INTEGRATED MODELLING AND DATA NEEDS: QUANTIFICATION BASED ON THE INTERACTIONS WITHIN THE WASTEWATER SYSTEM

J.G. Langeveld*, F.H.L.R. Clemens* and G.J.R. Henckens**

*Section of Sanitary Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology (DUT), P.O. Box 5048, 2600 GA Delft, the Netherlands
(E-mail: *J.G.Langeveld@citg.tudelft.nl*)
**Witteveen+Bos Consulting Engineers, P.O. Box 233, 7400 AE Deventer, the Netherlands

Abstract
The scientific pioneering with respect to integrated models has resulted in the widespread availability of integrated models. The practical application of integrated models, however, is still rather limited. This is partly due to the lack of (reliable) data necessary for calibration of the integrated models and the high expenses necessary to obtain this data. The design of monitoring networks for obtaining the data necessary is, however, a non-trivial task.
This paper is a first step towards a design method for monitoring networks for calibration of integrated models. The method builds on the knowledge on and experience with a design method for monitoring networks for the calibration of hydrodynamic models (Clemens, 2001a; 2001b; 2002) and the general recommendations for setting up measuring campaigns give by (Vanrolleghem et al.,1999). The original design method for hydrodynamic models is extended to monitoring networks for wastewater systems (comprising sewer systems and wastewater treatment plant) by introducing insight and research results from the aforementioned ongoing research at Delft University of Technology on the interactions within the wastewater system (Langeveld, in prep.).

Keywords
Integrated modelling, interactions within wastewater system, design of monitoring networks

INTRODUCTION

Integrated modelling has been and will be a topic receiving a lot of attention in urban water research. A large number of integrated models have been developed, ranging from research software to commercial model packages (Leinweber, 2002).
At the end of the 1990s, it has been recognised that simply connecting existing submodels results in over-complicated integrated models (Rauch et al.,1998). Instead, they proposed a problem oriented modelling approach that theoretically results in models taking only the dominant processes into account. This approach has been applied by Leinweber (2002), identifying (but not quantifying) requirements for sewer models. Langeveld (in prep.) continue on the same path and quantifies these requirements, thus giving measurable requirements for sewer process models.

Recently, (Rauch et al., 2002) concluded that the practical application of integrated models does not develop as expected. The lack of (reliable) data necessary for calibration of the integrated models and the high expenses necessary to obtain this data was seen as the main reason for the hampering of the widespread use of integrated models.

The problem of the high expenses is exarcebated by a lack of knowledge on the data needs themselves. In literature, a lack of data is often claimed, but no clear requirements for this data are formulated. Consequently, it is not known exactly what type and quality of data are lacking, which hampers the development of a satisfying monitoring network. Basically, the question to be answered is: what data is needed to be able to calibrate an integrated model? This question involves all aspects related to the design of monitoring networks: location, duration, frequency and parameter selection.

This paper is a first step towards a design method for monitoring networks for calibration of integrated models. The method builds on the knowledge on and experience with a design method for monitoring networks for the calibration of hydrodynamic models (Clemens, 2001a; 2001b; 2002) and the general recommendations for setting up measuring campaigns give by (Vanrolleghem et al.,1999). The original design method for hydrodynamic models is extended to monitoring networks for wastewater systems (comprising sewer systems and wastewater treatment plant) by introducing insight and research results from the aforementioned ongoing research at Delft University of Technology on the interactions within the wastewater system (Langeveld, in prep.).

METHODS AND MATERIALS

Monitoring networks for the calibration of integrated models

The calibration of (integrated) models requires measurement data to be able to adjust the model parameters correctly. Often, this data is either not available or does not contain sufficient information. Clemens (2001) developed a method for the design of monitoring networks for hydrodynamic models. Theoretically, this method can also be applied to monitoring networks for integrated models, as the method deals with basic issues of monitoring network design:

- required measuring accuracy
- measuring locations
- measuring frequency

Compared to monitoring networks for hydrodynamic models, the selection of the parameters to be measured is added. This selection depends on the required output of the integrated model, which is determined by emission standards or receiving water quality objectives.
Recent research (Langeveld, in prep.) has shown ammonium to be a very important parameter with respect to the interactions within the wastewater system. As ammonium is a dissolved substance which is dominantly contributed by the dwf, the ammonium concentration in a sewer system (and wwtp influent) during storm events is mainly determined by dilution. Describing dilution only is, however, not sufficient with respect to the simulation of the pollution load discharged by CSOs, as a large proportion of the pollutant load is associated with particles. Current knowledge on suspended solids transport, however, is still limited. Nevertheless, it is well known that suspended solids transport is correlated with hydraulic parameters like shear stress (Langeveld et al., subm.). Most available models describing suspend solids transport use transport formulae, e.g. the Ackers-White equation applied in Hydroworks, which incorporate a relation with the hydraulics. Consequently, even though current suspended solids transport models incorporated in commercial software do not necessarily provide good results, they can be used for monitoring network design as they will likely give a high transport activity at the right locations. (provided that the hydraulic performance of the sewer is described well).

As a first step towards the design of monitoring networks for an integrated model, the implications of the design of a monitoring network for sewer water quality models will be discussed. This discussion is based on a comparison of the results of designing a network for the three discussed processes:

- hydrodynamics
- solute transport/dilution
- suspended solids transport

Methodology for the design of monitoring networks

For the design of the monitoring network, the design method for monitoring networks developed by Clemens (2001; 2002) was used. This design method comprises the following working sequence:

1 Identify the locations in the system available for monitoring (selection on practical criteria like accessibility). This results in N_{max} locations.

2 Define the measuring accuracy and the sampling interval using time domain and frequency domain analysis. This step determines the selection of the measuring equipment.

3 Define the maximum number of gauging locations (normally limited by available budgets and manpower). This results in M_{max} locations, generally $N_{max}>>M_{max}$.

4 Identify the model parameters that are candidates for parameter optimisation, P_{max}. For a hydrodynamic model, these are typically parameters related to the inflow, i.e. runoff coëfficiënts and dwf, and hydraulics of special stuctures, i.e. weir coëfficiënts. For a water quality model, the number of parameters will likely increase. Calibrating solute transport may involve adjusting the dispersion coefficient, whereas with respect to sewer sediment of suspended solids transport parameters describing the erosion, deposition and transport may be included.

5 Run a model $P_{max}+1$ times. In each model run only one of the P_{max} parameters is adjusted with a fixed multiple (typically 5%) to be able to assess the sensitivity of the model result to parameter P.

6 Construct the Jacobean for N_{max} monitoring locations and P_{max} model parameters. (This Jacobean is referred to as $\underline{\underline{J}}_{total}$). The total Jacobean is constructed based on the Jacobean per potential measuring location. These Jacobeans, comprising the effect of model parameter variation on the model result have the following form:

$$\underline{\underline{J}}_i = \begin{bmatrix} \frac{\partial x_{t_0}}{\partial p_1} & . & . & . & \frac{\partial x_{t_0}}{\partial p_n} \\ \frac{\partial x_{t_1}}{\partial p_1} & . & . & . & \frac{\partial x_{t_1}}{\partial p_n} \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{\partial x_{t_{\max 0}}}{\partial p_1} & . & . & . & \frac{\partial x_{t_{\max}}}{\partial p_n} \end{bmatrix}$$

where x is the model result as a function of time and p_i is a model parameter. Effectively each row in the Jacobean contains information on the sensitivity of the model in a particular timestep for each individual parameter. The Jacobeans per potential measuring location are combined in a total Jacobean ($\underline{\underline{J}}_{total}$):

$$\underline{\underline{J}}_{total} = \begin{bmatrix} \underline{\underline{J}}_1 \\ \underline{\underline{J}}_2 \\ . \\ . \\ \underline{\underline{J}}_{N\max} \end{bmatrix}$$

7 Make a singular value decomposition for $\underline{\underline{J}}_{total}$ and judge the rank of $\underline{\underline{J}}_{total}$.

8 If the rank of $\underline{\underline{J}}_{total}<P_{max}$ or $\underline{\underline{J}}_{total}$ is nearly rank deficit for some parameters then identify these parameters and remove them from the parameter list, this results in $P_{max}-P_{rankdef}$ parameters.

9 Since M_{max} is normally significantly smaller than N_{max}, choices have to be made on which monitoring location are to be selected; this can be done in two manners.

a Study the contribution for each monitoring location for each parameter on the Jacobean and select the M_{max} locations having the largest contribution. Construct the Jacobean based on the selected monitoring locations and make a singular value decomposition in order to verify if this Jacobean is (nearly) rank deficit. If it is not, the parameterisation and the measuring setup match. If the Jacobean is (nearly) rank deficit, a further reduction in the number of parameters is needed. Alternatively step 2 can be reconsidered; an increase in measuring accuracy could help decrease the sampling interval making more information in the time domain available.

b Judge every possible combination of M_{max} out of N_{max} monitoring locations by the results of the singular value decomposition of the Jacobean associated with it. Pick the one combination with the largest minimum singular value. If this combination is still judged to be nearly rank deficit, either M_{max} should be increased (resulting in an increase of investments and operational costs) or a further reduction of the number of parameters must be induced. Alternatively step 2 can be reconsidered; an increase in measuring accuracy could help decrease the sampling interval making more information in the time domain available.

10 Perform auto calibration runs with different initial parameter vectors using the parameterization and the measuring setup obtained in the preceding 8 steps. If in the calibration process problems related to stability or non-uniqueness occurs, a further scrutiny of the steps 5-8 is advised.

This design method has been applied for the design of a hydraulic monitoring network for the sewer system of Loenen (the Netherlands), shown in figure 1 (Clemens, 2001; 2002).

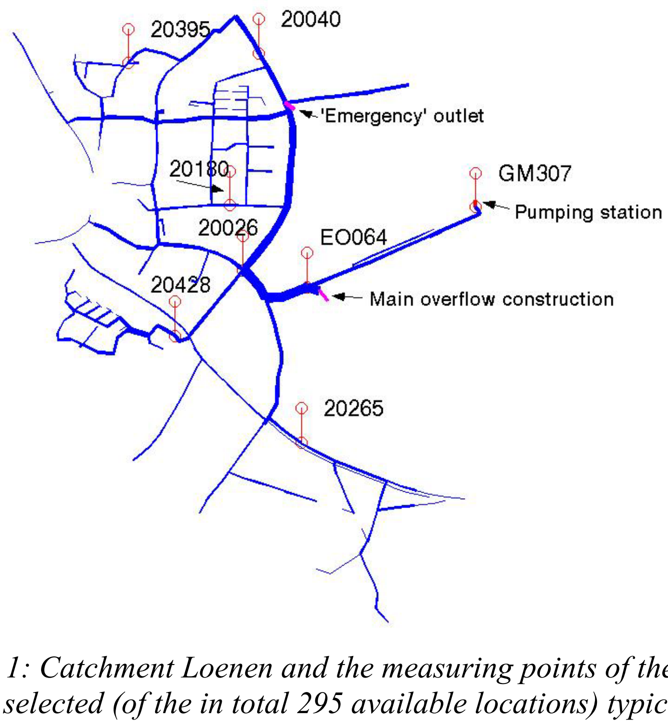

Figure 1: Catchment Loenen and the measuring points of the hydraulic monitoring network. The 8 gauges selected (of the in total 295 available locations) typically give 20 to 40 %, depending on the calibration parameter, of the total information content which can be obtained by measuring in all 295 locations (Clemens, 2002).

RESULTS AND DISCUSSION

Figure 2 shows the water level, dilution rate and sediment transport rate simulated in gauge 20026.The water level rises fast during the filling phase of the storm event and falls slowly during the emptying phase. Consequently, the water level changes constantly during the storm event. The dilution rate and sediment transport rate, however, only change significantly during the beginning and final phase of the storm event.

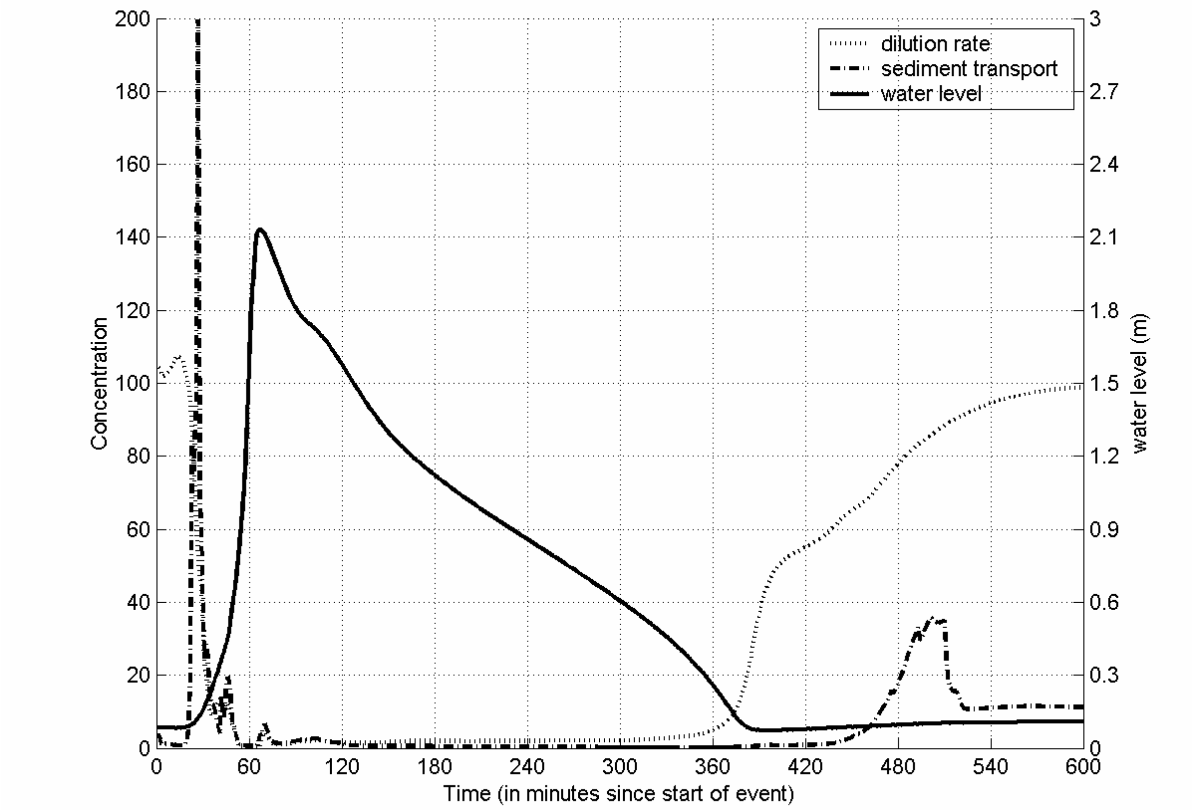

Figure 2: Water level, dilution rate and sediment transport manhole 20026, Loenen

The method for the design of monitoring networks is in its essence a sensitivity analysis. In this example, only the dwf was used as the model parameter to be varied, which causes minor changes in the simulated water level, as shown in figure 3. As only 1 model parameter was analysed, the Jacobean constructed comprises only 1 column and 600 rows (the number of time steps used) and the total Jacobean 1 column and 295 rows (the number of potential measuring locations). By applying a singular value decomposition, the information contained in the simulation result was analysed for the water level, the dilution rate and the sediment transport. In addition, a correlation analysis was applied to decorrelate the results, (see Henckens and Clemens (in press) for the decorrelation method applied). Decorrelation is necessary to prevent valuating the same information twice (this is due to the fact that e.g. 2 adjacent manholes contain to a large extent the same information).

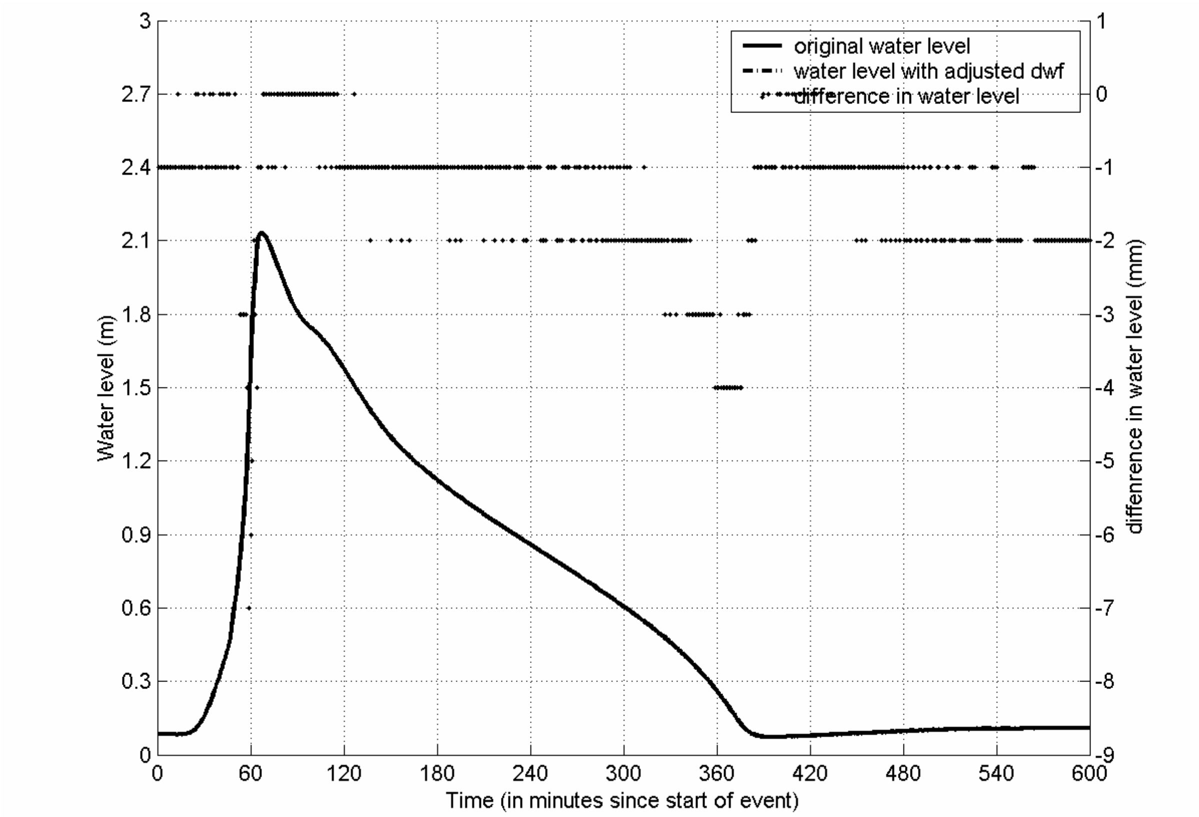

Figure 3: Sensitivity of simulated water level to the dwf, manhole 20026, Loenen

Table 1 gives an overview of the information content of the original network designed by Clemens (2001). With respect to the dwf, the hydraulic network almost delivers 40% of all available information, which is proven to be sufficient for calibrating hydrodynamic models (Clemens, 2002). With respect to dilution and sediment transport, the same monitoring network delivers only respectively 15 and 8 % of the total information content.

Table 1 Information content with respect to the model parameter dwf of the network shown in figure 1

process	**information content**
hydrodynamics	40%
dilution	15%
sediment transport	8%

The variation in information held by the monitoring network may be due to the fact that the concentration of the sewage does not change significantly over a long period of time during the event, which reduces the available information content. Another explanation is the fact that at the beginning of the storm event the concentration in the sewage is not yet stable. This can be seen in figure 2, where the concentration of the solutes is higher than 100% at the beginning of the storm event. In other words: the applied software introduces numerical problems, which affect the assessment of the information content.

Nevertheless, it can be concluded that the optimal measuring locations in networks designed for calibrating hydrodynamic models are not necessarily the optimal for water quality networks. Consequently, designing water quality monitoring networks in sewer systems (which is to be seen as a first step towards monitoring networks for wastewater systems aiming at the calibration of integrated models) likely results in a different optimal network (given the information content) per assessed water quality process.

In addition, it can be concluded that the design of networks for water quality can be based on the same procedure as developed for hydrodynamic models. This may enhance the development of knowledge on in-sewer processes, like suspended solids transport, via bootstrapping. Current, proven incorrect model approaches for suspended solids transport can be improved by the data from a monitoring network designed using these models. By going through this loop a number of times it may be possible to significantly improve current models.

In this paper the design of a monitoring network for sewer systems was used to illustrate the applicability of the procedure for the design of monitoring networks developed by (Clemens, 2001). However, with respect to the design of a network for an integrated model, the same procedure could be used, as this will only involve the inclusion of more potential measuring locations and a model for the wwtp. A practical limitation at this point in time is the fact that the software to design the network and to calculate the information content is not yet able to deal with other file formats than Hydroworks output files. Another practical limitation is the fact that the optimal monitoring network depends to a certain extent on the storm event selected. In order to eliminate this impact, the design has to be tested for a number of storm events, which may result in a large number of time-consuming simulations.

CONCLUSION

This paper discussed the applicability of the procedure for designing monitoring networks for the calibration of hydrodynamic models developed by (Clemens, 2001) to the design of water quality networks. Based on the first results it is concluded that the described procedure is also suited for the design of monitoring networks.
In addition, it is concluded that for each process or water quality parameter studied, the optimal configuration of the monitoring network can be different. This complicates implementing measuring networks in practice, as an increasing number of measuring locations inevitably increases the total costs.

ACKNOWLEDGEMENTS

The first author would like to gratefully acknowledge the funding by Arcadis in cooperation with HKV and Vertis, Grontmij and Witteveen+Bos.

REFERENCES

Clemens F.H.L.R. (2001a) *Hydrodynamic models in urban drainage: application and calibration.* PhD-thesis Delft University of Technology, Delft

Clemens, F.H.L.R. (2001b). A desing method for monitoring networks in urban drainage. In: Brashaer, R.W. and Maksimovic, C. Urban Drainage Modeling. Proceedings of the specialty symposium held in conjunction with the world water and environmental resources congress. May 20-24, 2001. Orlando, Florida

Clemens, F.H.L.R. (2002). Evaluation of a method for the design of monitoring networks in urban drainage. In: 9ICUD Ninth International Conference on Urban Drainage, Portland, Oregon, USA, 8-13 september 2002. CD-ROM, 17 pages

Henckens, G.J.R. and Clemens, F.H.L.R. (in press) Design and optimisation of monitoring networks in urban drainage. Accepted for the 19th European Junior Scientist Workshop.

Langeveld, J.G. (in prep.). *Interactions within the wastewater system*. PhD-thesis Delft University of Technology, Delft

Langeveld, J.G., Veldkamp, R.G. and Clemens, F.H.L.R. (subm.). Suspended solids transport: an analysis based on shear stress and turbidity measurements. submitted to: SPN'04, Madeira.

Leinweber, U. (2002). *Anforderungen an die Modellierung bei der integrierten Betrachtung von entwässerung und Kläranlage.* Dissertation Kaiserslautern. Germany

Rauch W., Aalderink H., Krebs P., Schilling W. and Vanrolleghem P. (1998). Requirements for integrated wastewater models – driven by receiving water objectives. *Wat.Sci.Tech*., **38**(11), 97-104.

Rauch, W., Bertrand-Krajewski, J.-L., Krebs, P., Mark, O., Schilling, W., Schütze, M. and Vanrolleghem, P.A. (2002). Deterministic modelling of integrated urban drainage systems. *Wat.Sci.Tech*. **45**(3), pp 81-94

Vanrolleghem, P.A., Schilling, W., Rauch, W., Krebs, P. and Aalderink, H. (1999): Setting up measuring campaigns for integrated wastewater modelling. *Wat.Sci.Tech.* Vol. 39. No.4, pp. 257-268