Detection of the most optimal measuring points for water quality variables: application to the river water quality model of the River Dender in ESWAT

V. Vandenberghe, A. van Griensven and W. Bauwens
Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Pleinlaan 2 B-1050 Brussels, Belgium (E-mail: vvdbergh@vub.ac.be; avgriens@vub.ac.be; wbauwens@vub.ac.be)

Abstract This paper presents a methodology for the definition of an optimal set of sampling data for the calibration of a river water quality model. Starting with an extensive set of measurements, it is the aim to reduce those data to obtain just as much data as necessary for a calibration with an acceptable uncertainty in the parameters. The method requires a model for the river under examination and the availability of samples for a first calibration of the model. With the model, synthetic time series are generated, which can be used as virtual observations. In the next step, the method of D-optimal design is applied. The amount, frequency, period, place and kind of variables measured of the water samples that gives the most reliable estimates of the parameters of the model are considered to be the best observations that can be made for that river. Also, the percentage of improvement of the reliability can be defined, as a function of the observations. The method is applied to the river Dender.

Keywords Modelling; optimal experimental design; parameter estimation; reliability analysis; river water quality

Introduction
In many countries, samples of river water are taken to identify the present state of the water pollution. The sampling places, frequency and the type of quality variables to measure are mainly determined by practical and cost limitations. Typically, the frequency will be limited to 1 or 2 samples per month. Such a low frequency may prove to be insufficient, even if only a general statistical evaluation of the water quality is aimed at, for rivers with fast and complex dynamics (Vandenberghe et al., 2001). The problem becomes even worse, if data are to be used for the calibration of dynamic water quality simulation models – e.g. to predict the impact of future pollution abatement scenarios – as the process dynamics cannot be investigated on low-frequency data.

As the current water quality models are only partly physically based, several model parameters do not necessarily represent a real physical quantity and are, consequently, not universal. The identifiability of the model parameters and the reliability of the model depend on the amount and quality of the data available for calibration. The decision on the kind of measurements, the space distribution, the duration or periodicity and time resolution of the sampling is hereby not straightforward, as, in most cases, the amount of measurements is limited by the limitation of resources.

This paper presents a methodology for the definition of an optimal set of sampling data for the calibration of a water quality model. The methodology is applied to the Dender basin in Belgium.

The Dender Basin
The River Dender, a tributary of the River Scheldt, drains an area of 1,384 km². The flow of the river is very irregular with high peak discharges (100 m³/s) during rainfall storms and very low discharges (1 m³/s) during dry periods. To suit navigation and to temper the high
flows, the Dender is canalised and regulated by 14 sluices. Due to this, during dry periods the river reacts as a succession of reservoirs with a typical depth of 3 to 5 m, a width of 12 to 50 m and lengths of 2 to 8 km. In periods of high flow, all sluices are opened and the river regains its natural stream profile (Bervoets et al., 1989). The river is heavily polluted by domestic, industrial and agricultural pollution (Demuynck et al., 1997).

ESWAT
The model of the River Dender is implemented in ESWAT, an extended version of SWAT (Soil Water and Assessment Tool; Arnold et al., 1996). In ESWAT, an in-stream water quality model based on QUAL2E has been implemented. Moreover, the processes are represented on a sub-daily time base, to allow for applications to smaller river basins and for the simulation of the impacts of eutrophication (van Griensven and Bauwens, 2001).

The model was calibrated, using a multi-objective calibration technique, based on high frequency water quality observations during the last four months of 1994 (van Griensven et al., 2002). The water quality calibration objectives included oxygen, BOD, ammonia, nitrate and phosphate at three locations (upstream, middle and downstream) in the basin.

A sensitivity analysis showed that the water quality module implemented in ESWAT for the Dender requires the calibration of eight parameters (Vandenberghe et al., 2000). These parameters are the BOD de-oxygenation rate coefficient, the oxygen re-aeration rate, the benthic oxygen demand rate, the rate constant for biological oxidation of NH₄ to NO₂ and 4 parameters related to algae activity. As the latter could not be well identified during the calibration – due to a lack of measurements during periods with algae bloom – the latter parameters are not considered in what follows.

Optimal Experimental Design (OED)
Different experiments (sampling schemes) will reveal more or less information and more or less reliability, e.g. schemes that lack dynamics will provide less information than schemes with more curvature. Optimal sampling design techniques aim at the identification of sampling schemes to improve different facets of the mathematical modelling process, according to explicitly stated objectives. The objective considered here is to increase the precision of the parameters for the water quality module of ESWAT.

The applied OED method is based on the D-optimal criterion (Goodwin and Payne, 1977; Walter and Pronzato, 1999).

Hereby, the precision of the parameters is assessed by considering the determinant of the inverse of the covariance matrix of the parameter estimates (C) or Fisher Information Matrix (FIM) (Godfrey and Distefano, 1985):

\[ C(b) = \sigma^2(J'QJ) \]
\[ FIM(b) = C^{-1}(b) \]

with \( b \) representing the model parameter vector, \( Q \) a diagonal matrix, the elements being the squares of the observation weights and \( J \) the Jacobian matrix. Calculation of the covariance matrix based on the Jacobian matrix instead of the Hessian is acceptable when assuming linearity and assuming observations with constant standard deviations (Bard, 1974).

The determinant of the FIM, \( \text{Det}(\text{FIM}) \) is proportional to the volume of the confidence region. Thus, by maximising \( \text{Det}(\text{FIM}) \), the volume of the confidence ellipsoids, and, correspondingly, the geometric average of the parameter errors are minimised. D-optimal experiments also have the advantage of being invariant with respect to any scaling of the parameters (Petersen, 2000).

The FIM is obtained after a calibration of the model with optimisation software. For the research presented here, the PEST optimisation program, based on the Gauss-
Marquardt-Levenberg technique (Anonymous, 1994) was used for the calculation of the FIM.

The OED technique thus requires an initial data set to calibrate the model. One should hereby keep in mind that non-accurate parameter estimates may lead to an inefficient experimental layout. The design can only be approached by an iterative process of data collection and design refinement, known as a “sequential design” (Casman et al., 1988).

For a limited amount of experimental conditions that are well defined, it is possible to find the experiment that provides the maximum det(FIM), by considering all the possible experiments (Baetens, 2001). However, for a river water quality problem, the possible experimental schemes are unlimited. Therefore, in a second optimisation process, the design parameters of the sampling schemes (the amount, the place and frequency of sampling, the period of sampling, the kind of variables measured...) are being optimised by searching for the combination of design parameters that maximises det(FIM). For this study, this optimisation was performed with the shuffled complex optimisation method (Duan et al., 1992).

**Generation of a synthetic observation series**

The evaluation of different sampling schemes requires the availability of a long series of high frequency water quality data series at different places along the river. As such historic series were not available, synthetic “observation” series were generated by simulations with the ESWAT model and subsequently altered by addition of pseudo-random noise terms. The noise terms were generated, considering a normal distribution and variations that are consistent with the accuracy of the measuring devices used to measure the variables (Bols, 1999): 3% for Dissolved Oxygen (DO); 10% for Biological Oxygen Demand (BOD) and 5% for NO₃ and NH₄.

**Uncertainty analysis**

To evaluate final uncertainty in the model results due to the uncertainty of the parameters after calibration, an uncertainty analysis is performed with a Monte Carlo method. The 95% confidence bounds on the model results are calculated by sampling 100 parameter sets with the Latin Hypercube technique (McKay, 1988), considering uniform distributions for the parameters.

**Analysis of the results**

**The search for the optimal experimental design**

As an illustration of the applicability of the method, a simple case will be considered first, whereby only DO is considered at one specific location. The synthetic “observation” series consisted of 1 year of hourly data. The optimisation is thus limited to the time step, the amount of samples and the period of the year during which samples are taken. The sampling time step was allowed to vary between one hour and two days; the minimum amount of samples is 1 and the maximum amount is 365 × 24. Samples could be taken during winter, summer or a mixed summer–winter period, depending on the start of the period and the total amount of samples that are taken.

In Figure 1, the optimisation process is shown. The shuffled complex method used 136 runs to find the optimum for which the (Det(FIM))⁻¹ is smallest.

As could be expected, the results (Figure 2 and Figure 3) show that the uncertainty in the parameters was minimal for the smallest sampling time step, a very large amount of samples and a large period, mainly spring and summer months. A sample every hour, starting in February and ending August 30th, representing a total of 5,804 samples appears to provide the best results.
A second example shows a more complex case, whereby in addition to the previous variables, also the data type (only DO or combined DO-NO₃, DO-NO₃-BOD or DO-NO₃-BOD-NH₄) and the sample location (4 possible combinations of 3 possible locations: upstream, halfway, downstream) are considered as variables.

A substantial increase in the number of iterations for the optimisation is observed (Figure 4).

The best way to take samples is on an hourly time base, over nearly the whole year (8,730 samples), on two locations and with measurement of the four variables. This is again a very logical result.

However, looking at Figure 5 and Figure 6, one observes that other sampling schemes could be defined that provide a quasi similar accuracy, with fewer samples or a lower frequency. As can be seen in Figure 6, e.g. the confidence regions around the parameters do not differ very much in the range of 5,000 to 8,000 samples. This is explained by other factors that influence the accuracy, such as the period of the year during which the sampling takes place.
Alternatively, some sampling schemes clearly appear as non-optimal (such schemes are indicated by squares in Figure 6): these schemes require a lot of samples, but due to the wrong choice of other factors, the information content of these schemes is poor. More details on these schemes are given in Table 1.

The reason for the bad performance of these schemes is related to the sampling place (upstream) and to the fact that the sampling period does not include the spring period, which seems here to be important for the calibration process.

The search for the optimal experimental design including practical considerations

In order to evaluate sampling schemes, it is not so much the parameter uncertainty that is of importance, but the final uncertainty in the model results that is of importance. This uncertainty can then be further related to cost and other practical implications of the different schemes.

To illustrate the procedure, three sampling schemes from the first test case are considered (indicated by squares in Figure 2 and Figure 3). The schemes were chosen because they give clearly different results for the reliability in parameter estimations. More details about the schemes are given in Table 2.

The model outputs and the 95% confidence intervals for the considered schemes for an arbitrarily chosen day (22 February) are given in Figures 7, 8 and 9. The results of the uncertainty analysis show that the average width of the confidence interval in the model output is reduced by 45% for scheme 2 when compared to scheme 1 and by 60% if scheme 3 is compared to scheme 1.

Table 1 Non-optimal sampling schemes

<table>
<thead>
<tr>
<th>Sampling time step (hours)</th>
<th>Amount of samples</th>
<th>Period</th>
<th>Place</th>
<th>Observed variables</th>
<th>1/det(FIM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5,972</td>
<td>16 Apr.–31 Dec.</td>
<td>Upstream</td>
<td>DO-NO₃</td>
<td>2.45E-18</td>
</tr>
<tr>
<td>1</td>
<td>4,902</td>
<td>11 May–31 Dec.</td>
<td>Upstream</td>
<td>DO-NO₂-BOD</td>
<td>1.69E-21</td>
</tr>
<tr>
<td>1</td>
<td>4,154</td>
<td>22 May–15 Nov.</td>
<td>Upstream</td>
<td>DO-NO₂-BOD</td>
<td>8.41E-20</td>
</tr>
</tbody>
</table>

Table 2 Selected sampling schemes

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Sampling time step (hours)</th>
<th>Amount of samples</th>
<th>Period</th>
<th>1/det(FIM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>42</td>
<td>26 Oct.–31 Dec.</td>
<td>2.03.10E-15</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>818</td>
<td>23 Oct.–31 Dec.</td>
<td>5.9.10E-21</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>8,008</td>
<td>2 Feb.–30 Aug.</td>
<td>1.04.10E-23</td>
</tr>
</tbody>
</table>
The results illustrate the possibilities of the method to define a dedicated sampling strategy, in view of a given modelling accuracy.

Conclusions
It has been shown that OED methods can be used for an iterative, sequential design of a strategy for measuring water quality variables on a river, in view of the calibration of water quality models. In a first stage a relatively extensive set of measurements is needed to set up the model for the river. Using the model, the OED method enables the definition of efficient measurements strategies, to find better model parameter estimates and reduce the uncertainty in those estimates. In subsequent stages, the measurement strategy can be updated in an iterative way.

For the River Dender, it was shown that the method is able to define the most logical solution to the problem, if a maximal accuracy is aimed at; the optimal sampling strategy will be the one with the highest amount of samples and the highest sampling frequency, at the maximal number of locations and whereby a maximal number of variables are measured.

The usefulness of the method, however, resides in its ability to evaluate sub-optimal sampling strategies, whereby strategies are evaluated in view of the limitations of costs and
other practical considerations. By extending the OED method with a procedure for the definition of the modelling uncertainty, it thus becomes possible to define the optimal sampling strategy to obtain a given modelling accuracy.

Acknowledgements
The authors give special thanks to the Fund for Scientific Research (G.0102097), the EU CHESS project (ENV4-CT97-0440) and the Council of Research (OZR) of the VUB for the financial support.

References