Fault tree analysis for urban flooding

J.A.E. ten Veldhuis[1]*, F.H.L.R. Clemens[1], P.H.A.J.M. van Gelder[1],
[1] Delft University of Technology, Department of Water Management, PO Box 5048, 2600 GA Delft, the Netherlands

Corresponding author, e-mail j.a.e.tenVeldhuis@tudelft.nl

ABSTRACT
Traditional methods to evaluate flood risk mostly focus on storm events as the main cause of flooding. Fault tree analysis is a technique that is able to model all potential causes of flooding and to quantify both the overall probability of flooding and the contributions of all causes of flooding to the overall flood probability. This paper gives the results of a fault tree analysis for urban flooding for the case of Haarlem, a city of 105.000 inhabitants. Data from a complaint register, rainfall data and hydrodynamic model calculations are used to quantify the probabilities of the basic events in the fault tree. The flood probability that is calculated for Haarlem is 0.78/week. Gully pot blockages make the main contribution to flood probability: 79%, storm events contribute only 5%. This implies that in this case an increased efficiency of gully pot cleaning is a more effective strategy to reduce flood probability than to increase the drainage system capacity. Whether this is also the most cost-effective measure can only be decided if the risk calculation is completed with a quantification of the consequences of both types of events. To do this will be the next step in this study.

KEYWORDS
Fault tree, risk analysis, uncertainty, urban flooding

INTRODUCTION
The term risk has been subject for years to different interpretations that never resulted in a generally accepted definition (Kaplan, 1997). In this paper the interpretation by Kaplan and Garrick (1981) is adopted who describe risk analysis as a search for the answers to 3 questions: What can happen, how likely is it to happen and given that it occurs, what are the consequences?

Urban flood risk analysis, traditional approach
Traditionally, urban drainage systems have been designed to cope with a design rainfall intensity or a design storm with an estimated return period that defines the protection level against flooding. The protection level and thus the rainfall intensity that is used to evaluate the design depend on the conditions of the terrain, like hilliness, population density and infiltration capacity. The most commonly applied method to evaluate a design is to feed a theoretical model with the design storm and check whether flooding occurs. The main disadvantage of design storms is that they have been defined for a limited number of return periods and this does not allow for an accurate design evaluation.

There is a more important deficiency in the traditional approach: the methods assume a perfect system in the sense that every component works as it should and that the hydrodynamic model that is used in the calculations is an exact representation of the real

system. Neither defective components nor incompleteness of the model or uncertainties in the model calculations are addressed in the evaluation for flooding. A fault tree analysis reveals these additional failure mechanisms and their contributions to the overall probability of flooding.

Fault tree analysis and quantification of risk
Probabilistic risk analysis aims to quantify the probability and consequences of undesired events. The quantification is preferably based on historical records, but if recorded data are lacking it must be based on extrapolation of the available data and on expert judgement. This implies uncertainty in the estimation of both the probability and the consequences that constitute risk. Fault tree analysis is an effective technique to support probabilistic risk assessment, especially for complex systems, because it can model sequences of events that lead to system failure and interdependencies between events and their probabilities.
The Fault Tree Handbook NUREG-0492 issued by the Nuclear Regulatory Commission in the USA in 1981 has been a leading technical information source for fault tree analysis in the USA (US Nuclear Regulatory Commission, 1981). In 2002 NASA issued a handbook for aerospace applications that contains additional information on recent techniques (NASA, 2002). Ang and Tang provide a short introduction for applications in the field of structural engineering (Ang and Tang, 1984). Risk-based decision making in water resources matured as a professional niche in the US in the 1980's (Haimes, 1998). Probabilistic techniques have had applications in urban drainage in research projects in Denmark (Harremoes and Carstensen, 1994) and Belgium (Thorndahl and Willems, 2008), amongst others. Practical applications remain rare. In the UK, flood risk assessment and management have received much attention recently and the approach has been applied to several cases in the UK (Flood Risk Management Consortium, 2007). No references to applications of fault tree analysis in urban drainage have been found.

This paper outlines a method to quantify urban flood frequencies for all possible causes of flooding. The method is demonstrated in a case study for the urban drainage system of Haarlem.

METHOD: FAULT TREE ANALYSIS FOR URBAN FLOODING
The objective of the fault tree analysis is to identify all possible failure mechanisms that can lead to urban flooding, related to the functioning of an urban drainage system. The analysis aims to quantify the overall probability of flooding and the probabilities of the individual contributing failure mechanisms. The top event of the fault tree is the failure event that is the subject of the analysis. The top event "urban flooding" is defined as the occurrence of a pool of water on the surface somewhere in an urban area lasting long enough to be detected and cause disturbance. The top event includes the appearance of water on the surface as a result of rainfall that is not properly drained and of water that flows out of the drainage system onto the surface due a particular component failure. These failure mechanisms are analysed in detail whereas the occurrence of pools on the urban surface due to failure of other urban water systems: drinking water, groundwater or surface water, are not analysed in detail here.

Independent events
Probabilistic analysis of the fault tree is more straightforward if successive events are independent: probability distributions like the binomial, the Poisson and the exponential distribution are only applicable if the events are independent. Successive events are independent when the total urban drainage system has returned to its initial conditions before

the second event starts. Since in practical cases insufficient data are available to check if initial conditions have been restored for all system components at the start of an event, a safe and practical assumption is made here for the analysis of data on flood events. A period of at least 24 hours without rainfall or flooding is chosen as a criterion to separate independent events. Also, many data that are used in the analysis are only available on a daily basis.

To find a criterion for the independency of events on a spatial scale is less straightforward. Hydraulic relationships determine the flow patterns throughout sewer systems which may lead to dependency between flood events at separate locations. For this reason it is more convenient to evaluate the fault tree for an urban drainage system as a whole and to derive the consequences of a flood event from the number of flooded locations per event. Other information on the extent of the flooding, if available, can be added to quantify the consequences. There is no longer a need to define a criterion to separate events at different locations, because the consequences can be calculated on a gradual scale.

Quantitative analysis of fault trees
The quantitative analysis of the fault tree provides the probabilities of occurrence of the basic events and the top event and it gives quantitative rankings of the contributions of the basic events to system failure. A failure probability model must be chosen that suits the type of failure processes in the fault tree. In this analysis the occurrence of events is assumed to be a Poisson process, which implies that the probability that an event will occur in any specified short time period is approximately proportional to the length of the time period, the occurrences of evens in disjoint time periods are statistically independent and events do not occur exactly simultaneously. Under these conditions, the number of occurrences x in some fixed period of time is a Poisson distributed variable:

$$p_X(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \tag{1}$$

Where: $p_X(x)$: probability of x occurrences in a period of time t

λ : average rate of occurrence of events per time unit

The rate of occurrence λ is derived from failure data over a certain period of time. In a homogeneous Poisson process, the event occurrence rate λ is constant. In a nonhomogenous Poisson process, λ is modelled as a function of time. This model is useful to analyse trends, e.g. due to ageing processes. In this fault tree analysis a constant failure rate has been assumed.

Since failure occurs due to the occurrence of 1 or more events, the probability of failure can be calculated from :

$$P(X \geq 1) = 1 - p_X(0) = 1 - e^{-\lambda t} \tag{2}$$

Where: $P(X \geq 1)$: probability of one or more events

$p_X(0)$: probability of no events

The time period t can be chosen at will; the longer t, the higher the probability of occurrence. The time scale preferably fits the repeat frequency of events. In the case of urban flooding flood events typically occur up to several times per month and the duration of events is in the order of several days. A time period of 1 week has been chosen for the fault tree analysis of urban flooding.

DATA NEED AND AVAILABILITY

A complete analysis of flood events requires a full register of all flood events at every location in an urban area. The register must include information on failures that caused the flood events and on their consequences. In practice, direct data on flood events are scarce and generally incomplete with respect to event causes and consequences. The data source that has proved to provide most direct information on flood events for 5 investigated cases (Haarlem, Breda, Apeldoorn, Utrecht, Delft) is a complaint register. It provides data on flood events due to heavy rainfall, blockages of pipes, gully pots, gully pot manifolds and non-connected surfaces. This data source is more difficult to control than an automated monitoring system, since the behaviour of the "sensors", city inhabitants, is far less predictable. Consistency checks must be carried out to provide an estimate of the reliability of such data. For an application in practice, a comparison between rain events derived from rainfall data and the number of registered flood events, in a complaint register can give a first impression of the completeness and quality of the available data and whether fault tree analysis is a sensible way to proceed.

Other data that improve the quality of a fault tree analysis for urban flooding are: rainfall data, data from a sewer monitoring network, a geometrical sewer database and sewer inspection data.

Case study: city of Haarlem

In this study for the city of Haarlem, complaint data are used to derive probabilities of flooding due to blockages of pipes, gully pots, manifolds etc. Rainfall data are used to define rain events, to select heavy storm events that are likely to cause flooding and to compare complaints with rainfall data. Hydrodynamic model calculations provide data on the theoretical storage and transport capacity of the sewer system and on the theoretical occurrence of flooding for the selected heavy storm events. These data are used to calculate the probability of flooding due to overloading of the sewer system. Figure 1 shows a map of Haarlem with the location of the rain gauges. Table 1 gives a summary of the available complaint and rainfall data and some general characteristics of Haarlem.

Table 1 *Summary of data for Haarlem case study: complaint data in municipal complaint register, rainfall data from 5 rain gauges*

Data case study Haarlem	General data
Number of inhabitants	147.000
Length of sewer system (% combined)	460 km (98%)
Total impervious surface	1225 ha
Total surface connected to sewer system	1110 ha
Theoretical storage in the sewer system	8.1 mm
Complaint register city of Haarlem	
Period of complaint data	12-06-1997 to 02-11-2007
Total number of complaints on urban drainage	6361
Length of data series	3788 days
Missing data:	
- call center closed on weekend and holidays[1]	- 30% of total number of days
- events that do not generate a complaint	- See data analysis

---

[1] Complaints generated in weekend days are likely to be entered next working day: e.g. in 2004-2005 83 out of 104 Mondays hold complaints (80%), while 303 out of 521 working days hold complaints (58%)

Rain gauges in Haarlem	Period of rainfall data	Data interval
P01, P03, P04: in Haarlem (see figure 1)	17/11/2004-24/07/2005	2 minutes
Leiduin: 3 km SW of Haarlem	1/1/2003–2/10/2007	day
Schiphol: 10 km SE of Haarlem	1/1/2003-2/10/2007	day
Correlation between P01, P03, P04	P01-P03: 0.97 P01-P04: 0.96	P03-P04: 0.96
Correlation between data Leiduin and Schiphol (2003-2007):	0.635	
Correlation between data rain gauge P01 and Leiduin (18/11/04-23/07/05):	0.51 (daily rainfall from 0 to 24H for P01) 0.81 (daily rainfall from 8 to 8H for P01)	
Correlation between data rain gauge P01 and Schiphol (18/11/04-23/07/05):	0.89 (daily rainfall from 0 to 24H for P01) 0.59 (daily rainfall from 8 to 8H for P01)	

Figure 1 Map of Haarlem that shows the location of the rain gauges P01, P03 and P04 within the city area and the approximate location of the rain gauges Leiduin and Schiphol

RESULTS

Construction of the fault tree

The occurrence of urban flooding can be a result of 7 events, 3 of which are a related to failure of the urban drainage system:

1) Rainwater that falls on the urban surface at a certain location cannot flow to a drainage facility and as a result forms pools on the surface;
2) Rainwater that falls on a sloping surface flows downhill, does not enter the intended entrance to a drainage facility, but flows to a location where it remains on the surface;
3) Water from the sewer system flows onto the surface;
4) Drinking water flows onto the surface as a result of a pipe burst;
5) The groundwater table rises above ground level;
6) The surface water level rises and flows over overflow weirs, via an urban drainage system onto the surface or flows over river banks onto the urban surface;
7) An amount of water is discharged onto the surface, e.g. extracted groundwater from a building excavation or water from a swimming pool that is replenished.

Figure 1 shows a fault tree for urban flooding as a result of these 7 events. The events form a first level of intermediate events in the tree, because they in their turn result from other events. Each of the events can lead to flooding; therefore an "OR-gate" connects the top event to this first level of intermediate events. The 3 events that concern failure of components of an urban drainage system are analysed in further detail, the other 4 events are outside the scope of this fault tree analysis. They are included in the fault tree as undeveloped events.
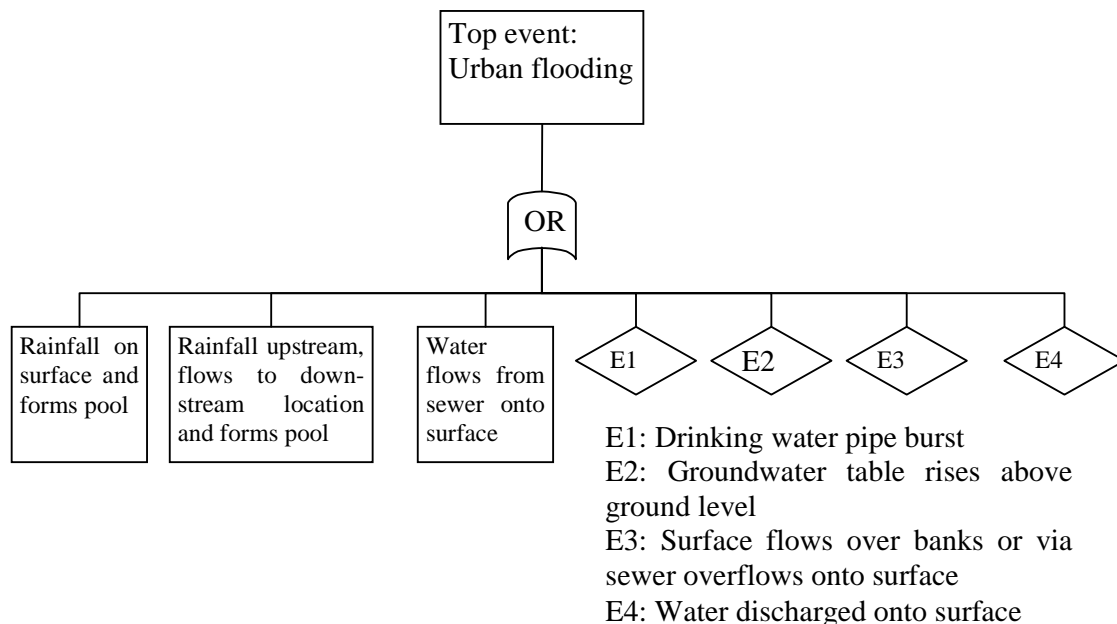

Figure 2 Fault tree for urban flood event. Seven immediate causes for the top event have been identified and are depicted in the tree. The 'OR'-gate between the top event and the underlying causes signifies that each individual cause can lead to the top event.

In a step by step analysis we seek the basic events of the failure mechanisms behind the 3 intermediate events that concern failure of drainage system components. In this way 22 basic events are identified.

Fault tree analysis for urban flooding

Qualitative analysis
Since the first gate below the top event "urban flooding" is an OR-gate, the occurrence of the top event can be a result of 1 basic event or of a combination of basic events. An example of a single event that can cause flooding is pipe blockage: if a pipe in a main transport route gets blocked, the water level will back up and eventually rise above ground level. Looped networks offer more safety in this situation than branched networks, because alternative flow routes are available if one route gets blocked. In the case of a main transport route, however, alternative routes have a smaller capacity that may not be sufficient to convey the flow. Especially if a main route leads to a sewer overflow or a pumping station and an alternative overflow or pump is far away, flooding is likely to occur. Other single event cut sets are those that relate to the blockage of flow routes: if water that falls on the surface cannot reach a rainwater facility, pools will directly occur.

Data analysis case study Haarlem
Reliability of data in complaint register Three aspects have been analysed to check the reliability of the data in the complaint register: accuracy, completeness with respect to rain events and consistency. The accuracy depends on the information in the memo that is attached to a complaint. The memo contains a description of the event that generated the complaint and of the action that has been undertaken to solve it. 99% of all complaints has a memo attached to it that describes the complaint and 93% of all complaints also has information on how it has been handled. Complaints that have no attached memo are still in the handling process.
The completeness is derived from the coverage of rainfall events, that are potential causes of flooding, in the complaint register. In the period 2004-2006, rainfall has been registered on 480 days and on 277 out of these 480 days the complaint register contains 1 or more complaints; on 371 days out of these 480 days the complaint register contains 1 or more complaints on the same or the next day. On 420 out of 480 days a complaint is registered on the same or next day or on the next Monday, if rainfall occurred during the weekend. In this same period 172 independent rain events have occurred, 143 of which have generated a complaint. Of the 29 events that did not generate a complaint, 19 have a smaller volume than 4 mm and 7 events are on Fridays, weekend days or holidays.
The consistency of the data in the complaint register is checked in 2 ways. The correlation between the event volume and the number of complaints per event is determined. This results in a correlation coefficient of 0.7 for the Leiduin rainfall data, 2003-2007. The streets where flooding occurs according to the results of an Infoworks calculation with a stationary rainfall event of 90 l/s/ha are compared with street names that appear in the complaint database. Complaints have been registered for all these streets. The number of complaints per street has no relation with the amount of flooding in the model calculation results; it does have a relation with the length of the street.
With the available data it is not possible to check whether the total number of complaints per event is complete with respect to real flood occurrences. The number of complaints per event does not influence the calculation of the probability of flooding.
Subcategories in the complaint register The complaint register contains 4 subcategories that contain complaints related to flooding: "Flooding", "Blocked gully pots", "Blockages" and "Sewerage". Most complaints fall into the subcategory "blocked gully pots: 79%. Of these complaints, 22% (for period 2004-2007) has an explicit remark of flooding in the attached memo.

The complaints for the year 2004 are checked in detail in order to classify each complaint according to the basic events in the fault tree. The basic event occurrence rates and

probabilities of occurrence are calculated based on these data. The probabilities are used as input for the fault tree analysis Table 2 gives 6 examples of basic events and their probabilities of occurrence.

Table 2 *Six examples of basic events in the fault tree. The fault tree has a total of 22 basic events. The second column gives the results of the calculation of the event occurrence rate and the probability of occurrence of basic events.*

Basic events in fault tree for urban flooding.	Event occurrence rate λ and probability of occurrence P[2]
Rainfall intensity is larger than evaporation + infiltration	$\lambda = 0.93$/week $P=0.60$
No outflow route is available from a pool to an infiltration, storage or transport facility	$\lambda = 0.15$/week $P=0.09$
Blocked or full gully pot	$\lambda =1.6$/week $P=0.61$
Gully pot connection is blocked or broken	$\lambda =0.13$/week $P=0.09$
Water level at inflow point of sewer system is at ground level	$\lambda =0.04$/week $P=0.04$
Sewer pipe is blocked during rainfall	$\lambda =0.12$/week $P=0.06$

Quantitative analysis: Monte Carlo simulation of fault tree

The basic event occurrence probabilities are used to calculate the probability of occurrence of the top event and the contributions of the basic events to the overall failure probability. The quantitative analysis is based on Monte Carlo simulation: the occurrences of the basic events are simulated with the use of a random number generator. Each simulation that results in failure is stored, with the combination of basic events that caused the failure.

A Monte Carlo simulation for the case of Haarlem results in 7800 failures out of 10.000 simulations. The probability of the top event, the occurrence of a pool somewhere in the urban area, is 0.78 per week. Gully pot blockages make the main contribution to system failure: these have a contribution of 6160 out of 7800 or 79% to the overall probability of failure. The second largest contribution comes from drinking water pipe bursts that are responsible for 1620 out of 7800 occurrences of the top event, 21%. The failure mode "water level in sewer system is at ground level" and "capacity of sewer system is as constructed", which corresponds with urban flooding due to heavy rainfall in a well-functioning sewer system, is responsible of 400 out of 7800 failures. The contribution of this failure mode is to the overall failure probability is 5%.

Sensitivity analysis for fault tree calculation

The sensitivity of the quantitative fault tree analysis to the probabilities of the basic events is tested by changing the probabilities of the basic events within an estimated range of uncertainty that is based no expert judgment. The probability that is derived from the complaint register is assumed to be a minimum, since the likeliness of a false positive in the register after cross-checking with rainfall data and memo information is assumed to be small. A maximum estimate is based on the number of basic events that could occur under

---

[2] The event occurrence rate is the number of events associated with an individual cause in a year (2004) divided by the number of weeks in a year. In this case the inter-arrival time $\theta \neq 1/\lambda$, because the duration of events is not negligible.

Fault tree analysis for urban flooding

unfavourable conditions, with a minimum if maintenance and a maximum of human errors. For gully pot blockage, for instance, the maximum expected event occurrence probability has been set equal to the probability of occurrence of a rain event. For "no outflow route " the maximum estimate has been set equal to the average number of road reconstruction and traffic barrier projects, assuming that in every project a mistake is made that leads to a "no outflow" situation. The mistake is assumed to be repaired after the first rain event.

The probability of the top event rises to 0.97 when maximum occurrence probabilities are entered for all basic events. The contribution of most individual basic events to the overall probability of failure increases; nevertheless gully pot blockages still contribute 75% to the top event probability. Drinking water pipe bursts and blocked or broken gully pot manifolds are now in the same range, with a contribution of 50 and 44 % to the top event probability, respectively. The contribution of heavy rainfall events to the top event has increased from 5 to 15 %[3].

The configuration of the fault tree that has been used in this analysis is not the only possible configuration. The quantitative analysis has been carried out for a number of different configurations that incorporate the same failure modes in order to test the sensitivity of the outcomes to the tree configuration. The results of a comparison of the outcomes reveal the importance of a correct interpretation of the data that are used to derive the basic event probabilities in relation to the tree configuration. This illustrated in figure 3 that shows 3 different configurations of a fault tree branch. For the left and middle configuration separate data on gully pot blockages and rainfall must be fed into the tree. The outcomes of the fault tree analysis are similar. The right configuration, however, requires direct data on gully pot blockages that occurred during rainfall.
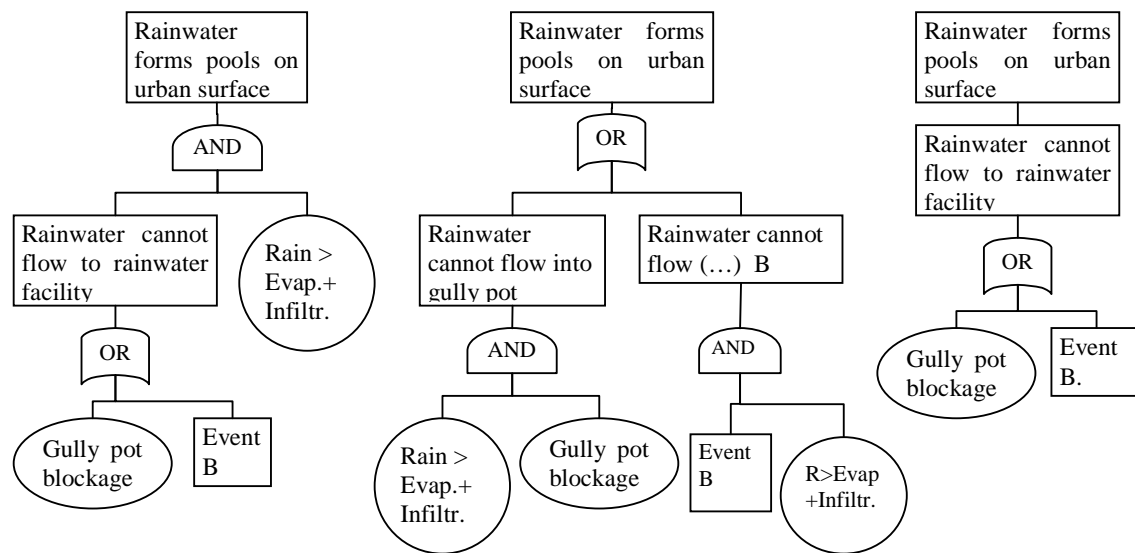

Figure 3 *Different configurations of a branch in the fault tree for urban flooding. The configurations on the left and in the middle give the same result. The configuration on the right gives the same result if the basic event occurrence rate is derived from those gully pot blockages that occurred during rain events.*

---

[3] The percentage contributions of the basic events do not sum up to 100%, because basic events can contribute to the top event in various failure modes, i.e. combinations of basic events. The percentage indicates the ratio of the failures in which the basic event is involved to the total number of failures.

CONCLUSIONS

A quantitative fault tree analysis for urban flooding for the case of Haarlem shows that the probability of flooding is 0.78 per week, for the year 2004. Gully pot blockages contribute 79% to the overall probability of flooding. The contribution of heavy storm events to the flood probability is 5%. These results show that degraded urban drainage system conditions make a much larger contribution to the probability of urban flooding than the occurrence of heavy storm events. Other potential causes of flooding are design faults, like the absence of an outflow route for rainwater, and failures of other water systems, like the drinking water pipe bursts.

An important prerequisite for the quantitative fault tree analysis is a correct interpretation of the data that are used to calculate the basic event probabilities: these must match the configuration of the fault tree. For urban flood analysis the distinction between the conditional probability of a basic event, given that it rains, and the probability of a basic event independent of rainfall is particularly important.

The results of the fault tree analysis provide a basis for decision making on the improvement of the functioning of urban drainage systems: for the case of Haarlem the urban flood probability can be effectively reduced by an increased efficiency of gully pot cleaning. An alternative to this preventive approach can be a more reactive strategy that increases the efficiency in complaint handling in order to reduce the inconvenience due to blocked gully pots to a minimum.

An economical evaluation of this type of decisions can only be made if a complete risk assessment is available. The information on flood probabilities that the fault tree provides, must then be completed with information on consequences of flood events. This requires not only information on material damage, but also on less tangible consequences, like traffic delay and inconvenience for pedestrians in crossing the street or reaching parked cars. This will be the next step in this study.

ACKNOWLEDGEMENT

The authors would like to thank the municipality of Haarlem for making available the data in their complaint register.

REFERENCES

Ang A.H.S. and Tang W.H. (1984). Probability concepts in Engineering Planning and Design. Volume II: Decision, Risk and Reliability. John Wiley and Sons, Inc.

Flood Risk Management Consortium (2007). Year Three Progress Report. Flood Risk Management Consortuim website, http://www.floodrisk.org.uk

Haimes, Y.Y. (1998). Preface. In: Haimes, Y.Y., Moser D.A., Stakhiv, E.Z. (eds.): Proc. 8th Conf. on Risk-based Decision making in water resources, Santa Barbara, California, 12-17 October 1997.

Harremoës, P., Carstensen, J. (1994). Deterministic versus stochastic interpretation of continuously monitored sewer systems. *European Water Pollution Control*, **4**(5), 42-48.

Kaplan S., (1997). The words of risk analysis. *Journal of Risk Analysis*, **17**(4), 407-417

Kaplan S., Garrick B.J. (1981). On the quantitative definition of risk. *Journal of Risk Analysis,* **1**(1), 11-27

Thorndahl S., Willems, P. (2008). Probabilistic modelling of overflow, surcharge and flooding in urban drainage using the first-order reliability method and parameterization of local rain series. *Water Research*, **41**(455-466)

Vesely, W., Dugan, J., Fragola, J., Minarick, J., Railsback, J. (2002). Fault Tree Handbook with Aerospace Applications. Version 1.1. NASA Headquarters, Washington.

Vesely W., Goldberg, F.F., Roberts, N.H., Haasl, D.F. (1981). Fault Tree Handbook. NUREG-0492. US Nuclear Regulatory Commission, Washington.