Technical University of Denmark

DTU

Strategies for monitoring of priority pollutant emission barriers

Pettersson, Maria; De Keyser, Webbey; Birch, Heidi; Lützhøft, Hans-Christian Holten; Gevaert, Veerle; Cerk, Matej; Benedetti, Lorenzo; Vanrolleghem, Peter A.

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

ScorePP

Strategies for Monitoring of Priority Pollutant Emission Barriers

Deliverable No: D7.5, Date: 2010-02-22
Dissemination level: Public

Maria Pettersson[1], Webbey De Keyser[2], Heidi Birch[3], Hans-Christian Holten Lützhøft[3], Veerle Gevaert[2], Matej Cerk[4], Lorenzo Benedetti[2] and Peter A. Vanrolleghem[5]

[1] Environment and Health Administration, City of Stockholm, Sweden
[2] Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Belgium
[3] Department of Environmental Engineering, Technical University of Denmark, Denmark
[4] Institute of Sanitary Engineering, Faculty of Civil and Geodetic Engineering, University of Ljubljana, Slovenia
[5] modelEAU, Department for civil engineering, Université Laval, Canada

Deliverable number:	D7.5
Deliverable title:	Strategies for Monitoring of Priority Pollutant Emission Barriers
Authors:	Maria Pettersson, Webbey De Keyser, Heidi Birch, Hans-Christian Holten Lützhøft, Veerle Gevaert, Matej Cerk, Lorenzo Benedetti and Peter A. Vanrolleghem
Review and Assessment:	Boris Kompare and Peter A. Vanrolleghem
Date submitted to project coordinator:	2010-02-22
Approved by (Work package leader) :	2010-02-16
Approved by (Coordinator) :	2010-02-25

Abstract

The objective of Task 7.5 was to develop tools for model-based planning of sampling campaigns in the design of monitoring strategies for priority pollutant emission barriers.

Using integrated urban wastewater system (IUWS) models, measurement campaigns can be designed to improve the calibration of the model's parameters (Optimal Experiment Design for Parameter Estimation (OED/PE)). Furthermore, the knowledge contained in the IUWS models can be used to optimize the planning of sampling campaigns aiming at assessing the efficiency of emission control strategies. To do this, the emission barriers need to be implemented in the IUWS model, as well as the sampling and measuring devices that will be used. The simulation results are presented as a Substance Flow Analysis (SFA). These SFAs can be compared with empirical SFAs and can also be used to set up measurement campaigns aiming at gathering information to establish or improve empirical SFAs (OED/SFA).

Moreover, the national Danish environmental monitoring programs were compared to the US programs under the Clean Water Act and the proposed environmental monitoring of the European water courses as described in the European Water Framework Directive. Also, kinetic passive sampling methods for the measurement of average water concentrations are presented.

Table of Contents

1 Introduction and aims

This report was conducted within the scope of the EU project entitled Source Control Options for Reducing Emissions of Priority Pollutants (ScorePP). The aim of the project is to develop comprehensive and appropriate source control strategies that authorities, cities, water utilities and chemical industries can use to reduce priority pollutant emissions to urban waterways. The project focuses on the 33 priority and priority hazardous substances identified in the European Water Framework Directive (WFD) (European Commission, 2000).

The objective of the task presented in this report (Task 7.5) has been to develop tools for model-based planning of sampling campaigns in the design of strategies for monitoring of priority pollutant emission barriers.

The models implemented in Task 7.3 (*Multimedia model at river basin scale*) and 7.4 (*Integrated dynamic urban scale sources-and-flux model for priority pollutants*) of the project were used to develop strategies for monitoring the effects of applying the emission barriers identified and developed in work package (WP) 4 (*Limiting releases of priority pollutants*) and 5 (*Treatment options*) of the project. The models are used to understand/predict variations in discharges of priority pollutants from urban water systems in order to optimise frequency, duration and location for sampling campaigns.

A monitoring strategy can include different methods for sampling and analyses, and its design can benefit from the use of Substance Flow Analysis (SFA) and by adopting a life cycle perspective to avoid exclusion of important sources for priority pollutants. In Optimal experiment design (OED) a set of measurements forming a measurement campaign is proposed. The proposed campaign is then optimised by iteratively calculating a certain objective function and minimising its value by exploiting one or more degrees of freedom. In this Task, OED strategies have been developed for three distinct objectives: parameter estimation, SFA and for the assessment of emission barriers.

2 Review of monitoring strategies and sampling techniques

This chapter of the report deals with existing monitoring programmes and tools for monitoring.

In the first section the national Danish environmental monitoring programs are compared with the US programs under the Clean Water Act and the proposed environmental monitoring of the European water courses as described in the European WFD. A few proposals for monitoring programs published in the peer-reviewed literature have been included as well. Some considerations on the sampling techniques and uncertainty associated with environmental sampling will conclude this section.

The second section presents some kinetic passive sampling methods for the measurement of average water concentrations of substances, with a special focus on the priority substances listed in the WFD.

2.1 Selected national and regional monitoring programmes

2.1.1 NOVA and NOVANA

In Denmark national surveillance of watercourses started in 1974, latest with the two programmes Nationalt Program for Overvågning af Vandmiljøet (National Programme for the Surveillance of the Aqueous Environment), NOVA, from 1998 to 2003, and Det nationale program for overvågning af vandmiljøet og naturen (The national programme for the surveillance of the aqueous environment and

the nature), NOVANA, from 2004 to 2009. It was only with these last two programmes that focus has been put on priority pollutants, as the former programmes mainly have dealt with nutrients.

The two programmes are very similar, with the main difference that NOVANA is slightly reduced in volume and slightly changed in selection of monitored substances. Under the NOVA programme about 68 pesticides, 21 PAHs, 13 other xenobiotics and 8 heavy metals have been analysed, whereas under the NOVANA programme it is about 18 pesticides, 9 other, 19 PAHs and 5 heavy metals.

2.1.1.1 *Purpose of the NOVA and NOVANA programmes*

The aims with NOVA have been to assess the environmental status and the development of the environmental status as well as to assess the effects on various water courses through a systematic monitoring campaign. The aim was also to assess the effects of various efforts made according to earlier programmes as well as to assess if new efforts have to be introduced.

With NOVANA also the WFD was established and therefore one of the aims was to obey the MS obligations described in the WFD. Still an aim was to assess earlier efforts made.

2.1.1.2 *Monitoring frequencies within NOVA and NOVANA*

Under the NOVA programme (1998-2003) a number of water courses were selected to undergo monitoring. A range of pesticides, PAHs, other xenobiotics, e.g. plasticisers and heavy metals were in general analysed 12 times per year. For some locations the analysis frequency for pesticides was only 6 times per year but for some other locations it was 16 times per year.

Under the NOVANA programme five larger locations will be analysed 12 times per year (2004-2007) with respect to PPs. During the years 2007-2009 only screening analyses will be performed, but on a larger number of water courses.

2.1.2 The European Water Framework Directive

Within the European Union (EU) the WFD was put in force in the year 2000 (European Commission, 2000). The EU WFD aims at protecting the aquatic environment from deterioration due to pollution with individual or groups of priority substances. The ultimate aim is to achieve background levels for naturally occurring substances and close to zero levels for man-made substances. This aim shall be fulfilled within 20 years after adoption at community level.

The monitoring programme proposed within the WFD (European Commission, 2000) is operating on three levels:

1. Control surveillance
2. Operational surveillance
3. Investigative surveillance

2.1.2.1 *Control surveillance*

The control surveillance serves to supply information to and validate the procedure for assessment of the environmental effects, to contribute to an efficient design of future monitoring programmes, to assess long term changes of the natural conditions as well as assessing the long term changes due to human activities.

Each EU member state has to select a number of catchments such that a sufficient assessment of the environmental status can be performed. Those catchments are selected on the following criteria: catchments where the water flow is significant, i.e. certain catchments at large water courses, and where the catchment is larger than 2.500 km2; where the water body is significant, i.e. large lakes and reservoirs; where significant water bodies cross borders between member states; catchments selected according to the decision on a common procedure for the exchange of information on the quality of

surface fresh water in the Community (77/795/EEC) (European Economic Community, 1977) and catchments where it is necessary to assess the quantity of pollution that is transferred from one member state to another or to the ocean.

The control surveillance is performed for each of the catchments selected according to the above criteria and during a period of time of one year within the time horizon of the river basin management plan analysing for the prioritised substances, i.e. priority pollutants that are emitted within the catchment. However, if a previous control surveillance has shown a good environmental status, the control surveillance is allowed to take place only every third river basin management plan.

2.1.2.2 *Operational Surveillance*

The operational surveillance serves to assess the environmental status of the catchment that is suspected not to obey the environmental quality standards (EQSs) and to assess the effects of a given effort to reduce the emissions of certain priority pollutants.

The catchments to undergo operational surveillance are selected based on the member state performed assessment of the priority pollutant stress in a particular catchment, with respect to for example point sources, diffuse sources and human activities. If necessary, modelling can be used to predict the environmental releases within a certain catchment. Also catchments that during the above mentioned control surveillance have shown trends not to be able to obey the EQSs have to undergo operational surveillance. And thirdly, when known releases of PPs are taking place, the catchment has to undergo operational surveillance.

The locations within a catchment to undergo operational surveillance are selected according to the EQS Directive (European Commission, 2008). If no specific guideline is found in the EQS Directive, the following procedure is followed: Sufficient amounts of sampling locations are selected within the catchments in order to assess the volume and consequences of both point and diffuse pollution sources.

The substances to undergo such analyses are the priority pollutants that are released and other substances released in significant amounts.

2.1.2.3 *Investigative Surveillance*

The investigative surveillance is performed when the reasons for exceeding EQSs are unknown, when it has been documented through control surveillance that a certain catchment most likely will not obey the EQSs or to confirm the volume and consequences of an accident. The investigative surveillance will then form the basis for the efforts to decrease emissions to the particular catchment.

2.1.2.4 *Water Framework Directive Surveillance Frequencies*

For water courses like rivers, streams, lakes, estuaries and coastal waters the analysis frequency for *control surveillance* is *monthly*, unless technical or expert judgement justifies less frequent analysis.

For *operational surveillance* the frequency must be *more often than selected for control surveillance*.

The analysis frequency must be selected in a way that both reliability and precision are acceptable.

Additionally, both the natural and human affected variation must be considered. Seasonal variations must be minimised and it must be ensured that the analyses reflect changes within the catchment due to human activities.

2.1.3 Clean Water Act

The Clean Water Act or more correctly the "Federal Water Pollution Control Act" in the United States of America (USA) was established in the late 1940ies with substantial changes in 1972 and

amendments in 1977. The aim was to eliminate discharges and achieve a pollutant free water environment not later than the mid 1980ies. Substantial amounts of money were set aside for various purposes to fulfil the aim of the Act. For instance 175 mill $ were devoted to training during a 13 year period from 1978 to 1990. Focus was also put on special geographical areas like the Chesapeake Bay, the Hudson River, the Lake Tahoe and Alaskan villages, where amounts were invested ranging from 0,5 mill $ for the Alaskan project, 200 mill $ over a five year period for the Chesapeake Bay project and more than 300 mill $ for the Great Lakes over a 20 year period. Investigations of stormwater best management practices were funded with 9 mill $ for 2002 to 2004.

Effluent limitations should be achieved in 1977 for all point sources, except publicly owned treatment works. The National Pollution Discharge Elimination System (NPDES) was established. According to the NPDES any source seen as a point source, e.g. pipes or man-made ditches must obtain a discharge permit. The focus was on concentrated animal feeding operations (CAFOs), combined sewer overflows (CSOs), pre treatment, sanitary sewer overflows (SSOs) and stormwater.

The Clean Water Act is a framework for the entire USA. This means that it sets out guidelines for the individual states. However, the Act only demands each state to establish emission limit values, EQSs and to ensure that monitoring programmes are performed in a way that the state can live up to the established EQSs. The Act also demands the states to report the results of the monitoring programmes to the US Environmental Protection Agency (US EPA), but the reported data are not to be compared, as they have been established on different grounds.

Thus, monitoring programmes are performed in order to check the EQSs set up by the states them selves, and neither the monitoring programmes nor the EQSs are centrally controlled.

2.1.4   Peer-reviewed contributions to sampling and monitoring strategies

The following are a few examples on how alternative approaches will contribute to the improvement of the analytical results of a monitoring campaign.

According to Battaglin and Lauren (1996) monthly sampling of pesticides in rivers gave the most accurate picture of the pesticide levels in the rivers. The monthly sampling strategy was compared with 9 other strategies; for example quarterly (as recommended by the US EPA during the time the study was performed), once in June (all other set to zero), once in May and June (all other set to zero), once in April, May and June (all other set to zero) and once in April, May, June and July (all other set to zero), bimonthly, once in the second quarter. The analyses were used in Monte Carlo simulations to estimate the annual mean pesticide levels. Interestingly, compared with the EPA recommendation of sampling quarterly, the monthly sampling showed the most reliable results closely followed by the single measurements in two, three or four of the months from April to July. Monthly sampling is more expensive than quarterly, but the single sampling campaigns in the spring/summer months are tempting, as they will be cheaper and give more reliable results. Only 30% of the simulations with quarterly sampling were within 0,1 µg/L of the annual mean concentration, whereas monthly and single sampling ranged from 44-54% at the same uncertainty level.

One issue is the temporal variance as described above, another issue is the spatial variance. This is of course more evident for solid phases like sediment and soil. However, the spatial variance in aqueous samples will also be existing, though to a smaller extent. The following is an example on the use of GPS and GIS combined with pre-existing knowledge to design a sampling campaign with focus on the spatial variance (Caeiro et al., 2003). The estuary of Sado (Portugal) was divided in small areas and random assignment of the sampling spots was done. Sampling spots were located using GPS. The advantage of using this approach was the good correlation and interpolation among results and as a further advantage, the use of GPS and association with a GIS framework allows future comparison with the same data on a much more precise spatial level.

2.2 Uncertainty and representativeness

2.2.1 <u>Uncertainty</u>

When you consider a water sample with the naked eye, it is tempting to assume it is homogeneous and when a substance is truly dissolved in pure water, the sample is verging on being homogeneous. However, natural and urban water bodies very often contain dissolved organic matter, various organic and particulate matter or even suspended particles. As such organic matter and particles almost never have the exact same shape and density, there is a large chance that two samples of such a water will have different constituents with respect to organic matter and particles. Substances sorbed to such constituents will therefore be represented to different extents in two different water samples.

With these issues in focus almost no samples are homogeneous and especially any environmental sample must be regarded as heterogeneous (Pitard, 1993).

Yet another issue is the variation in time. Natural and urban watner systems have inflowing and out flowing water, which not only contribute to the content of organic matter, suspended particles etc., but also contribute to variation over time.

Analytical results are always associated with some degree of variance, whether it is connected to the external calibration curve (analyte dissolved in an organic solvent) produced in the lab, replicates of the same sample, replicates from the same sampling site or replicates from the same sampling site over time representing the environmental compartment under study; of course the variance increases in the mentioned order.

According to Esbensen and Heydorn (2004) at least three factors have influence on the variance associated with analytical results; the *analytical procedure*, the *sampling process* and the *matrix* itself. Ignoring any other factors with influence on the uncertainty and assuming a linear relationship without interactions among the individual factors the variance can be expressed as:

$$u_{tot}^2 = u_a^2 + u_s^2 + u_m^2$$

where u^2 represents the variance and the subscripts *tot*, *a*, *s* and *m* represent total, analysis, sampling and matrix, respectively. Analysis uncertainty is related to for example instrument, dilution/concentration, pre-treatment and purification. Sampling uncertainty is related to for example sampling equipment and the overall sampling strategy. Matrix uncertainty is related to for example homogeneity/heterogeneity or constituents of the matrix.

To illustrate this the following example takes its off spring in Table 1 where 10%, 30% and 50% apply to the analytical, the sampling and the matrix variances, respectively, for a basic setting. The total variance can then be calculated according to the equation above. To improve the total variance, one can choose to put effort into reducing either one of the three respective parts of the total variance. In Table 1 a 50% reduction is shown for either the analytical, the sampling or the matrix components and the total variance is consequently calculated and illustrated in Figure 1.

Table 1. Overview of consequences in reducing the variance of one of the components constituting the total variance.

Scenario	Total variance	Analytical variance	Sampling variance	Matrix variance
Basic	59	10	30	50
Analytical	59	5	30	50
Sampling	53	10	15	50
Matrix	40	10	30	25
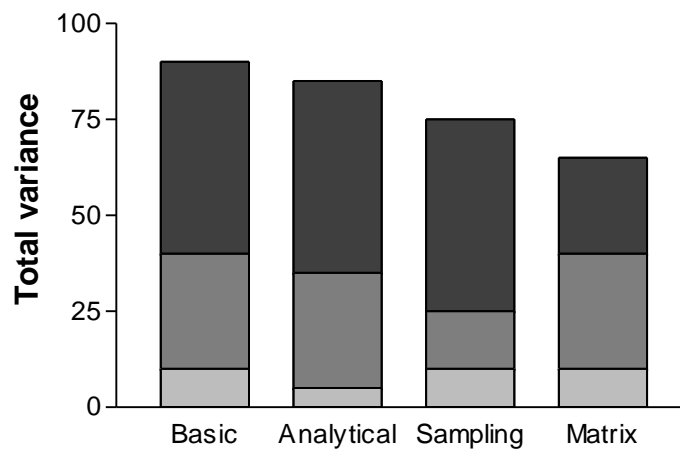
Figure 1. Illustration of the effect in reducing uncertainty in either the analytical procedure, the sampling procedure or by considering matrix effects. ▢: analytical variance, ▨: sampling variance and ■: matrix variance. In each case the variance of the considered component has been reduced by 50%, see also Table 1

Figure 1 illustrates that reducing the variance of one of the individual components of the total variance is best paid off when the particular component has a high variance as a starting point. In this example the variance associated with the matrix effects (50%) was much higher than with the analytical procedure (10%). So by reducing the matrix effects from 50% to 25% compared with reducing the analytical variance from 10% to 5%, resulted in a total variance of 40% compared with 59% for both the basic setting as well as the scenario with reducing the analytical variance.

A study on PCBs in marine sediment showed that the short-term analytical variance was in the order of 10% and the long-term analytical variance was in the order of 20-30%, whereas the field (matrix) variance was in the order of 20-50% (Kelly et al., 994). It is further stated that the effort to reduce the analytical variance will have little influence on the precision of environmental analyses.

In an interlaboratory test of river water sampling according to serve the aims of the EU WFD (Strub et al., 2009), it was reported that the total variance (analytical as well as sampling) was 49% and 81% for tetrachloroehtylene and di(2-ethylhexyl) phthalate (DEHP), respectively. The analytical variance was 22% and 28%, respectively. Using the same propagation of error as illustrated in the equation above, but only taking analysis and sampling into account, the resulting sampling variance was calculated to

43% and 76%, respectively. This clearly illustrates that sampling variance dominates over sampling variance and that it almost accounts for the total variance.

Groundwater from two different boreholes has been studied to investigate sources of variability with respect to measuring heavy metals and chlorinated solvents (e.g. trichloroethylene, tetrachloroethylene) (Roy and Fouillac, 2009). The result was that 73% of the total variability originated from inter and intra borehole variability (spatial variations, flow patterns), 12% was related to a temporal variability in bacterial activity and as little as 6% was related to the entire analytical process (sampling and analysis).

In fact, when the total sampling error can be up to 1000 times larger than the total analytical error it only adds to the cost of the analysis to decrease the uncertainty associated with the analytical procedure itself (Petersen et al., 2005). Therefore sampling correctness (especially sampling of heterogeneous materials) must never be compromised and only correct mass reduction, i.e. the act of reducing a sample lot into a sample size suitable for the analytical procedures, will ensure representative sampling (Petersen et al., 2005).

2.2.2 Reducing uncertainty

Most literature regarding theory of sampling concerns solid materials, for example Gy (1999), Pitard (1993), Petersen et al. (2005), Petersen et al. (2004), but sampling of liquid samples is also addressed, for example by Pitard (1993), Gy (1999), Holm-Nielsen et al. (2006), Esbensen and Heydorn (2004). Even though it is easier to comprehend the principles when considering solid materials, the concept of sampling theory can be applied to liquid samples as well. The term sampling correctness covers various aspects of the sampling procedure. According to Petersen et al. (2005) it can be summed up in the following 7 sampling unit operations:

Optimising the sampling:

1. Mix well before all further sampling steps
2. Use composite sampling instead of premature focus on the final analytical mass
3. Only use representative mass reduction
4. Homogenise whenever necessary
5. Perform heterogeneity characterisation
6. Perform variographic characterisation of 1-D heterogeneity
7. Whenever possible turn 2-D and 3-D lots into 1-D equivalents

Whereas items 1-4 are strongly recommended to be considered whenever a sampling plan is produced and can be applied to both solid as well as liquid samples, items 5-7 mainly concern solid samples and will not be discussed further in this report.

Re. 1. and 4. It should not be necessary to discuss further that proper homogenisation and mixing is a prerequisite for proper analysis. It is stressed here just to draw the attention to bring that in mind whenever sampling is conducted.

Re. 2. Environmental sampling has the challenges that a sample not only should be representative in time but certainly also in scale. Scale meaning that one sample of for example 100 g or 100 mL is going to represent a huge compartment of either sediment, soil or water. As the general trend is that analytical equipment requires smaller and smaller sample amount/volume (Petersen et al., 2005) there is this schism that the smaller and smaller sample still should represent an enormous environmental compartment. And as the analytical equipment may not require a large sample, one can be tempted to collect a sample just large enough to be satisfactory for the analytical procedure, which to some extent is justified by easier sampling and less sample to be transported (less time spent and less energy spent on transport). However, the pitfall is that the sample is not representative for the sampled environmental compartment. This is especially valid for suspended matter, particles etc. in liquid

samples. Therefore it is a necessity to collect a sample large enough to be representative, which then has to be correctly reduced in mass/volume afterwards. This could for instance lead to sampling 10 L instead of 100 mL.

Additionally regarding the analytical equipment, as the result in principle only shows the quantity of substance actually contributing to the analytical signal metrological quality is a prerequisite for representativeness (Esbensen and Heydorn, 2004).

Re. 3. Mass reduction is to reduce the collected sample from the (environmental) compartment into a representative subsample that is suitable for the analytical procedure. Again in the literature the focus has been on solid material, where a key issue is the likelihood that the sample is split in randomly even size and content. As it often is the case that the sample is reduced with a factor of 1.000 to 100.000 it is not only a question of reducing the weight, but certainly also to obtain an analytical sample representing the content of the collected sample, focal substance(s) as well as particulate matter (Petersen et al., 2004). It is thus of utmost importance to homogenise whenever necessary and possible.

With respect to liquid samples the main concern is the suspended solids or the particulate matter. One way to overcome this issue is of course to filter the sample before further analysis. But then the question is which size fraction has been allowed to pass the filter. And when analysing chemicals capable of sorbing to these particles, the question is how much these particles (with sorbed chemicals) small enough to pass the filter contribute to the final analytical determination of the chemical content of the liquid sample. Sampling techniques (solid phase micro-extraction) measuring the free concentrations have been developed to partly overcome this issue. However, when the laboratory does not have access to such kind of techniques or if the chemical in focus is not analysed by such techniques other means have to be considered.

Going back to the 10 L vs. 100 mL sampling issue, how is it assured that a sample is as homogeneous as possible? It can be imagined, that shaking or stirring 10 L and then drawing a sample of 100 mL is not that feasible. For instance, if the stirring mechanism is traditional magnetic stirring, where one has to lift the flask or beaker from the stirring apparatus in order to pour the sample into a measuring cylinder/flask, suspended solids or particulate matter can easily begin to settle and thus not be homogeneously distributed in the total volume. An alternative method is for instance to use a container equipped with a tap in the bottom, which allows continuous stirring, even when the subsample is drawn from the total volume. Another suggestion could be just to divide the 10 L into ten 1 L-samples which can be shaken or stirred more easily. From these ten 1 L-samples sub samples of 100 mL are drawn and pooled into a 1 L-sample which then can be shaken or stirred and from which a representative 100 mL sample can be drawn.

A concluding remark on reducing the volume of the sample is that one should consider how the sample gets representative, both with respect to the entire compartment, but also with respect to the constituents of the sample, i.e. suspended solids and particulate matter.

A study on various mass reduction methods performed by Petersen et al. (2005) and Petersen et al. (2004) included the grab sampling technique. Various methods to reduce the 2000 g sample size of an artificial sample consisting of glass spheres, rape seed and wheat seed to a 100 g sample were compared and it was shown that the overall representativeness of grab sampling was at least a factor of 50 worse than the best performing mass reduction methods. Though tempting to use due to its ease, grab sampling will not bring reliable results and samples conducted in this way are thus not worth spending time and money on.

This conclusion was also drawn by Hazelton (1998) studying the river water quality with respect to conventional parameters like biological oxygen demand, nitrogen and dissolved oxygen.

2.2.3 Chemometrically assisted sampling

Chemometrically assisted sample composition producing supersaturated experimental design matrices have been used to prepare composite samples for the analysis of PCBs in water samples (Martinez et al., 2002). The strategy is based on only a few really contaminated specimens, compared to the total number of specimens and that the experimental design is build up in an effective manner.

A number of samples are taken and then optimal experiment design is used to predict which samples are contaminated and to what level.

2.3 Comparison of and recommendations for monitoring programmes

At this moment a true comparison among the Danish, the European and the US monitoring programmes is not possible. The reason for this is that the Danish programmes are currently much relying on the guidelines set out by the EU WFD and due to the fact that the US monitoring programmes differ from state to state.

However, comparing with the work performed by Battaglin and Lauren (1996), it seems that the strategy of sampling once per month will give an accurate estimate of the annual mean concentrations. And, when this is established a less comprehensive monitoring campaign can be introduced, for example the ones with 2-4 samples during April to July. This is much in line with the recommendations set out within the EU WFD, i.e. the control and operational surveillances.

Besides the frequencies it is also important to consider the actual sampling, both with respect to location and how the sample is taken.

Both EU wide and US wide it is much up to the individual member state and states, respectively, to select appropriate locations to be monitored. Of course, this relies on the fact that local knowledge is required to select the correct locations. But, as stated by Caeiro et al. (2003), one could benefit from including advanced techniques like GPS to locate the precise locations and to use a GIS framework to store data for later comparison. It could be argued, that based on previous knowledge also a statistical component when selecting sampling locations should be introduced, both spatial and temporal.

Moreover, as shortly touched on in this section, using optimal experimental design to select a few samples that will be able to represent a larger data set could be of benefit in the sense that less samples are needed and thus the expenses will be lower OR a more comprehensive monitoring campaign can be performed.

2.4 Tools for monitoring

A monitoring 'toolbox' for the European WFD including all relevant emerging sampling techniques for water sampling has been published by Allan et al. (2006b). In a later publication they discussed how the tools can be used according to the different monitoring purposes outlined in the WFD (Allan et al., 2006a). Some of the tools are suitable for measuring instant concentrations of chemicals, such as spot sampling, equilibrium passive samplers or sensors, while some are suitable for measuring average concentrations, such as flow-proportional samplers and kinetic passive samplers.

This section will focus on kinetic passive sampling methods for measuring average concentrations of compounds. Special focus will be on the priority substances listed in the WFD.

A range of reviews and a book have been written on passive sampling in general. The more recent reviews are by Namiesnik et al. (2005), Vrana et al. (2005b), Stuer-Lauridsen (2005) and Seethapathy et al. (2008). Stuer-Lauridsen (2005) focuses on the use of the passive samplers in relation to

monitoring for the WFD. The book by Greenwood at al.(2007) includes chapters on the presently most used passive sampling techniques.

2.4.1   Passive samplers

The principle behind passive sampling is that a receiving phase, which has a high affinity for the compounds of interest, is placed in the medium to be sampled; e.g. water. In this chapter, only sampling from water will be considered since the scope of this is sampling for the WFD. The sampler can be with or without a membrane separating the receiving phase and the sampled phase. The passive sampler can then, based on chemical or physical properties of the membrane, receiving phase and analytes, selectively accumulate chemical substances (Stuer-Lauridsen, 2005).

Many different designs of kinetic passive samplers have been described in literature. In this chapter kinetic samplers, which have been used quantitatively in the field and which are presently commercially available (see Appendix 1), will be described in more detail. These passive samplers include: semi permeable membrane devices (SPMDs) used for hydrophobic substances, polar organic chemical integrative samplers (POCISs) used for polar organic substances, diffuse gradient in thin films (DGTs) used to sample metals and Chemcatcher which have different versions for sampling of hydrophobic and polar organic substances as well as metals.

2.4.1.1   *Semi Permeable Membrane Devices (SPMD)*

The passive sampler called SPMD is the most frequently used passive sampler to date. In 1990 the SPMD was first described (Huckins et al., 1990), and literature on this sampler includes a review (Petty, et al., 2000), and a book written by the developers of the SPMD technology (Huckins et al., 2006).

A standard SPMD consists of a polyethylene membrane tube, filled with triolein (Huckins et al., 2006). The surface to volume ratio is high (460 cm^2/ml), which ensures a higher concentration factor in the SPMD than in other passive samplers (Greenwood et al., 2007). A device for deploying SPMDs is seen in Figure 2.

The SPMD sampler is a sampler for hydrophobic compounds, and thus can be used for compounds with $\log K_{ow} > 3$. To pass the membrane the molecules have to be less than 10 Å large (Huckins et al., 1990), and therefore only dissolved molecules will be sampled. Compounds which have been sampled by the SPMD sampler include PAHs (Huckins et al., 1999; Luellen and Shea, 2002), non-polar pesticides such as DDT, Lindane, Dieldrin, Endrin (Alvarez et al., 2008; Vrana and Schuurmann, 2002), PCBs (Booij et al., 2003; Meadows et al., 1998) and other organic compounds such as PBDE, tributyltin and nonylphenol (Harman et al., 2008).

SPMD   POCIS   Chemcatcher
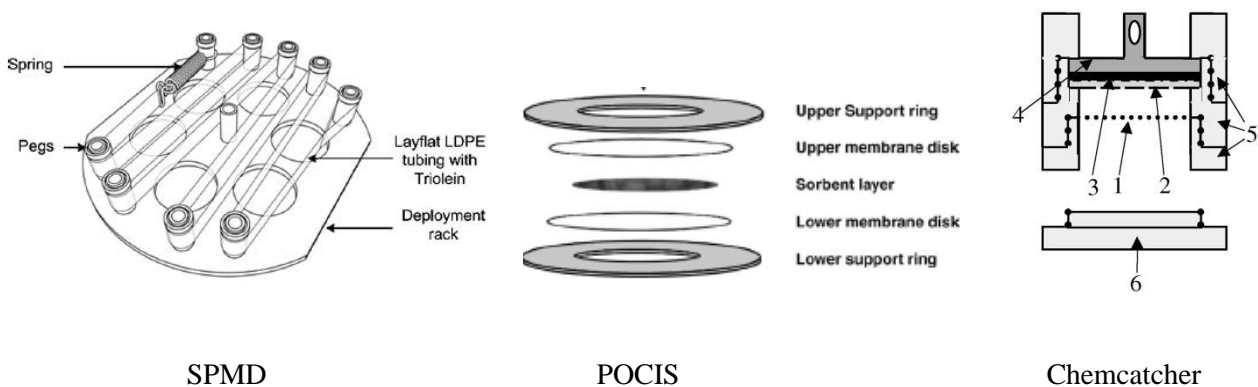
Figure 2. SPMD (Seethapathy et al., 2008), POCIS (Alvarez et al., 2005) and Chemcatcher: PTFE body (5), solid phase (3), rate limiting membrane (2), protective mesh (1), fastening lug (4), screw cap (6)

2.4.1.2 Polar Organic Chemical Integrative Sampler (POCIS)

The POCIS was developed by Alvarez et al. (2004). This sampler was designed especially to capture polar substances, which are not efficiently sampled using e.g. SPMDs.

The POCIS consists of a solid medium sandwiched between two microporous membranes held together by two stainless steel washers (Alvarez et al., 2004), see Figure 2. Two different sorbents are used in the POCIS; one for pharmaceuticals and one for pesticides and other polar organic chemicals (Alvarez et al., 2004), but Mazella et al. (2007) showed that the pharmaceutical configuration can also be used for sampling basic, neutral and acidic herbicides.

POCISs can accumulate a range of polar compounds and has been calibrated in the laboratory for herbicides such as atrazine, simazine, diuron, isoproturon, alachlor, terbuthylazine, lindane (Alvarez et al., 2004; Harman et al., 2008; Mazzella et al., 2007), and nonylphenol and nonylphenolethoxylates were found to accumulate in POCISs (Alvarez et al., 2005; Jones-Lepp et al., 2004; Petty et al., 2004).

2.4.1.3 Chemcatcher

The Chemcatcher passive sampler has been developed by Kingston et al. and was first described in 2000 (Kingston et al., 2000). It consists of a polytetrafluoroethylene (PTFE) body, which holds a solid receiving phase and a rate-limiting membrane (see Figure 2). The sampler is designed so that the choice of membrane and solid receiving phase can vary and depend on the target substance.

Chemcatcher is not a single passive sampler, but a range of samplers with similar housing, because each time sorbent and membrane is changed, properties and even uptake mechanisms are altered. Different versions of the Chemcatcher have been proposed for a broad range of chemicals including polar organic chemicals, non-polar organic chemicals and heavy metals. An overview of the different configurations can be seen in Table 2.

Table 2. Configurations of the Chemcatcher for different target compounds. CA= cellulose acetate, PS=polysulfone, PE=polyethylene, PES=polyethersulfone, LDPE=low density PE

Target compounds	Sorbent	Membrane	Reference
Polar organic substances	C_{18} Empore disk	PS	(Kingston et al., 2000)
Hydrophobic substances	C_{18} Empore disk	PE or LDPE	(Kingston et al., 2000) (Vrana et al., 2005a)
Metals	Chelating disk	Nafion coated CA	(Blom et al., 2002; Persson et al., 2001)
Inorganic mercury	Chelating disk	PES	(Aguilar-Martinez et al., 2008a)
Organotin	C_{18} Empore disk	CA	(Aguilar-Martinez et al., 2008a; Aguilar-Martinez et al., 2008b)
Polar and semi-polar pesticides	SDB-XC Empore disk	-	(Schafer et al., 2008)
Pharmaceuticals	SDB-RPS Empore disk	-	(Vermeirssen et al., 2008)

The Chemcatcher has been calibrated for a range of non-polar priority organic pollutants including PAHs, lindane, endosulfan, pentachloro phenol, hexachlorobenzene and dieldrin (Vrana et al., 2006). Calibration data for tetrabrominated diphenyl ether and hexabrominated diphenyl ether were obtained by de la Cal et al. (2008).

2.4.1.4 Diffuse Gradient in Thin Film (DGT)

The diffuse gradient in thin film technique, DGT, is a technique for measuring speciation of metals. It was first described in 1994 (Davison and Zhang, 1994; Zhang and Davison, 1995). It consists of a pre-filter, a gel layer, typically polyacrylamide hydrogel, and a resin-gel, e.g. Chelex 100 gel, which is held together in a plastic sampling device (Greenwood et al., 2007). The principle of DGT is the same as for the other passive samplers, but instead of a membrane, diffusion is controlled by the diffusive gel layer, which can be varied in thickness.

2.4.1.5 Theory

The mechanism of uptake of analytes by the passive sampler is diffusion. In Figure 3 the diffusion over an aqueous boundary layer, over a membrane and over the receiving phase boundary layer is illustrated. Depending on the turbulence of the water in which the sampler is placed and on the properties of the compound, the diffusion limiting layer can be either the membrane or the aqueous boundary layer. For DGT, the thickness of the diffusive gel is chosen in order to be rate limiting.

A model for the uptake of a substance over time in a sampler is also shown (Figure 3).
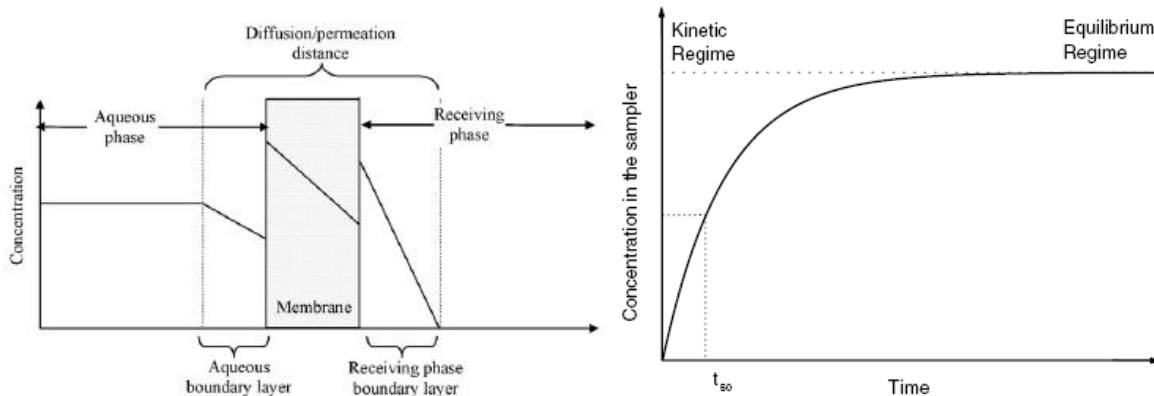
Figure 3. Schematic concentration profile of an analyte in a passive sampler with a membrane (left) (Seethapathy et al., 2008), theoretical concentration of an analyte in a passive sampler over time (right) (Vrana et al., 2005b)

The uptake can be described as an exponential function:

$$C_s = C_w K_{SW}\left[1 - \exp\left(-\frac{R_s t}{K_{SW} V_s}\right)\right] \tag{1}$$

Where C_s is the concentration in the sampler, C_w is the concentration in the water, K_{sw} is the equilibrium constant between the sampler and the water, V_s is the volume of the sampler, R_s is the sampling rate and t is the sampling time (Huckins et al., 1993). The sampling rate, R_s, has to be found

from calibration experiments. The unit is volume per time, e.g. L/d, and it is equivalent to how much water has been 'cleared' by the sampler per unit of exposure time.

For SPMDs, the K_{SW} can be estimated based on an empirical relationship between the $\log K_{ow}$ and the partitioning equilibrium constant between an SPMD and water:

$$\log K_{sw} = a_0 + 2.321 \log K_{ow} - 0.1618 (\log K_{ow})^2 \text{, n = 45, s =0.25, r = 0.97}$$

Where a_0 is -2.61 for PCBs, PAHs, 4,4´-DDE and -3.20 for pesticides such as HCHs, dieldrin, chlorpyrifos, heptachlor and trifluralin (Huckins et al., 2006).

For the hydrophobic Chemcatcher the K_{SW} can be estimated from:

$$\log K_{SW} = 1.382 \cdot \log K_{OW} - 1.77 \log, \text{ n = 31, s = 0.13, r = 0.97) (Greenwood et al., 2007)}$$

In the beginning of the deployment the uptake is almost linear. This area of the curve is called the kinetic sampling regime. The period in which the passive sampler remains in the kinetic sampling regime, $t_{1/2}$, is determined by:

$$t_{1/2} = \frac{0.693 K_{sw} \cdot V_s}{R_S} \tag{2}$$

Where $t_{1/2}$ is the time required to reach 50% of the equilibrium concentration of the sampler (Greenwood et al., 2007).

As both K_{SW} and R_s have influence on the kinetics, the kinetic sampling regime depends on both substance and environmental conditions (e.g. turbulence around the sampler), and thus is different from case to case. For PAHs with $\log K_{OW} > 4.5$, $t_{1/2} \geq 30$ d have been observed, whereas for PAHs with $\log K_{OW} < 4.5$ equilibrium was reached within 15 d; $t_{1/2} \approx 2$ d (Luellen and Shea, 2002). A similar result was shown for naphthalene with a $t_{1/2}$ of 4.5 d and benzo(g,h,i)perylene with a $t_{1/2}$ of 21 d (Ke et al., 2006).

Kinetic sampling can be used to find Time Weighted Average concentrations (TWA) of the freely dissolved substances including any labile complexes, C_W, using the following equation:

$$C_W = \frac{M_s}{R_s \cdot t} \tag{3}$$

Where M_s is the mass of the pollutant accumulated in the sampler. This equation is valid if the deployment time is less than $t_{1/2}$, otherwise the full equation (1) should be used.

A similar equation is used for DGT, taking into account the fact that the thickness of the sampler can be varied:

$$C_w = \frac{M_{DGT} \Delta g}{D_{gel} t A} \tag{4}$$

Where A is the area of gel/filter which is exposed to the water, Δg is the thickness of the filter and gel, D_{gel} is the diffusion coefficient for the compound in the gel, t is the sampling time and M_{DGT} is the mass of metal accumulated in the resin (Davison and Zhang, 1994; Greenwood et al., 2007). In this case the D_{gel} is found in laboratory tests instead of the R_s. For DGT the diffusive boundary layer is neglected, while for the other samplers it is incorporated in the R_s. Applying DGT with two different

gel layer thicknesses can be used to correct for the thickness of the diffusive boundary layer (Warnken et al., 2006).

A fundamental property of passive samplers is that the sampling rate does not depend on the concentration. This has been shown to be true for the different types of passive samplers (Huckins et al., 1993; Huckins et al., 1999; Mazzella et al., 2007; Togola and Budzinski, 2007). Another important property for passive samplers is that sampling rates vary linearly with the surface area of the sampler (Alvarez et al., 2004; Huckins et al., 2006; Zhang et al., 2008). This fact can be used when comparing different types of passive samplers.

There are three main environmental factors which have an influence on the sampling rate and which complicate the interpretation of results from passive samplers: temperature, biofouling and turbulence (Greenwood et al., 2007). First of all, temperature is a factor which has to be taken into account because the diffusion of analytes increases with temperature. Biofouling is important because it can decrease the diffusion, and in some cases even degrade some of the substances (Gunold et al., 2008). Turbulence has an influence on the thickness of the aqueous boundary layer. In most cases the limiting barrier for diffusion is the aqueous boundary layer at stagnant situations. In some cases the membrane becomes the limiting barrier at turbulent flow regimes. The effect of these factors varies depending on the type of passive sampler and the pollutant of interest. Studies which deal with these factors are summarized   in

Table **3** for each passive sampler.

In principle the samplers should be calibrated for each set of conditions, but this is not economically feasible. One way to get around the uncertainties is to use Performance Reference Compounds (PRCs). PRCs are compounds which do not occur in the environment and are added to the sampler before deployment.

Table 3. Studies of the factors influencing the sampling rate, R_s, in the different passive samplers

Sampler	Temperature		Fouling		Velocity	
	ΔT, °C	ΔRs (factor)	Change	ΔRs	Change	ΔRs (factor)
SPMD	10→20	↑ 1-4.5[a]	Lab→ river	↓ 31 %[b]	Laminar→ turbulent	↑ 14[c]
	2→30	↑ 3[c]	Excess fouling	No change[d]	Small laminar change	Not significant[e]
POCIS	5→21	↑ 0-100%[f]	Tendency is low, but can be important[g]		Quiescent → turbulent	↑ 10[g,h,]
Chemcatcher - organics	4→20	↑ 2-10[i]	Not observed in a harbour[i]		Depends on the layer determining the diffusion[i]	
Chemcatcher - metals	4→11	yes[j]	Lab→ river	↓ 10-25%[k]	0→70 rpm	↑ 1-7[j]
	11→18	no[j]				

[a]: (Huckins et al., 2006); [b]: (Lebo et al., 1996); [c]: (Booij et al., 2003); [d]: (Booij et al., 2006); [e]: (Vrana and Schuurmann, 2002); [f]: (Togola and Budzinski, 2007); [g]: (Alvarez et al., 2004); [h]: (Macleod et al., 2007); [i]: (Kingston et al., 2000); [j]: (Aguilar-Martinez et al., 2008b); [k]: (Aguilar-Martinez et al., 2008a)

The off-loading of the PRCs from the sampler is subject to the same environmental conditions as the uptake of analytes. Therefore the release rate of the PRCs can be used to compensate for variations in the conditions under the actual deployment. If sampling rates of the PRCs and the compounds of interest in another exposure are known, then the exposure adjustment factor can be used to adjust the sampling rates of the compounds of interest:

$$\frac{R_{i,1}}{R_{i,2}} = EAF ,$$

where $R_{i,1}$ is the sampling rate of compound i at exposure 1 and $R_{i,2}$ is the sampling rate of compound i at exposure 2.

The technique of using PRCs can only be used when the uptake is determined by partitioning and not when it is determined by sorption (to a solid medium). This is because the uptake and offload kinetics have to be similar, which is not necessarily the case for sorption mechanisms. If the uptake is determined by diffusion over the membrane, this problem can in some cases be overcome as found by Mazzella et al. (2007) for the POCIS sampler and by Tran et al. (2007) for the Chemcatcher. An approach to determine diffusion and to correct for turbulence is to use two passive samplers with different membrane thickness as described in Persson et al. (2001).

Sampling rates are compound specific. Using data from nine studies Huckins et al. (2006) found that the compound specific effect on the sampling rate could be modeled by the equation:

$$\log \alpha_i = 0.013 \log K_{ow}^3 - 0.3173 \log K_{ow}^2 + 2.244 \log K_{ow} \text{ (Huckins et al., 2006)}$$

The standard deviation of the fit corresponds to an uncertainty factor of 1.5. The sampling rate of compound i in exposure j can then be found by:

$$R_{i,j} = R_{PRC,j} \frac{\alpha_i}{\alpha_{PRC}} \text{ (Huckins et al., 2006)}$$

A validation of this empirical equation on other compounds showed that the equation fits well for 7 of 9 compounds (tributyltin, pyrene, benzo(a)pyrene, PCB-77, DDT, PCb-153, PBDE 47) but did not fit well to data from lindane and nonylphenol, which had a much lower sampling rate than predicted (Harman et al., 2008).

Likewise the Chemcatcher has been calibrated for a range of non-polar priority organic pollutants (Vrana et al., 2006) and an empirical model was obtained depending on a correction factor for environmental conditions P_i (Vrana et al., 2007):

Log R_S = P_i + 22.755 log K_{ow} − 4.061 $\log^2 K_{ow}$ + 0.2318 $\log^3 K_{ow}$ (R=0.92, s=0.22, n=134)

This relationship is applicable for compounds with log K_{ow} in the range of 3.7-6.8 (Vrana et al., 2007). The constant P_i can be found from PRCs at each deployment (Vrana et al., 2007). This empirical model is valuable for substances for which no calibration data has been obtained, and has an uncertainty factor of approximately two (Vrana et al., 2007).

Sampling with SPMDs, POCISs and Chamcatchers for organic compounds have mainly been done in rivers and streams (Alvarez et al., 2004; Alvarez et al., 2008; Kolok et al., 2007; Lebo et al., 1996; Matthiessen et al., 2006; Petty et al., 2004; Schafer et al., 2008; Sharpe and Nichols, 2007; Vrana et al., 2007) the sea (Aguilar-Martinez et al., 2008b; Booij et al., 2006; Kingston et al., 2000), and recently in more complex matrices such as wastewater treatment plants (Alvarez et al., 2005; Augulyte and Bergqvist, 2007; Jones-Lepp et al., 2004; Katsoyiannis and Samara, 2007).

Sampling with DGT and Chemcatcher for metals has been done in estuarine and seawater (Dunn, et al., 2003; Dunn, et al., 2007; Munksgaard and Parry, 2003; Warnken et al., 2004), rivers (Allan et al., 2007; Allan et al., 2008; Aung et al., 2008; Cleven et al., 2005; Divis et al., 2007; Roig et al., 2007), and wastewater (Buzier et al., 2006)

2.4.2 Discussion

The requirements in the WFD are based on spot sampling. The WFD sets the frequency for surveillance monitoring of priority pollutants in surface waters to one month unless other choices are justified. This can be sufficient in water bodies with slow variations in concentrations, but especially in dynamic water bodies with high chemical fluctuations or spatial variety of water composition, such as estuaries, rivers or streams, many samples are necessary in order to establish the representative level of pollution (Allan et al., 2006a; Ort and Gujer, 2006). Both flow-proportional samplers and kinetic passive samplers can be used as time-integrative samplers, and in this way reduce the number of samples (and thus the cost) needed in order to establish the level of pollution in water bodies with temporal variations.

For all sampling methods there are advantages and disadvantages related to the practical issues of sampling. Spot sampling is a low-tech flexible method, but since it is manual, there can be time and logistic difficulties. Flow proportional sampling is automated but requires expensive equipment and installation of electricity on site. Passive samplers do not have these drawbacks, but vandalism and loss of samplers have been reported. For both spot sampling and flow-proportional sampling large

volumes of water have to be handled in order to get low detection limits. For passive samplers, a high pre-concentration on site reduces the detection limit. The high pre-concentration in the samplers means that special care should be taken when handling the samplers in order to avoid contamination. Especially for the SPMD sampler, which is a very effective sampler in air, the exposure to air should be minimized. Likewise, the handling of samplers for metal analysis up to the point of deployment should be in ultra-clean lab facilities.

The difference between grab-samples, flow-proportional samples and passive samplers is not only the time-frame of the sampling but also the species which are sampled. For analysis of grab and other active water sampling techniques, the total amount of the compound is typically measured in the sample (unless particles and colloidal material is removed by filtering), while the passive samplers only extract the freely dissolved fraction of the compound. The advantage of this is that it can be more representative of the concentration which is actually available for uptake in flora/fauna than the total concentrations. The disadvantage of obtaining time-integrative concentrations of freely dissolved and labile species is that passive samplers cannot be directly compared to the traditional types of measurements. Current regulations on substances have been based on total concentrations. Regulations have to be reformulated, and quality assurance procedures for passive samplers have to be established before passive samplers can be used for checking compliance with regulations. Instead of reference materials, reference sites or pelletized reference compounds have to be developed (Roig et al., 2007). This means that passive samplers are presently appropriate to use when assessing which pollutants are problematic for a certain water body, to evaluate trends and for pollution tracking, but that there are some barriers before they can be used for legal purposes. Seethapathy et al. (2008) points out that even though there are a range of protocols and regulatory guidelines developed for passive sampling in air, few such protocols exist for passive sampling in water and soil.

When choosing which passive sampler to use, there are many considerations to take into account. Kot et al. (2000) defined the ideal passive monitor as being 'inexpensive, easy to deploy, small enough to be mailed to remote locations, sensitive to analytes and insensitive to interfering matrix components'. First of all obviously, the passive sampler has to be able to sample the compounds of interest (polar, non-polar or metals). For sampling of both hydrophobic compounds, polar compounds and metals, there exists more than one possible passive sampler. Each one has both advantages and disadvantages compared to the others. The degree of description of the sampler in literature and of earlier experience with the sampler can influence the choice, but other more technical features should also be considered.

One of the differences is the sampling rates of the samplers. For example SPMD has much higher sampling rates than Chemcatcher for polar organic compounds, mainly due to the much higher surface area of SPMDs. The advantage of a high sampling rate is that the concentration of compounds in the sampler will then be higher and lead to lower detection limits. On the other hand the sampler will approach equilibrium faster, and therefore the assumption of linear uptake is only valid for a shorter time interval. If long time-intervals are wanted, a sampler with a lower uptake rate is preferable.

Chemcatcher provides a low-cost sampler body which can be mass produced and house different combinations of membranes and receiving phases, which means that configurations for different compounds can be fitted in the same type of housing. It can be easier to handle because the size is smaller and because it is not as efficient for air sampling as the SPMD, which makes contamination during installation less likely. The POCIS sampler is easy to handle because it can be stored dry and does not need to be conditioned (as the Chemcatcher) before deployment but is readily wetted.

The sample extraction and clean-up following deployment is different for the different samplers. For example Vermeirssen et al. (2005) noted that sample preparation of POCIS is much easier than SPMD. The fact that recovery of analytes from the triolein in the SPMD sampler is a very complex procedure (including dialysis, size-exclusion chromatography and further clean-up such as silica gel) using large quantities of organic solvent, is one of the main disadvantages of this sampler (Gorecki

and Namiesnik, 2002). Alternative extraction methods which are being tried include accelerated solvent extraction (ASE) (Wenzel et al., 2004), sonication enhanced extraction (Setkova et al., 2005) and microwave-assisted extraction (Yusa et al., 2006), which can reduce the time and solvent amount needed for the extraction procedure. New research into more green chemistry is also being conducted involving exchanging the toxic solvent triolein with for example ionic liquid (Zhao et al., 2006). A review by Esteve-Turrillas et al. (2008) goes into detail with all the different configurations and handling methods which have been used recently. The Empore disk used in Chemcatcher can be extracted in an ultrasonic bath and needs no further clean up.

In relation to this, the MESCO passive sampler is an interesting type of passive sampler which is under development. It has only recently been described, and is therefore not fully developed and commercialized yet, but has some interesting advantages. It consists of a membrane with a silicone rod or a Twister® stir-bar inside. This method has the advantage that the collector phase can easily be separated from the membrane and loaded directly into the injection port of a GC-MS for thermal desorption so that no further clean-up is necessary. PRCs can also be used in this method (Greenwood et al., 2006).

3 Models

3.1 Using models built in Task 7.4

Dynamic urban scale source-and-flux models were developed. The models are used for quantifying the release of micropollutants from urban sources and the fate of micropollutants within different treatment systems. These models – which can be linked to simple, river basin scale multimedia models used in ecological risk assessment – enable "what-if" scenarios to assess the effect of emission barriers as well as to evaluate their potential in enabling monitoring systems and sampling programmes to be optimised. An integrated, dynamic model is able to predict the dynamic fate of micropollutants and therefore it assesses the compliance with EQSs (annual averages and peak concentrations).

The integrated urban wastewater system (IUWS) at urban scale consists of several coupled entities:

- Urban sub-catchment sources

- Urban surface (of each sub-catchment) where pollutant accumulation takes place, followed by rain-induced wash-off (stormwater flow) and soil infiltration processes

- Main sewer network (including combined sewer overflows (CSOs) and retention tanks) receiving the different stormwater flows generated in each of the sub-catchments and/or the dry weather flows corresponding to the same geographical areas

- Stormwater treatment options accepting stormwater flows from separate sewer systems

- Wastewater treatment systems (including physical, chemical and/or biological units) receiving input from the main sewer network

- Sludge treatment options accepting residues from wastewater treatment systems

- Receiving surface water (e.g. river) accepting direct emissions (each sub-catchment) and sewage water flows from the other IUWS entities

Each part of the IUWS can be represented by a dynamic model, in which the physical, chemical and biological behaviour of the herein present components can be described as function of time by mathematical equations. These unit fate models are explained in detail in the deliverable for Task 7.2 (Vezzaro et al., 2009a).

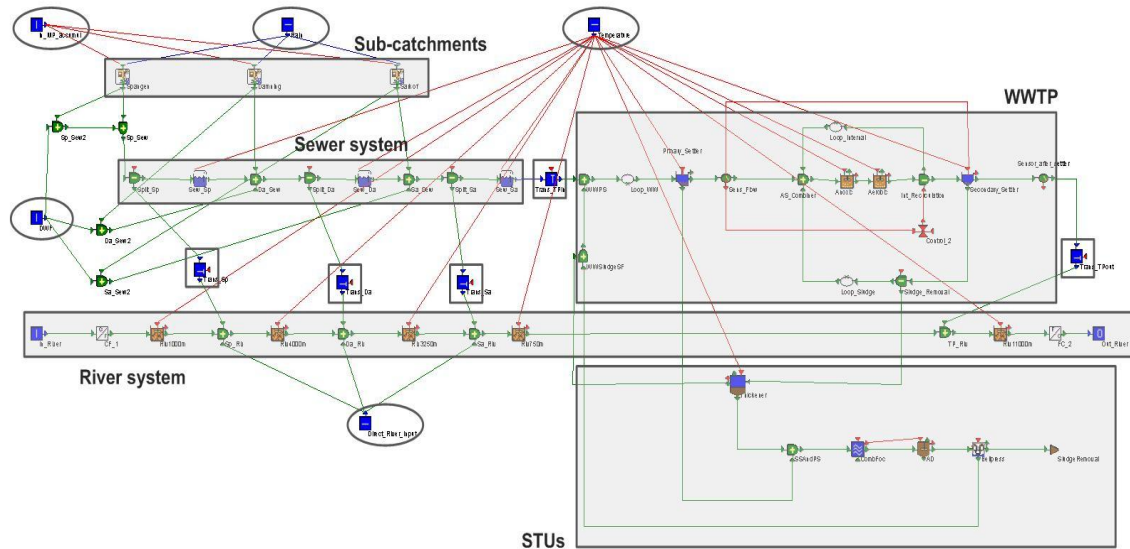In Figure 4 an example of the model configuration of an integrated urban wastewater system in WEST is shown.

Figure 4: Model configuration of IUWS at urban scale in WEST

The **input data** for the integrated model are marked with circles on Figure 4, and include:

- Rainfall [mm/d];

- Priority pollutant mass fluxes [g/d] in the dry weather flow, the direct river discharges and accumulating on the catchment surface, for each sub-catchment;

- River background flow rate [m³/d] and pollutant concentrations [g/m³];

- Temperature [°C].

This information was collected in different columns in a text file and loaded as such in the integrated model (in the first column time [d] is included). More details about the input data is described in section 3.3.

For the sewer catchment and network, the KOSIM_MP model was applied (Solvi, 2007). The model can be divided in two environments where the water is passing through: (a) the (sub-)catchment and (2) the sewer system.

The **sub-catchment** blocks (see Figure 4) are coupled sub-models: (1) potential evaporation sub-model, (2) sub-model for transformation of incoming rainfall into run-off flow (run-off sub-model), (3) sub-model for mimicking flow over the catchment surface and in the local sewer networks, and (4) dry weather flow generator.

The *dry weather flow generator* (De Keyser *et al.*, 2010) generates dry weather flow rates and traditional pollutant (COD, N, P, etc.) charges. Hereby, it (a) uses a mean daily quantity of DWF produced per person per day, (b) applies a daily user-defined flow and pollution pattern (interpolation between hourly values), and (c) takes into account lower flows and pollution during weekends and non-touristic seasons by multiplying them by fixed factors, to generate DWF flow and pollution loads

for each sub-catchment. Furthermore, the generator calculates the quantity of unpolluted drainage water based on a mean flow per total connected area and a yearly user-defined infiltration pattern (interpolation between monthly values). The *run-off sub-model* allows to calculate the amount of rainfall run-off (stormwater) based on catchment area, proportion of impervious and pervious catchment area, and rainfall data (see input data). Pollutants accumulated on the catchment surface during dry weather periods are washed-off with the run-off flow only from impervious surfaces. The pollutant mass accumulating on the catchment surface (in kg/ha/d) can be set at a mean value, defined by the user (the case for the traditional pollutants), or an input time series can be used (the case for the selected priority pollutant).

The **sewer system** in this example (Figure 4) is composed of sewer pipes and combined sewer overflows (CSOs). The sewer model simulates pipe flow and the transport and fate of pollutants within the system. The CSO model is a special kind of splitter, where a part of the incoming flow is transferred over the overflow weir to the receiving water when it reaches a critical value. The CSOs can be replaced by retention tanks which are modelled as storage tanks that overflow when their capacity exceeds a defined critical value. These structures temporarily store sewage and in the meantime also treat the water by allowing coarse particles to settle. In case of separate sewer systems, stormwater can be treated before discharge into the receiving water by stormwater treatment options, which are described by an extension of the Universal Stormwater Treatment Model with a priority pollutant fate sub-model (Vezzaro et al., 2009b).

The **wastewater treatment plant (WWTP)** in this example (Figure 4) is composed of 2 clarifiers and 2 activated sludge units. The clarifiers were modelled with the Takács_MP model; while for the activated sludge units the ASM2_T_MP model (carbon, nitrogen and phosphor removal) was selected. This WWTP layout can be extended with a sand filter as tertiary treatment step, described by an extension of the SimpleDownFlow model with a priority pollutant fate sub-model. Also the addition of iron hydroxide or alum to a mixing tank in order to allow soluble phosphorus to precipitate can be modelled.

A gravity thickener, a mesophilic anaerobic digester and a dewatering unit (e.g. belt filter press) are grouped in the **sludge treatment units (STUs)** (Figure 4). The selected models for these units are the extended sludge thickening model with a priority pollutant fate sub-model, Siegrist_MP AD model and the extended sludge dewatering model with a priority pollutant fate sub-model, respectively.

The **river system** is represented by a sequence of river stretches (Figure 4), each described by the RWQM1s1_MP model, which includes carbon, nitrogen and phosphor removal and algae dynamics.

The **transformers** are marked with squares on Figure 4. Connector models (Deliverable 7.4 (Benedetti et al., 2009)) were built to link the sewer system (KOSIM_MP state variables) and the wastewater treatment plant (ASM2_T_MP state variables), the sewer system (KOSIM_MP state variables) and the river system (RWQM1s1_MP state variables), and the wastewater treatment plant (ASM_T_MP state variables) and the river system (RWQM1s1_MP state variables).

The transformers to connect the ASM2_T_MP state variables (gravity thickener) to the Siegrist_MP AD state variables (anaerobic digester) and the Siegrist_MP AD state variables (anaerobic digester) to the ASM2_T_MP state variables (dewatering unit), are implemented within the Siegrist_MP AD model itself, causing no transformer blocks to be visible to link the anaerobic digester to the 2 other sludge treatment units.

Priority pollutants can have the tendency to occur in more than one environmental compartment outside the urban water cycle: air, groundwater, etc. Interactions with these 'external' compartments can affect some processes within the urban area. For this, a multimedia model at river basin scale was developed within the ScorePP project and implemented in WEST (D7.3, De Keyser et al., 2008b). The unit fate models at urban scale (D7.2, Vezzaro et al., 2009a) can be connected to this multimedia

model. The direct input of pollutants from sources within the urban area towards the 'external' compartments (e.g. air) is provided by the dynamic model input generator (D7.1, De Keyser et al., 2008a). The linking between the two models and the development of the models for representing the multimedia compartments are discussed in detail in the Deliverable for Task 7.3 (D7.3, De Keyser et al., 2008b).

3.2 Modelling sampling devices

In order to test monitoring strategies, models of the most useful sampling devices have been coded in MSL to be included in the WEST model library. This way, such devices can be incorporated in the models of the case studies.

The following devices have been included in the model library:

1. Grab sampler

2. Flow-proportional sampler

3. Time-proportional sampler

4. Flow-proportional "time-based" sampler

5. Passive sampler – kinetic (see Chapter 2.2)

6. Passive sampler – equilibrium (see Chapter 2.2)

All models of the above devices have a set of common features:

- Analysis accuracy and sampling uncertainty, both defined by the mean and by the maximum variation (defined as the interval including $+\backslash-3$ times the standard deviation) of the Gaussian noise added to the measurement.

- Total cost during the monitoring campaign, calculated as unit cost of each measurement (data point produced) multiplied by the number of data points.

- Output both as concentration and as load, both with and without addition of noise due to measurement uncertainty. The quality of the load estimation depends on the sampling device, i.e. with a flow proportional sampler the load estimation is better than with a grab sampler.

The **grab sampler** measures the concentration at time points defined by a sampling frequency. A detection limit is applied.

The **flow-proportional sampler** model requires inputting the detection limit, the volume of each bottle in the sampler (e.g. 2l), the minimum flow above which sampling starts, the volume of each sample (e.g. 100ml) and the volume of water between each sample (e.g. 1000m^3).

The **flow-proportional "time-based" sampler** is the same as the flow-proportional sampler except for the fact that there is only one bottle which never fills up completely, so that the composite sample is not taken when a bottle is full but with a given frequency (e.g. daily).

The **time-proportional sampler** is similar to the flow-proportional sampler except for the sampling time, which is not driven by the volume of water between each sample, but by a sampling frequency.

The **passive samplers** have been implemented as described in Chapter 2.2. They both include uncertainty (noise) on the sampling rate. In particular, as the common equation to derive the concentration in the sampler is:

$$\frac{dCs}{dt} = \frac{Rs}{Vs} \cdot \frac{Cw - Cs}{Ksw}$$

the specific behaviour of the **kinetic** sampler is given by:

$$Cw = \frac{Ms}{Rs \cdot t}$$

and of the **equilibrium** sampler by:

$$Cw = \frac{Cs}{Ksw}$$

3.3 Input data to the models

First, information about *rainfall* is necessary. The time series must be as long as the modeller wants to simulate and as detailed as the preferred time step of the input file. This data is then used by the KOSIM_MP model to calculate run-off from the catchment surface.

Secondly, the *sources* of the selected *priority pollutant* (soluble and particulate form) in each sub-catchment should be known. For each source, the release dynamics, the receiving compartments and the release factors can be retrieved from the ScorePP emission string database (Holten Lützhøft et al., In preparation). The combination of this generic information with case-specific data (e.g. number of inhabitants), is used by the model input generator (D7.1, De Keyser et al., 2008a) to create a time series for releases to the different receiving compartments. For the integrated model, the quantity released on the urban surface (U), expressed in kg/ha/d; the amount present in the dry weather flow (WI), expressed in g/d; and the direct river discharges (WD), expressed in g/d; should be known.

Also knowledge about the *physico-chemical properties* of the *priority pollutant* is necessary, e.g. molecular weight, biodegradation rate in each sub-system of the IUWS, etc. This information can partly be retrieved from the database developed in Task 3.1 (*Compilation of priority pollutant inherent properties*, (Holten Lützhøft et al., 2008).

The *river background concentration* of each component and *flow rate* is also necessary to collect before modelling.

Last, the *characteristics of the different systems of the IUWS*: temperature at air/water interface and in the system, length, width, volume, porosity, etc. (depending on the system).

4   Substance flow analysis (SFA)

4.1 Introduction on empirical and model-based SFA

4.1.1   Empirical SFA

A Substance Flow Analysis (SFA) is an analysis of the flows of a substance or substance group, including inflow, emissions and outflow, within a defined system (usually a geographic region) during a specified time (usually a year) (Jonsson et al., 2008). Sometimes a SFA also includes stock and environmental distribution.

The inflow describes how much of the substance enters the system. A substance may enter a city via different routes, for example with raw materials to industries within the city, with building materials to construction sites and with goods to consumers. Production of the substance within the city is another form of inflow.

If the products and materials in which the substance enters the city has a service life which is longer than the period which the analysis is based on, the substance will accumulate in the system. The stock describes the accumulated amount present in the system. Sometimes one application of the substance may be represented in the stock, but not in the inflow. This could be because the use of the substance for this application has been phased out, but the material or products in which the substance was used has a long service life, and therefore some of the amounts accumulated in the system before the phase-out are still there. There may also be applications with very short service lives, for which there is no stock, only inflow and outflow.

The outflow of the substance consists of several fluxes: products may be exported from production facilities within the city or they may leave the city as waste, or the substance may be emitted to soil, water or air during different stages of the product's life-cycle. Also processes like combustion, wear and tear of road pavement etc, result in releases that are part of the outflow of the substance. The relative contributions from different applications of the substance may be different in the outflow than in the stock or the inflow. One explanation to this is that different applications will be exposed differently, and hence different applications emit different amounts of the substance to the environment.

The routes and environmental distribution of the releases depend very much on what type of sources that are relevant for the substance and on the chemical properties of the substance. For example industrial emissions of a volatile substance will be emitted to air while industrial emission of a water-soluble substance will be emitted to surface water via wastewater, similary, emissions from products and/or materials used indoors will most probably enter the wastewater system while emissions from products and/or materials used outdoors may be distributed to soil or stormwater.

4.1.2 Model-based SFA

An SFA based on the models developed and implemented in Tasks 7.2 (*Unit process models for priority pollutant emission control measures*), 7.3 and 7.4 consist of an analysis performed on a detailed integrated model of the urban water system (IUWS). This allows logging data at any input/output of the IUWS and of its components, at any sampling frequency. Therefore, the model-based SFA estimates the real substance flows as close as the model gets to reality.

With the models of sampling devices described in chapter 3.2 it is possible to evaluate the performance of monitoring strategies designed to produce an SFA of the IUWS.

A monitoring strategy is implemented as the combination of a number of sampling devices that produce data at a number of sampling points in the IUWS, with a given sampling frequency. A monitoring strategy can be used to produce an SFA of the system.

A model-based SFA obtained with data logged at any unit input and output in the IUWS at high frequency (e.g. 15 minutes sampling interval), one can assume that to be the "real" SFA of the IUWS.

The quality of a monitoring strategy can be assessed as the closeness of its SFA to the "real" SFA, e.g. measured as the average square difference for each substance flow in the system.

4.2 Methodology for model-based SFA

In an integrated urban water system (IUWS) model, all pollutant mass fluxes from one compartment to another are calculated during a simulation. The challenge is thus not to estimate the fluxes, but to retrieve the needed information from the simulation software platform and translate it into an SFA – typically presented in a kind of flow charts or so-called Sankey diagrams.

In this chapter, a methodology is presented to perform this data transformation semi-automatically. The principles are illustrated on a simple example (see Figure 5), i.e. an IUWS model layout consisting of only one catchment, one sewer pipe, a basic wastewater treatment plant (primary settling, anoxic tank, aerated tank, secondary settling) and a river modelled as two completely mixed compartments as the receiving surface water. The mathematical models developed in Task 7.2 and implemented in Tasks 7.2 and 7.4 were configured to simulate the fate of DEHP as an example substance.
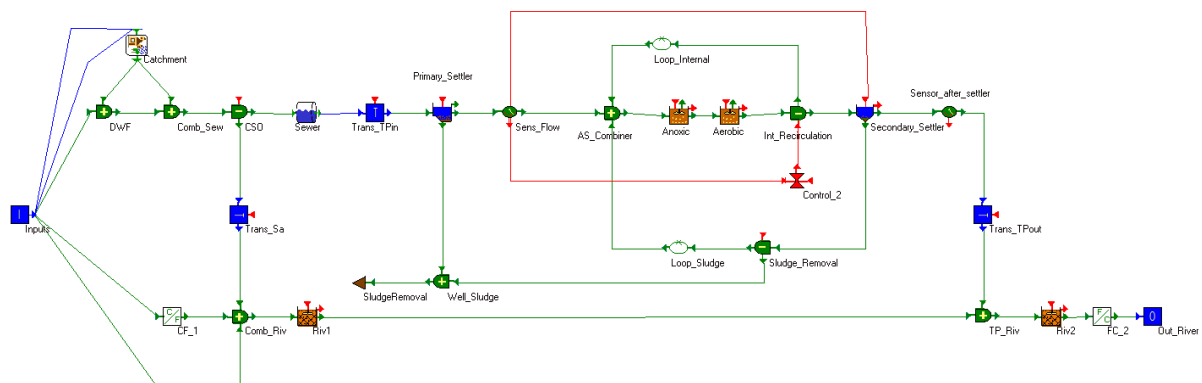

Figure 5. An overview of the simple IUWS model layout as it was implemented in WEST

WEST simulations provide fully dynamic mass flux data, but in SFA, the mean fluxes (e.g. over one year) are typically studied. Multiple possibilities exist to obtain averaged data. The first one is to write the desired variable values to an output file and to post-process the file (e.g. with software like Matlab or Excel) to calculate the column averages. This approach however has some disadvantages. For example, to obtain reliable results, the temporal resolution should be fine enough not to miss short dynamic events like peaks. As a result, the ouput file length and the time needed to process it increase considerably. Further, the column index in the output file might change when additional output variables are selected. This is certainly an issue with simulations run in Tornado, due to alphabetical sorting of the output variables.

A second approach is to have the mean values calculated in the model itself during the simulation. Like this, no output file is needed, which shortens the simulation duration, and all simulated dynamics are incorporated in the calculated mean. Post-processing is still necessary, as Sankey diagrams can not be created in WEST, but it is less computationally demanding (see below).

To calculate the means of certain variables in WEST, a new model class was introduced in the modelbase: AverageCalculatorMP. The node accepts input in four possible formats: WWTPTerminal, SewTerminal, RivTerminal and Real numbers. From the three vector inputs, only the water flowrate, S_MP and X_MP elements are used. The mean values are calculated with the help of an integral:

$$\bar{x} = \frac{\int_{t_0}^{t} dx}{t - t_0} \qquad (5)$$

At each moment t during the simulation, the running average of the input variables is calculated, expressed in m³·d⁻¹ (water flow rate) or g·d⁻¹ (S_MP and X_MP).

A drawback of this approach is that the necessary mass fluxes need to be linked to the AverageCalculatorMP node, most probably in the WEST configuration editor, causing a number of additional links in the layout (Figure 6). If desired, these can be hidden. A small modification of the node library was necessary to allow for more than one link to certain input or output ports.
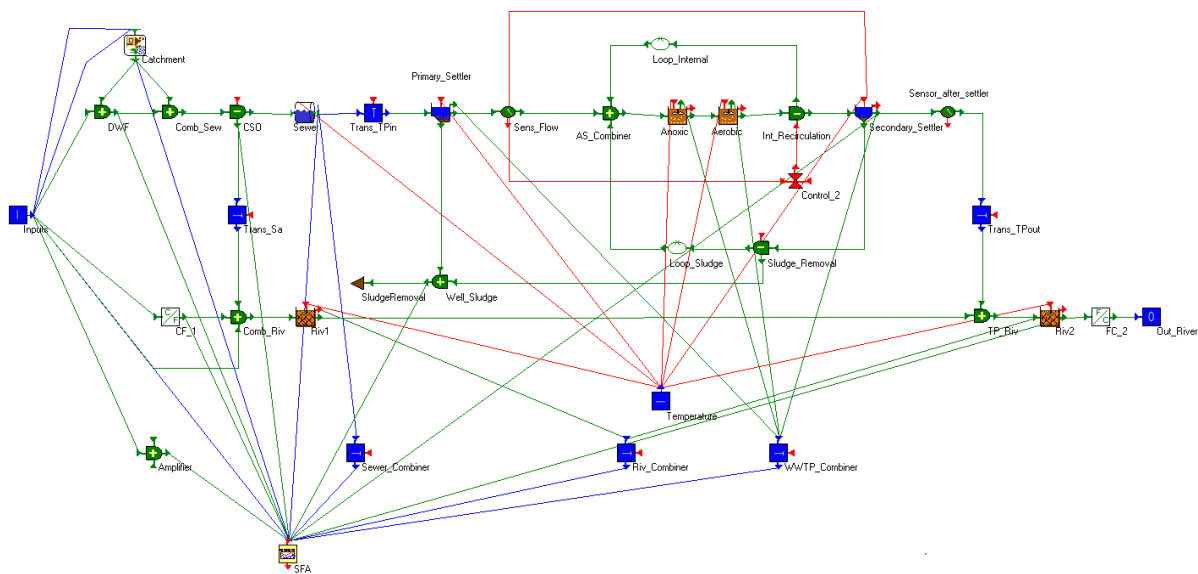

Figure 6. The necessary mass flux variables need to be connected to the average calculator block

The experimenter needs to keep a manual list of which mass fluxes are connected to which input acceptor of the average calculator node, since these input variables have generic names such as Input_RivTerminal_x, Input_SewTerminal_x, Input_WWTPTerminal_x and Input_Real_x (with x a number from 01 to 10).

At the end of the simulation (e.g. 365 days), the average mass fluxes can be retrieved from the average calculator node (Figure 7) and be saved to an Excel file. To plot an SFA chart (Sankey diagram) from these values, a software tool called e!sankey was used (http://www.e-sankey.com). Once the sources, sinks and fluxes are drawn, the thickness of each individual arrow (mass flux) can be linked to a value in the Excel file containing the averages calculated in WEST. When these values are updated in Excel, e.g. because new simulation results become available, the arrow thicknesses in e!sankey are updated as well.

The methodology to obtain the model-based SFA chart (Figure 8) can be summarized as follows:

1. Extending the WEST model configuration with an AverageCalculatorMP class node and connecting all needed mass fluxes to this node.
2. Running the simulation for the desired time.
3. Exporting the state variables values to an Excel file.

4. Creating an SFA diagram in e!sankey with all sources, sinks and fluxes.
5. Copying the mass flux quantities from Excel and pasting them as arrow thickness values in e!sankey to create a permanent automatic link.
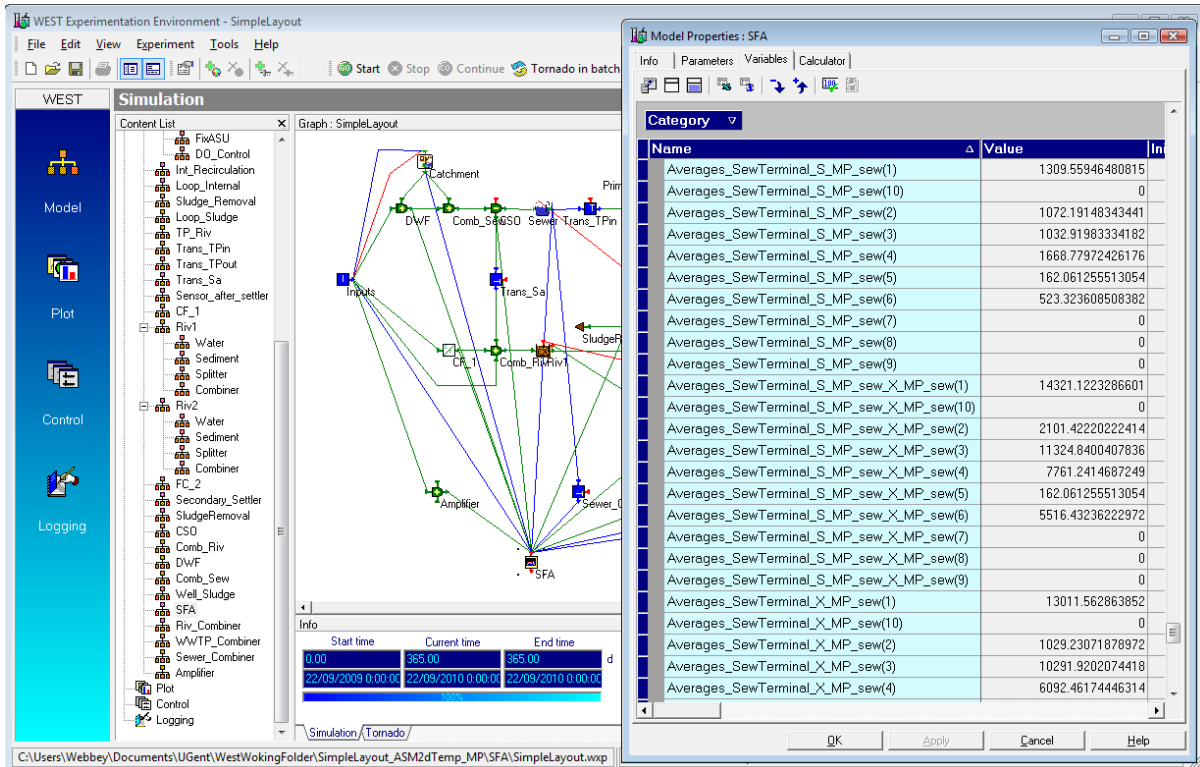
Figure 7. At the end of the simulation, the calculated average mass fluxes can be saved to an Excel file
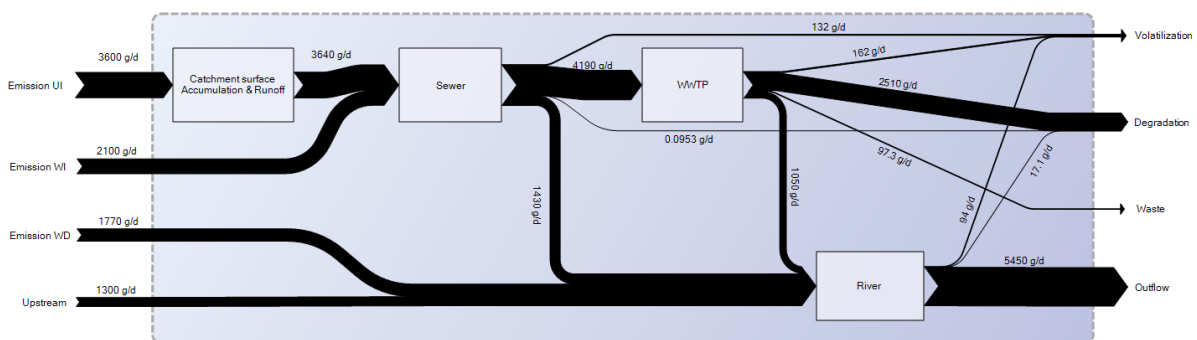
Figure 8. Sankey diagram representing the priority pollutant flows in the modelled urban system

The visual representation of the model-based SFA allows for assessing both the model implementation and the real system. For example, it can easily be noticed from the Sankey diagram if the mass balance over a subsystem is unclosed. Possible causes for this are that the model implementation contains an error, or for example that the pollutant mass assumed to be initially present in the subsystem is unrealistically low or high relative to the mass fluxes during the simulation, causing a large pollutant accumulation or washout.

Another model anomaly that could easily be detected with the help of the SFA representation is an unrealistic ratio of certain mass fluxes. For example, if the average pollutant flux from sewer to river (via occasional sewer overflow events) is large compared to the flux from sewer to WWTP, this might indicate a malfunctioning buffer tank, an incorrectly designed sewer system, or maybe an unrealistic frequency of heavy rainfall events. On the other hand, if the model proves to be correct, model-based SFA allows for much to be learnt about the bottlenecks in the real system. In the example of Figure 8, it can be noticed that almost one quarter of the mass flux entering the sewer system is conveyed directly to the river through a combined sewer overflow (CSO). Studying the simulation output learns that a lot of rain events indeed cause CSO events and that this storm water indeed contains quite a high priority pollutant concentration.

5   Optimal experiment design

In the ScorePP project, modelling is used as a tool to assess the impact of different source control options on the environmental quality of urban water systems. The various unit process models developed in Task 7.2 and the integrated urban wastewater system model developed in Task 7.4 need to be calibrated to some possible extent. Data retrieved from case studies can help in this calibration process. Once a reliable model is available for a certain urban area, it can be used in combination with measuring strategy models (chapter 3) to determine when, where and how the measurements of a measurement campaign should be performed to yield the most informative data on the efficiency of implemented source control options.

This chapter first focuses on the automated model-based design of measurement campaigns for model calibration. In the second part of the chapter, the concept of using measuring strategy models and SFA to design measurement campaigns for emission barrier efficiency monitoring is explained.

5.1   Optimal experiment design for parameter estimation

This section discusses how to design a measurement campaign so that it yields as much valuable information as possible on the model parameters to be calibrated. This discipline is called Optimal Experiment Design for Parameter Estimation (OEDPE) and is not completely new. It arose first in the analysis of experiments in agriculture and its first steps can be traced back to the beginning of the 20th century. Especially the work of Fisher in the 1930's formed the basis for further developments. Initially most attention was given to experimental design for linear regression models. Later, also non-linear models which had an analytical solution were examined. Only since the 1980's and thanks to the development of computer science, optimal experiment design was performed on models which could only be solved numerically, like the complicated dynamic environmental models used in this WP.

In order to design a measurement campaign (commonly referred to as "an experiment"), different choices have to be made. It has to be decided whether, where and how the system under study will be manipulated and where, how and when measurements will be performed on this system. The procedures discussed in this chapter provide a quantitative evaluation method to assess the information content of a proposed experiment with respect to the model parameters one wants to calibrate. The search for the best experiment can be automated in optimization algorithms, leading to automated optimal experiment design.

First, the mathematical background for OEDPE is summarized in chapter 5.1.1. Next, chapters 5.1.2 and 5.1.3 describe how the existing modelling and simulation software WEST®/Tornado was used as

the basis for a software tool to assess the information content of proposed experiments. Finally, a basic example (chapter 5.1.4) illustrates the use of the tool.

5.1.1   Mathematical background

Note: this section was based on Munack (1991), Vanrolleghem *et al.* (1995), Petersen (2000), Dochain and Vanrolleghem (2001), De Pauw (2005) and Franceschini and Macchietto (2008).

5.1.1.1   *From objective function to Fischer Information Matrix*

The aim of calibration of a model is to find a certain parameter set θ that has an optimal value so that the agreement between the measured data and the model prediction is as good as possible. To quantify the deviation of the model from the data, an objective function needs to be specified. A frequently used function is the least squares or sum of squares objective function given by Equation (6). The objective $J(\theta)$ is the sum over all N data points of the squared difference between the model prediction $y_i(\theta)$ and the measurement y_i.

$$J(\theta) = \sum_{i=1}^{N} \left(y_i - y_i(\theta) \right)^2 \tag{6}$$

Note that the N data points can be measurements of one variable at N different time instances, or measurements of N variables at one moment in time, or a combination of both (Figure 9).
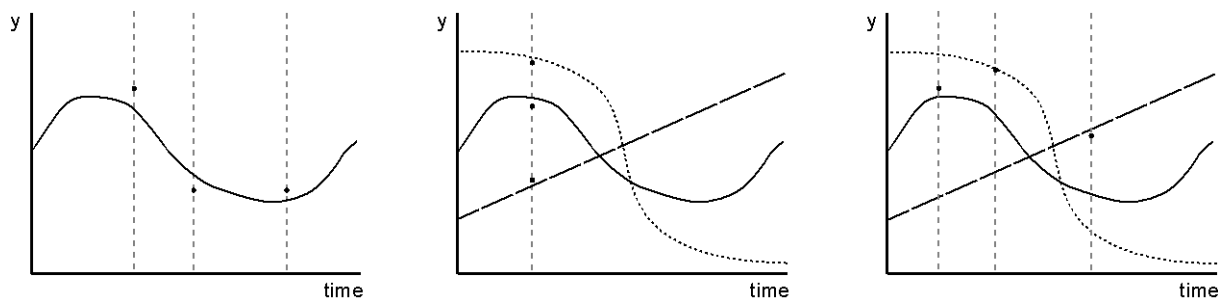
Figure 9. The data points can either be chosen as N different measurements of one variable (left), or as a simultaneous measurement of N different variables (centre), or as a set of N (variable, time) pairs (right).

If parameters are chosen close to the true model parameters (which are unknown), the value of $J(\theta)$ will tend to become small. The parameters of the model thus need to be adjusted in order to minimize the objective function. Up to now, the objective function does not take into account measurement error or noise. It is a well-known fact that all data is corrupted by measurement error to a certain extent which means that a model will never perfectly fit to the measurements even if the model structure is correct. This also means that uncertainty is associated with each measurement point an. Measurements which are accompanied by large measurement errors can be regarded as uncertain, while measurements with small measurement errors are more to be trusted. This notion of uncertainty can be added to the parameter estimation objective function resulting in a weighted least squares objective function:

$$J(\theta) = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \left(y_i - y_i(\theta) \right)^2 \tag{7}$$

In this objective function σ_i^2 is the variance (square of the standard deviation σ_i) associated with the measurement errors and $1/\sigma_i^2$ acts as a weight in order to compensate for the possible uncertainty on the measurements. Small measurement error variances correspond with large weights which will amplify the difference between model and measurement and result in significant contributions to the final objective function value. On the other hand, deviations from the model predictions for measurements with a large variance (large measurement error) will contribute less to the overall objective function because of smaller weights. As will be outlined further, also other, user-specified weighting factors can be used instead of the variance.

The sum of squares can also be written as a multiplication of two vectors, causing the weights also to be written in matrix format:

$$J(\theta) = \left(y - y(\theta)\right)^{\mathrm{T}} \cdot Q^{-1} \cdot \left(y - y(\theta)\right) \tag{8}$$

With $(y - y(\theta))$ a column vector of size N containing the elements $y_1 - y_1(\theta)$ to $y_N - y_N(\theta)$ and Q^{-1} a diagonal square N×N matrix containing the weighting factors $1/\sigma_i^2$ on the diagonal.

The expected value of the objective function for a parameter set slightly different from the optimal one can be written as:

$$E[J(\theta + \delta\theta)] = \left(y - y(\theta + \delta\theta)\right)^{\mathrm{T}} \cdot Q^{-1} \cdot \left(y - y(\theta + \delta\theta)\right) \tag{9}$$

Introducing a linearization of the model with respect to the parameters:

$$y_i(\theta + \delta\theta) \approx y_i(\theta) + \left[\frac{\partial y_i(\theta)}{\partial \theta}\right]_\theta \cdot \delta\theta \tag{10}$$

leads to the following expression of the expected value:

$$E[J(\theta + \delta\theta)] = E[J(\theta)] + \delta\theta^T \cdot \left[\left(\frac{\partial y(\theta)}{\partial \theta}\right)^{\mathrm{T}} \cdot Q^{-1} \cdot \left(\frac{\partial y(\theta)}{\partial \theta}\right)\right] \cdot \delta\theta \tag{11}$$

In order to get a clearly defined minimum for $J(\theta)$ it is necessary that the difference between $J(\theta)$ and $J(\theta+\delta\theta)$ is maximized. If this is the case, the minimum will be located in a deep pit rather than a small depression on a flat objective function plane. In practice the maximization of this difference means that the model fit of a non-optimal parameter set is significantly worse, even though the parameter set is only slightly different from the optimal parameter set. Maximizing the difference between $J(\theta)$ and $J(\theta+\delta\theta)$ can be achieved by maximizing the last term between square brackets in Equation (11). This matrix is known as the Fisher Information Matrix or FIM.

$$FIM = \left(\frac{\partial y(\theta)}{\partial \theta}\right)^{\mathrm{T}} \cdot Q^{-1} \cdot \left(\frac{\partial y(\theta)}{\partial \theta}\right) \tag{12}$$

The FIM quantifies the information content of an experiment related to the model parameters. The effect of the FIM maximization on the parameter estimation objective function surface is illustrated in Figure 10 for a single parameter estimation problem. This figure shows the objective function surfaces for a parameter estimation based on data of two experiments, a non-optimal experiment (outer curve) and an optimal experiment (inner curve). The parameter estimation based on the data of the optimal experiment (after FIM maximization) results in a much narrower objective function profile and a smaller confidence interval.

Note that the FIM is sometimes written as

$$FIM = \sum_{i=1}^{N} \left[\left(\frac{\partial y_i(\theta)}{\partial \theta} \right)^{\mathrm{T}} \cdot Q_i^{-1} \cdot \left(\frac{\partial y_i(\theta)}{\partial \theta} \right) \right] \tag{13}$$

In this notation, it is assumed that the matrices within the square brackets contain information about one or multiple variables at one certain time instance *i*. The summation is necessary to integrate the information over N time instances. This also implies that all variables should be measured at all N moments. The notation used in Equations (8) to (12) avoids this problem and thus allows for a flexible combination of measured variables and measurement times. As a consequence, the matrices grow larger, but this has no significant influence on the calculation times compared to summing the N smaller matrices.

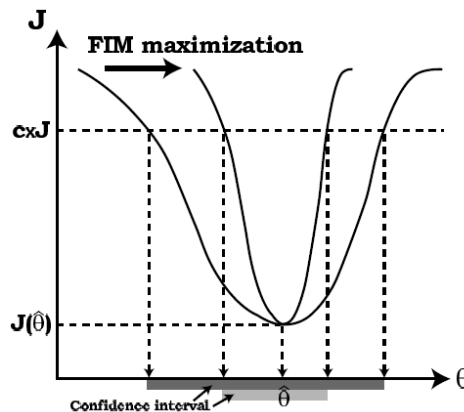**Figure 10. Influence of the FIM maximization on the objective function surface J(θ) and confidence intervals for a single parameter estimation problem; the objective function cut-off value for constructing the confidence interval is represented by cₓJ**

5.1.1.2 *Sensitivity matrix*

The FIM is clearly function of two matrices. The first one is the sensitivity matrix:

$$\left(\frac{\partial y(\theta)}{\partial \theta}\right) = \begin{pmatrix} \dfrac{\partial y_1(\theta)}{\partial \theta_1} & \dfrac{\partial y_1(\theta)}{\partial \theta_2} & \cdots & \dfrac{\partial y_1(\theta)}{\partial \theta_p} \\ \dfrac{\partial y_2(\theta)}{\partial \theta_1} & \dfrac{\partial y_2(\theta)}{\partial \theta_2} & \cdots & \dfrac{\partial y_2(\theta)}{\partial \theta_p} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial y_N(\theta)}{\partial \theta_1} & \dfrac{\partial y_N(\theta)}{\partial \theta_2} & \cdots & \dfrac{\partial y_N(\theta)}{\partial \theta_p} \end{pmatrix} \tag{14}$$

in which N is the number of measurement data and p is the number of parameters to be calibrated. Each element represents the local sensitivity of variable y_i to parameter θ_j and should be determined by performing a sensitivity analysis with the model.

In a sensitivity analysis, the dependence of model predictions on the parameter values is quantified. Figure 11 illustrates this for a simple logistic growth model consisting of one equation:

$$W(W_0, W_f, \mu, t) = \frac{W_0 W_f}{W_0 + (W_f - W_0)e^{-\mu t}} \tag{15}$$

In this simple model, the yield of biomass W is function of the initial biomass W_0, the maximum achievable biomass W_f, a growth constant μ and the time t. The left graph in Figure 11 shows the biomass yield in function of time in case $W_f = 0.1$ t.ha^{-1} and $W_f = 10$ t.ha^{-1}. In the base scenario (solid line), the parameter μ is estimated to be 0.07 d^{-1}. To obtain the sensitivity $\partial W/\partial \mu$, the simulation is run a second time, but with a slightly perturbed parameter value $\mu + \Delta\mu$ (dashed line). The resulting 'forward' sensitivity function, calculated according to Equation (16), is shown on the right (dashed line).

$$\frac{\partial W}{\partial \mu} \approx \frac{W(W_0, W_f, \mu + \Delta\mu, t) - W(W_0, W_f, \mu, t)}{\Delta\mu} \tag{16}$$

The same can be done with a parameter value slightly smaller than the original one ($\mu - \Delta\mu$), resulting in a 'backward' sensitivity (dotted line). It is clear from the right plot in Figure 11 that both sensitivities are similar but not equal. Therefore, also a 'central' sensitivity can be defined as the average of the forward and backward sensitivies (solid line).
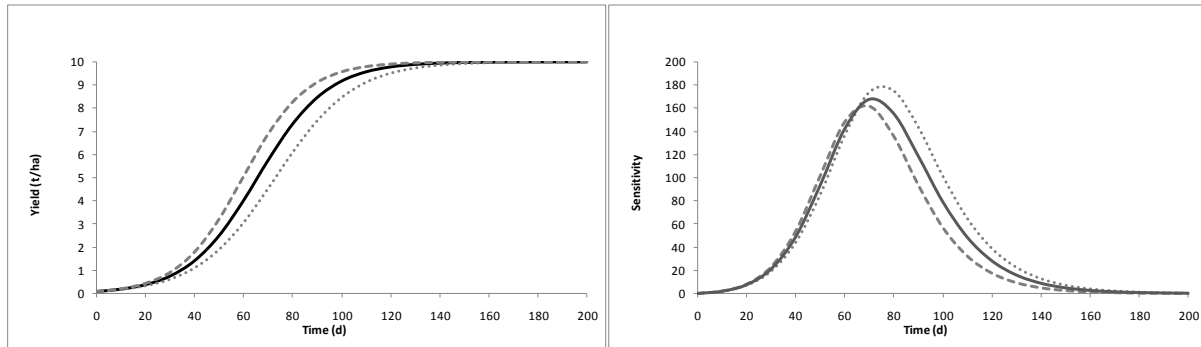
Figure 11. Sensitivity analysis of a simple logistic growth model; left: yield as a function of time for a primary estimation of the parameter *μ* (solid line), a forward perturbation (dashed line) and a backward perturbation (dotted line); right: forward (dashed line), backward (dotted line) and central (solid line) sensitivity functions ∂W/∂μ

For each of the p parameters, such a sensitivity analysis is performed. This means that in total 1+2p simulations have to be run: a base scenario and two runs with perturbed parameter values for each parameter. The elements of the matrix in Equation (14) are obtained from the resulting sensitivity functions.

In theory, when the perturbation factor approaches zero, the backward and forward sensitivities converge. Numerically, this can never be achieved because of the limited precision of the calculations. Taking the perturbation factor too small will result in numerical inaccuracies. On the other hand, choosing the perturbation factor too large will also influence the sensitivity calculations due to the nonlinearity of the model. Optimal perturbation factors are parameter dependent. Consequently, applying a fixed perturbation factor to all parameters is not advisable. De Pauw and Vanrolleghem (2006) discussed the usage of quality criteria for the selection of an appropriate perturbation factor.

5.1.1.3 Covariance Matrix

The second matrix appearing in the FIM is the inverse of the measurement error variance-covariance matrix:

$$
Q^{-1} = \begin{pmatrix}
\sigma^2_{y_1} & \sigma^2_{y_1y_2} & \cdots & \sigma^2_{y_1y_N} \\
\sigma^2_{y_2y_1} & \sigma^2_{y_2} & \cdots & \sigma^2_{y_2y_N} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma^2_{y_Ny_1} & \sigma^2_{y_Ny_2} & \cdots & \sigma^2_{y_N}
\end{pmatrix}^{-1}
\tag{17}
$$

As explained above, the elements on the diagonal ($1/\sigma_i^2$ after inversion) are weighting factors to reduce the contribution of uncertain measurements in the FIM. If the measurements of the different variables are correlated, this can also have a negative influence on the experiment's information content. For example, if two variables are always measured on one grab sample, they will be more correlated than if they are measured completely independently. Therefore the covariance between measurements (σ^2_{yiyj}) can be taken into account as an additional weighting factor.

The variance on a type of measurement (σ_i^2) and the covariance between two types of measurements (σ^2_{yiyj}) can be determined in a dedicated experiment, before the actual measuring campaign takes place (Petersen, 2000 (p. 184)). In case the measurement noise is white (i.e. normally distributed random errors with zero mean), the variance can be estimated unbiased with Equation (18):

$$\sigma_i^2 \approx VAR(y_i) = \frac{\sum_{a=1}^{N}(y_{i,a} - \overline{y_i})^2}{N-1} \tag{18}$$

With *VAR(y$_i$)* the sample variance of *N* measurements of the variable *y$_i$* and $\bar{y}_i$ the mean of the sample. The covariance σ^2_{yiyj} can be estimated unbiased as

$$\sigma_{y_i y_j}^2 \approx COV(y_i y_j) = \frac{\sum_{a=1}^{N}(y_{i,a} - \overline{y_i})(y_{j,a} - \overline{y_j})}{N-2} \tag{19}$$

With *COV(y$_i$y$_j$)* the covariance between the two samples (*N* measurements of variable *y$_i$* and *N* measurements of variable *y$_i$*).

Note that only measurement errors are incorporated in the covariance matrix, no (possible) errors in the model structure. This approach thus can only be used to design experiments that could provide a better estimate for model parameters in case the model structure is known to be correct. In case this is unknown, techniques like Optimal Experiment Design for Model Descrimination could be used, but these are not discussed here (see e.g. Donckels, 2009).

5.1.1.4   *Relevance of the FIM*

The approximation of the objective function in Equation (11) allows drawing lines of constant objective functional *J* values in the parameter space, and the delimited regions give confidence regions around the best parameter estimates for different confidence levels. In case a two-parameter problem is addressed, these lines form ellipses (Figure 12). The orientation of the ellipse axes is given by the eigenvectors of the FIM, and their lengths are proportional to the square root of the inverse of the corresponding eigenvalues λ. It can be proven that if the weighting matrix Q^{-1} is chosen as the inverse of the measurement error covariance matrix, the inverse of the FIM is the parameter estimation error covariance matrix V:

$$\mathrm{V} = \mathrm{FIM}^{-1} = \begin{pmatrix} \sigma_{\theta_1}^2 & \sigma_{\theta_1\theta_2}^2 & \cdots & \sigma_{\theta_1\theta_N}^2 \\ \sigma_{\theta_2\theta_1}^2 & \sigma_{\theta_2}^2 & \cdots & \sigma_{\theta_2\theta_N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\theta_N\theta_1}^2 & \sigma_{\theta_N\theta_2}^2 & \cdots & \sigma_{\theta_N}^2 \end{pmatrix} \tag{20}$$

The diagonal elements of this matrix are the variances (σ_θ^2) of the errors of the parameter estimates and the off-diagonal elements are the covariances between the parameter estimation errors.
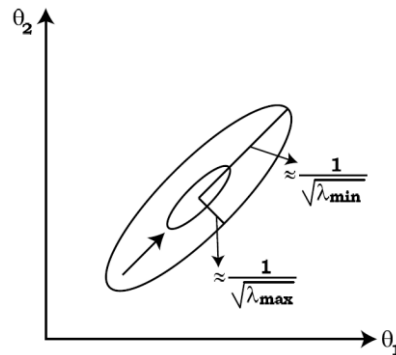
Figure 12. Confidence ellipses in a two-dimensional parameter space

The point of OEDPE is to manipulate these confidence regions so that the parameter estimation errors become smaller, or the parameter estimates more certain. Different approaches to achieve this goal are presented below.

The D-optimal criterion: max(Det(FIM))

First, the volume of the confidence region could be reduced (indicated by the arrow in Figure 12) by maximizing the product of the FIM's eigenvalues, which is the same as maximizing the determinant. This is called the D criterion. By maximising Det(FIM) one minimises the volume of the confidence ellipsoids, and, correspondingly, one minimises the geometric average of the parameter errors. Moreover, D-optimal experiments possess the property of being invariant with respect to any rescaling of the parameters, as was outlined in Petersen (2000).

The A-optimal criterion: $min(Tr(V)) = min(Tr(FIM^{-1}))$

In this criterion the focus is on a minimisation of the trace of the inverse FIM, and thereby the sum of eigenvalues, i.e. the squares of the lengths of the axes of the confidence ellipsoids, of the inverse FIM. This is equivalent to minimisation of the arithmetic average of the parameter errors. Note that this criterion is based on an inversion of the FIM. Thus, this may lead to numerical problems in case the FIM is ill-conditioned, i.e. close to singular.

The Modified A-optimal criterion: max(Tr(FIM))

This criterion is similar to the A criterion, only is the trace of FIM maximised. The problem with this criterion is, however, that in case an unidentifiable experiment is conducted, i.e. a case where FIM becomes singular and one of the eigenvalues becomes zero, which means that the confidence region goes to infinite in a certain direction, the trace may still be optimised and the problem of unidentifiability will not be noticed (Goodwin and Payne, 1977). This is less of a problem with the A criterion since an inversion will not be possible if the FIM becomes singular and the problem of unidentifiability will thereby be exposed.

The E-optimal criterion: $max \lambda_{min}(\lambda_{min}(FIM))$

The E criterion maximises the smallest eigenvalue of the FIM, λ_{min}, and thereby minimises the length of the largest axis of the confidence ellipsoids ($\lambda_{min}^{1/2}$). Thus, this design aims at minimising the largest parameter error and thereby at maximising the distance from the singular, unidentifiable, case.

The Modified E-optimal criterion: $min(\lambda_{max}(FIM)/\lambda_{min}(FIM))$

This criterion is also related to the shape of the confidence region. Here, the focus is on a minimisation of the condition number, i.e. the ratio between the largest and the smallest eigenvalue. The minimum

of this ratio is 1 indicating the case where the shape of the confidence ellipsoids is a (hyper)sphere. The ratio $\lambda_{max}(FIM)/\lambda_{min}(FIM)$ expresses the stiffness of the FIM. The more important the stiffness becomes, the more problematic it becomes numerically to invert FIM until finally a singular matrix is obtained ($\lambda_{min}(FIM)=0$) and the information content becomes zero, i.e. the ratio is infinite.

5.1.2   Iterative optimal experiment design procedure

Performing optimal experiment design in practice involves a series of steps summarized in Figure 13. A first and essential requirement of the experiment design procedure is the availability of a preliminary model. This model may have been calibrated on the basis of data of an initial experiment or default or literature parameter values can be assumed. Another important assumption in order for the design procedure to produce useful results is the validity of the model structure. The model structure should be an adequate representation of the system under study.
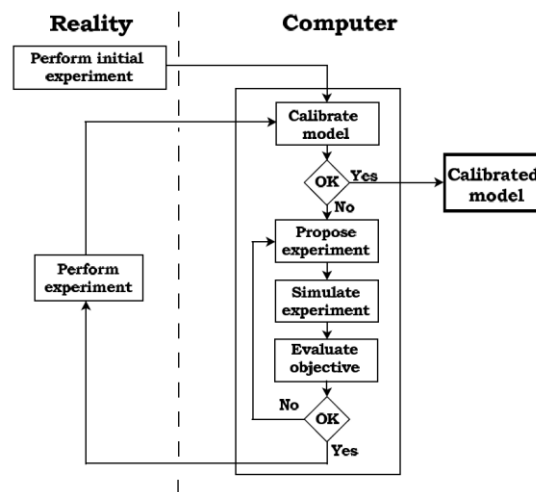

Figure 13. Schematic representation of the OEDPE procedure (Dochain and Vanrolleghem, 2001)

Essentially, the algorithm performing the design performs "simulated experiments" to quantify the potential effect of proposed experimental conditions on the objective function. An optimization algorithm proposes new experiments taking into account the experimental constraints. When an "optimal" experiment is found, it is applied in real life to yield new data for a more precise model calibration.

The experimental degrees of freedom the algorithm has to take into account can be subdivided into two categories: measurements and manipulations. Several questions can be posed concerning the measurements that will be performed during the experiment: (1) what to measure, (2) where to measure and (3) when to measure? The second category of experimental degrees of freedom is the available manipulations. These relate to the excitation signals that will act on the system in order to produce highly qualitative information. Examples of manipulations could be flow rate profiles, influent concentration profiles, etc. In order for the experimental degrees of freedom to be practically useful, constraints for each experimental degree of freedom should be defined. This will allow the experiment design procedure to come up with experiments which can be performed in practice.

To evaluate and optimise the potential information content of a proposed experiment, it needs to be expressed as a scalar quantity of the FIM. As discussed above, there are five criteria that can be used for this. The D-optimal criterion is used most because it aims at a reduction of the total confidence region, the FIM does not need to be inverted and there is no dependency on the units of the parameters.

5.1.3 Software implementation

The OED procedure depends on several important software entities: (1) the simulation platform in which the model is implemented, (2) a sensitivity analysis algorithm, (3) the calculation of the FIM and (one of the) criteria, and (4) an optimization algorithm.

5.1.3.1 The WEST/Tornado framework

Within the ScorePP project, the WEST® software (MOSTforWATER, Kortrijk, Belgium) is used as modelling and simulation platform. In the 2009 version of the software, the powerful Tornado back-end kernel is embedded, ensuring compatibility between model configurations built in earlier WEST versions and the C/C++ based Tornado algorithms.

At the time the work for this task started, this latest version of the WEST software was not available yet, but the source code of Tornado was available for internal use at BIOMATH. Like this, Tornado components like sensitivity analysis and optimization algorithms were available for the development of an OEDPE software tool – which was not included in Tornado so far.

The structure and functioning of Tornado is discussed thoroughly in Claeys (2008), and can be summarized as shown in Figure 14.
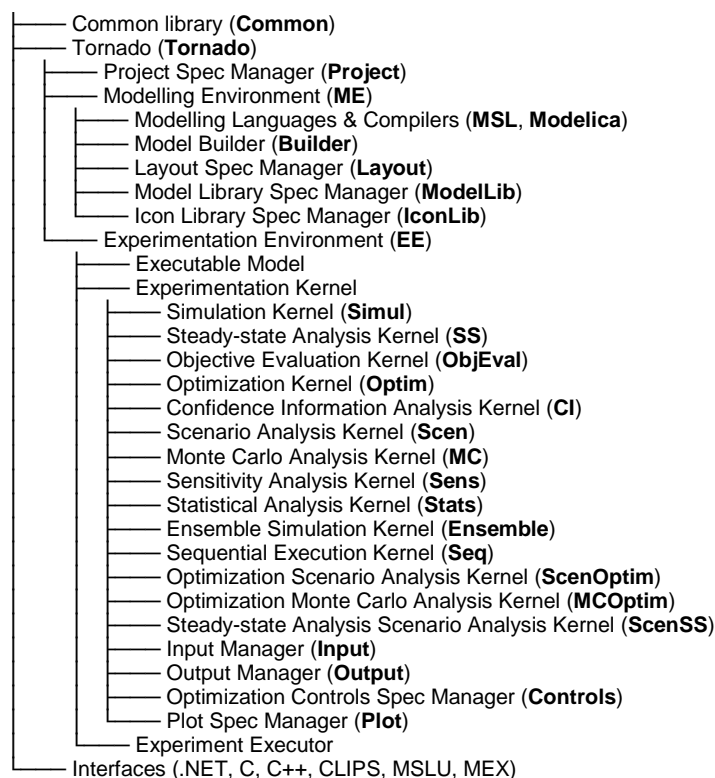

Figure 14. Structure of Tornado, the generic back-end behind WEST

The coherence between the parts of Tornado relevant for this task can be described as follows:

- First models need to be written, e.g. in the WEST MSL editor, resulting in a modelbase. In the case of WEST, a specific layout – a kind of supermodel of the whole modelled system – is composed in the WEST configuration editor.

- Next, the WEST configuration file is converted to MSL (wco2t command), and together with the MSL modelbase converted (with the tmsl command) into a c file and a symbolic model (SymbModel.xml file).

- The model builder creates a dll-file (tbuild command) containing the complete model, to serve as plug-in library for the Experimentation Kernel.

- A Simulation Experiment contains all the necessary information to run a simulation with the model (initial conditions, parameter values, input file location, etc). Either such an experiment is set up in the WEST experimentation environment and then transferred to Tornado (with the wxp2t command), or it is directly created in a Simul.Exp.xml file.

- The now available simulation experiment can be embedded into a sensitivity analysis experiment. The sensitivity analysis kernel performs reference and perturbed simulations and provides the sensitivity functions as an output.

- To perform OEDPE, the FIM and the different criteria need to be calculated. As mentioned in the introduction to this section, this feature was not available in Tornado.

- If FIM calculation would be available, one of the criteria could be used as the objective for the optimization kernel.

5.1.3.2 *Extension for experimental design*

It is clear that the only missing software link in the complete OEDPE procedure is the module that creates the FIM and calculates (one of) the criteria based on this FIM. Although this may sound as a minor task, it is a challenging one. First of all, because the structure of Tornado is advanced and quite complicated, second because the different kernels which are to be involved (simulation, sensitivity analysis, optimization) are no stand-alone modules which can easily be reused: in fact, they are wrapped around each other, so the new FIM module can not just be plugged in, but should be created as an intermediate shell in between the mentioned kernels.

The OEDPE tool was conceived as an application that first gathers the information needed to perform the OEDPE and then launches an optimization experiment in Tornado. The objective, whose value is iteratively reduced in the optimization experiment, is a custom function that returns the value of a criterion using the information provided by the FIM class. The FIM class is a newly implemented class that relies on a sensitivity analysis to get the values needed in building the sensitivity matrix. The functioning of the whole OEDPE procedure is summarized in Figure 15 and is explained in the sub-sections below.
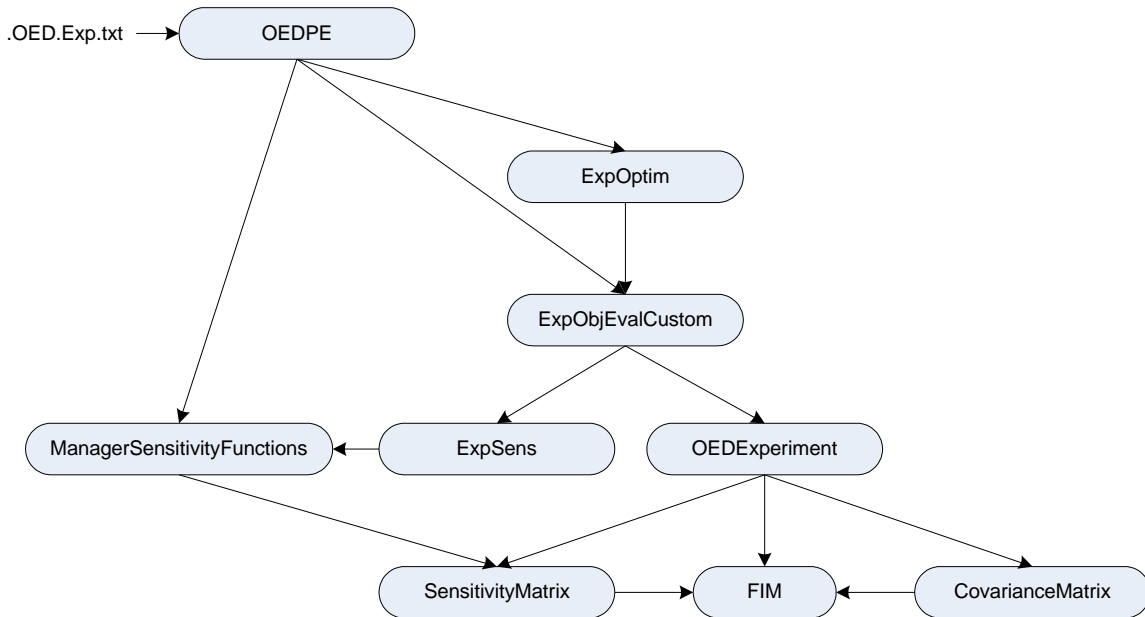
Figure 15. Chart representing the main actors in the OEDPE procedure and the links between them

After adding the folder containing the OEDPE.exe application to the PATH environment variable, the application can be executed from a command line prompt by typing OEDPE. This should be done in the folder that contains the tornado simulation experiment (<basename>.Simul.Exp.xml), the model (<basename>.dll, <basename>.ExecModel.xml, <basename>.SymbModel.xml) and the input file (.txt). The only argument needed is the name of the experiment settings file <basename>.OED.Exp.txt:

```
C:\MyPath> OEDPE MyModel.OED.Exp.txt
```

Loading the OED/PE experiment settings

When the application OEDPE is executed, it first loads the OED experiment settings from the file <basename>.OED.Exp.txt, which should also be located in the same folder. The settings file should have the same structure as the example shown in Figure 16.

```
# lines starting with the # symbol are comments and will not be processed


# ========================================================
ExperimentSettings
# ========================================================
#Name                                    Value
ObjName                                  D
CovarianceMatrixMethod                   specs
CovarianceMatrixPath                     C:\
Sensitivity                              CAS
SensitivityFunctionInterval              0.01
PertFactor                               0.0005
```

```
FallbackPertFactor                              0.05
SolveOptimMethod                                GA
SolveOptimProps.Accuracy                        0.01
SolveOptimProps.PopulationSize                  100
SolveOptimProps.ReplacementPercentage           60
SolveOptimProps.CrossoverPercentage             50
SolveOptimProps.MaxNoEvals                      10000
SolveOptimProps.CrossoverAlpha                  2
SolveOptimProps.Seed                            123456
SolveOptimProps.PrintLevel                      3


# ==========================================================
Parameters
# ==========================================================
#Name                                           Dummy Value
FirstParameterToBeEstimated                     1
SecondParameterToBeEstimated                    1


# ==========================================================
MeasurableVariables
# ==========================================================
# Remarks:
# 1/ put equal values for LowerBound, UpperBound and InitialValue in case the measurement
#    time should not be optimized
# 2/ use a comma separated list of var names for VarName in case the optimiser can choose
#    the variable to be measured
# 3/ set StepSize to 0 for continuous vars


#VarName      LowerBound   UpperBound   Variance      InitalValue     Autoscale     StepSize
X1            0            100          1.85          50              false         0.1
X2,X3,X4      30           30           1.32          50              false         1
X5            0            100          2.04          50              false         0


# ==========================================================
ManipulableQuantities
# ==========================================================
# Remarks:
# 1/ can be left blank if there are no quantities (parameters, initial conditions)
#    to be optimised


#Quantity     LowerBound   UpperBound   InitialValue   StepSize
P1            5.6          6.1          5.8            0.01
```

Figure 16. An example of a *.OED.Exp.txt file containing the settings for the OEDPE experiment

It contains four sections, titled ExperimentSettings, Parameters, MeasurableVariables and ManipulableQuantities. The first section (ExperimentSettings) contains a series of properties in the format "<PropertyName> tab <Value>". The following properties should be listed:

- ObjName: name of the criterion used in the objective function (choose between "A", "modA", "D", "E" and "modE").

- CovarianceMatrixMethod and CovarianceMatrixPath: specifying how the content of the variance-covariance matrix is provided. There are three possibilities: either (1) the complete variance-covariance matrix is provided by the user in a text file (in this case CovarianceMatrixMethod should be set to the value "file" and CovarianceMatrixPath should point to the complete path of this file), or (2) the variance-corvariance matrix is the unit matrix (set CovarianceMatrixMethod to the value "um"), or (3) the variance-covariance matrix is composed by the application according to the "variance" values in the MeasurableVariables section (set CovarianceMatrixMethod to the value "specs"). Note that the first option is the only one allowing to take into account covariances between the variables. The other options yield a variance-covariance matrix with only the diagonal elements (variances) different from zero.

- Sensitivity: the sensitivity function applied in the sensitivity analysis. Choose between the sensitivity functions available in the Tornado Sensitivity Analysis experiment type. Two common used possibilities are "CAS" (central absolute sensitivity) and "CRS" (central relative sensitivity).

- SensitivityFunctionInterval: the Interval property in the Tornado Sensitivity Analysis experiment type (interval between the points sampled from the sensitivity functions when they are stored in the ManagerSensitivityFunctions.

- PertFactor and FallbackPertFactor: the PertFactor (perturbation factor) and FallbackPertFactor property in the Tornado Sensitivity Analysis experiment.

- SolveOptimMethod: the optimization solver to be used in the Tornado Optimization experiment. Choose between the optimization solver plug-ins that are available in Tornado. A commonly used one is "GA" (genetic algorithm), next to "Simplex", "Praxis" and "SA" (simulated annealing).

- SolveOptimProps.*: the properties needed to configure the optimization solver (depending on the choice of SolveOptimMethod).

The second section of the OEDPE settings file contains a list of all the parameters to be estimated (the actual aim of OEDPE). The parameter names should be full names, but without the name of the base model <basename> e.g. ASU1.Kla (not .MyModel.ASU1.Kla). The dummy value in the second column (tab delimited) has no function but should be there because of the way the file is read.

The third section contains the list of possible measurements. A measurement is the combination of a variable and a time instance. Each line of this section contains the following (tab delimited) data:

- VariableName: either the name of the variable that has to be measured, or a comma separated list of variables that can be measured. In this second case, the optimiser can try all possible combinations. This method can be applied to optimise the choice of e.g. sampling locations. Note that the variable names should be full names but without the name of the base model <basename> (e.g. ASU.M(S_O), not .Mymodel.ASU.M(S_O)).

- LowerBound and UpperBound: the boundaries of the time interval (in days) during which the variable can be measured. In case only the name of the variable should be optimised and the measurement time is fixed, lowerbound and upperbound should both be set at the measurement time.

- Variance: the variance on the measurement. This value will be used to compose the variance-covariance matrix in case the property CovarianceMatrixMethod is set to "specs".

- InitialValue: the measurement time (in days) initially used in the optimization. This number should be $\geq$ LowerBound and $\leq$ UpperBound. The genetic algorithm optimization solver does not use this property, so in case the SolveOptimMethod is set to "GA", the InitialValue is not important (however a dummy value should be there because of the way the settings file is read and processed).

- Autoscale: a setting used in the optimization algorithm ("true" or "false")

- Stepsize: the time between two possible measurements in case the chosen optimization algorithm can handle discrete variables. Put this value to 0 to treat it as a continuous variable. Note that the GA optimization algorithm plug-in included in Tornado 0.36 is only able to use continuous variables. An adapted version of the plug-in also capable of handling discrete variables was created and is included in the OEDPE package.

The fourth and last section of the settings file is called ManipulableQuantities and contains a list of quantities (parameters or initial values of derived state variables) that can be altered by the experimenter. Starting the experiment from different initial conditions or performing the experiment with different parameter values, can yield a more informative experiment. The optimizer will search for the best value. The properties LowerBound, UpperBound, InitialValue and Stepsize (in tab delimited columns) have the same meaning as in the MeasurableVariables section.

Preparing the OED/PE experiment

After the OED/PE settings file has been read and processed, the OEDPE application creates an object to keep track of the sensitivity functions (ManagerSensitivityFunctions), an objective to minimize (ExpObjEvalCustom) and an optimization experiment (ExpOptim). Next, the optimization experiment ExpOptim is run. For each set of manipulable quantities, measurable variables and measurement times the optimization solver (e.g. GA) proposes, the objective ExpObjEvalCustom is evaluated. If the sensitivity functions needed to build the sensitivity matrix are not in the ManagerSensitivityFunctions yet, a sensitivity analysis experiment (ExpSens) is launched. Next, an OEDExperiment object is created and evaluated. This involves the creation of a SensitivityMatrix, CovarianceMatrix and FIM object. The SensitivityMatrix is composed with the elements from the sensitivity functions (kept in ManagerSensitivityFunctions) and the CovarianceMatrix is composed according to the settings file (unit matrix, import from a user-defined text file or composition based on the variable-specific variances). The SensitivityMatrix and CovarianceMatrix objects are used to calculate the FIM and the criteria. Finally, the desired criterion value is passed on to the optimization solver via the OEDExperiment and the ExpObjEvalCustom.

5.1.4 Example

As an example to illustrate the OED/PE procedure and the use of the software tool, a simple fermentor model is considered:

$$\frac{dS}{dt} = \frac{Q}{V} \cdot (S_{in} - S) - \frac{1}{Y} \cdot \mu_{max} \cdot \frac{S}{S + K_S} \cdot X \tag{21}$$

$$\frac{dX}{dt} = -\frac{Q}{V} \cdot X + \mu_{max} \cdot \frac{S}{S + K_S} \cdot X \tag{22}$$

This model represents a reactor with a volume V (l) where substrate S (g/l) is consumed by biomass X (g/l) according to Monod kinetics: μ_{max} is the maximal specific growth rate (h^{-1}), K_S is a saturation constant (g/l) and Y is the yield coefficient. The reactor is fed with an influent stream with flow rate Q (l/h) and substrate concentration S_{in} (g/l). The simulation results are also determined by the initial conditions (initial substrate concentration S_0 (g/l) initial biomass concentration X_0 X_0 (g/l) in the reactor).

The model was implemented in Modelica (Figure 17) and converted to Tornado with the "`mof2t Fermentor.mof`" and "`tbuild Fermentor`" commands, resulting in a Fermentor.c, Fermentor.SymbModel.xml and Fermentor.dll file. An empty experiment (Fermentor.Simul.Exp.xml) is created with the "`tmain ExpCreateSimul`" command. The initial conditions can be set in the Quantities section of the Fermentor.Simul.Exp.xml file:

```
<Quantities>
    <Quantity Name=".X" Value="0.00005"/>
    <Quantity Name=".S" Value="0.02"/>
</Quantities>
```

```
fclass Fermentor

  parameter Real V = 20;
  parameter Real Y = 0.67;
  parameter Real mumax = 0.4;
  parameter Real Ks = 0.015;


  parameter Real Q = 2;
  parameter Real Sin = 0.02;


  Real S;
  Real X;


  output Real yS;
  output Real yX;


equation

  der(S) = Q * (Sin - S) / V - mumax * S * X / (Y * (S + Ks));
  der(X) = - Q * X / V + mumax * S * X / (S + Ks);


  yS = S;
  yX = X;


end Fermentor;
```

Figure 17. Implementation of the fermentor model in Modelica (file Fermentor.mof)

The input variables are *Q* and S_{in} but were supposed to be constant and were therefore defined as parameters. In case the model uses input variables specified in an input file, this should be set in the Inputs section of the Simul.Exp.xml file. Also the start and stop times of the experiment should be specified. The stop time in this example was set to 100 hours.

The aim of the OED exercise is to create the most informative experiment in order to calibrate the parameters μ_{max} and K_S. The values 0.4 and 0.015, specified in the modelica model, are default values taken from literature as a starting point. Possible degrees of freedom to optimize the information content of the experiment are the initial substrate and biomass concentrations S_0 and X_0, as well as the choice of measured variable (*S* or *X*) and measurement time, varying from 0 to 100 hours.

The OED/PE experiment is set up in a Fermentor.OED.Exp.txt file based on the example in Figure 16, with the Parameters section containing mumax and Ks, the MeasurableVariables section containing as many lines as there are measurements to be done (in this example five measurements of which two should be substrate measurements and three can be either substrate or biomass) and the ManipulableQuantities section containing the initial substrate concentration (Figure 18).

```
# ===========================================================
MeasurableVariables
# ===========================================================


#VarName        LowerBound      UpperBound      Variance        InitalValue     Autoscale       StepSize
yS              0               100             1               0               false           1
yS              0               100             1               0               false           1
yS,yX           0               100             1               0               false           1
yS,yX           0               100             1               0               false           1
yS,yX           0               100             1               0               false           1


# ===========================================================
ManipulableQuantities
# ===========================================================


#Quantity       LowerBound      UpperBound      InitialValue    StepSize
S               0.010           0.050           0.020           0.010
```

Figure 18. The MeasurableVariables and ManipulableQuantities section in the Fermentor.OED.Exp.txt file

Now the experiment is ready to start the OED/PE procedure. In a command window, the "`OEDPE Fermentor.OED.Exp.txt`" command launches the application. After reading the experiment specifications and before launching the optimization experiment, the optimization problem is translated into a set of nine degrees of freedom (Table 4): the initial value of S, the three measurements "yS,yX" where a variable is to be chosen (given the internal names Set1, Set2 and Set3) and the five sampling times (given the prefix "Var#." to obtain unique variable names).

Table 4. The nine degrees of freedom manipulated in the optimization experiment

Name	Meaning	Lowerbound	Upperbound	Interval
.S	Initial value S	0.01	0.05	0.01
Set1	Choice of variable 3rd measurement	0	1	0 = yS, 1 = yX
Set2	Choice of variable 4th measurement	0	1	0 = yS, 1 = yX
Set3	Choice of variable 5th measurement	0	1	0 = yS, 1 = yX
Var1.yS	Time 1st measurement	0	100	1
Var2.yS	Time 2nd measurement	0	100	1
Var3.yS,yX	Time 3rd measurement	0	100	1
Var4.yS,yX	Time 4th measurement	0	100	1
Var5.yS,yX	Time 5th measurement	0	100	1

For each combination of these nine degrees of freedom the optimizer proposes, the ExpObjEvalCustom checks if the sensitivity functions needed to compose the SensitivityMatrix are available in the ManagerSensitivityFunctions. If not, a sensitivity analysis experiment ExpSens is first set up and launched to determine the sensitivity functions $\partial S/\partial \mu_{max}$, $\partial X/\partial \mu_{max}$, $\partial S/\partial K_S$ and $\partial X/\partial K_S$ for the chosen value of the initial substrate concentration (ManipulableQuantities). The resulting sensitivity function is stored in the ManagerSensitivityFunctions and the creation of an OEDExperiment, SensitivityMatrix, CovarianceMatrix and FIM can continue. Next, the SensitivityMatrix retrieves the needed sensitivities from the ManagerSensitivityFunctions. Together with the CovarianceMatrix, which is composed from the variance information in the MeasurableVariables section of the OED.Exp.txt file, the FIM is established. The requested criterion value is calculated and passed on to the ExpObjEvalCustom, which inverts the criterion value if needed (the Modified A, D, and E criteria need to be maximised, while the optimization algorithms minimise the objective) and passes the objective value on to the optimizer.

The output files Fermentor.Optim.out.txt and Fermentor.Optim.log.txt provide a list of all the experiments the optimizer proposed, together with the associated objective value. The best solution is mentioned at the end of the log file:

```
.S     Set1 Set2 Set3   Var1.yS    Var2.yS    Var3.yS,yX    Var4.yS,yX    Var5.yS,yX

0.05   0    0    0      45         34         43            31            31
```

This means the substrate concentration should be measured at t = 31, 34, 43 and 45 hours, with a duplicate measurement at 31h and an initial substrate concentration of 0.05 g/l. This means that measuring the substrate concentration is more informative than measuring the biomass concentration due to a higher sensitivity (Figure 19). The duplicate measurement reduces the uncertainty and is therefore apparently more informative than placing the measurement at another time.
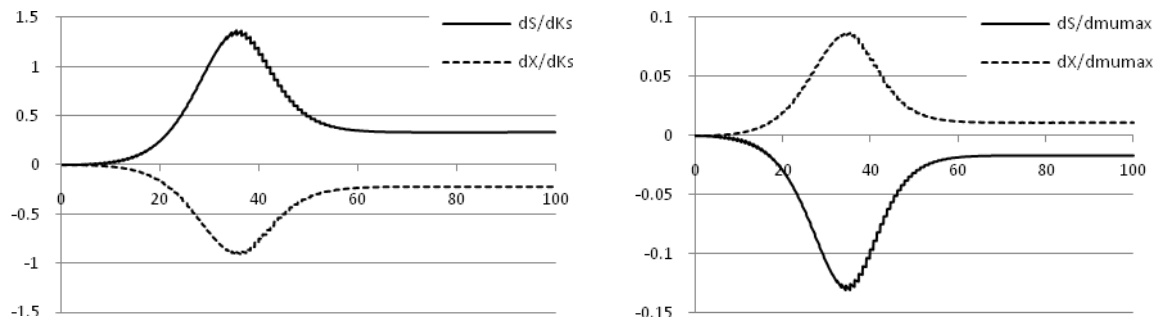
Figure 19. Sensitivity functions of S and X to the parameters Ks and mumax for an initial substrate concentration of 0.05 g/l

5.1.5 Full-scale application

For didactic reasons, the fermentor example explained in the previous section used a very simple model and optimized only a very limited number of measurements and initial conditions. In principle, the OED/PE tool is able to design more extensive measurement campaigns with much more complicated models. The main bottleneck to application in practice is that the integrated models constructed in Task 7.4 perform too slowly. Especially when there are initial conditions or model parameters to be optimized, the multiple execution of sensitivity analysis experiments is very time consuming. Moreover, the storage of all sensitivity functions in the ManagerSensitivityFunctions can be a problem due to memory limitations.

The aim of OED/PE is to determine which measurement campaign is the most informative to calibrate certain model parameters. To focus the search, it might be considered to perform a global sensitivity analysis on the model first, to determine which the key parameters to calibrate in an OED/PE-based measurement campaign are. Also for this initial global sensitivity analysis, the simulation speed of the model is a bottleneck.

5.2 Optimal experiment design for SFA

As was concluded in the previous section, calibrating the overparameterized integrated priority pollutant models can be challenging due to the number of parameters, the model complexity, the simulation speed, the numerous possibilities for placing measurements and the costs and reliability of priority pollutant measurements. A different question rises when the aim of the measurement campaign is defined as obtaining a good overview of the priority pollutant fluxes in the modelled urban system instead of obtaining a well-calibrated model with which these fluxes can be assessed. The overview of priority pollutant sources, fluxes and sinks in the urban environment is called an SFA and is described in chapter 4.

In the design of measurement campaigns for the purpose of SFA, the objective can be formulated as "defining the times, places and variables to measure in order to gather data that allow for establishing an SFA which maps the priority pollutant sources, fluxes and sinks as realistic as possible". Translated to an OED problem, this means that the information content of different possible sets of measurements should be compared and optimised in terms of minimising the difference between the "real" SFA (SFA_r) and the measurement-based SFA (SFA_m).

In model-based OED, the SFA_r, serving as reference, can be obtained from high-resolution simulation results with the integrated urban model. Thanks to the implementation of sensor and sampling models (see chapter 3.2), it is also possible to model the data-acquisition that would be possible with certain

sampling devices. Chapter 4.2 introduced a methodology to obtain from the model the average pollutant fluxes typically shown in an SFA. This can be done both for mass flux data coming directly from the model (SFA_r) and for data coming from the modelled measuring devices (SFA_m). The only missing step to conduct OED/SFA is an objective function to calculate the difference between SFA_r and the SFA_m, to judge how similar they are. In case automation is desired, an extra requirement is an optimization algorithm that tries different combinations of measurements and optimizes the value of the objective function.

In principle, both issues could be solved within WEST with an end value optimization experiment (Figure 20). The objective function can be defined as a weighted sum of squared differences between the mass fluxes in both SFAs and can be implemented via calculator variables. To automate the comparison of different possible combinations of measurements, it is necessary to implement all possible measurements in the model and to use a 'selector' node to route a particular set of measurement signals to the AverageCalculator node based on a parameter value that can be altered by the optimization algorithm.
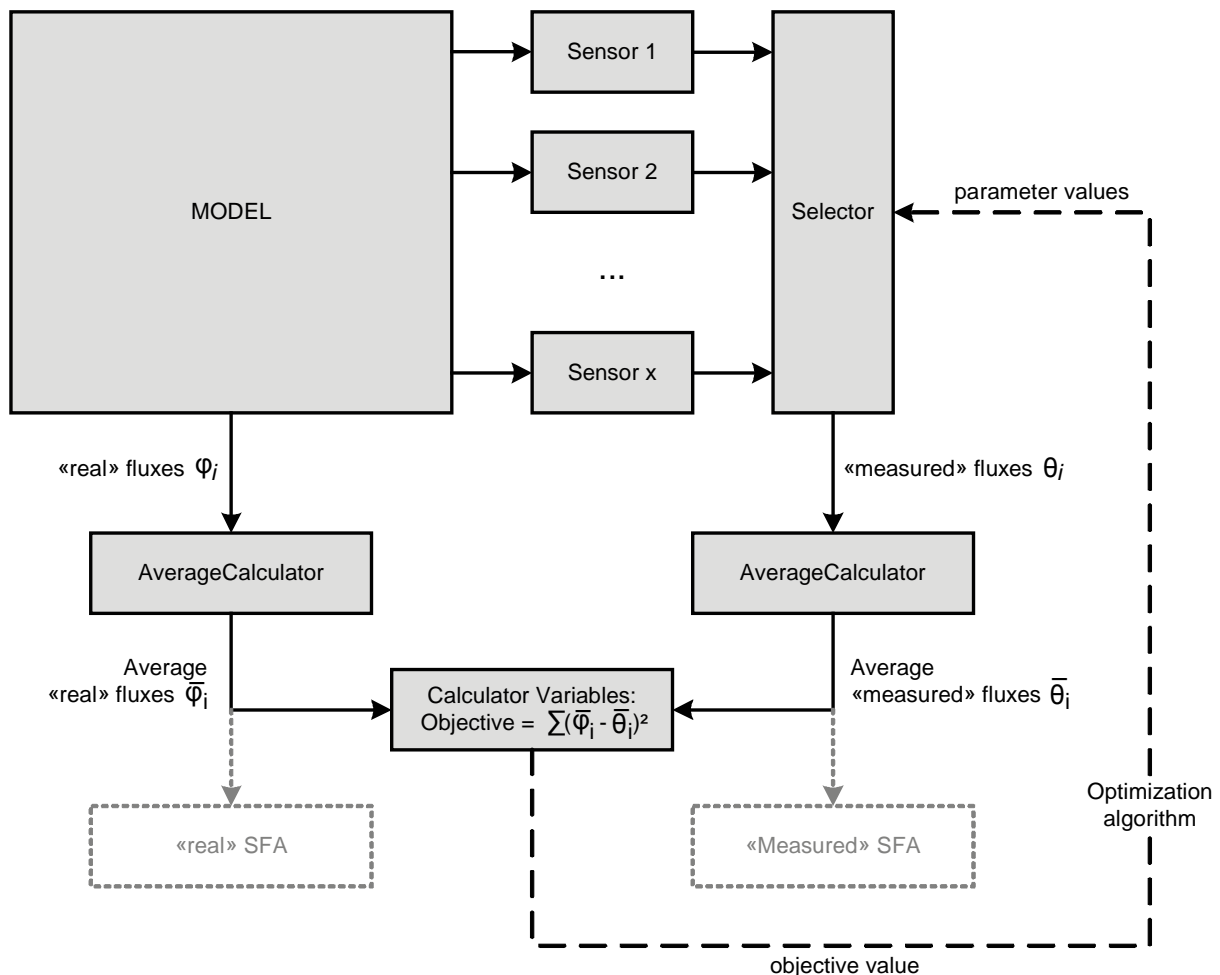

Figure 20. Scheme of the proposed methodology to conduct OED/SFA in WEST

5.3 Optimal experiment design for assessment of emission barriers

The methodology for model-based SFA that was presented in chapter 4.2 and used in the previous chapter on OED/SFA, can also be applied to perform OED for the assessment of emission barriers. In this case, the objective is to design a measurement campaign such that the effect of a certain emission control strategy – whatever this effect might be – can be monitored in the field as accurately as possible.

The potential effect of an emission control strategy can be predicted quantitatively in a model-based assessment by implementing the control strategy in the integrated urban water system model. The results in terms of total yearly pollutant fluxes can be visualised in a SFA-chart as demonstrated in chapter 4.2. The difference between this SFA and the original SFA (without the emission barrier) is an indication of the emission barrier effect and is comparable to the sensitivity functions used in OED/PE (see chapter 5.1.1): a measurement campaign will be more informative when measurements are performed on the most 'sensitive' locations in the system. An intuitive example is that the efficiency of a tertiary treatment step in a WWTP will probably better be assessed with measurements in the WWTP effluent than with measurements somewhere downstream in the river, while it will be useless to place measurements anywhere "upstream" of the WWTP. To perform SFA-based OED for assessment of emission barriers, an objective function can be defined as a weighted sum of squared differences between mass fluxes in both SFAs. The optimal placing of (ideal) measurements can be found by maximising the objective function, similar to the maximization of the Fischer Information Matrix in OED/PE.

If only the impact of the emission control strategy on the yearly pollutant fluxes is to be assessed, this SFA-based methodology is sufficient. However, SFAs summarize information over a certain time period, causing dynamics to be lost. In many situations, a dynamic assessment will be desired, e.g. to study the potential of an emission barrier to flatten pollution peaks in a river ecosystem. This implies that in the design of measurement campaigns, not only the location but also the time of measuring is important. An intuitive example is that the efficiency of a stormwater infiltration pond should be assessed during rain events and not in long dry periods.

To determine the most suited measuring times and locations, a kind of sensitivity functions can be derived from the simulation output (Figure 21). Allowing an optimizer to sample measurements from these sensitivity functions, leads to the optimal measurement campaign.

In the conceptual example shown in Figure 21, a rain event occurs at t = 0, causing a peak load in the river a bit later. Two locations in the river are available to measure, modelled as river nodes x and x+1. River node x+1 is downstream, so the peak occurs later than in node x, while the total load is also higher due to other emissions between nodes x and x+1. Using the same model but with a certain emission control strategy (ECS) implemented, yields similar load profiles, although the loads are lower thanks to the ECS. From these load profiles, a kind of sensitivity function can be derived, indicating for each possible measuring location separately at which time the load – which can be determined as a combination of concentration and flow rate – is influenced the most by the applied ECS. Note that only the magnitude (absolute value) and not the sign of the sensitivity functions should be taken into account, similar as in OED/PE. Now an optimization algorithm can be used to determine the most informative (or sensitive) set of measurements. A measurement is the combination of a model variable to measure (in practice often a combination of a variable and a location in the system) and a moment in time to measure. Suppose two discrete measurements can be performed in this example, then the algorithm will consider the load measurements in river node x at t = 0.5 and 0.6 days after the rain event as the best measurement campaign. Of course, it would also be possible to configure the algorithm such that not all measurements are at the same location (e.g. to avoid a systematic measurement error due to a badly positioned sampling device). In this case, the outcome should be a measurement at river node x at 0.6 days and at river node x+1 at 0.5 days after the rain event.
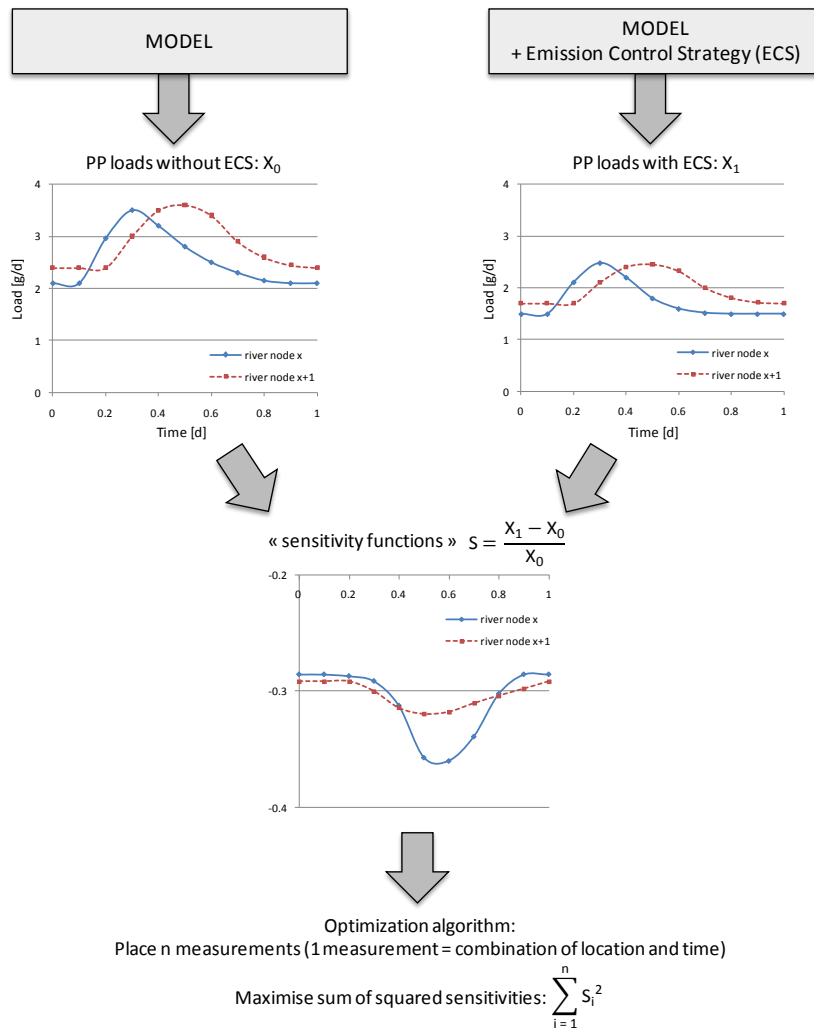
Figure 21. Conceptual overview of the procedure to perform dynamic OED for assessment of Emission Control Strategies

A possible extension would be to include the sampling and measuring models (chapter 3.2) to take into account the influence of applying a certain sampling strategy or using a certain measuring device. Similar to OED/PE, the estimated measurement error could be used as a weighting factor in the objective function:

$$\sum_{i=1}^{n} \frac{1}{\sigma_i^2} S_i^2$$

6 Visualisation of monitoring data

As models needs and generates grate amounts of tabular data, visualisation for end users is essential. This simplifies the user interpretation of the tabluar data. Since models can simulate what a potential monitoring campaign would measure, an interaction between the monitored dataset and the simulated dataset is needed. Therefore a visualization framework was developed.

6.1 Visualisation of input data for models

Models are proven to be very useful tools for estimation and simulation of substances dynamics as well as for better understanding of the processes that influence these dynamics. However, establishing a good model is a difficult task, due to the complex processes that take place in the modeled systems. When talking about theoretical (knowledge-driven, also called conceptual) models, based on the physical, chemical and biological principles, selecting the 'correct' model complexity is of key importance. Simple models are efficient and easy to cope with, but they sometimes do not describe the system according to our needs. The more complex the model is, the more accurately it describes the system, but it also contains many unknown parameters, which makes them numerically unstable and difficult to cope with. Note here that the systems are complex, and our knowledge incomplete. So, regardless of how complex the model is, it can never represent a complete picture of the reality, and it will always be limited by our knowledge. Thus, the quality of the obtained models greatly depends on the modeler's skills and experience.

Due to the nature of conceptual models, data from many fields are needed for its construction and later for using the model. For example, when simulating the behavior of micropollutants in an urban system the following is needed:

- specification of modelled catchments, which contains locations of sources of emission, sewer network locations, and lengths, to simulate the distribution of micropollutants, sewer overflows, permeable and impermeable areas in the catchments,
- quantities and dynamics of emissions,
- characteristics of WWTP(s),
- physical and chemical characteristics of a micropollutant,
- weather and terrain characteristics.

Even though, such models are useful and have big descriptive power, they can be a bottleneck when constructing an integrated decision support system, where models are part of the system. First, they need a lot of data to be simulated, if too complex they can be numerically unstable, and finally they can be computationally very expensive, which makes the DSS inefficient. So, it is reasonable to simplify the model.

One of the ways to approach this task is to learn the behavior of a complex model from the simulated data by the model, and provide a simplified model for the same tasks, i.e., to apply data-driven modeling techniques. In contrast to the knowledge-driven modeling, data-driven modeling aims at building models from observations. This approach enables to tackle various problems without the necessity to introduce any domain knowledge in the process of model construction. As a result many of these methods, e.g. statistical methods, produce so called 'black-box' models, which cannot be explained, or rather vaguely, by domain experts. Yet, a branch of artificial intelligence, i.e., machine learning (ML), tends to produce so called "grey-box" models, which can be partly explained and understood by a domain expert. ML algorithms include different methodologies to search for patterns in data and thus many times successfully identify simple relations between the input and the output data (Kompare et al., 1997; Atanasova and Kompare, 2002a, 2002b).

A data-driven model can be set in a way that it mimics the behavior of a conceptual model. In that case it has to be constructed from the selected input data that enter the conceptual model and corresponding output simulations. This method was successfully shown in Task 6.4 (*Implementation of mathematical models in a distributed manner*) on a hypothetical city example of simulating DEHP concentrations in the city, and is more closely described in the deliverable of that Task (D6.4, Cerk et al., 2009).

When constructing a data-driven model, which mimics the behavior of a conceptual model, a large and representative dataset of the conceptual model's simulated output is needed. The conceptual model has to be run multiple times, i.e. it has to simulate sufficient number of scenarios, so that the simulated data captures its behavior under different conditions. In this way we provide a sufficient data set including input data (different scenarios of changing e.g. external conditions) and corresponding simulated data for the data-driven algorithm to produce a model that learns how to connect the inputs and the outputs in a simple way, i.e. without the necessity to run the complex conceptual model. When the learning dataset is representative enough the data-driven model can mimic the conceptual model with high accuracy.

Common for both types of models is the weather data (precipitation) and the loads data, i.e., data that describes quantities and dynamics of emissions (releases). The loads data is obtained from the Emission Strings (literature data about emissions classified and composed in a database) and the site specific data presented in an Adaptation Matrix (locations and sizes of emitting sites), see also D6.4 (Cerk et al., 2009). The loads data can be described as Localised Releases or LOREs (see Figure 22).
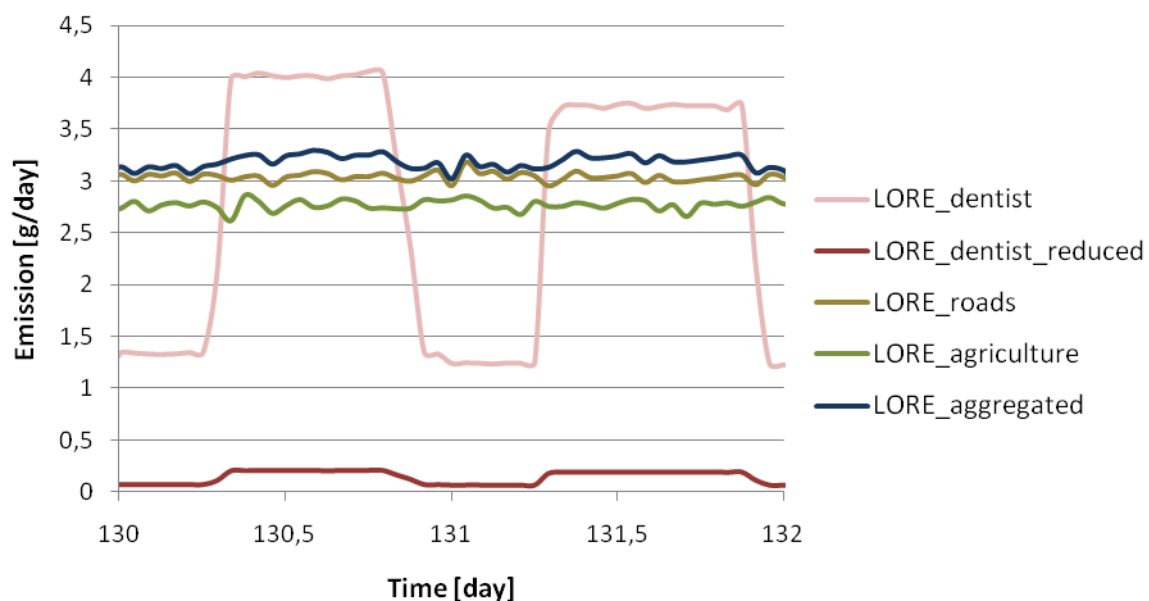

Figure 22. Localised Releases (LOREs), which are obtained by combining Emission Strings and Adaptation Matrix

The concept of Emission strings, Adaptation Matrices and LOREs are more closely described in the deliverables of Task 9.3 (*Data requirement analysis and definition of common data structures*) (Banovec et al., 2009) and Task 6.4 (Cerk et al., 2009).

6.2 Results from models

A model can be used to estimate different parameters. For comparison with the measurement campaigns, the most important are concentrations of pollutants. The concentrations are variant in both time and the location of the measurement in the system. For visualization of priority pollutant concentrations at various locations, the use of GIS tools is convenient. However, the maps can only display a limited number of concentrations at one location, i.e. only the minimum, maximum or average concentration, by representing it by different intensity, color or size of the map object. Therefore the concentration diagrams can be added to GIS maps in order to add the visualization of concentrations in time.
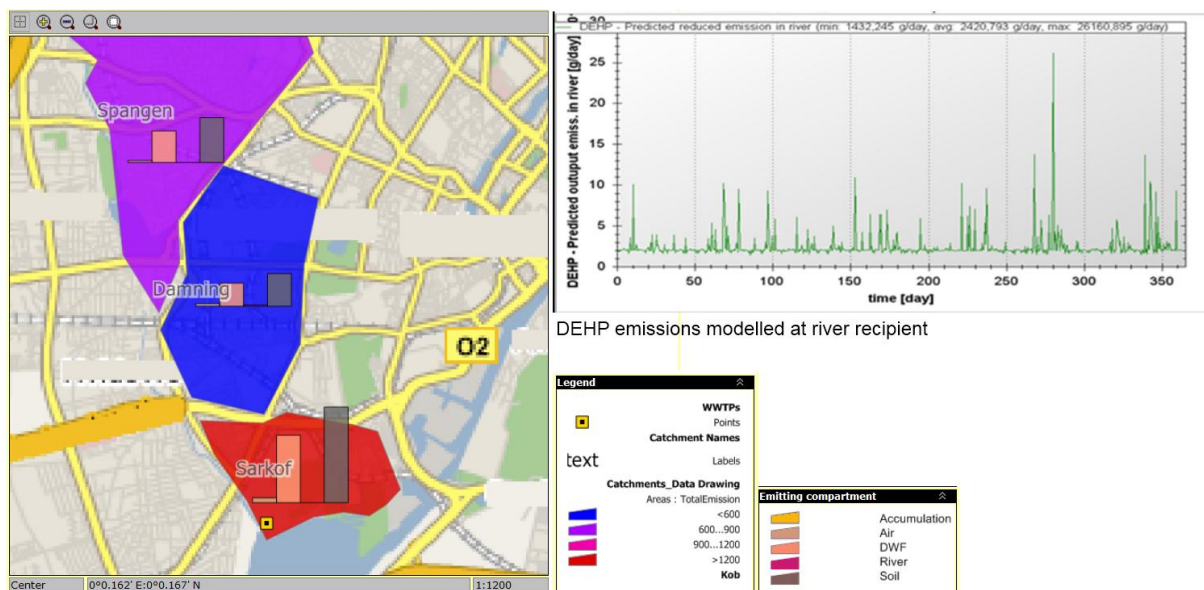

Figure 23. Representing a modelled area and modelled concentrations at selected location

A tool for visualizing model output was developed in Task 6.5 (*Induction of optimal control measures*) and is presented in Figure 23 above. At the moment the tool is used for visualizing concentrations only at the recipient - after the treatment in WWTP, but it could be easily modified to visualize the concentrations at any location.

The tool can visualize data from both conceptual models and data-driven models. Since the integration of conceptual models into the GIS and the database is hard to achieve, all the data must be generated beforehand and stored in the database. Therefore a limited number of scenarios can be prepared for the visualization.

However, data-driven models are much lighter, faster and easier to integrate to GIS and databases (Cerk et al., 2009). Therefore greater interaction with the visualizing tool can be achieved. For example, the input data can easily be modified and the model can be re-run while visualizing the output. The visualization tool will be more closely described in the deliverable of Task 6.5.

6.3 Results from monitoring campaigns

Measured data can be visualized the same way as the modeled data. The data is both location and time dependent. The measurement locations can be presented on a map as points. The measured

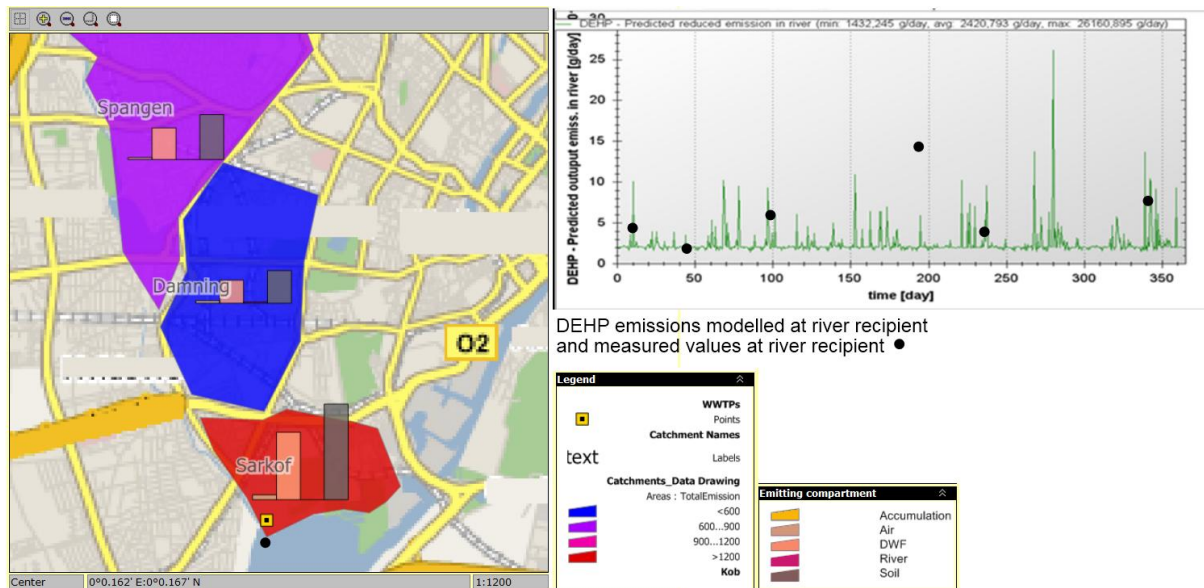concentrations can be further visualized on either the same or on different diagrams as the modeled data.

Figure 24. Modelled data and measured data presented together. The measured data is presented as black dots.

If the model is accurate enough, it is expected that the monitoring results are within the range of the modeled concentrations. The model can be used to find the locations in the systems that are interesting to measure. It can also be used in a way to predict the range of concentrations – assisting in the selection of a measurement method.

7   Conclusions

In WP 7, some milestones were achieved concerning priority pollutant modelling. Task 7.2 provided unit process models to simulate the behaviour of priority pollutants in urban sub-systems. Task 7.4 united these sub-system models to an integrated urban wastewater system (IUWS) model and coupled it with the multimedia model that was developed in Task 7.3, to be able to also take into account the wider surroundings of the urban system (i.e. the whole river catchment). In Task 7.1, an input generator tool was developed to provide these IUWS models with the input needed to run simulations, thereby constructing a bridge to WP 3 where knowledge on priority pollutant emissions was gathered.

Task 7.5, as the concluding task of WP 7, has shown some possible applications of the IUWS models with respect to measurement campaigns. The interaction is bidirectional. First, measurement campaigns can be designed with the help of the model serving the purpose of a better calibration of certain parameters in this IUWS model. This is called Optimal Experiment Design for Parameter Estimation (OEDPE) and was discussed in chapter 5.1. Second, the knowledge contained in the IUWS models can be used to optimize the planning of sampling campaigns that aim at assessing the efficiency of certain emission control strategies or emission barriers (chapter 5.3). To do this, several tools and methods can be helpful. First, the emission barriers need to be implemented in the IUWS model, as well as the sampling and measuring devices that could or will be used in the measurement campaign (chapter 3.2). Second, the presentation of the simulation results in the format of an SFA. This can be a very helpful way to wrap-up the results, as well as to detect anomalies in the model or in the real system (chapter 4.2). Furthermore, these SFAs can be compared with empirical SFAs (chapter

4.1), and the model-based SFAs can be used to set up measurement campaigns that aim at gathering information to establish empirical SFAs: this was called Optimal Experiment Design for SFA (chapter 5.2).

For chemical monitoring in relation to the WFD, there are a range of possibilities. Current requirements for monitoring are based on spot sampling. However, in dynamic water systems flow-, volume- or time-proportional ways of sampling are more appropriate than spot samples. Flow-proportional samplers are more costly than passive samplers, and are most often used to find total concentrations of pollutants, while passive samplers can measure freely dissolved concentrations. Passive samplers are appropriate for evaluating trends, identifying pollutants in water bodies and for pollution tracking when long time-frames are wanted. It is not appropriate for compliance monitoring since quality assurance procedures for passive samplers are not established yet.

8 References

Aguilar-Martinez, R., Greenwood, R., Mills, G. A., Vrana, B., Palacios-Corvillo, M. A., Gomez-Gomez M. M. (2008a) Assessment of Chemcatcher passive sampler for the monitoring of inorganic mercury and organotin compounds in water. International Journal of Environmental Analytical Chemistry, 88 (2), pp. 75-90.

Aguilar-Martinez, R., Palacios-Corvillo, M. A., Greenwood, R., Mills, G. A., Vrana, B., Gómez-Gómez, M. M. (2008b) Calibration and use of the Chemcatcher® passive sampler for monitoring organotin compounds in water. Analytica Chimica Acta, 618 (2), pp. 157-167.

Allan, I. J., Vrana, B., Greenwood, R., Mills, G. A., Knutsson, J., Holmberg, A., Guigues, N., Fouillac, A. M., Laschi, S. (2006a) Strategic monitoring for the European Water Framework Directive. Trends in Analytical Chemistry, 25 (7), pp. 704-715.

Allan, I. J., Vrana, B., Greenwood, R., Mills, G. A., Roig, B., Gonzalez, C, (2006b) A "toolbox" for biological and chemical monitoring requirements for the European Union's Water Framework Directive. Talanta, 69 (2), pp. 302-322.

Allan, I. J., Knutsson, J., Guigues, N., Mills, G. A., Fouillac, A. M., Greenwood, R. (2007) Evaluation of the Chemcatcher and DGT passive samplers for monitoring metals with highly fluctuating water concentrations. Journal of Environmental Monitoring, 9 (7), pp. 672-681.

Allan, I. J., Knutsson, J., Guigues, N., Mills, G. A., Fouillac, A. M., Greenwood, R. (2008) Chemcatcher(R) and DGT passive sampling devices for regulatory monitoring of trace metals in surface water. Journal of Environmental Monitoring, 10 (7), p. 821.

Alvarez, D. A., Petty, J. D., Huckins, J. N., Jones-Lepp, T. L., Getting, D. T., Goddard, J. P., Manahan, S. E. (2004) Development of a passive, in situ, integrative sampler for hydrophilic organic contaminants in aquatic environments. Environmental Toxicology and Chemistry, 23 (7), pp. 1640-1648.

Alvarez, D. A., Stackelberg, P. E., Petty, J. D., Huckins, J. N., Furlong, E. T., Zaugg, S. D., Meyer, M. T. (2005) Comparison of a novel passive sampler to standard water-column sampling for organic contaminants associated with wastewater effluents entering a New Jersey stream. Chemosphere, 61 (5), pp. 610-622.

Alvarez, D. A., Cranor, W. L., Perkins, S. D., Clark, R. C., Smith, S. B, (2008) Chemical and toxicologic assessment of organic contaminants in surface water using passive samplers. Journal of Environmental Quality, 37 (3), pp. 1024-1033.

Atanasova, N. and Kompare, B. (2002a) The use of decision trees in the modelling of a wastewater treatment plant. Acta hydrotech., 20(33), 351-370.

Atanasova, N. and Kompare, B. (2002b) Modelling of wastewater treatment plant with decision and regression trees. In: Proc. of the Workshop in Binding Environmental Sciences and Artificial Intelligence, ECAI, Lyon, France, 23 July 2002.

Augulyte, L. and Bergqvist, P. A. (2007) Estimation of water sampling rates and concentrations of PAHs in a municipal sewage treatment plant using SPMDs with performance reference compounds. Environmental Science and Technology, 41 (14), pp. 5044-5049.

Aung, N. N., Nakajima, F., Furumai, H. (2008) Trace metal speciation during dry and wet weather flows in the Tama River, Japan, by using diffusive gradients in thin films (DGT). Journal of Environmental Monitoring, 10 (2), pp. 219-230.

Banovec, P., Cerk, M., Atanasova, N., Kompare, B., Lützhøft, H.-C. H., Donner, E., Bessat M.-C. (2009) Data Requirement Analysis and Definition of Common Data Structures. ScorePP project deliverable D9.3.

Battaglin, W. A., Hay, L. E. (1996) Effects of Sampling Strategies on Estimates of Annual Mean Herbicide Concentrations in Midwestern Rivers. *Environmental Science and Technology* 30 (3), pp. 889-896.

Benedetti, L., De Keyser, W., Vezzaro, L., Atanasova, N, Gevaert, V., Verdonck, F., Vanrolleghem, P.A., Mikkelsen, P.S. (2009) Deliverable 7.4. Integrated dynamic urban scale sources-and-flux model for PPs. ScorePP project deliverable. 32p.

Blom, L. B. Morrison, , G. M., Kingston, J., Mills, G. A., Greenwood, R., Pettersson, T. J. R., Rauch, S. (2002) Performance of an in situ passive sampling system for metals in stormwater. Journal of Environmental Monitoring, 4 (2), pp. 258-262.

Booij, K., Hofmans, H. E., Fischer, C. V., van Weerlee, E. M. (2003) Temperature-dependent uptake rates of nonpolar organic compounds by semipermeable membrane devices and low-density polyethylene membranes. Environmental Science and Technology, 37 (2), pp. 361-366.

Booij, K., van Bommel, R., Mets, A., Dekker, R. (2006) Little effect of excessive biofouling on the uptake of organic contaminants by semipermeable membrane devices. Chemosphere, 65 (11), pp. 2485-2492.

Brown, D. G., Riolo, R., Robinson, D. T., North, M., Rand, W. (2005) Spatial process and data models: Toward integration of agent-based models and GIS. Journal of Geographical Systems, 7(1), 25-47.

Buzier, R. Tusseau-Vuillemin, , M. H., Meriadec, C. M. D., Rousselot, O., Mouchel, J. M. (2006) Trace metal speciation and fluxes within a major French wastewater treatment plant: Impact of the successive treatments stages. Chemosphere, 65 (11), pp. 2419-2426.

Caeiro, S., Painho, M., Goovaerts, P., Costa, H., Sousa, S. (2001) Spatial sampling design for sediment quality assessment in estuaries. *Environmental Modelling and Software* 18 (10), pp. 853-859.

Cerk, M., Atanasova, N., De Keyser, W. (2009) Implementation of distributed mathematical models. ScorePP project deliverable D6.4.

Claeys F. (2008) A generic software framework for modelling and virtual experimentation with complex environmental systems. PhD thesis. Faculty of Bioscience Engineering. Ghent University, Belgium. 303p.

Cleven, R., Nur, Y., Krystek, P., van den Berg, G. (2005) Monitoring metal speciation in the rivers meuse and Rhine using DGT. Water Air and Soil Pollution, 165 (1-4), pp. 249-263.

Davison, W. and Zhang, H. (1994) In situ speciation measurements of race components in natural waters using thin-film gels. Nature, 367, pp. 546-548.

De Keyser, W., Gevaert, V., Benedetti, L., Verdonck, F, Vezzaro, L., Mikkelsen, P.S. & Vanrolleghem, P.A. (2008a) Deliverable 7.1: Mathematical formulations of pollution generating processes. ScorePP project deliverable. 33p.

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

De Keyser, W., Verdonck, F, Gevaert, V., Benedetti, L. & Vanrolleghem, P.A. (2008b) Multimedia model at river basin scale. ScorePP project deliverable D7.3. 43p.

De Keyser, W., Verdonck, F, Gevaert, V., Benedetti, L., Vanrolleghem, P.A. (2008b) Deliverable D7.3. Multimedia model at river basin scale. ScorePP project deliverable. 43p.

de la Cal, A., Kuster, M., de Alda, M. L., Eljarrat, E., Barcelo, D. (2008) Evaluation of the aquatic passive sampler Chemcatcher for the monitoring of highly hydrophobic compounds in water. Talanta, 76 (2), pp. 327-332.

De Pauw D. (2005) Optimal experimental design for calibration of bioprocess models: a validated software toolbox. PhD thesis. Faculty of Bioscience Engineering. Ghent University, Belgium. 246p.

De Pauw D.J.W. and Vanrolleghem P.A. (2006) Practical aspects of sensitivity function approximation for dynamic models. Math. Comp. Modell. of Dyn. Sys., 12, 395-414.

Divis, P., Docekalova, H., Brulik, L., Pavlis, M., Hekera, P. (2007) Use of the diffusive gradients in thin films technique to evaluate (bio)available trace metal concentrations in river water. Analytical and Bioanalytical Chemistry, 387 (6), pp. 2239-2244.

Dochain D. and Vanrolleghem P.A. (2001) Dynamic modelling and estimation in wastewater treatment processes. IWA Publishing, London, UK. 342p.

Donckels B. (2009) Optimal experimental design to discriminate among rival dynamic mathematical models. PhD thesis. Faculty of Bioscience Engineering. Ghent University, Belgium. 304p.

Dunn, R. J. K., Teasdale, P. R., Warnken, J., Schleich, R. R. (2003) Evaluation of the diffusive gradient in a thin film technique for monitoring trace metal concentrations in estuarine waters. Environmental Science and Technology, 37 (12), pp. 2794-2800.

Dunn, R. J. K., Teasdale, P. R., Warnken, J., Arthur, J. M. (2007) Evaluation of the in situ, time-integrated DGT technique by monitoring changes in heavy metal concentrations in estuarine waters. Environmental Pollution, 148 (1), pp. 213-220.

Esteve-Turrillas, F. A., Yusa, V., Pastor, A., de la Guardia, M. (2008) New perspectives in the use of semipermeable membrane devices as passive samplers. Talanta, 74 (4), pp. 443-457.

European Economic Community (1977) 77/795/EEC, Council Decision of 12 December 1977 establishing a common procedure for the exchange of information on the quality of surface fresh water in the Community

European Commission (2000) Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, Official journal of the European Commission L 327/1-72.

European Commission (2008) Directive 2008/105/EC of the European Parliament and the Council of 16 December 2008 on environmental quality standards in the field of water policy, Official journal of the European Commission L 348/84.

Franceschini G. and Macchietto S. (2008) Model-based design of experiments for parameter precision: State of the art. Chem. Eng. Sci., 63, 4846-4872.

Goodwin G.C. and Payne R.L. (1977) Dynamic system identification. Experiment design and data analysis. Acadamic Press. New York. 291p.

Gorecki, T. and Namiesnik, J. (2002) Passive sampling. Trac-Trends in Analytical Chemistry, 21 (4), pp. 276-291.

Greenwood, R., Mills, G. A., Vrana, B. (2006) Improving environmental monitoring. A report on the 2nd International Passive Sampling Workshop and Symposium, Bratislava, Slovakia, 3-6 May 2006, and associated satellite workshops. Trends in Analytical Chemistry, 25 (8), pp. 751-754.

Greenwood, R., Mills, G., Vrana, B. (2007) Comprehensive Analytical Chemistry, Passive Sampling Techniques in Environmental Monitoring, first edn, Elsevier B.V., Amsterdam, The Netherlands.

Gunold, R., Schafer, R. B. Paschke, A. Schuurmann, G., Liess, M. (2008) Calibration of the Chemcatcher passive sampler for monitoring selected polar and semi-polar pesticides in surface water. Environmental Pollution, 155 (1), pp. 52-60.

Harman, C., Boyum, O., Tollefsen, K. E., Thomas, K., Grung, M. (2008) Uptake of some selected aquatic pollutants in semipermeable membrane devices (SPMDs) and the polar organic chemical integrative sampler (POCIS). Journal of Environmental Monitoring, 10 (2), pp. 239-247.

Holten Lützhøft, H.-C., Eriksson, E., Scholes, L., Donner, E., Wickman, T., Lecloux, A., Ledin, A. (2008) Deliverable D3.1. Database Presenting Basic Information about EU WFD Priority Substances. ScorePP project deliverable. 32p

Holten Lützhøft, H.-C., Donner E., Gevaert V., De Keyser W., Wickman T., Cerk M., Eriksson E., Banovec P., Lecloux A., Mikkelsen P.S., Ledin A. In preparation. Classifying Urban Sources of Priority Pollutants: A Source Classification Framework.

Huckins, J. N., Manuweera, G. K., Petty, J. D., Mackay, D., Lebo, J. A. (1993) Lipid-Containing Semipermeable-Membrane Devices for Monitoring Organic Contaminants in Water. Environmental Science and Technology, 27 (12), pp. 2489-2496.

Huckins, J. N., Petty, J. D., Orazio, C. E., Lebo, J. A., Clark, R. C., Gibson, V. L., Gala W. R., Echols, K. R. (1999) Determination of uptake kinetics (Sampling rates) by lipid-containing semipermeable membrane devices (SPMDs) for polycyclic aromatic hydrocarbons (PAHs) in water. Environmental Science and Technology, 33 (21), pp. 3918-3923.

Huckins, J. N., Petty, J. D., Booij, K. (2006) Monitors of organic Chemicals in the Environment Springer, New York, NY.

Huckins, J. N., Tubergen, M. W., Manuweera, G. K. (1990) Semipermeable-Membrane Devices Containing Model Lipid - A New Approach to Monitoring the Bioavailability of Lipophilic Contaminants and Estimating Their Bioconcentration Potential. Chemosphere, 20 (5), pp. 533-552.

Jones-Lepp, T. L., Alvarez, D. A., Petty, J. D., Huckins, J. N. (2004) Polar organic chemical integrative sampling and liquid chromatography-electrospray/ion-trap mass spectrometry for assessing selected prescription and illicit drugs in treated sewage effluents. Archives of Environmental Contamination and Toxicology, 47 (4), pp. 427-439.

Jonsson, A., Fridén, U., Thuresson, K., Sörme, L. (2008) "Substance flow analysis of organic pollutants in Stockholm". Water, Air and Soil Pollution: Focus 8:433-443.

Katsoyiannis, A. and Samara, C. (2007) Comparison of active and passive sampling for the determination of persistent organic pollutants (POPs) in sewage treatment plants. Chemosphere, 67 (7), pp. 1375-1382.

Ke, R. H., Xu, Y. P., Wang, Z. J., Khan, S. U. (2006) Estimation of the uptake rate constants for polycyclic aromatic hydrocarbons accumulated by semipermeable membrane devices and triolein-embedded cellulose acetate membranes. Environmental Science and Technology, 40 (12), pp. 3906-3911.

Kingston, J. K., Greenwood, R., Mills, G. A., Morrison, G. M., Björklund Persson, L. (2000) Development of a novel passive sampling system for the time-averaged measurement of a range of organic pollutants in aquatic environments. Journal of Environmental Monitoring, 002 (5), pp. 487-495.

Kolok, A. S., Snow, D. D., Kohno, S., Sellin, M. K., Guillette, L. J. (2007) Occurrence and biological effect of exogenous steroids in the Elkhorn River, Nebraska, USA. Science of the Total Environment, 388, pp. 104-115.

Kompare, B., Steinman, F., Cerar, U., Džeroski, S. (1997) Prediction of rainfall runoff from catchment by data analysis with machine learning tools within artificial intelligence tools. Acta Hydrotechnica, 15(17), 79-94.

Kot, A., Zabiegala, B., Namiesnik, J., (2000) Passive sampling for long-term monitoring of organic pollutants in water. Trac-Trends in Analytical Chemistry, 19 (7), pp. 446-459.

Lebo, J. A., Zajicek, J. L., Orazio, C. E., Petty, J. D., Huckins, J. N., Douglas, E. H.. (1996) Use of the semipermeable membrane device (SPMD) to sample polycyclic aromatic hydrocarbon pollution in a lotic system. Polycyclic Aromatic Compounds, 8 (1), pp. 53-65.

Luellen, D. R. and Shea, D., (2002) Calibration and field verification of semipermeable membrane devices for measuring polycyclic aromatic hydrocarbons in water. Environmental Science and Technology, 36 (8), pp. 1791-1797.

Macleod, S. L., Mcclure, E. L., Wong, C. S., (2007) Laboratory calibration and field deployment of the polar organic chemical integrative sampler for pharmaceuticals and personal care products in wastewater and surface water. Environmental Toxicology and Chemistry, 26, pp. 2517-2529.

Matthiessen, P., Arnold, D., Johnson, A. C., Pepper, T. J., Pottinger, T. G., Pulman, K. G. T. (2006) Contamination of headwater streams in the United Kingdom by oestrogenic hormones from livestock farms. Science of the Total Environment, 367 (2-3), pp. 616-630.

Mazzella, N., Dubernet, J. F., Delmas, F., (2007) Determination of kinetic and equilibrium regimes in the operation of polar organic chemical integrative samplers - Application to the passive sampling of the polar herbicides in aquatic environments. Journal of Chromatography A, 1154 (1-2), pp. 42-51.

Meadows, J. C., Echols, K. R., Huckins, J. N., Borsuk, F. A., Carline, R. F., Tillitt, D. E.. (1998) Estimation of uptake rate constants for PCB congeners accumulated by semipermeable membrane devices and brown trout (Salmo trutta). Environmental Science and Technology, 32 (12), pp. 1847-1852.

Munack A. (1991) Optimization of Sampling. In: Biotechnology, a Multi-volume Comprehensive Treatise, Vol. 4. Measuring, Modelling and Control. Ed. Schügerl K., VCH, Weinheim, 251-264.

Munksgaard, N. C. and Parry, D. L., (2003) Monitoring of labile metals in turbid coastal seawater using diffusive gradients in thin-films. Journal of Environmental Monitoring, 5 (1), pp. 145-149.

Namiesnik, J., Zabiegala, B., Kot-Wasik, A., Partyka, M., Wasik, A. (2005) Passive sampling and/or extraction techniques in environmental analysis: a review. Analytical and Bioanalytical Chemistry, 381 (2), pp. 279-301.

Ort, C. and Gujer, W., 2006. Sampling for representative micropollutant loads in sewer systems. Water Science and Technology, 54 (6-7), pp. 169-176.

Persson, L. B., Morrison, G. M., Friemann, J. U., Kingston, J., Mills, G., Greenwood, R. (2001) Diffusional behaviour of metals in a passive sampling system for monitoring aquatic pollution. Journal of Environmental Monitoring, 3 (6), p. 639.

Petersen B. (2000) Calibration, identifiability and optimal experimental design of activated sludge models. PhD thesis. Faculty of Bioscience Engineering. Ghent University, Belgium. 364p.

Petty, J. D., Orazio, C. E., Huckins, J. N., Gale, R. W., Lebo, J. A., Meadows, J. C., Echols, K. R., Cranor, W. L. (2000) Considerations involved with the use of semipermeable membrane devices for monitoring environmental contaminants. Journal of Chromatography A, 879 (1), pp. 83-95.

Petty, J. D., Huckins, J. N., Alvarez, D. A., Brumbaugh, W. G., Cranor, W. L., Gale, R. W., Rastall, A. C., Jones-Lepp, T. L., Leiker, T. J., Rostad, C. E., Furlong, E. T. (2004) A holistic passive

integrative sampling approach for assessing the presence and potential impacts of waterborne environmental contaminants. Chemosphere, 54 (6), pp. 695-705.

Roig, B., Valat, C., Allan, I. J., Greenwood, R., Berho, C., Guigues, N., Mills, G. A., Ulitzur, N. (2007) The use of field studies to establish the performance of a range of tools for monitoring water quality. Trends in Analytical Chemistry, 26 (4), pp. 274-282.

Schafer, R. B., Paschke, A., Vrana, B., Mueller, R., Liess, M. (2008) Performance of the Chemcatcher (R) passive sampler when used to monitor 10 polar and semi-polar pesticides in 16 Central European streams, and comparison with two other sampling methods. Water Research, 42 (10-11), pp. 2707-2717.

Seethapathy, S., Gorecki, T., Li, X. J., (2008) Passive sampling in environmental analysis. Journal of Chromatography A, 1184 (1-2), pp. 234-253.

Setkova, L., Hajslova, J., Bergqvist, P. A., Kocourek, V., Kazda, R., Suchan, P. (2005). Fast isolation of hydrophobic organic environmental contaminants from exposed semipermeable membrane devices (SPMDs) prior to GC analysis. Journal of Chromatography A, 1092 (2), pp. 170-181.

Sharpe, A. J. and Nichols, E. G., (2007) Use of stable nitrogen isotopes and permeable membrane devices to study what factors influence freshwater mollusk survival in the Conasauaga river. Environmental Monitoring and Assessment, 132 (1-3), pp. 275-295.

Solvi, A-M. (2007) Modelling the sewer-treatment-urban river system in view of the EU Water Framework Directive. PhD Thesis, Ghent University, Belgium, 218p.

Stuer-Lauridsen, F., (2005) Review of passive accumulation devices for monitoring organic micropollutants in the aquatic environment. Environmental Pollution, 136 (3), pp. 503-524.

Togola, A. and Budzinski, H., (2007) Development of polar organic integrative samplers for analysis of pharmaceuticals in aquatic systems. Analytical Chemistry, 79 (17), pp. 6734-6741.

Tran, A. T. K., Hyne, R. V., Doble, P., (2007) Calibration of a passive sampling device for time-integrated sampling of hydrophilic herbicides in aquatic environments. Environmental Toxicology and Chemistry, 26 (3), pp. 435-443.

Vanrolleghem P.A., Van Daele M., Dochain D. (1995) Practical identifiability of a biokinetic model of activated sludge respiration. Water Reserch, 29, 2561-2570.

Vermeirssen, E. L. M., Korner, O., Schonenberger, R., Suter, M. J. F., Burkhardt-Holm, P. (2005) Characterization of environmental estrogens in river water using a three pronged approach: Active and passive water sampling and the analysis of accumulated estrogens in the bile of caged fish. Environmental Science and Technology, 39 (21), pp. 8191-8198.

Vermeirssen, E. L. M., Asmin, J., Escher, B. I., Kwon, J. H., Steimen, I., Hollender, J. (2008) The role of hydrodynamics, matrix and sampling duration in passive sampling of polar compounds with Empore (TM) SDB-RPS disks. Journal of Environmental Monitoring, 10 (1), pp. 119-128.

Vezzaro L., Eriksson E., Ledin A.,, Mikkelsen P.S. (2009b) Modelling removal of micropollutants in stormwater treatment systems. In: Proceedings of the 8th International Conference on Urban Drainage Modelling (8UDM), Tokyo, Japan, 7-11 September 2009.

Vezzaro, L., Gevaert, V., Benedetti, L., De Keyser, W., Verdonck, F., Vanrolleghem, P.A., Boisson, P., Mikkelsen, P.S. (2009a) Deliverable 7.2. Unit process models for fate of PPs. Internal ScorePP project deliverable.

Vrana, B. and Schuurmann, G., (2002) Calibrating the uptake kinetics of semipermeable membrane devices in water: Impact of hydrodynamics. Environmental Science and Technology, 36 (2), pp. 290-296.

Vrana, B., Mills, G., Greenwood, R., Knutsson, J., Svenssone, K., Morrison, G. (2005a) Performance optimisation of a passive sampler for monitoring hydrophobic organic pollutants in water. Journal of Environmental Monitoring, 7 (6), pp. 612-620.

Vrana, B., Mills, G. A., Allan, I. J., Dominiak, E., Svensson, K., Knutsson, J., Morrison, G., Greenwood, R. (2005b) Passive sampling techniques for monitoring pollutants in water. Trac-Trends in Analytical Chemistry, 24 (10), pp. 845-868.

Vrana, B., Mills, G. A., Dominiak, E., Greenwood, R. (2006) Calibration of the Chemcatcher passive sampler for the monitoring of priority organic pollutants in water. Environmental Pollution, 142 (2), pp. 333-343.

Vrana, B., Mills, G. A., Kotterman, M., Leonards, P., Booij, K., Greenwood, R. (2007) Modelling and field application of the Chemcatcher passive sampler calibration data for the monitoring of hydrophobic organic pollutants in water. Environmental Pollution, 145 (3), pp. 895-904.

Warnken, J., Dunn, R. J. K., Teasdale, P. R. (2004) Investigation of recreational boats as a source of copper at anchorage sites using time-integrated diffusive gradients in thin film and sediment measurements. Marine Pollution Bulletin, 49 (9-10), pp. 833-843.

Warnken, K. W., Zhang, H., Davison, W. (2006) Accuracy of the diffusive gradients in thin-films technique: Diffusive boundary layer and effective sampling area considerations. Analytical Chemistry, 78 (11), pp. 3780-3787.

Wenzel, K. D., Vrana, B., Hubert, A., Schuurmann, G. (2004) Dialysis of persistent organic pollutants and polycyclic aromatic hydrocarbons from semipermeable membranes. A procedure using an accelerated solvent extraction device. Analytical Chemistry, 76 (18), pp. 5503-5509.

Yusa, V., Pastor, A., de la Guardia, M. (2006) Microwave-assisted extraction of polybrominated diphenyl ethers and polychlorinated naphthalenes concentrated on semipermeable membrane devices. Analytica Chimica Acta, 565 (1), pp. 103-111.

Zhang, H. and Davison, W. (1995) Performance-Characteristics of Diffusion Gradients in Thin-Films for the In-Situ Measurement of Trace-Metals in Aqueous-Solution. Analytical Chemistry, 67 (19), pp. 3391-3400.

Zhang, Z. L., Hibberd, A., Zhou, J. L. (2008) Analysis of emerging contaminants in sewage effluent and river water: Comparison between spot and passive sampling. Analytica Chimica Acta, 607, pp. 37-44.

Zhao, W. Y., Han, M., Dai, S. U., Xu, J. A., Wang, P. (2006) Ionic liquid-containing semipermeable membrane devices for monitoring the polycyclic aromatic hydrocarbons in water. Chemosphere, 62 (10), pp. 1623-1629.

Appendix 1

How to purchase information for passive samplers

SPMDs:

Can be purchased from Environmental Sampling Technologies (EST) in St. Joseph, Missouri, MO, USA.

POCIS:

Can be purchased from Exposmeter, Tavelsjö, Sweeden and EST labs St. Joseph, MO, USA.

Chemcatcher:

Can be purchased from Alcontrol AB, Linkoping, Sweden.

DGT:

Can be purchased from DGT Research Ltd. in the UK from http://www.dgtresearch.com. On this website instructions for deployment and diffusion coefficients of metals in the gel at relevant temperatures can be found.