Influence of model parameter uncertainties on decision-making for sewer system management

H. Korving[*, **], P. van Gelder[***], J. van Noortwijk[, ****] and F. Clemens[*]**

*Department of Sanitary Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O.Box 5048, 2600 GA, Delft, The Netherlands.
j.l.korving@citg.tudelft.nl
**HKV Consultants, P.O.Box 2120, 8203 AC, Lelystad, The Netherlands.
***Department of Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O.Box 5048, 2600 GA, Delft, The Netherlands.
****Department of Mathematics, Faculty of Information Technology and Systems, Delft University of Technology, P.O.Box 5031, 2600 GA, Delft, The Netherlands.

Abstract Decisions on the rehabilitation of a sewer system are usually based on a single computation of CSO volumes using a time series of rainfall as system loads. Therefore, uncertainty in knowledge of model parameters is not taken into account. Moreover, statistical uncertainties are left aside. This paper presents the effect of uncertainties in model parameters on overall model results taking into account statistical uncertainties. It could even be argued that it does not matter whether the predictions from the model are uncertain. What matters is whether the decisions informed by these predictions are insensitive to such uncertainties. As an example the sewer system of 'De Hoven' (the Netherlands) is used. CSO volumes per storm event are computed using Monte Carlo simulations. In each Monte Carlo simulation of 1000 runs a different combination of fixed and variable model parameters is used. Probability distributions of computed CSO volumes are estimated taking into account the model uncertainties involved. The extent to which uncertainty in individual model parameters influences model results is quantified using relative confidence intervals of the statistical parameters of estimated distribution functions. The results show that the uncertainty in overall model results decreases the most if contributing areas are exactly known. Also the gain of model calibration is shown.

Keywords sewer system, parameter uncertainty, Monte Carlo analysis

Introduction

Decisions on investments in sewer rehabilitation often have to be made with uncertain information about the structural condition and the hydraulic performance of a sewer system. Because of this, decision-making involves considerable risks. A number of examples are known of sewer rehabilitations, which appeared to be dimensioned too large or too small or even unnecessary later on.

Decision-making on sewer system management requires the use of models to predict compliance of the system with performance criteria, i.e. Combined Sewer Overflow (CSO) volumes and flooding. However, in every model prediction lies uncertainty (see also Harremoës and Madsen, 1999). Besides, the decisions are usually based on a single model run. Uncertainty in model results mostly arises from model structure and estimation of calibration parameters, or can be linked to input observation errors or numerical errors due to the calculation method. Propagation of uncertainties in models is usually studied by means of

Hydroinformatics 2002: Proceedings of the Fifth International Conference on Hydroinformatics, Cardiff, UK

first-order variance propagation or Monte Carlo simulation analysis. But as stated by Reda and Beck (1997), the sensitivity to uncertainties of decisions taken on the basis of model predictions provides even more valuable information than the uncertainties alone.

This paper discusses the effect of uncertain model parameters on calculated CSO volumes. For the uncertainty analysis Monte Carlo simulation is used.

Uncertainties

Sources of uncertainty in sewer system modelling

Usually, the assessment of sewer system performance is based on calculated series of CSO volumes and flooding events. The question arises which elements in the modelling of sewer systems can be acknowledged as being uncertain and to what extent the modelling results are sensitive to these uncertainties.

Uncertainties in rainfall are the result of variability in time, spatial differences (catchment averaged rainfall) and measurement errors. For example, spatial variation in rainfall input accounts for approx. 30% of overall variability in model results (Willems, 2000). Dry weather flow (dwf) consists of sewage and leaking groundwater. It may show substantial variation in time due to varying dwf from households and industry during the day and leakage, which varies due to fluctuating groundwater levels and may add up to 50% of the dwf (Clemens, 2001).

The data set applied in a sewer model is never entirely perfect. Errors in this database (sewer system geometry, catchment area, runoff parameters, etc.) considerably influence calculation results of hydrodynamic models, especially errors in contributing areas and the structure of the sewer system (Clemens, 2001).

Models in urban drainage consist of two separate process descriptions: rainfall runoff process and in-sewer hydraulic processes. The rainfall runoff model comprises wetting of dry surface, infiltration, depression storage, evaporation and overland flow. The processes are strongly simplified in the hydrologic model and the uncertainties stem from variability of runoff in time, local differences in surfaces, lack of data and insufficient knowledge of the processes. The output of the rainfall runoff model is the input of the hydraulic model. In general, uncertainties in results of the hydraulic model stem from an incomplete or incorrect description of processes, errors in the database of the sewer system and numerical or software errors. Calibration of the model reduces uncertainties in model results.

Coping with the above-mentioned uncertainties requires classification, since reduction of each type of uncertainty requires its own approach.

Types of uncertainty

According to Van Gelder (2000) uncertainties can primarily be divided in two categories:

- Inherent uncertainty (Table 1): Uncertainties that originate from variability in known (or observable) populations and therefore represent randomness in samples (e.g. measured rainfall volumes). For example, even in the event of sufficient data, one cannot predict the maximum rain intensity that will occur next year. The two main types of inherent uncertainty are inherent uncertainty in time (e.g. variations of rainfall intensities in time) and inherent uncertainty in space (e.g. fluctuations in local terrain slope).
- Epistemic uncertainty (Table 1): Uncertainties that originate from limited knowledge of fundamental phenomena, e.g. rainfall-runoff processes or in-sewer processes

(Ashley et al., 1998). The two main types of epistemic uncertainty are model uncertainty (due to lack of understanding of the physics) and statistical uncertainty (due to lack of sufficient data). In general, epistemic uncertainties can be reduced as knowledge increases and more data becomes available.

Table 1. Types of uncertainty. Uncertainties can primarily be divided in inherent and epistemic uncertainty (Van Gelder, 2000). The latter consists of model and statistical uncertainty.

uncertainty	inherent		time
			space
	epistemic	model	model parameter
			model structure
		statistical	statistical parameter
			distribution type

Model

In this case study, the sewer system is modelled as shown in Figure 1, a reservoir with an external weir and a pump. The rainfall runoff part consists of the standard rainfall runoff model in the Netherlands (NWRW 4.3 model). In this model evaporation, infiltration, storage on street surfaces and overland flow are modelled as described in (among others) Clemens (2001). As system loads a 10-year rainfall series (1955-1964) from De Bilt (the Netherlands) is used. Dwf is ignored in the model.

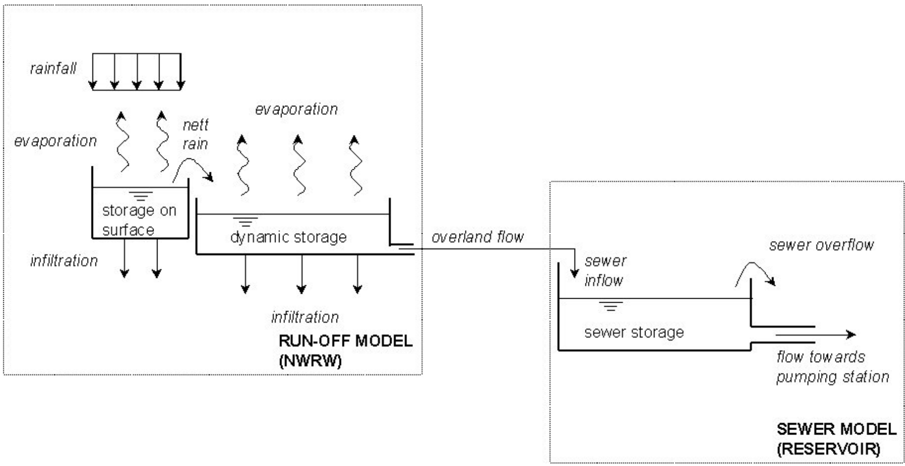

Figure 1. Sewer model. It comprises a rainfall runoff model and a reservoir model with an external weir and a pump.

The sewer system ‘De Hoven’ (12.69ha) is situated in the Netherlands on the banks of the river IJssel in the city of Deventer. The sewer system (865m^3) is of the combined type and comprises one pumping station (119m^3/h) transporting the sewage to a treatment plant and three CSO structures.

Analysis of uncertainties in model parameters

The model predicts CSO volumes as a function of time. The results are affected by uncertainty, i.e. variability of rainfall in time, model parameter uncertainty due to variability in parameter values of the reservoir model, model structure uncertainty due to the modelling (rainfall runoff model and reservoir model), statistical parameter uncertainty due to the small number of observed CSOs each year and distribution type uncertainty due to the distribution type to be chosen. Uncertainty analysis in water quality modelling is reviewed by Beck (1987).

In this paper, the influence of model parameter uncertainty on overall model results is analysed with Monte Carlo simulation. This is a simple and straightforward (but computer intensive) method that does not require the model to be linear as first-order analysis does. The Monte Carlo method has been applied in the field of urban drainage by several authors e.g. Lei and Schilling (1994), Reda and Beck (1997) and Grum and Aalderink (1999).

The Monte Carlo technique implies sampling from a priori distributions of the input variables and successive simulation for all sampled inputs. An estimate of the overall uncertainty is obtained by statistical analysis of the model outputs of all runs. A major disadvantage of the technique is that a large number of simulation runs is needed to obtain a reliable estimate of overall uncertainty. In particular for complex models this may result in long calculation times. To reduce the number of runs importance sampling, Latin Hypercube Sampling or parallel computing can be used.

Monte Carlo simulations

Monte Carlo simulations are conducted for each of the 12 parameter combinations in Table 2 using the reservoir and rainfall runoff model (Figure 1) and a 10-year rainfall series (1955-1964) from De Bilt, The Netherlands. Each simulation consists of 1000 runs, with parameters selected randomly from the ranges presented in Table 3, assuming normal distributions. The ranges selected from values reported by Clemens (2001). The following model parameters are made either variable or fixed: storage volume (S), pumping capacity (pc), contributing area (A) and weir coefficient (CC).

Table 2. Combinations of variable and fixed model parameters as applied in the Monte Carlo simulations.

Parameter Combination	S (m^3)	pc (m^3/h)	A (ha)	CC ($m^{0.5}$/s)
A	Fixed	Fixed	Fixed	Fixed
B	*Var.*	*Var.*	*Var.*	*Var.*
C1	*Var.*	Fixed	Fixed	Fixed
C2	Fixed	*Var.*	Fixed	Fixed
C3	Fixed	Fixed	*Var.*	Fixed
C4	Fixed	Fixed	Fixed	*Var.*
D1	Fixed	*Var.*	*Var.*	*Var.*
D2	*Var.*	Fixed	*Var.*	*Var.*
D3	*Var.*	*Var.*	Fixed	*Var.*
D4	*Var.*	*Var.*	*Var.*	Fixed
Residuals [1]	Fixed	Fixed	Fixed	Fixed
DeBilt 55-79 [2]	*Var.*	*Var.*	*Var.*	*Var.*

[1] residuals of calibrated model instead of variable model parameters.
[2] 25-year rainfall series (De Bilt, 1955-1979) instead of 10-year series.

Model parameters are changed in different combinations of fixed and variable parameters (Table 2), which enables not only quantification of the joint influence of all parameters being variable, but also of the influence of one variable parameter. In addition to these simulations with the 10-year rainfall series a simulation with a 25-year series is made. Moreover, the benefits of model calibration are studied using residuals of a calibrated model instead of the above-mentioned variability in model parameters. Calibration combines the uncertainties from different sources (e.g. database of sewer system, dwf, runoff parameters) in the resulting uncertainties of the calibration parameters.

Table 3. Variations in model parameters (from: Clemens, 2001).

Model parameter	μ	σ	CV (=σ/μ)
S (m3)	865.0	43.25	0.05
pc (m3/h)	119.0	5.95	0.05
A (ha)	12.69	0.64	0.05
CC (m0,5/s)	1.40	0.35	0.25

Results and discussion

The calculated CSO volumes from the Monte Carlo simulations are summed over the storm events and analysed statistically. Using Bayes weights the distribution function with the best fit to these CSO data is chosen. This Bayesian approach to distribution type selection is applied e.g. in Van Noortwijk *et al.* (2001). It quantifies both inherent and statistical uncertainty. Exponential, Rayleigh, normal, lognormal, gamma, Weibull and Gumbel distributions are considered. The Weibull distribution appears to fit best with the data (largest Bayes weight) and is chosen to describe the CSO volumes per storm event statistically (Korving *et al.*, subm.).

Given the CSO data $\mathbf{x} = (x_1,\ldots,x_n)$ the shape parameter a and the scale parameter b of a Weibull distribution,

$$f(x) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} \exp\left\{-\left(\frac{x}{b}\right)^{a}\right\} \quad a > 0, b > 0 \tag{1}$$

are estimated using the Maximum Likelihood (ML) method. Those values of a and b are chosen for which the likelihood function is maximised.

Table 4. Averages and confidence intervals of the parameters of Weibull distributions fitted to CSO data resulting from Monte Carlo simulations with 12 different parameter combinations (Tabel 2).

	A	B	C1	C2	C3	C4	D1	D2	D3	D4	Resid	DeBilt 55-79
μ(a)	0.826	0.786	0.781	0.809	0.784	0.823	0.787	0.784	0.779	0.787	0.802	0.832
$\frac{a_{97,5\%}-a_{2,5\%}}{\mu(a)}$		0.284	0.199	0.170	0.242	0.044	0.259	0.264	0.234	0.282	0.168	0.242
μ(b)	4.489	4.327	4.309	4.454	4.345	4.476	4.343	4.331	4.300	4.333	4.397	5.191
$\frac{b_{97,5\%}-b_{2,5\%}}{\mu(b)}$		0.405	0.274	0.145	0.357	0.046	0.346	0.417	0.306	0.400	0.142	0.319

On the basis of estimated parameters a and b in the Weibull distributions their relative 95% confidence intervals are computed for each combination of fixed and variable model parameters. This interval is considered indicative for the uncertainty in overall model results due to the variability in the model parameters. The results in Table 4 show that:

- $\mu(a)$ and $\mu(b)$ decrease when comparing combination A (all parameters fixed) to B (all parameters variable). This means that the results of a model with mean values of the model parameters do not correspond with the average effect of parameter variation on these model results.

- the relative confidence interval is reduced the most when only weir coefficient (CC) is varied or only contributing area (A) is fixed. From this we can conclude that uncertainty in CC has the smallest influence on model results, whereas uncertainty in A has the largest influence.
- a calibrated model reduces uncertainties the most (combination Resid.). The need to calibrate sewer models has already been stressed by e.g. Price and Catterson (1997).
- using 25 years of rainfall data (combination DeBilt 55-79) reduces uncertainties in model results to the same extent as using a fixed (known) contributing area (combination D3).

Conclusions

The objective of this paper is to describe the influence of parameter uncertainties on overall model results. The analysis performed takes into account the uncertainties in the parameters of a reservoir model of a sewer system. It leads to the conclusion that, in the particular case study, the uncertainty in model results decreases the most if the contributing areas are exactly known. Also the gain of model calibration is shown from the perspective of decision-making for sewer management. Since the distribution type is chosen on the basis of Bayes weights statistical distribution type uncertainties are minimised.

Acknowledgements

This paper describes the results of a research, which is financially supported by and carried out in close co-operation with HKV Consultants and the RIONED Foundation. The authors would like to thank HKV and RIONED for their support.

References

Beck M.B. (1987) Water quality modeling: a review of the analysis of uncertainty. *Water Resources Research.* **23**(8), 1393-1442.

Clemens, F.H.L.R. (2001) *Hydrodynamic Models in Urban Drainage: Application and calibration.* PhD thesis. Technische Universiteit Delft, Delft.

Grum, M. and Aalderink, R.H. (1999). Uncertainty analysis of combined sewer overflow effects using embedded Monte Carlo simulations. *Water Science and Technology.* **39**(4), 233-240.

Harremoës, P. and Madsen, H. (1999) Fiction and reality in the modelling world – Balance between simplicity and complexity, calibration and identifiability, verification and falsification. *Water Science and Technology.* **39**(9), 47-54.

Korving, H., Clemens, F., Van Noortwijk, J. and Van Gelder, P. (submitted) Bayesian estimation of return periods of CSO volumes for decision-making in sewer system management. Submitted to: *9th Int. Conf. on Urban Drainage.* September 2002, Portland, Oregon, USA.

Lei J.H. and W. Schilling (1994) Parameter uncertainty propagation analysis for urban rainfall runoff modelling. *Water Science and Technology.* **29**(1-2), 145-154.

Price R.K. and Catterson, G.J. (1997) Monitoring and modelling in urban drainage. *Water Science and Technology.* **36** (8-9), 283-287.

Reda, A.L.L. and Beck, M.B. (1997) Ranking strategies for stormwater management under uncertainty: sensitivity analysis. *Water Science and Technology.* **36**(5), 357-371.

Van Gelder, P.H.A.J.M. (2000) *Statistical Methods for the Risk-Based Design of Civil Structures.* PhD thesis. Technische Universiteit Delft, Delft.

Van Noortwijk, J.M., Kalk, H.J., Duits, M.T. and Chbab, E.H. (2001) The use of Bayes factors for model selection in structural reliability. In: *Proc. of 8th Int. Conf. on Structural Safety and Reliability (ICOSSAR).* June 2001, Newport Beach, California, USA.

Willems P. (2000) *Probabilistic Immission Modelling of Receiving Surface Waters.* PhD thesis. Katholieke Universiteit Leuven, Leuven.