.3 Constructing and evaluating the regression tree model

To build a regression tree model we employ the M5P algorithm for induction of regression trees, which is implemented in the WEKA modelling software (Witten and Frank, 2000). Using the prepared training data set the algorithm induced a regression tree model presented in Figure 6.6.

![Regression Tree Model](image)

**Figure 6.6. – Regression tree model**

The model is composed of 4 linear models (LM), which are used according to the values of the attributes in the internal nodes. For example, LM 2 is used if DEHP-air-6 is less then 103.417 and more then 41.374. Linear models represent linear dependencies of the target variable DEHP-river and all input attributes.

The correlation coefficient on the training data set is 0.78. The model was validated by running it on the entire data set, i.e. all five scenarios. Each scenario contains data that was used for training as well as data ‘unknown’ to the model (Figure 6.5). In Figure 6.7 we present a fraction of the simulated data with the regression tree model, together with the training data (data simulated by the IUWS model).
Figure 6.7 – Comparison of the simulated data by the regression tree (RT) model and conceptual IUWS model for the scenarios: a) DFW increased by 50% b) ACC decreased by 20% and c) basic load reduced by 50%
The simulations show that the regression tree model successfully identifies the peaks, but it sometimes fails to correctly simulate the amplitudes. The correlation coefficients for all five scenarios range from 0.56 to 0.67. Possible reason for not achieving higher accuracies is that the training data is not representative enough, i.e. more attention should be put in the selection of the training data as well as more simulated data should be provided. However, verification of the model showed good and stable behaviour by correctly responding to the change of the input attributes, i.e., input loads (Figure 6.8). The highest peaks appear at increased loads, whereas at reduced loads we observe lower DEHP loads in the river section. Additionally, we simulated a dry year event by decreasing the amount of rain from 800 mm/year to 540 mm/year, which was not included in the training data, and the model reacted according to expectations.

Thus, although not achieving high accuracy compared to the simulated data from the IUWS model, the RT model correctly responds to changes in the input parameters. The computational time of the RT model is much shorter compared to the IUWS model and its structure is expressed in terms of linear equations. As such, the model allows for successful integration in other (computationally demanding) decision support tools, such as GIS and relational databases.

![Figure 6.8 – Simulating the RT model for different loads](image-url)
7 Conceptual development of a simplified version for education

To facilitate the use of the integrated model for educational purposes, a simplified layout has been implemented in WEST. This system configuration includes one urban draining catchment, one pipe, one CSO structure, one primary clarifier, two activated sludge units (anoxic and aerobic), one secondary clarifier and two river stretches (one before the CSO outlet and one before the WWTP outlet).

The simplicity of the layout allows faster data entering – e.g. when changing chemical’s properties or physical system parameters – and faster simulation.

The model and experiment in WEST are available.
8 Conclusions

The combination of integrated urban water system (IUWS) models with multimedia fate and transport models (MFTM) can either be done by creating an interface between two existing models or by merging them into one ‘supermodel’. Both approaches can be challenging, depending on the flexibility of the simulation platform. The added value of wrapping a MFTM around an IUWS model is twofold: (1) the multimedia model provides boundary conditions to the IUWS, and (2) it allows to make a holistic assessment of the overall environmental status of the modelled system, beyond urban surface water quality. A combined MFTM/IUWS model was presented and used to simulate the fate of the micropollutant DEHP in two scenarios: a reference and a scenario with stormwater infiltration ponds. The main effect of the stormwater infiltration ponds on the outer-urban environment was shown to be that mass fluxes towards downstream water and soil have decreased, while fluxes to the groundwater and the air compartment have increased.

Outflow concentrations can be modelled from the input load only, i.e. it is possible to simplify the conceptual model if sufficient data set exists. This means sufficient number of simulated scenarios by combining different precipitations and loads in the catchment, so that most of the potential situations are covered.
9 References


Quinlan, J.R., Learning with continuous classes. In: Adams & Sterling (Eds.), proceedings AI’92 (Australian Conference on AI), Singapore, World Scientific 1992, 343-348


Appendix A – Model transformers

Names in square parenthesis are state variables, names ending with _In are parameters for variables not existing in the source model, names starting with f_ are fraction parameters, names starting with i_ are elemental composition fraction parameters.

Sewer to WWTP

\[
\begin{align*}
[H_2O] &= [H_2O\_sew] \\
[S\_O] &= S\_O\_In \\
[S\_ALK] &= S\_ALK\_In \\
[S\_NO] &= S\_NO\_In \\
[S\_N2] &= S\_N2\_In \\
[S\_NH] &= [NH4\_sew] \\
[S\_PO] &= [PO4\_sew] \\
[X\_PP] &= X\_PP\_In \\
[S\_F] &= [COD\_sol] \times f\_S\_F \\
[S\_A] &= [COD\_sol] \times f\_S\_A \\
[S\_I] &= [COD\_sol] \times (1 - f\_S\_F - f\_S\_A) \\
[X\_S] &= [COD\_part] \times f\_X\_S \\
[X\_H] &= [COD\_part] \times f\_X\_H \\
[X\_AUT] &= X\_AUT\_In \\
[X\_PAO] &= X\_PAO\_In \\
[X\_PHA] &= X\_PHA\_In \\
[X\_I] &= [COD\_part] + [X\_H] + [X\_AUT] + [X\_PAO] + [X\_PHA] + [X\_S] \\
[X\_MEOH] &= X\_MEOH\_In \\
[X\_MEP] &= X\_MEP\_In \\
[X\_TSS] &= [X\_TSS\_sew] \\
[S\_MP] &= [S\_MP\_sew] \\
[X\_MP] &= [X\_MP\_sew]
\end{align*}
\]

Sewer to river

\[
\begin{align*}
[rH2O] &= [H2O\_sew] \\
rS\_O2 &= rS\_O2\_In \\
rS\_ALK &= rS\_ALK\_In \\
rS\_NO &= rS\_NO2\_In
\end{align*}
\]
\[ rS_{-NO3} = rS_{-NO3\_In} \]
\[ rS_{-NH} = [NH4\_sew] \]
\[ rS_{-PO} = [PO4\_sew] \]
\[ rX_{-P} = [PO4\_sew] + i_P_S_I \ast [rS_I] + i_P_S_S \ast [rS_S] + i_P_X_I \ast [rX_I] \]
\[ + i_P_X_S \ast [rX_S] + i_P_BM \ast [rX_{N1}] i_P_BM \ast [rX_{N2}] + i_P_BM \ast [rX_H] \]
\[ rS_S = [COD_{sol}] \ast f_S_S \]
\[ rS_I = [COD_{sol}] \ast (1 - f_S_S) \]
\[ rX_S = [COD_{part}] \ast f_X_S \]
\[ rX_H = rX_H\_In \]
\[ rX_{N1} = rX_{N1}\_In \]
\[ rX_{N2} = rX_{N2}\_In \]
\[ rX_I = [COD_{part}] + [rX_S] + [rX_H] + [rX_{N1}] + [rX_{N2}] \]
\[ rX_{ALG} = rX_{ALG\_In} \]
\[ rX_{ND} = rX_{ND\_In} \]
\[ rX_{II} = rX_{II\_In} \]

**WWTP to river**

The following pages are extracted from Solvi (2007), describing the method adopted for the WWTP to river transformer.
To create an integrated sewer-WWTP-river model, submodels need to be interfaced, in this case KOSIM-WEST variables need to be linked to ASM and RWQM variables, as well as ASM to RWQM variables. Problems arise from the following three arguments (Vanrolleghem et al. (2005b)):

- some state variables used in one model do not exist in the connected model
- the 'meaning' of a state variable in one system may not hold for the other system (e.g., components can be considered as inert in one system but may be biodegradable in another)
- the elemental composition of a component variable in one model is not identical with the component variable in the connected model.

A connector model, respecting closed mass and elemental balances, was proposed in a case study on the river Lambro (Italy) that links the states of the ASM1 and RQWM1 (Meißen et al. (2001), Benedetti et al. (2004)). The continuity-based interfacing method (CBIM) (Vanrolleghem et al. (2005b)), creates a formalised frame on the basis of this work. The main idea of the interfaces is that one constructs a set of algebraic transformation equations on the basis of a Petersen matrix description of the two models to be interfaced (i.e., from origin model $P$ to destination model $Q$). Through this approach it is possible to maintain the continuity of elements $C$, $H$, $N$, $O$, $P$, charge and COD, while the two models remain unaltered. The methodology consists of the following steps (Benedetti (2006)):

1. Formulation of elemental mass fractions and charge density.
2. Set-up of the composition matrix.
3. Definition of the transformation matrix.
4. Implementation of the transformation equations.

3.4.2.1 Formulation of elemental mass fractions and charge density

The main hypothesis in this phase is that the mass of each component $k$ is made up of constant fractions of the elements $C$, $N$, $O$, $H$ and $P$. The elemental mass fractions $\alpha_k^C$, $\alpha_k^N$, $\alpha_k^O$, $\alpha_k^P$, $\alpha_k^D$ and $\alpha_k^P$ are given in grams of element per gram of component. For the components for which the elemental composition is known, the calculation of the mass fractions is straightforward. For the components for which the elemental composition is not known, it is necessary to make some assumptions and to use data provided by literature (e.g., Reichert et al. (2001)). As a result, for component $k$,

$$\sum_{E \in \mathcal{E}} \alpha_k^E = 1 \quad \text{for} \quad \mathcal{E} = \{C, N, O, H, P\}. \quad (3.40)$$

Then, also $\alpha_k^{COD}$ and $\alpha_k^{CH}$ (CH stands for charge) can be calculated. The COD equivalent of a component is defined as the grams of oxygen that are consumed during oxidation of a mass unit.
of the component to $NH_4^+$, $CO_2$, $H_2O$, $H^+$ and $PO_4^{3-}$. The COD equivalent of a component $k$ is related to the mass fractions of elements and charge through the relationship (Reichert et al. (2001)):

$$\alpha_k^{COD} = 32 \frac{\alpha_k^C}{16} + 8 \alpha_k^H - 16 \frac{\alpha_k^O}{16} - 24 \frac{\alpha_k^N}{14} + 40 \frac{\alpha_k^P}{31} - 8 \alpha_k^{Ch}$$  \hspace{1cm} (3.41)

Using the charge $Ch_k$ of a component $k$ and the molecular weight $m_k$, the charge density is

$$\alpha_k^{Ch} = \frac{Ch_k}{m_k}$$  \hspace{1cm} (3.42)

Such formulation of mass fractions and charge density is done both for the components of the origin and the destination model.

3.4.2.2 Set-up of the composition matrix

All fractions of all components from both origin and destination model can be placed into a matrix $\alpha = \alpha_k^{E}$, where $k = 1, \ldots, p, \ldots, p+q$ where $p$ and $q$ are the number of components in the origin and destination matrices $P$ and $Q$ respectively. To set-up the final composition matrix $i$ used to connect the models, $\alpha$ needs further conversion. The ASM, as well as the RWQM model components are expressed in various stoichiometric units like $M(COD)/L^{-3}$, $M(N)/L^{-3}$, $M(P)/L^{-3}$ or $M(H)/L^{-3}$ depending on the kind of transformations they are submitted to. An element of the composition matrix $i$ is therefore given by:

$$i^E_k = \alpha_k^E \cdot M_k$$  \hspace{1cm} (3.43)

where $M_k$ is expressed in grams of component $k$ per gram of stoichiometric unit ($COD$, $N$, ...)

and can be calculated using the molecular weights of the stoichiometric units. Also,

$$i_k^{COD} = \alpha_k^{COD} \cdot \sum_{All E} i_k^E$$

and

$$i_k^{Ch} = \alpha_k^{Ch}$$

3.4.2.3 Definition of the transformation matrix

The main concept behind the transformation matrix $\theta$ is that the components of the origin model $P$ are transformed completely into the variables of the destination model $Q$. To ensure this, a number of transformations have to be specified. The definition of these equations depends on the knowledge available on the processes. Usually the number of transformations $i$ to be defined is equal to the number of state variables $p$ of the origin model. Each transformation converts a number of components of the origin model to a number of components of the destination model. Every transformation $j$ for component $k$ is characterized by its stoichiometry $\theta_{j,k}$. While stoichiometry coefficients of the origin components are set to an arbitrary value (with negative sign in order to maintain the right direction of the transformation), the coefficients of
the destination state variables are set so that each transformation maintains the COD content. For each transformation \( j \) the elemental continuity must be guaranteed, which is easily checked by the equation (where \( k \) are the components and \( E \) the elements):

\[
\sum_{k} \theta_{j,k} \cdot i_{k,E} = 0
\]  

(3.44)

Each transformation \( j \) is also characterized by its transformation rate \( \rho_j \) which, together with the stoichiometry coefficient, specifies the amount of the component \( k \) transformed per unit of time, equal to \( \theta_{j,k} \cdot \rho_j \).

### 3.4.2.4 Implementation of the transformation equations

The set of interface unknowns consists of the stoichiometric coefficients \( \nu_{j,k} \) and the transformation rates \( \rho_j \). Together they enable the calculation of the outflux from the destination model. In order to solve the unknowns it is necessary to set up a system of two sets of equations taking into account the fluxes in and out from the interface.

\[
\Phi_{k}^{in} = - \sum_{j=1}^{N} \nu_{j,k} \cdot \rho_j \quad \text{for} \quad k = 1, \ldots, p
\]  

(3.45)

\[
\Phi_{k}^{out} = \sum_{j=1}^{N} \nu_{j,k} \cdot \rho_j \quad \text{for} \quad k = p + 1, \ldots, p + q
\]  

(3.46)

where \( \Phi_{k}^{in} \) is the known positive influx of a component \( k \) of the origin model, \( \Phi_{k}^{out} \) is the unknown outflux of a component \( k \) of the destination model, \( p \) is the number of origin state variables and \( q \) is the number of destination state variables. It is important to check that all transformation rates \( \rho_j \) are positive, in order to assure that the transformations are in the right direction (origin model to destination model). In case this is not verified, the transformation equations should be modified.