

Probabilistic modelling of combined sewer overflow using the First Order Reliability Method

Thorndahl, Søren; Schaarup-Jensen, Kjeld; Jensen, Jacob Birk

Published in:
Sewer Processes and Networks 5th International Conference

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Thorndahl, S., Schaarup-Jensen, K., & Jensen, J. B. (2007). Probabilistic modelling of combined sewer overflow using the First Order Reliability Method. In M. V. D. Meulen, J. A. E. T. Veldhuis, & R. P. S. Schilperoot (Eds.), *Sewer Processes and Networks 5th International Conference: August 29-31, 2007 : Delft, the Netherlands : Conference proceedings* (pp. 91-100). Delft University of Technology, Department of Sanitary Engineering.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Probabilistic modelling of combined sewer overflow using the First Order Reliability Method

S. Thorndahl, K. Schaarup-Jensen, and J. B. Jensen

Department of Civil Engineering, Aalborg University, Sohngaardsholmsvej 57, 9000 Aalborg, Denmark
(E-mail: st@civil.aau.dk; ksj@civil.aau.dk; jbj@civil.aau.dk)

Abstract This paper presents a new and alternative method (in the context of urban drainage) for probabilistic hydrodynamical analysis of drainage systems in general and especially prediction of combined sewer overflow. Using a probabilistic shell it is possible to implement both input and parameter uncertainties on an application of the commercial urban drainage model MOUSE combined with the probabilistic *First Order Reliability Method (FORM)*. Applying statistical characteristics on several years of rainfall, it is possible to derive a parameterization of the rainfall input and the failure probability and return period of combined sewer overflow to receiving waters can be found.

Keywords Combined sewer overflow; First order Reliability Method (FORM); Uncertainties; Monte Carlo Sampling; Urban drainage modelling; MOUSE.

INTRODUCTION

Hydrodynamic urban drainage models for load prediction of drainage systems are frequently used by consulting engineers to determine if the system in question maintains the requirements defined by the authorities. The purpose of modelling is mainly to determine the number of failures in an urban drainage system during a given period of time, i.e. to attach return periods to different occurrences in the system, e.g. surcharge, flooding, or combined sewer overflow to receiving waters. However, inputs (boundary conditions), parameters, model structure, etc. are encumbered with uncertainties causing model outputs to be uncertain which affects the reliability of the return periods.

Defining an occurrence of a combined sewer overflow as a system failure, the aim of the paper is to determine the system failure probabilities and return periods. To quantify these, standard approaches make use of simulation of design storms or long historical rainfall series in a hydrodynamic model of the urban drainage system. In this paper, an alternative probabilistic method, the *First Order Reliability Method (FORM)*, is investigated. To apply this method, a long rainfall time series is divided in rain storms (rain events) and each rain storm is conceptualized to a synthetic rainfall hyetograph by a Gaussian shape with the parameters rain storm depth and duration (Thorndahl and Willems, 2007; Willems, 2001). Using a hydrodynamic simulation model, the failure conditions for each set of variables are predicted. The method takes into account the uncertainties involved in the rain storm parameterization and uncertainties related to the measurement of the rain as well as the geographical variation. In addition to these input uncertainties, a number of hydrological and hydrodynamical variables are selected and handled stochastically. In order to validate the *FORM* approach the analysis is also conducted using a Monte Carlo Direct Sampling (*MCDS*) technique and a Monte Carlo Importance Sampling (*MCIS*) technique.

FORM has been extensively applied within the area of structural engineering and building technology (Ditlevsen and Madsen, 1996; Madsen et al., 1986; Melchers, 1999), and to some extent

within the area of groundwater and river modelling as well as water quality modelling (Sørensen and Schaarp-Jensen, 1995, 1996; Schaarp-Jensen and Sørensen, 1996; Portielje et al., 2000), but as far as the authors know only in the context of urban drainage in Thorndahl and Willems (2007).

METHODOLOGY

The concept of *FORM* is to find the probability of failure of a component in a given system. In the predefined probability distributions for each variable, the *FORM* algorithm searches for the combination of variable values which are most likely to cause failure of the system.

This approach is unique compared to traditional long term simulations of drainage systems as a parameterization of the rainfall input is conducted. Thus, it is possible to determine the frequency of combined sewer overflow (with uncertainty assessment) using much less computation time. Moreover, it is possible to add statistically based uncertainties to the rainfall input, which is traditionally difficult to apply to real measured rainfall input time series.

The First Order Reliability Method

The present paper does not present the specific details of the *FORM* algorithm. For further details see Thorndahl and Willems (2007). In a model setup with i random variables the limit of the failure space is also called the failure function and is defined as an i -dimensional surface. In *FORM* this multidimensional failure surface is approximated by a hyperplane in a standard normal space. The point on the hyperplane in which the failure probability is the highest (corresponding to the vector of variable values which is most likely to occur) is labelled the design point (\mathbf{x}^*). In Figure 1 a theoretical example of a two-variable *FORM* analysis is shown. As two variables are applied, the failure surface is approximated by a line. Failure is defined whenever the maximum water level in the combined sewer overflow structure exceeds the overflow crest level:

$$g(\mathbf{x}) = H_{crit}(\mathbf{x}) - H_{max}(\mathbf{x}) = 0 \quad (1)$$

$H_{max}(\mathbf{x})$ is the maximum water level in the overflow structure, and $H_{crit}(\mathbf{x})$ is the crest level. \mathbf{x} is a vector of random variables. From this, a value of the failure function g smaller than zero corresponds to failure (overflow). The failure probability, P_f , within one rainfall event in the observation period P_t is defined as (Melchers, 1999):

$$P_{f,FORM} = P(g(\mathbf{x}) \leq 0) = \Phi(-\beta) \quad (2)$$

Φ is the standard normal distribution function, and β is the Hasofer & Lind reliability index, which is the minimized distance perpendicular from the linearized failure surface (the point with the highest joint probability density) to the origin in a standard normal space (cf. Figure 1 left). β represents the point with the largest failure probability, given the probability distributions of \mathbf{x} . *FORM* is based on standard normal independent variables, and all variables in the standard normal space (\mathbf{u}) are transformed into in the real space (\mathbf{x}) using inverse transformation. From an initial guess of \mathbf{u} , a transformation to the \mathbf{x} -space is performed and the MOUSE model is evaluated with these values. *FORM* is based on an iteration procedure in which new values of \mathbf{u} are calculated until convergence of β and \mathbf{x} (or \mathbf{u}) is obtained. The gradient vector of the failure surface is found using a central finite difference approximation for every variable. This means that the model must be executed twice for every variable. The return period of the event corresponding to the design point is calculated by:

$$T = \left(P_f \cdot \frac{E}{P_t} \right)^{-1} \quad (3)$$

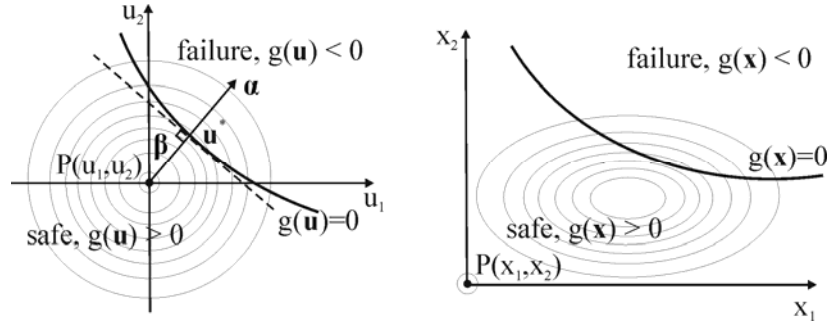


Figure 1 Example of failure surface, design point and contours of the joint probability density function in *FORM* with *two* variables. Left: Failure surface and linear approximation in a standard normal space. Right: Failure surface in the real space.

E is the number of rainfall events in a given period of time P_t and the number of failures per time period f_p (most often in years) is calculated by:

$$f_p = T^{-1} \quad (4)$$

A disadvantage of *FORM* is that it requires good initial guesses of variable values (especially if more than two variables are implemented). The algorithm often finds a local minima on the failure surface instead of the global one. Therefore, *FORM* is tested using a Monte Carlo direct sampling (*MCDS*) technique. Random values are sampled from the standard normal distribution and transformed to the real space as explained above. The failure probability is then calculated (Melchers 1999):

$$P_{f,MCDS} = \frac{1}{N} \sum_{n=1}^N I(g(\mathbf{x}) \geq 0) \quad (5)$$

N is the total number of simulations, and I is an indicator function ($I=1$ if failure and $I=0$ if no failure). Simulations are performed until convergence on P_f . This approach can be used to validate the linear approximation of *FORM*, as the whole variable space is simulated. The *MCDS* approach samples the whole variable space, but in order to test the linear approximations in *FORM* it can be advantageous only to sample around the design point (\mathbf{u}^*) with a specified standard deviation (σ), using a Monte Carlo Importance Sampling (*MCIS*) technique (Sørensen 2004):

$$P_{f,MCIS} = \frac{1}{N} \sum_{n=1}^N I(g(\mathbf{x}) \geq 0) \cdot \frac{f_U(\sigma \cdot \mathbf{u}_n + \mathbf{u}^*)}{f_U(\mathbf{u}_n)} \cdot \sigma^n \quad (6)$$

f_U is the joint density function of the standard normal density functions

Sensitivity measures

Using *FORM*, it is possible to define two different sensitivity measures to determine the relative sensitivity of every variable regarding the model output (Melchers, 1999):

- The α -vector is a unit normal vector to the failure surface at the design point, cf. Figure 1, left. α_i^2 is a measure of the percentage of the total uncertainty associated with the stochastic variable i . The sum of all α_i^2 equals 1.
- The omission sensitivity factor (ζ_i) determines the relative importance of the failure probability by assuming that the stochastic variable i is fixed, i.e. it is considered deterministic (Madsen, 1988):

$$\zeta_i = \frac{1}{\sqrt{1 - \alpha_i^2}} \quad (7)$$

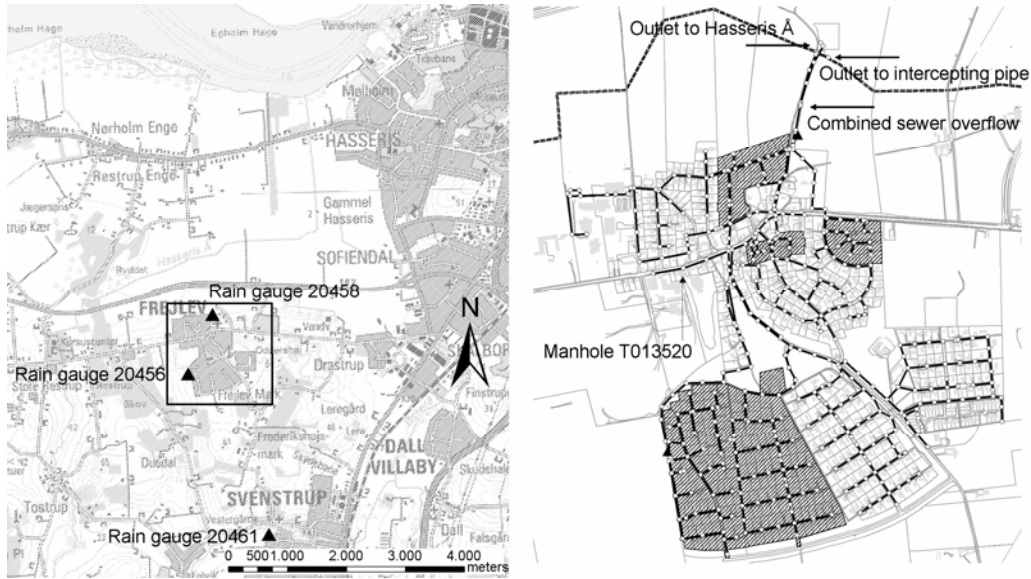


Figure 2 Left: The Frejlev catchment and surroundings. The black triangles are rain gauges. Right: Close up of the Frejlev catchment. The hatching marks the areas with separate sewer system and the rest is a combined system.

Setup of the MOUSE model and randomization of variables

The commercial urban drainage model MOUSE 2005 (DHI 2005) features advanced hydrological and hydraulic simulations of a complete urban catchment and drainage system. The model setup applied in this paper is based on a well calibrated setup of the Frejlev catchment in the northern part of Denmark, as described in Thorndahl et al. (2006); Thorndahl and Schaarup-Jensen (2007); Schaarup-Jensen et al. (2005). The choice of variables in this paper is also based on these references. The model is divided in two sub models, the surface runoff model and the pipe flow model. The hydrological part of the surface runoff sub model is governed by two parameters: 1) the *hydrological reduction factor* (ϕ) determining the part of the impervious area contributing to the runoff and 2) the *initial loss* (i) which is the hydrological loss due to wetting and filling of terrain depressions. These two parameters are considered global variables, i.e. the same value is implemented for every catchment. The *hydrological reduction factor* is also used to implement an error term on the rainfall input, see paragraph: *Conceptualization of rainfall input*. The flow routing on the surface can be modelled in different ways using the MOUSE model. In this paper the *Time Area model* is applied. This is based on a constant *concentration time* on the surface (DHI, 2005). Values of this variable are sampled from a uniform distribution based on Thorndahl (2007).. All variables and distributions are presented in Table 1. The geometrics of the drainage system are based on technical maps from the Municipality of Aalborg. In the model setup these values are kept deterministic. Parameters related to the loss of energy are made stochastic, i.e. the friction loss in pipes (*the Manning number*) and the headloss in outlets from manholes. The pipes in the drainage system are of different materials with different roughness, e.g. plastic, smooth or normal concrete. The Manning number is considered a global variable, and therefore these are drawn fully correlated and normally distributed. Values of the *headloss factor* are also drawn fully correlated depending on whether the outlet is round edged or sharp edged, but as no preferences are given to the distribution of the variable, a uniform distribution is applied.

Conceptualization of rainfall input

In Thorndahl and Willems (2007) the rainfall event duration t_d and the rainfall event depth d are parameterized by two-component exponential distributions using 18 years of data from the Svenstrup rain gauge (no. 20461, Figure 2). It is shown that these two variables are the most decisive in modelling of combined sewer overflow. This is reasonable, at least in the Frejlev

catchment, as a large inline retention basin is located just upstream from the overflow structure. This pipe basin fills up slowly and smoothes out the hydrographs, neglecting the peaks. On the contrary Thorndahl and Willems (2007) shows that rain intensity peak values for different aggregation levels are decisive for modelling of surcharge and flooding in manholes.

By sampling correlated values from the exponential distributions it is possible to generate synthetic rain storm events with a truncated Gaussian shape (Willems, 2001). As this synthetic event is obviously a simplification of the real events, Thorndahl and Willems (2007) investigated the errors in the maximum water level prediction (E_{Hmax}), introduced by this conceptualization. It was found that the errors could be parameterized by a normal distribution, in which the data was transformed with a Box-Cox transformation to account for heteroscedasticity. One of the advantages of modelling with synthetic rainfall events is the possibility to implement uncertainties on the rainfall input. Two types of rainfall input uncertainty are considered in this paper. The first is the

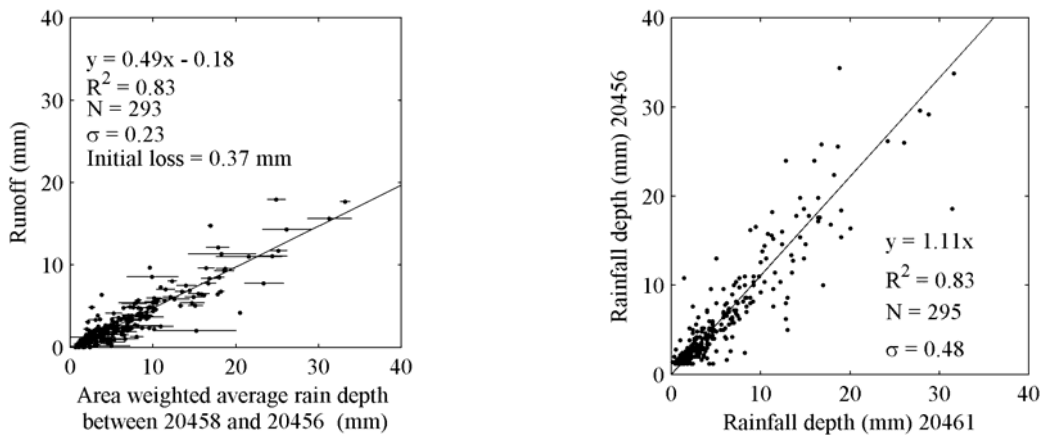


Figure 3 Left: Calculation of the hydrological reduction factor and initial loss from an area weighted rain depth of two rain gauges. The runoff is calculated as the runoff volume per event divided by the impervious area. Horizontal lines indicate the individual values of the two gauges. Right: Event depth correlation between rain gauge 20456 (within the Frejlev catchment) and gauge 20461 (3.5 km from the centre of Frejlev).

Table 1 Variables and chosen probability distributions.

No.	Variable	Distribution	Parameters
1	Rainfall event duration, t_d (min)	2-comp. exp.	$\beta_1=160$, $\beta_2=50$, $p=0.57$
2	Rainfall event depth, d (mm)	2-comp. exp.	$\beta_1=7.0$, $\beta_2=1.8$, $p=0.20$
3	Error on rainfall event depth, E_d (mm)	Normal	$\mu=0$, $\sigma=0.48$
4	Water level error, overflow structure, E_{Hmax} (m)	Normal	$\mu=0.002$, $\sigma=0.003$
5	Hydrological reduction factor, ϕ (-)	Normal	$\mu=0.49$, $\sigma=0.23$
6	Initial loss, i (mm)	Uniform	$x_{min}=0$, $x_{max}=0.001$
7	Surface concentration time, t_c (min)	Uniform	$x_{min}=1$, $x_{max}=10$
8	Manning number, M ($m^{1/3}/s$)		
	Smooth concrete	Normal	$\mu=85$, $\sigma=5$
	Normal concrete	Normal	$\mu=75$, $\sigma=5$
	Rough concrete	Normal	$\mu=68$, $\sigma=5$
	Plastic	Normal	$\mu=80$, $\sigma=5$
9	Headloss factor, K_m (-)		
	Round edged outlet	Uniform	$x_{min}=0$, $x_{max}=0.5$
	Sharp edged outlet	Uniform	$x_{min}=0.25$, $x_{max}=0.75$

uncertainty introduced by not applying a geographical variability over the catchment. This uncertainty is implemented implicitly within the hydrological reduction factor (Figure 3), as the scatter around the regression line obviously is due to imperfectly uniform distributed rainfall events. The scatter is fitted to a normal distribution, cf. Table 1. The second type of rainfall input uncertainty is the uncertainty introduced when using a rain gauge which is not located within the catchment. The Svenstrup rain gauge (20461, Figure 2) is used for the parameterization of the rain as it is the longest of the local series. Using this gauge entails a small uncertainty due to its placement approx. 3.5 km from the centre of Frejlev. This uncertainty is investigated in Figure 3 (right), in which the rainfall depths from gauge 20461 are plotted against the depths from 20456, in a 9 year period. It is obvious that there is a small bias as well as some scatter. This error (E_d) is implemented as an additional variable added to the synthetic rainfall depth d , as the scatter can be fit to a normal distribution. This variable clearly accounts for some of the error in the geographical distribution of the rain fall input, thus this type of input uncertainty is treated as a lumped uncertainty.

RESULTS

The first order reliability method finds the design point (\mathbf{x}^*), i.e. the set of variable values with the highest failure probability in terms of combined sewer overflow, corresponding to the values found in Table 2. It is seen that the rainfall event with the highest failure probability or smallest return period has a duration of 53 min. and a depth of 3.9 mm. The optimum of the hydrological reduction factor ($\varphi = 0.54$) is somewhat larger than the mean value ($\varphi = 0.49$). This is due to the large correlation with the rain depth. In order to maintain the same runoff volumes a small rain depth, which has the highest probability (cf. the exponential distribution), will cause a high value of the reduction factor. Examining the two sensitivity measures, the most important variable is by far the rainfall depth, which constitutes 92 % of the total uncertainty. Subsequently, the rainfall duration and the hydrological reduction factor represent approx. 2 % and 6 % of the total uncertainty, respectively. The other variables are negligible in terms of combined sewer overflow. However, if this analysis was conducted on flooding in a specific manhole instead of overflow, the first and second order variables, concentration time, Manning number, and headloss is expected to be more important, as they are more decisive for the hydrograph peaks.

During the analysis it was observed that *FORM* requires a good choice of initial values in order to find the design point as the global minimum. Especially for the two variables concerning the rainfall input, as small changes in the values cause a great change in the probability due to the exponential distributions.

Table 2 Results of the *FORM* analysis. x_i^* is the variable values in the design point and α_i^2 and ζ_i are sensitivity measures.

No.	Variable	x_i^*	α_i^2	ζ_i
1	Rainfall event duration, t_d (min)	52.60	0.0186	1.0094
2	Rainfall event depth, d (mm)	3.91	0.9163	3.4564
3	Error on rainfall event depth, E_d (mm)	0.058	0.0095	1.0048
4	Water level error, overflow structure, E_{Hmax} (m)	0.002	0.0001	1.0001
5	Hydrological reduction factor, φ (-)	0.543	0.0553	1.0289
6	Initial loss, i (mm)	0.000	0.0001	1.0001
7	Surface concentration time, t_c (min)	5.49	0.0000	1.0000
8	Manning number (smooth concrete) M (m ^{1/3} /s)	85.1	0.0000	1.0000
9	Headloss factor (Round edged outlet) K_m (-)	0.249	0.0000	1.0000

Table 3 Failure probabilities etc. with *FORM*, *MCDS*, and *MCIS*.

	<i>FORM</i>	<i>MCDS</i>	<i>MCIS</i>
Failure probability, P_f	0.1045	0.1098	0.1075
Return Period, T	0.0487	0.0464	0.0473
Failures per year, f_p	20.5	21.5	21.1
No. of iterations	13*	-	
No. of model simulations	247	3000	500

* The number of iterations is very dependent on the initial values of \mathbf{u} .

DISCUSSION

Using *FORM*, a failure probability of 0.105 (corresponding to 20.5 failures per year) is predicted (Table 3). Validating the method applying the *MCIS* simulations, a failure probability of 0.108 (corresponding to 21.1 failures per year) is found. This indicates that the fit of the hyperplane to the nine-dimensional failure surface is a valid method of finding the failure probability, despite the small deviation. The *MCDS*, which is considered the most reliable method of the failure predictions (as the whole variable space is sampled), deviates insignificantly from *MCIS*. One might consider using the second order reliability method (*SORM*) instead, which is based on a multidimensional second order polynomial approximation of the failure surface, but as the errors introduced by linear approximations in *FORM* are small, this is not of interest. Furthermore, convergence of *SORM* is empirically even more difficult compared to *FORM*, due to the numerical assessment of the second order derivatives.

Since 2004 the municipality of Aalborg has registered the number of combined sewer overflows and their durations in Frejlev 22, 17, and 25 overflows were registered in the years, 2004, 2005, and 2006, respectively (overflow events with less than 1 hour in between are counted as one). This is in the same order of magnitude as predicted with the three methods and the measurements can therefore not be used to accentuate if one of the methods predicts the failure better than the other.

Despite the consistency between the predicted failures per year and the observed, some uncertainty is still associated with conceptualisation of the rainfall events, as the rainfall events are treated individually. In reality two small rainfall events within a short span of time might induce an overflow which is not considered in the present. The choice of variables and their distributions are indeed empirical. This will affect the results of this analysis. However, some indication of a good and representative choice of variables and distributions is present, as the three techniques predict in the same order of magnitude as observed.

The return periods in this paper are only associated with the rainfall variables, i.e. all other variables are kept fixed in time. One might consider if a return period should be added to some other variables as well, e.g. the hydrological reduction factor, as this might also vary in time. This is, not investigated in the present paper.

Thorndahl and Willems (2007) showed that the method presented is very applicable in prediction of surcharge and flooding as well, but it is beyond the limits of this paper to describe this further. Nevertheless, this represents a potential alternative to simple design methods such as synthetic rain generation based on *intensity-duration-frequency (IDF)* curves or *Chicago Design Storm* (Kiefer and Chu, 1957).

CONCLUSION

It is concluded that the presented conceptualization of the rainfall input in an urban drainage model without crucial affects on the modelling accuracy, can be used as an alternative to traditional long term predictions of combined sewer overflow. The First Order Reliability Method has been

validated methodically using Monte Carlo Direct Sampling and Monte Carlo Importance Sampling, showing similar results. Thus, the simplifications in *FORM* are negligible in terms of predicting occurrences of combined sewer overflow. Moreover, it is observed that both *FORM* as well as the Monte Carlo sampling methods predict the number of overflows per year in the same order of magnitude as observed. Using *FORM* it is possible to reduce the simulation time to approx. 10 % of the simulation time using traditional long term simulations.

The prediction of combined sewer overflow is shown to be very dependent on the rainfall input variables and to some extent on the hydrological surface variables, which are also the variables that contain the highest level of uncertainty. However, the variables that are governing the temporal flow variations in both surface runoff model and pipe flow model are shown to be negligible in prediction of overflow.

ACKNOWLEDGEMENT

The author would like to thank the Municipality of Aalborg for use of the Frejlev catchment and for the overflow registration data.

REFERENCES

- DHI (2005) MOUSE reference manuals. DHI Water, Environment, Health.
- Ditlevsen, O. and Madsen, H.O. (1996) Structural Reliability Methods., John Wiley & Sons.
- Kiefer, C. J. and Chu, H. H. (1957) Synthetic storm pattern for drainage design. Journal of the Hydraulic Division of the ASCE 83 (4).
- Madsen, H. O., Krenk, S., and Lind, N. C. (1986) Methods of structural safety, Prentice-Hall.
- Madsen, H. O. (1988) Omission sensitivity factors. Structural Safety 5 35-45.
- Melchers, R. E. (1999) Structural Reliability - Analysis and Prediction, Ellis Horwood Limited.
- Portielje, R., Hvitved-Jacobsen, T., and Schaarup-Jensen, K. (2000) Risk analysis using stochastic reliability methods applied to two cases of deterministic water quality models. Water Research 34 (1).
- Schaarup-Jensen, K., Johansen, C., and Thorndahl, S. Uncertainties Related to Extreme Event Statistics of Sewer System Surge and Overflow. Proceedings of 10th International Conference on Urban Drainage, Copenhagen, Denmark August 21-26 2005.
- Schaarup-Jensen, K. and Sørensen, J. D. Randomization of river geometry in hydrodynamical modeling. 2. 1996. Aalborg University.
- Sørensen, J. D. (2004) Notes in Structural Reliability Theory and Risk Analysis, Aalborg University.
- Sørensen, J. D. and Schaarup-Jensen, K. Application of reliability methods in hydrodynamical modelling. 1. 1995. Aalborg University.
- Sørensen, J. D. and Schaarup-Jensen, K. (1996) Stochastic Modelling of River Geometry. Aalborg University.
- Thorndahl, S. (2007) Event based uncertainty assessment in urban drainage modelling applying the GLUE methodology. Paper in preparation.
- Thorndahl, S., Johansen, C., and Schaarup-Jensen, K. (2006) Assessment of runoff contributing catchment areas in rainfall runoff modelling. Water Science and Technology 54 (6-7), 49-56.
- Thorndahl, S. and Schaarup-Jensen, K. (2007) Comparative analysis of uncertainties in urban surface runoff modelling. Novatech 2007: 6th International Conference on Sustainable Techniques and Strategies in Urban Water Management, June 25-28 2007, Lyon, France, Accepted.
- Thorndahl, S. and Willems, P. (2007) Probabilistic modelling of overflow, surcharge, and flooding in urban drainage using the First Order Reliability Method and parameterization of local rain series. Accepted to Water Research.
- Willems, P. (2001) A spatial rainfall generator for small spatial scales. Jour. of Hydrology 252 (1-4),