URBAN RUNOFF POLLUTION

MODELLING AND UNCERTAINTY IN RETURN PERIOD ANALYSIS

Promotor:

Prof.dr.L.Lijklema
Hoogleraar waterkwaliteitsbeheer
Wageningen Universiteit.

De promotiecommissie bestaat uit:

Prof.dr.P.Harremoës, Technische Universiteit Denemarken, Lyngby
Prof.dr.J.Grasman, Wageningen Universiteit
Prof.dr.ir.G.van Straten, Wageningen Universiteit
Prof.dr.ir.F.H.L.R.Clemens, Technische Universiteit Delft
Prof.dr.ir. P.A.Troch, Wageningen Universiteit

URBAN RUNOFF POLLUTION

MODELLING AND UNCERTAINTY IN RETURN PERIOD ANALYSIS

Morten Grum

PROEFSCHRIFT
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
prof.dr.ir. L. Speelman
in het openbaar te verdedigen
op maandag 26 november 2001
des namiddags te vier uur in de Aula.

Grum, M.
Urban Runoff Pollution: Modelling and Uncertainty in Return Period Analysis
Thesis Wageningen University. With ref. With summary in Dutch
ISBN 90-5808-500-7

NN08201, 3089

Stellingen / Theses

1. *Many processes may be deterministic. But when the end is more sensitive to the beginning and to the progression than we are able to establish these, then it is best to assume the processes stochastic. Within water quality this is more often so than not.*

2. *Focusing on the differences in the underlying assumptions concerning the deviation between "what we see" and "what we model" will contribute to bridging the gap between deterministic and stochastic modelling approaches. The strengths of each approach may be drawn to provide modelling solutions appropriate to the available resources, the available data, the nature of the system and problem at hand.*

3. *On evaluating the uncertainty in a return period analysis a distinction should be made between uncertainty expressing a lack of knowledge concerning the value of a given quantity and its inherent variation in time. This may be done using embedded error propagation methods such as Embedded Monte Carlo simulations.*

4. *Automated calibration will have limited success in water quality modelling if it does not quantitatively include the engineers experience and intuition, for example, in the form of a priori parameter distributions.*

5. *Urban water quality management has not the role of "protecting our natural waters against pollution". Within the perspectives of the resources prioritised by society, urban water management should, whilst fulfilling functional, hygienic and aesthetic demands, create the conditions which will permit our chosen aquatic ecosystem to flourish.*

6. *Water quality of combined sewer overflow is influenced by so many unknown factors and poorly understood processes that investigations aimed at improved prediction require the incorporation of stochastic processes.*

7. *With the ever-increasing scientific specialisation, there is too an increasing need and potential for research focused at knowledge transfer between disciplines.*

8. *Proof is needed to convince a scientist that one methodology is better than another. Convincing a water quality manager requires much more.*

9. *Given to us are merely the data of our consciousness. ... There is only one way from them to "reality," to wit, the way of conscious or unconscious intellectual construction, which proceeds completely free and arbitrarily. ... We are free to choose which elements we wish to apply in the construction of physical reality. The justification of our choice lies exclusively in our success.*
   A. Einstein.

Stellingen / Theses belonging to the dissertation:
"Urban Runoff Pollution: Modelling and Uncertainty in Return Period Analysis"
Morten Grum, Wageningen University, 2001

to Nicolai who joined when this had all just begun
to Jonas who joined as it was all meant to end
and to Claudia who was there all along

PREFACE

The research presented in this thesis was done at the Aquatic Ecology and Water Quality Management Group, Department of Environmental Sciences, Wageningen University under the supervision of Professor Lambertus Lijklema and co-supervision of Ir. R. Hans Aalderink. The research work formed part of the EU sponsored MATECH research network (European Centre for Mathematics and Technology of Urban Water Pollution). This thesis is presented as one of the requirements in obtaining the degree of Ph.D. in Environmental Science. The thesis concerns urban runoff pollution in the context of integrated urban water management with special focus on deterministic and stochastic modelling and uncertainty in return period analysis.

ACKNOWLEDGEMENTS

Like most research today, the work presented in this thesis is the result of interchange of ideas, inspiration and constructive criticism between scientists. Alone in an office, this thesis would be something else. I would firstly like to acknowledge the efforts of my supervisor Professor Lambertus Lijklema and co-supervisor Ir. R. Hans Aalderink. Their scientific and professional ballast, curiosity and persistent questioning played an important role in identifying the relevant and interesting aspects, and in guiding the work as a whole.

The research was done as a sub-project under the EU's HCM Network MATECH. The project partners are thanked for their interest, comments, suggestions and discussions during the project. A special thanks is directed to Per Jacobsen and Eric Warnaars for their help in the MATECH coordination and with the regular research reporting to the EU commission.

There are many others who have either directly or indirectly help in the present research and thesis writing. I would like to thank all my former colleges at the Leerstoelgroep Aquatische ecologie en waterkwaliteitsbeheer. A special thanks to Michiel, Hristina, Jeroen, Rob and Elisabeth with whom discussions on methods and their applicability contributed to the present research. Thanks also to the visitors Morio Haneda and Peter Thyregod for the interesting and thought-provoking discussion on variograms and time series analysis. Nor will I forget the seven kiketjes in de bouren sloute.

Thanks to Henrik Madsen at the Institute for Mathematical Modelling, at the Technical University of Denmark for the possibility of participating remotely in his PhD course on Advanced Time Series Analysis and for the possibility of using the open source software CTLSM. I would also like to direct special thanks to Eligius M. T. Hendrix at the mathematics department for tips on optimisation and for letting me have the source code of his optimisation program which I gradually extended to handle the vast variety of optimisation and parameter statistic problems in this research work. And from Rasko Jovin I have received small but vital tips on programming (and soccer) that I am unlikely ever to forget. Indeed research such as the present is dependent on working computers and a local network that is up practically all the time. Thanks to Marianne Fontine and Henk Karsch for this and for their patient with those of us that drove the computers beyond their limit.

To visitors in a foreign land being well received outside work is as important as being well received at work. My wife and I would like to thank Ingrid and Bert Lijklema for the their exceptional hospitality during the stay. And once back in Denmark, thanks to Nannan and Bedstefar, John and later to Jens Jørgen Linde and Torkild Jensen for granting hideouts with the piece and quiet that thesis writing cannot be without.

LIST OF CONTENTS

	PREFACE	vii
	ACKNOWLEDGEMENTS	ix
	LIST OF CONTENTS	xi
PART I	**OPENING**	1
1.	General Introduction	3
2.	The Data Sets	13
3.	Multivariate Analysis of Event Mean Concentrations	21
PART II	**MODELLING**	29
4.	Event Lumped Modelling of Combined Sewer Overflow	31
5.	Deterministic and Stochastic Modelling	41
6.	Modelling using Stochastic Differential Equations	83
7.	Random Coefficient Modelling	91
8.	Discussion on Modelling	99
PART III	**UNCERTAINTY IN RETURN PERIOD ANALYSIS**	107
9.	Uncertainty Analysis using Embedded Monte Carlo Simulations	109
10.	Discussion on Return Period Analysis and its Uncertainty	121
	SUMMARY	143
	SAMENVATTING (DUTCH)	147
	About the Author	151
APPENDIX A	Equations of the Linear Kalman filter.	155
B	Equations of the Extended Kalman filter.	159

PART I

OPENING

CHAPTER 1

GENERAL INTRODUCTION

ABSTRACT

With its offset in the turbulent history of urban surface waters, this chapter gives the background for the presented research. An introduction to the principal problems motivating the present research leads on to a brief presentation of the main conclusions. Two main results pertain to deterministic and stochastic modelling approaches, and to handling of uncertainty in return period analysis. Finally, the build-up of this thesis and the interrelationship between parts and chapters is outlined.

Parts of this chapter are based on Grum, M. and Aalderink, R. H. (1997). Trends and Traditions at the 7. International Conference on Urban Storm Drainage, Hannover 1996. *European Water Pollution Control*, Vol. 7, No. 1, pp 69-71, 1997.

URBAN WATERS

Surface waters in and around urban settlements have lived turbulent lives. Streams and river have often been straightened to drain-out low-lying areas or diverted to bring water to city fortifications or to improve waterway-access into the cities. Increased urbanisation has meant that many urban streams have been covered or even piped under roads and buildings. The fast draining roofs and paved surfaces have resulted in increased peak hydraulic loading during rainfall adding further stress to river and stream environments. Indeed many urban waters have ended up as drains or as part of the sewer system.

During the second half of the 19th century drastic steps were taken in most of Europe to challenge the hazards of water born decease. This was done subsequent to the pests that swept across Europe killing large portions of the population. Clean water was to be distributed under pressure in pipes and foul water was to be removed through gravity sewers and discharged into nearby surface waters. Water born toilets were gradually introduced to replace the then common night-soil system. This uncompromising separation of clean and foul water constituted a cornerstone in raising the living standard and establishing robustness against water borne epidemics (Harremoës, 1999). This separation still constitutes a cornerstone in today's public health. In some of the world's largest urban concentrations economic and managerial factors have meant that the hard-line separation of clean and foul waters has not been implemented or maintained and here water borne decease and epidemics persist (Butler and Parkinson, 1997, Briscoe and Garn, 1995).

The sewage was for many years discharged to the nearest surface waters, which posed yet a stress on the urban waters. This discharge took place during both dry and wet weather. Clean water was often supplied from ground water reservoirs and the lowering of the groundwater table sometimes led to reduced flow and low dilution rates in the already hard hit urban waters. Figure 1.1 illustrates typical changes in water-pathways that have come about as a result of urbanisation. The combined effects of reduced flows, low dilution rates and pollutant loading resulted in oxygen depletion, ammonia poisoning, eutrophication and the accumulation of heavy metals and organic micro pollutants in sediments. In some parts of Europe rural activity had already changed the water landscape completely. Starting with the reclamation of subsiding peat moors in the years 800 to 1250, the Netherlands has a long history of active surface water management and structural intervention aimed initially at draining low lying areas (Van de Ven, 1993).

With increased welfare and shifts in the priorities of society, interceptor pipelines were constructed and sewerage was led by gravity or pumped to treatment plants discharging treated sewerage primarily into the sea, ocean, larger lakes or rivers. Although this constituted a major improvement to the state of urban surface waters it did not lead to any significant improvements in the urban aquatic ecosystems. One of the reasons for this was that during large rainfall events the limited flow capacity of the pipes leading to the sewerage treatments plants meant that sewerage mixed with rainfall water was discharged into the surface waters. Since the seventies storage basins have been constructed at overflow structures to reduce combined sewer overflow and other alternatives are constantly sought, tested and implemented. These include the reduction of the runoff surfaces through porous surfaces and other forms of local infiltration, domestic storage and usage of roof runoff and overflow treatment. Many of the alternative approaches have been part of an overall wish to close the urban water cycle and thereby obtain long term sustainable water systems. The recognition of the negative effects that storage basins can inflict on treatment plant performance fruited further interest in examining possible alternatives.

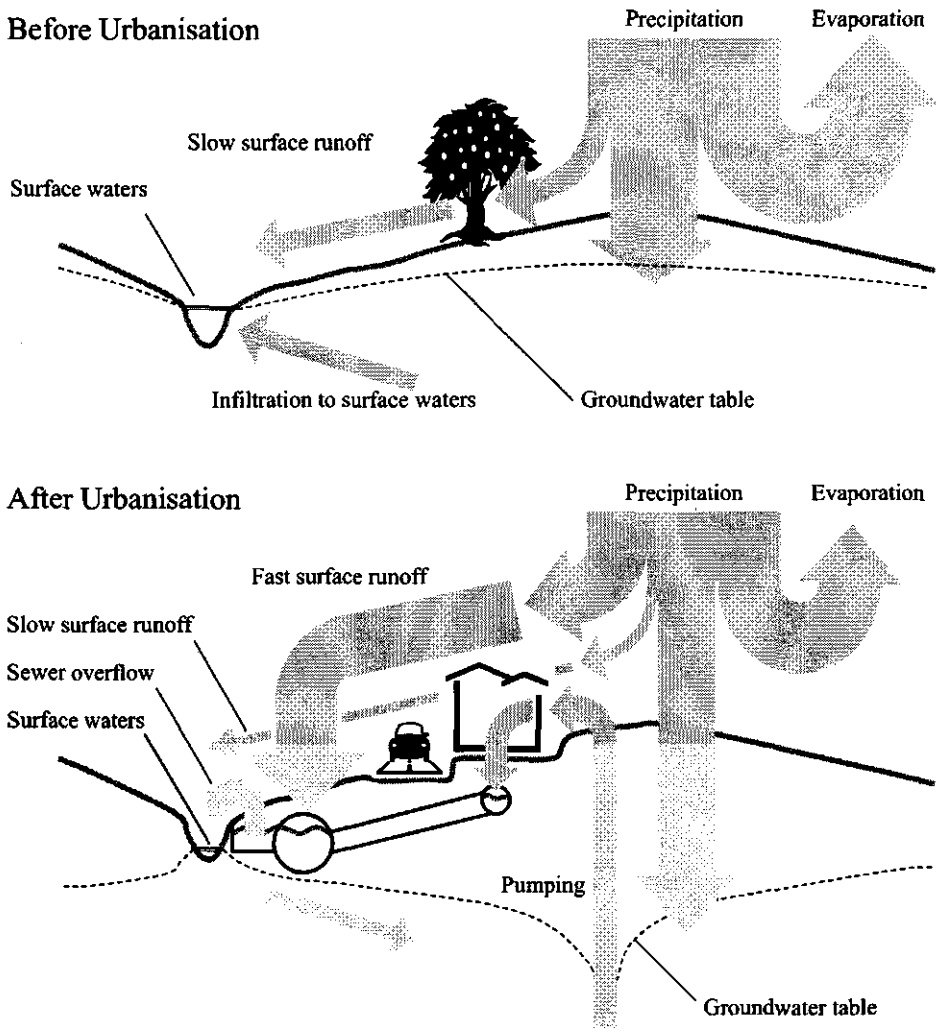


Figure 1.1. Illustration of typical effects of urbanisation resulting in low dry weather flow rates in urban streams and significantly reduced dilution rates in urban lakes, dams and ponds.

Pollution from combined sewer overflow was initially evaluated in terms of the total annual discharged loads and number of overflow events. However, as emphasised in Lijklema (1993), pollutant loads should be evaluated at time scales comparable to their rate of degradation in the surface waters. In the context of combined sewer overflow this led to the main distinction between acute and accumulative pollutants for which the effects are best evaluated using extreme statistic and annual averages respectively. This means that acute surface water effects such as oxygen depletion and ammonia poisoning should be evaluated using return period curves. Return period here being defined as the mean time between occurrence, for example, the mean time between the occurrence of a given low concentration of dissolved oxygen. Nutrient loading to lakes and bays however should be evaluated through annual averages due to longer characteristic time constants of the nutrient cycles which if overloaded ultimately lead to eutrophication.

Due to their different primary functions, surface waters, sewers and wastewater treatment plants developed to become legislatively and administratively independent. Combined sewer management decisions were most often taken irrespective of possible consequences to the treatment plant and possible lack of improvements to the receiving surface waters. The classical example being the increased discharge of nutrients from the treatment plant due to longer periods with maximum hydraulic loading resulting from basins installed with the aim of reducing combined sewer overflow. In the eighties it became increasingly clear that urban water quality management requires an integrated approach involving all three components, namely the surface waters, the treatment plant and the sewer system (INTERURBA, Lijklema et al., 1993). In spite of this recognition some years ago traditions and administrative and institutional barriers prevent a rapid transfer of the integrated approach to management practice.

Our perception of the essence of an integrated approach to urban water quality management has also changed from implying the need for an integrated evaluation to the need for an integrated optimisation of investments. Harremoës et al. (1996) found that water quality standards of a stream, which received discharge from combined sewer overflows during rain, were violated even during dry weather conditions. The funds, already allocated to the reduction of combined sewer overflow, would perhaps have been better spent creating a more robust stream by improving its dry weather state. Frequent or even permanent violation of the water quality standards is also common for many Dutch surface waters. Upholding an isolated policy of pollutant load reduction may often have the negative consequences of reducing dilution to almost nothing. Through groundwater exploitation and an increase in impervious areas, urbanisation has often led to a reduced base flow in urban streams and reduced dilution rates in urban dams and lakes. The consequences are comparable to direct water pollution and they ought therefore politically to be considered as such.

Considering the role of urban water quality management as having to *protect our natural waters* against urban runoff pollution is but a perception of the past. We must recognise the fact that most urban surface waters have been so tampered by the works of man that our task is to *create the urban aquarium of our choice*. Within the perspectives of resources prioritised by society, urban water quality management should, whilst fulfilling functional drainage, hygienic and aesthetic demands, create the conditions which will permit our chosen ecosystem to flourish.

MODELLING

There may be many reasons for modelling sewer systems, wastewater treatment plants, surface waters, ground water or the integrated system as a whole. In the following "reasons for modelling" are discussed in three categories: understanding, design and operation. Each of the different uses of models sets different, though overlapping demands to the models used.

Modelling for a Better Understanding

Models can be used to find cause-effect relationships explaining for instance unexpected flooding or fish kills. This would generally imply detailed modelling incorporating much physical, chemical and biological theory. The precise value of the simulated output would often be of subordinate importance compared to the ability to simulate the phenomena in question and thereby pinpoint bottlenecks or likely causes for an undesired behaviour.

Models can also be used to evaluate dominating ecological relationships and dependencies in a given recipient as presented in Scheffer (1998). In practice one may for a given recipient wish to evaluate the

expected consequences of changes in dilution rate, introduction of tree shading, encouragement of filter feeders or selective fishing. At a scientific level, models are often used to study complex postulated biological and physical mechanisms.

Models in Planning and Design

Models can be used for planning and design of the engineered elements of the integrated water system. The required precision will depend much on what stage the project is at. In the very early proposal stages of a project one may only require very rough estimates of the expected costs and water quality improvements. In such initial investigations it is often more important that the models are able to distinguish between proposed scenarios than the precise output values predicted. In the planning stages it is valuable if the models are computationally fast as the event lumped models presented in Chapter 4. At a later design stage more precision may be desired in order to be able to evaluate and compare a small selection of optional scenarios. High precision is generally required during the final stages of structural design where at most a handful of scenarios are modelled. The chosen precision will often also depend on the consequences of being wrong and on the resources available for the work. In the later detailed design stages of a project's life cycle it is common and often reasonable to model only isolated elements of the integrated system.

The characteristics of the models used for design will also depend on the type of design criteria used. In integrated water quality management design criteria may relate to average values (e.g. of nutrient discharges or nutrient flows) or return periods of given conditions (e.g. flooding or low oxygen concentrations in streams). These both require an evaluation of the integrated system's long-term behaviour and would often be done using historic rainfall series. In some cases it may be necessary and possible to boil down the long-term series approximating it by selecting a few events as representative of certain return periods. As discussed in Chapter 10 the appropriateness of this approximate method depends much on whether or not there is a linear relationship between the inputs (e.g. rainfall, sunlight, temperature and wind) and the evaluated effects (oxygen or ammonia concentration). Later in this introduction we will look more closely at modelling for return period analysis, which forms a major part of this thesis.

There are two main reasons why models should be used when calculating the return periods of given detrimental events. These can be summarised as follows:

1. Measuring in the field in order to find the return period of given detrimental effects would require monitoring periods several times longer than the return period of interest, which is not practical.
2. Analysis of proposed structural or operational modifications requires models in order to evaluate improvements. Downscaling of the real world system to a physical model in the laboratory is neither practical nor cost effective nor realistic.

As a consequence of the first point, long historic series of the primary driving force, rainfall, are used and the integrated systems behaviour is modelled and simulated. Amelioration projects are often followed by a period of a few years of intense monitoring of water quality conditions, such as surface water dissolved oxygen concentration, in order to evaluate improvements. Without models this evaluation cannot be made to any reasonable degree of certainty. Attempting to estimate the minimal oxygen concentration with a return period of one year on the basis of a two-year monitoring period is similar to estimating a population average on the basis of two observations. The result is highly uncertain. What the monitoring data *can* be used for, is the identification and calibration of hydraulic and water quality models which can then subsequently be used in combination with historic rainfall series several times longer than the return periods of interest. With models containing a sufficient degree of physical, chemical or biological theory it is then possible to evaluate the expected return periods for proposed improvement. In contrast, empirical models would not contain the

parameters and structure associated with the proposed improvements and would be of no used to such an extrapolation.

Models in Operation

Subsequent to developments in electronics, robust sensors and changing attitudes of operations personnel, models are becoming an integrated part of the operation of integrated urban water systems. This may be in the form of early warning systems or as automated real time control systems. The requirements of models used in operation differ considerably from those used in design. Models used in operation need to predict the water system's behaviour in the short-term and need not necessarily include much physical, chemical or biological theory. Models may also be used as software sensors for indirect monitoring of a variables value as a function of other monitored values.

Though this thesis does not deal with operation specifically, the studied approaches and methodologies have a high potential in both early warning systems and real time control.

APPROACHES TO MODELLING

Traditional deterministic approach to modelling has been deductive with models built up of physical, chemical and biological theory. Starting often with continuity or mass balance equations and proceeding to definitions of process rates. There has been a tendency to incorporate all thinkable processes into the models. This has often led to models with many very uncertain parameters. Calibration of such deterministic models is often done by fixing certain parameters based on intuition and engineering experience, and then adjusting other parameters to best fit the observed series. Often there are many more parameters in the model than can actually be identified from the data. The result is that the estimated set of parameters is non-unique and the parameters, which are being calibrated, cannot be identified from the available data. The continuous inclusion of more and more processes into the models is likely to lead to a situation where some of the included processes that are less important than phenomena that the model does even not describe.

Arising first in the sixties, the stochastic approach to modelling was based purely on statistical relationships between data and was initially deprived of any physical, chemical and biological theory. Associated to the highly data dependent empirical models of classical time-series analysis were methodologies such as automated parameter estimation, parameter statistics, identifiability, experimental design and residual analysis. These methodologies were all aimed at being able to simulate and forecast the observed data. Though stochastic modelling had a strong breakthrough in other fields such as economics, only very little interest was shown in water and water quality engineering.

It is not only on these peripheral aspects that deterministic and stochastic approaches differ. At the very core of modelling methodology, there is a clear cut difference in the underlying assumptions made during deterministic and stochastic modelling.

> *In the context of urban runoff pollution this thesis presents a new portrayal of the essential differences between deterministic and stochastic modelling which focuses on the different assumptions made concerning the source of the deviation between observed and modelled values.*

In deterministic modelling all deviation between modelled and observed variables is implicitly assumed to result from observation error alone. The model is assumed to contain a perfect description of the system behaviour. In stochastic modelling the deviation is implicitly assumed also to result from random behaviour within the system itself. Stochastic modelling thus recognises that the model is unable to describe all variations in the system.

This study and presentation of the essential differences is done in an attempt to bridge the gap between the deterministic and stochastic schools. By pointing to what is the only clear-cut distinction between deterministic and stochastic modelling the author hopes to ease the transfer of knowledge and methods between the two schools and thereby to strengthen water quality modelling as a whole.

UNCERTAINTY OF RETURN PERIOD ANALYSIS

Return periods of detrimental effects such as flooding and oxygen depletion form a major criterion in design of integrated urban water systems. The return period analysis can be made in a number of different ways with different forms of data and different types of models. This has been organised in a general framework in the methodology review of Chapter 10.

With return period analysis playing such an important role in integrated water engineering it is also relevant to look at the certainty with which the return periods are actually estimated. Substantial uncertainty is found in the return periods of the hourly minimal dissolved oxygen concentration in the pond of an integrated urban water system. Figure 1.2 illustrates what is understood by uncertainty in a return period analysis. In this thesis it is argued that water mangers and others involved in the decision making process would be better off with both the estimated return period and with the funnel shape in Figure 1.2 representing the 95% confidence region within which this curves may in fact be situated. Though not examined explicitly, it is in Chapter 10 also outlined how the uncertainty is often significantly lower once one is comparing proposed scenarios.

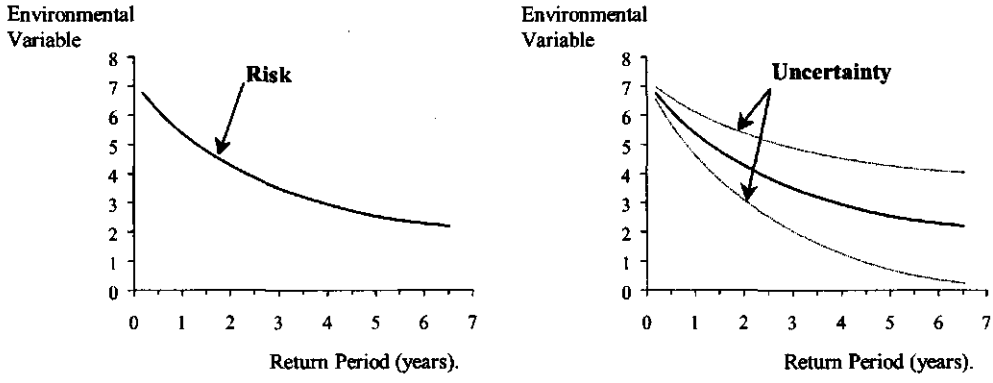


Figure 1.2. The return period curve (left) and its uncertainty (right). The curve give (fictitious) return periods for minimum values of an unspecified environmental variable of interest.

In this thesis a new methodology for evaluating uncertainty in return period analysis is presented:

> *On evaluating the uncertainty in a return period analysis a distinction should be made between uncertainty expressing a lack of knowledge and inherent variation in time.*

The evaluation is implemented as an embedded error propagation involving two layers of probability distributions: an inner layer consisting of quantities varying from event to event and an outer layer consisting of uncertain but constant quantities. Error propagation was in this work evaluated by Monte Carlo simulations. For each set of random realisations of values in the outer uncertainty layer a full Monte Carlo simulations of the inner inherent variations layer was performed. The general methodology is here described as Embedded Error Propagation and its implementation in this work using Monte Carlo simulations is described as Embedded Monte Carlo Simulations.

Often when uncertainty is included in return periods analysis the distributions of the uncertain parameters are treated in the same way as with the inherent variations from event to event. That inherent variations such as rainfall depth vary from event to event is in this context not uncertainty though how and how much it varies may well be.

In fact the whole purpose of the return period analysis is to find out how often the variations in the driving forces result in given rare effects. In this work it is shown how such pooling of uncertainty with inherent variation distorts the picture as it systematically increases the frequency of extreme events.

THIS THESIS

The chapters of this thesis have been arranged in three parts: an opening, a part on modelling and a part on uncertainty in return period analysis. This section briefly outlines the contents and relationship between the parts and the chapters that they contain. Figure 1.3 shows the interrelationship between parts and chapters of this thesis.

Part I opens the thesis with this introduction to the problem and brings attention to main findings of this thesis. With its focus on methodologies, the present research did not encompass any field monitoring or data collection. The research has been based on the analysis of existing water quantity and quality data. Chapter 2 introduces these main data sets that have been used and that are common to most of the studies presented in this thesis. The opening is rounded off with a multivariate analysis of event mean overflow concentrations from Dutch and Danish urban drainage catchments. The multivariate analysis is aimed at gaining preliminary insight into the underlying structure of variations in the event mean concentration data.

In Part II focus is on modelling of the combined sewer system. As primary driving force, rainfall is input to all the studied models. Water quantity and/or water quality variables of combines sewer overflow are the main model outputs. In Chapter 4 an evaporation dependent wetness index is identified from event lumped data of rainfall depth and duration, and overflow quantity. For the studied data sets event lumped rainfall variables are found to have little relation with the event mean overflow concentrations. Chapter 5 contains a detailed discussion on deterministic and stochastic modelling. Here a new portrayal of the essential differences between deterministic and stochastic modelling is presented and illustrated with a case study involving a combined sewer rainfall-runoff model. In Chapter 6 attempts have been made to use stochastic differential equations for both water quality and water quantity modelling of a combined sewer system.

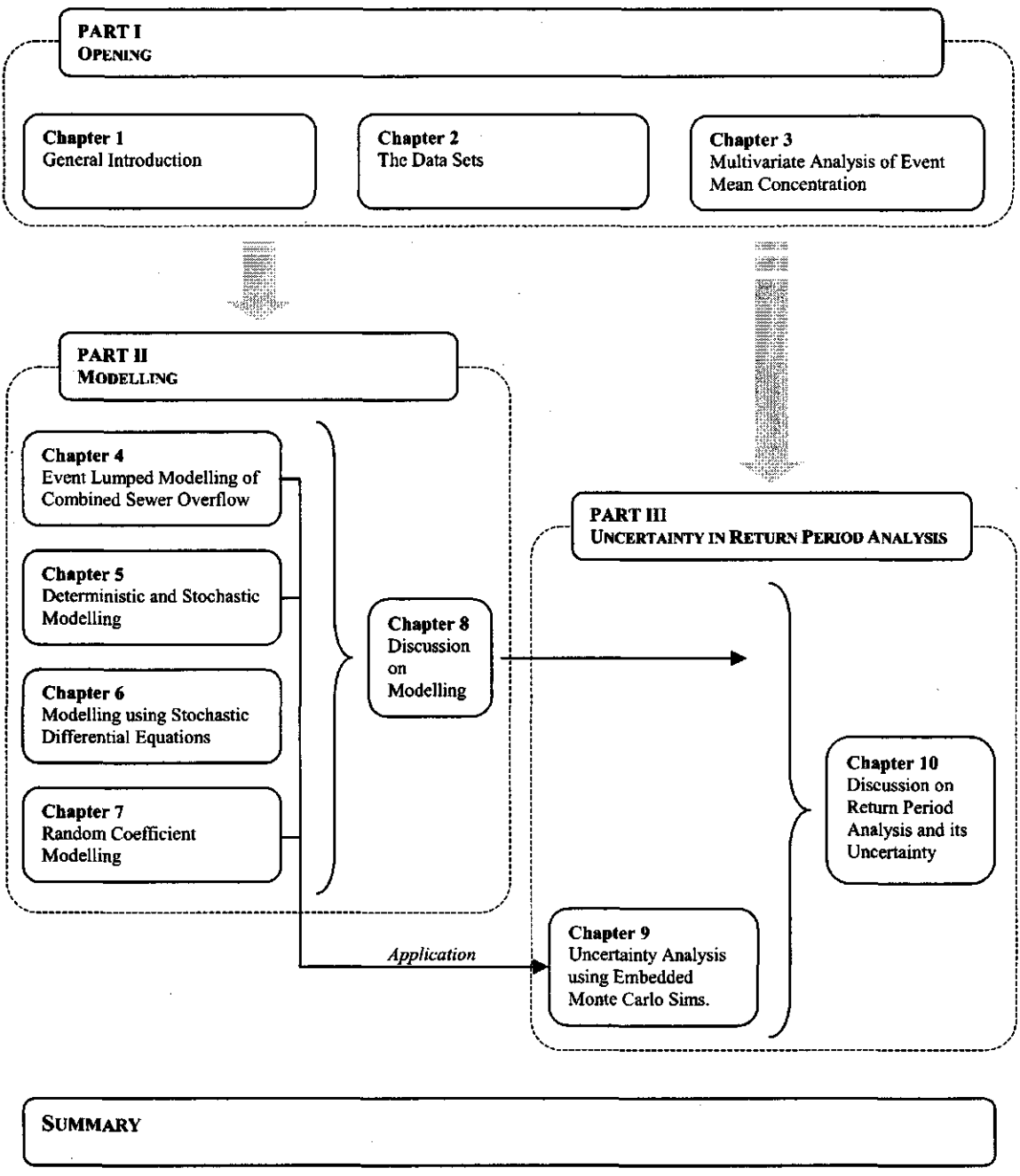


Figure 1.3. Relationship between parts and chapters of this thesis on Urban Runoff Pollution: Modelling and Uncertainty in Return Period Analysis.

One of the conclusions of the studies on stochastic modelling is that the data used has to fulfil certain minimum conditions with respect to the required sampling frequency needed for adequate system identification. Chapter 7 therefore looks at random coefficient modelling as an interesting alternative to stochastic modelling.

Part III encompasses application of models in its presentation of one of the key points of this thesis: a distinction should be made between uncertainty and inherent variation on evaluating the uncertainty of a return period analysis. In Chapter 9 this is presented within an integrated setting with a small pond as surface water receiving combined sewer overflow from an urban drainage system. The uncertainty in the return periods of minimum oxygen levels in the pond is here of primary interest. A methodological review on return period analysis and its uncertainty is presented in Chapter 10. A general framework of approaches to return period analysis in urban runoff pollution is presented and used to place the approach of Chapter 9 in a broader perspective.

REFERENCES

Briscoe, J. and Garn, H. A. (1995). Financing water supply and sanitation under Agenda 21. *Natural Resources Forum*, 19(1), pp. 57-70.

Butler, D. and Parkinson, J. (1997). Towards sustainable urban drainage. *Wat. Sci. Tech.*, 35(9), pp. 53-63.

Grum, M. and Aalderink, R. H. (1997). Trends and Traditions at the 7. International Conference on Urban Storm Drainage, Hannover 1996. *European Water Pollution Control*, Vol. 7, No. 1, pp 69-71, 1997.

Harremoes, P., Napstjert, L., Rye, C., Larsen, H. O. and Dahl, A. (1996). Impact of rain runoff on oxygen in an urban river. Wat. Sci. Tech., 34(12), pp 41-48.

Harremoes, P. (1999). Water as a transport medium for waste out of towns. *Wat. Sci. Tech.*, 39(5), pp 1-8.

Lijklema, L., Tyson, J. M. and Le Souef, A. (1993). INTERURBA - Interactions Between Sewers, Treatment Plants and Receiving Waters in Urban Areas. *Wat. Sci. Tech.*, 27(12), 1993, ISSN 0273-1223.

Scheffer, M. (1998). Ecology of Shallow Lakes. Chapman and Hall. London. 357 pp.

Van de Ven, G. P. ed. (1993). Man-made lowlands : history of water management and land reclamation in the Netherlands. Stichting Matrijs, P. O. Box 670, NL-3500 AR Utrecht, The Netherlands.

CHAPTER 2

THE DATA SETS

ABSTRACT

The assimilation of measured data into models is a central theme throughout this thesis. Observations of our physical, chemical and biological environment are used to strengthen our theoretical descriptions. Monitoring and data collection has however not been part of this research project. Selections from existing combined sewer data sets have been studied.

This chapter gives an overview of the catchment characteristics and the monitoring programs and the resulting data. References are made to the monitoring program's original documentation and to earlier studies in which the data has been used. Only information relevant to the present research work is presented. Readers are referred to the original documents for more details on the measuring campaigns, sampling techniques and analytical methodology.

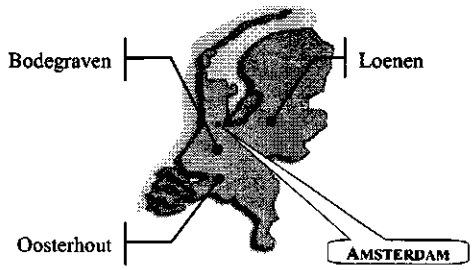


Figure 2.1. Map of The Netherlands showing the locations of the three studied combined sewer catchments.

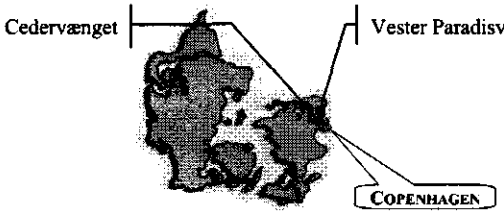


Figure 2.2. Map of Denmark showing the locations of the two studied combined sewer catchments.

INTRODUCTION

Observed data is a vital ingredient to most of the studies presented in the following chapters. Some of the studies have been performed using the same data sets. The descriptions of the urban catchments and the monitoring programs have therefore been concentrated in this chapter.

This chapter contains no description of methods or models. The chapter contains no discussion on the characteristics of the available data in relation to what would be required or desired in various modelling situations. Such a discussion at this point would be pre-emptive as these questions relate strongly to both the problem addressed and the methods applied. For discussions on data adequacy readers are referred to the chapters of the individual studies and to the more general discussion on identifiability in Chapter 8.

This chapter is divided in to two sections. The first section outlines the event lumped data sets and the respective catchments. The second section describes the time varying data set that has been used.

EVENT LUMPED DATA SETS

Event lumped data is in the present work used to describe data sets in which the considered variables take a single value for each event. This may be a totality, such as the total volume of rain or the duration of overflow, or it may be average variables such as pollutant event mean concentrations (EMC). Some variables had been measured directly as event lumped variables while others were converted into event lumped values from time varying data. Event lumped data has been used in the three chapters that follow. The multivariate analysis in Chapter 3 is conducted using data from the three Dutch and the two Danish urban catchments described below. The study on event lumped models in Chapter 4 and their application in Chapter 9 have been conducted using event lumped data from the Loenen catchment only.

Table 2.1. The five catchment areas and their impervious fractions.

Catchment	Units	Loenen	Bodegraven	Oosterhout	Cedervænget	Vester Paradisvej
		NL	NL	NL	DK	DK
Area	[ha]	56.5	48.0	22.5	5.29	17.2
Impervious fraction	[%]	28%	46%	49%	45%	23%

The three Dutch urban catchments were Loenen, located in the heaths (heather hills) of the eastern Netherlands, Bodegraven, located in the western lower lands of the old Rhine delta, and Oosterhout in the south west. Their geographical locations are shown on the map in Figure 2.1.

The data sets were collected from 1981 to 1986 as part of a study carried out by the National Working Group on Sewerage and Water Quality (NWRW/STORA 1990). Analyses of the Dutch data have earlier been made and presented in studies by Van der Heijden et al. (1986), Benoist and Lijklema (1989), van Sluis et al. (1991), van Walraven et al. (1985), Bakker et al. (1988) and by Bakker et al. (1989).

The study on event lumped models in Chapter 4 and their application in Chapter 9 have been conducted using event lumped data from the Loenen catchment only. An extract of the Loenen data set was used. Table 2.2 gives a summary of the variables used in Chapter 4. The particulate and dissolved pollutant components are more precisely described as fast settling and slow/non-settling parts. For the purpose of the study in Chapter 4 a new variable is introduced on this basis. This is the "slow settling fraction", f, which is defined as the sum of the slow and non-settling parts divided by the total concentration. Suppose a concentration of 280 mg COD/l was observed before settling and a concentration of 252 mg COD/l after settling. Then the slow settling fraction would be calculated as (280 – 252)/280 = 0.1 (with no dimensions). The "slow settling fraction", f, thus includes the entire dissolved component and the part of the particulate component which had not settled after one hour. Due to uncertainty in the measurement analysis a slow settling fraction greater than one is seen for a few observations where the absolute concentrations where very low.

The two Danish catchments are both located north-west of Copenhagen in the river basin of Mølleå (see Figure 2.2). The data was collected from 1979 to 1980 as part of an urban runoff study (the Mølleå study) carried out as a cooperation between the city council of greater Copenhagen (Hovedstadsrådet) and the Inst. of Env. Sci. & Eng., Tech. Univ. of Denmark (Johansen et al. 1981). The data have since been studied by Johansen et al. (1981), Johansen (1985), Hall et al. (1990), Jensen (1990), Arnbjerg-Nielsen et al. (1994) and others.

The areas of the five catchments can be seen in Table 2.1. The Danish data is based on volume proportional sampling from the combined sewer main stream during the rainfall-runoff event whereas the Dutch data set is exclusively based on multiple sampled combined sewer overflows with EMCs calculated by volume proportional weighting.

Table 2.2. Summary of the data variables studied.

	Variable	Symbol	Unit	N	Mean	Minimum	Maximum
Rainfall	Depth	H_{RAIN}	mm	284	8.1	1.3	57.0
	Duration	D_{RAIN}	min.	284	412	10	1986
	Intensity	I_{RAIN}	mm/min.	284	6.1	0.4	169.6
Overflow	Volume	$V_{OVERFLOW}$	m^3	63	1163	3	8591
EMCs	Kjeldahl Nitrogen (N-Kj)	$C_{N\text{-}KJ}$	mg/l	48	10	3	26
	Slow settling N-Kj fraction	$f_{N\text{-}KJ}$	-	44	0.83	0.03	2.94
	Chem. oxygen demand (COD)	C_{COD}	mg/l	48	262	62	873
	Slow settling COD fraction	f_{COD}	-	46	0.39	0.07	0.98
	Suspended solids (SS)	X_{SS}	mg/l	47	288	28	1196
	Slow settling SS fraction	f_{SS}	-	45	0.14	0.00	0.97

The exploratory multivariate analysis presented in Chapter 3 requires a data set without any missing values for any of the variables. Biological oxygen demand (BOD) had not been measured at the Danish catchments and, to obtain a complete data set for the multivariate analysis, total and dissolved BOD values of the Danish sets had to be estimated using regression models. These regression models expressed the deviation of the BOD concentration from the monthly averages (which were available for the Danish data sets) as a function of the deviations of the other EMC variables from their monthly average. The models were initially calibrated on the Dutch data sets and thereafter used to estimate the missing BOD values in the Danish sets. It is not thought that this has had much influence on the overall results of the multivariate analysis but it does pose limitations on conclusions that can be drawn in relation to the physical catchment characteristics.

TIME VARYING DATA SET

Time varying data set is here used in contrast to *event lumped* data and refers to data sets containing several observations of each variable during a single rainfall-runoff event. This could be water level every minute at a given point in the sewer system or the concentration of a pollutant in the overflow every five minutes or for every cubic meter of overflow. A Dutch time varying data set, Loenen, has been used in the present research work. Loenen is located as shown in Figure 2.1.

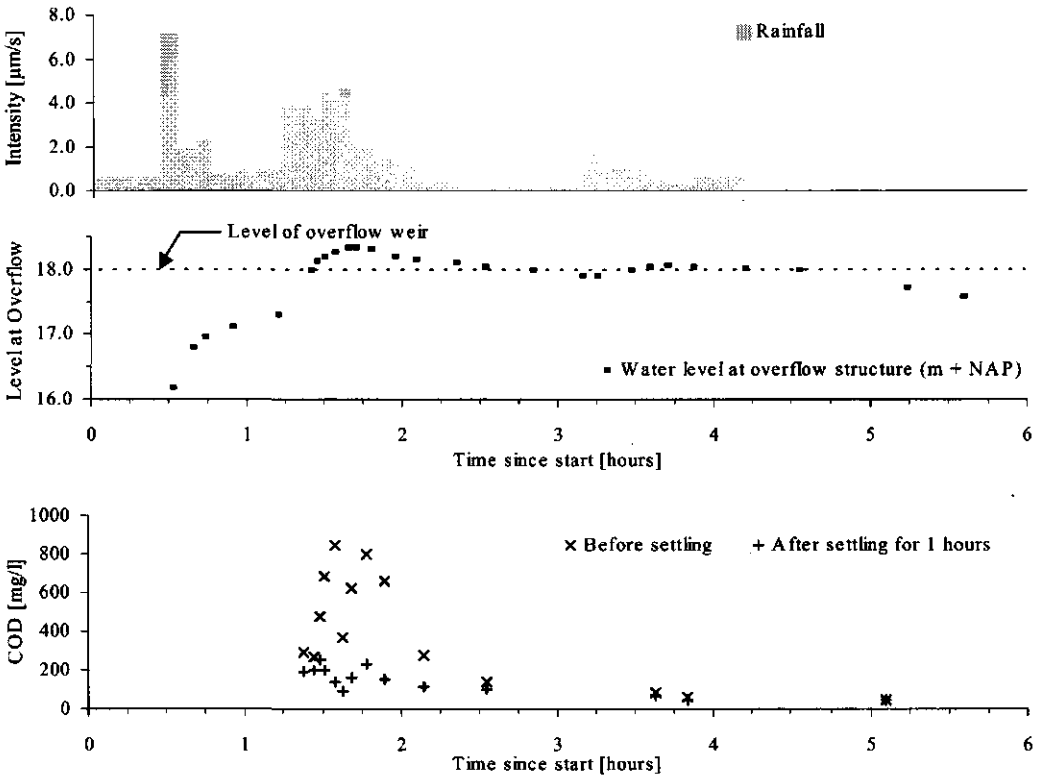


Figure 2.3. Extract of the time varying data from the Loenen catchment (Event 25).

Table 2.3. Summary of catchment characteristics (Loenen, The Netherlands).

Characteristic	Value	Unit
Total catchment area	56.5	ha
Impervious area	28	%
Impervious area	15.8	ha
Mean pipe gradient	3.3	‰
Volume below weir	895	m^3
Depth below weir	5.7	mm

Loenen

The Loenen sewer is a looped system with a single overflow structure at which water level and water quality variables have been monitored continuously over a period of four years from 1982 to 1986. Water is pumped to the treatment plant from the lowest point in the sewer. During actual overflow, samples for chemical analysis were taken roughly at volume proportional intervals. Thus no water quality samples were taken before and after the actual overflow. Only rainfall events that resulted in combined sewer overflow were used in the studies with time varying data. Table 2.3 gives a summary of the main catchment characteristics.

The rain gauge was situated a few hundred meters east of the 15.8 ha catchment. The tipping bucket raingauge had a cup corresponding to 0.1 mm and was connected to paper-roll-writer. The paper role recordings were later typed manually into a digital form.

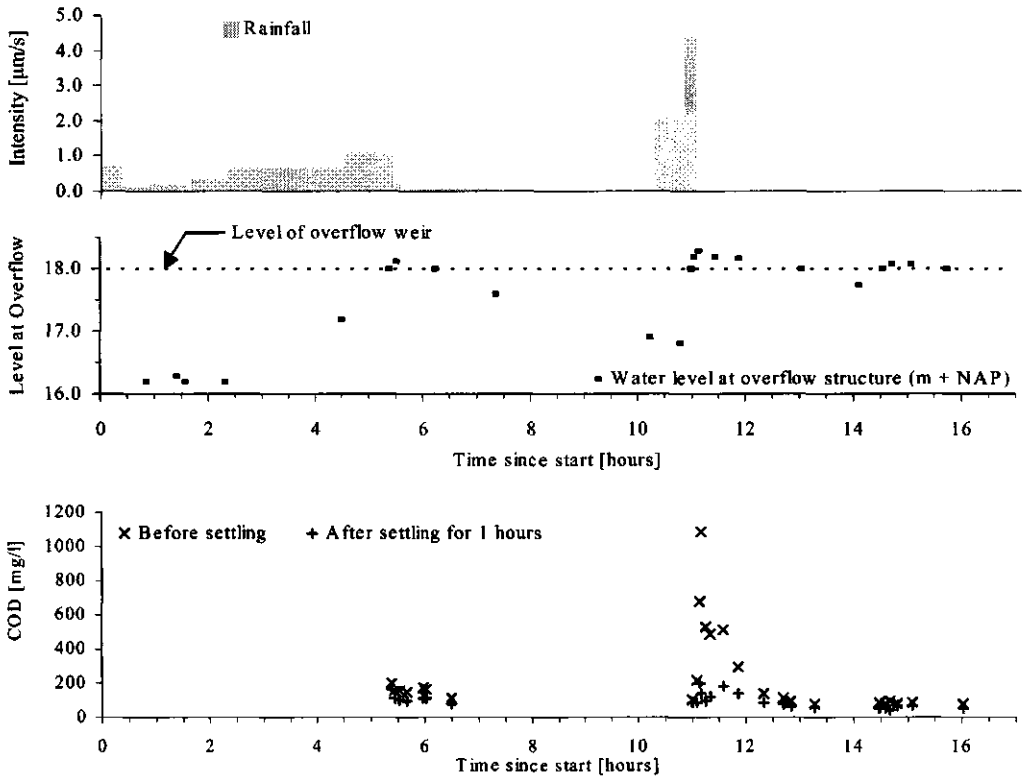


Figure 2.4. Extract of the time varying data from the Loenen catchment (Event 29).

Chemical analysis for biological oxygen demand (BOD), chemical oxygen demand (COD), suspended solids, Kjeldahl nitrogen and total phosphorous had been made on samples before and after one hour settling. Thus the variables available corresponded roughly to total concentrations and dissolved concentrations. A rough value for the corresponding particulate concentration was calculated by subtracting the value obtained after one hour's settling from that obtained before settling. In a number of cases one or the other value was missing and the difference could not be calculated. In other cases the concentration after settling was higher than the concentration before settling. Thus the number and frequency of the water quality observations varied greatly and a number of rainfall-runoff events were screened out as their data information content compared to modelling time was too low. The plots in Figure 2.3 and 2.4 give a visual impression of the kind of time varying data available in the Loenen data set.

CONCLUSION

The data selection criteria depended on the problem being addressed and the methodology being studied. For the multivariate analysis it was important to have or to be able to drive the event mean concentrations of a large number of pollutants common to all or most of the included data sets. Potential explanatory event lumped variables such as rainfall depth, discharged volume of water and preceding dry weather period should also be available along with a general catchment characteristics. Data from three Dutch and two Danish catchments was used for the multivariate analysis presented in Chapter 3.

An event lumped modelling study such as presented in Chapter 4 requires event lumped overflow data including volume of water discharged, duration of overflow and event mean concentrations of the pollutants of interest. It was also necessary to have event lumped rainfall data such as rainfall depth, duration and maximum intensity from a rain gauge in or near the studied catchment.

For the remainder of the studies in this thesis the data from the Loenen catchment in the Netherlands was preferred due to the time varying data of both water level and a number of pollutants, and because Loenen is a catchment with only one overflow structure. Ideally, pollutant sampling for the purpose of dynamic modelling should include more frequent and regular sampling of pollutants and sampling should take place from the onset of the rainfall event until the combined sewer overflow has stopped. Aspects of the information content of time series data sets are discussed further Chapter 5 and 8.

REFERENCES

Arnbjerg-Nielsen, K., Harremoës, P. and Spliid, H. (1994). Prediction of Hydrological Reduction Factor and Initial Loss in Urban Surface Runoff from Small Ungauged Catchments. International Workshop Closing the Gap Between Theory and Practice in Urban Rainfall Applications, St. Moritz, Switzerland, November 1994.

Bakker, K., Ten Hove, D., De Ruiter, M. A. And Van Walraven, J. H. A. (1988). Stormwater overflows from sewer systems: reduction of pollution emission by storage sedimentation tanks. *H2O* **21**(19), 568-572.

Bakker, K., Timmer, J. L. and Wensveen, L. D. M. (1989). De vuiluitworp van gemengde rioolstelsels. NWRW/STORA Report no. 5.2, Min. Of Housing, Phys. Planning and Env., SDU Publs., P.O. Box 20014, 2500EA 's-Gravenhage.

Benoist, A. P. and Lijklema, L. (1989). A methodology for the assessment of frequency distributions of combined sewer overflow volumes. *Wat. Sci. Tech.*, **23**(4), 1.

Hall, M. J., Ellis, J. B., and Brizio, M. G. (1990). On the Statistical Distribution of Event Mean Concentrations of Pollutants in Stormwater Runoff. Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Jensen, M. (1990). Rain-Runoff Parameters for Six Small Gauged Urban Catchments. *Nordic Hydrology*, **21**, pp.165-184.

Johansen, N. B. (1985). *Discharge to Receiving Waters from Sewer Systems During Rain*. Ph.D. Thesis, The Inst. of Env. Sci. and Eng., TUD, B 115, DK-2800 Lyngby, Denmark.

Johansen, N. B., Harremoes, P. and Dahi, E. (1981). *Regnafstrømning i Mølleåsystemet: Samlerapport (Rainfall-Runoff in the Mølleå Basin: Summary Report)*, The Inst. of Env. Sci. and Eng., The Tech. Univ. of Denmark, B 115, DK-2800 Lyngby, Denmark.

Van der Heijden, R. T. J. M., Lijklema, L. and Aalderink, R. H. (1986). A statistical methodology for the assesment of water quality effects of storm water discharges. In: Urban Runoff Pollution, NATO ASI series G; Ecol. Sci. Vol. 10, Springer.

Van Sluis, J. M., Ten Hove, D. and de Boer, B. (1991). Final Report of the 1982-1989 NWRW Research Prog. Concls. & Recom. Min. Of Housing, Phys. Plan. and Env., SDU Publs., P.O. Box 20014, 2500EA 's-Gravenhage.

Van Walraven, J. H. A., Bakker, K. and Wensveen, L. D. M. (1985). The STORA results on discharge from sewer systems into surface waters.

CHAPTER 3

MULTIVARIATE ANALYSIS OF EVENT MEAN CONCENTRATIONS

ABSTRACT

Urban runoff pollution loads can essentially be characterised by fluid quantities and pollutant concentrations. Although several pollutant storage and transport mechanisms have been postulated there remains substantial unexplained variation in event mean concentrations of combined sewer overflow pollutants. Through a series of well-established multivariate pattern recognition techniques the present study has aimed at disclosing the underlying structure of systematic variations in the event mean concentrations (EMC) of pollutants in combined sewers during rainfall. The statistical methods that have been applied to the pollutant concentration variables are factor analysis, cluster analysis, distribution analysis and correlation analysis. The event mean runoff data considered includes eleven pollutant variables originating from five combined sewer catchments in Denmark and in the Netherlands. The combined results of the analyses support earlier findings that EMCs are best described by bimodal or mixture distributions, and further suggest that event based pollutant modelling could be improved through a recognition of these characteristics.

This chapter is based on Grum, M., Aalderink, R. H., Lijklema, L. and Spliid, H. (1997). The Underlying Structure of Systematic Variations in the Event Mean Concentrations of Pollutants in Urban Runoff. *Wat. Sci. Tech.* **36**(8-9), 135-140. Publication after oral presentation at 7th International Conference on Urban Storm Drainage, Hanover, September 1996.

INTRODUCTION

Adequate water quality management requires mathematical models that can be used to estimate and predict overflow impacts under different engineering scenarios (House et al., 1993). It has been possible to construct models accounting for variations in runoff quantities with some success. However, although several pollutant storage and transport mechanisms have been postulated there still remains substantial unexplained variation in pollutant concentrations (Harremoës, 1994; Driver and Troutman, 1989). In an attempt to increase understanding and insight into the relevant pollutant runoff processes, the present study examines the underlying structure of EMC variations.

THE DATA

The data used in the present study originate from the five combined sewer catchments described in Chapter 2 of this thesis. These are the three Dutch catchments, Loenen, Bodegraven and Oosterhout, and the two Danish catchments, Cedervænget and Vester Paradisvej. See Chapter 2 for more detail on the catchments and data.

Of particular interest to the study in this chapter is the definition of a new variable used in the multivariate analysis. Under the assumption that all Biological Oxygen Demand (BOD) is included in the Chemical Oxygen Demand (COD), a "Non-Biological Oxygen Demand" variable has been defined as the COD value minus BOD value of the same sample.

It is assumed that BOD_5 is here roughly equal to $BOD_{infinity}$. In sewer overflows the BOD-decay rate constant is generally much higher than in pure sewage and wastewater treatment plant effluents and subsequently the $BOD_5/BOD_{infinity}$ ratio is close to 1. This suggests that this assumption is reasonable.

Prior to the analysis all the event mean concentrations were standardised with respect to the catchment's mean and standard deviation. Events with many missing EMC values were omitted from the data set and events with few missing values had these reconstructed using linear regression models on the non-missing values in order not to loose the valuable information present in the non-missing variables.

In the Danish data sets the total and dissolved biological oxygen demand (BOD) values have been estimated using regression models that had been calibrated on the Dutch data sets (see Chapter 2).

METHODOLOGY

This section contains a brief outline of the essential concepts of each of the statistical methods: cluster analysis, factor analysis, distribution analysis and correlation analysis.

Cluster analysis is designed to make an objective grouping of multivariate observations on the basis of the values of the observed variables. Considering each observation in an n-dimensional vector space (where n is the no. of variables), groups or clusters are formed by finding the events with the shortest distance between

Table 3.2. The distribution of the three rainfall-runoff pollution event types (the three clusters) for each of the combined sewer catchments, for their country of origin and for the entire set of events. The values in parenthesis are the percentages that have fallen within each of the three clusters.

Cluster	Combined Sewer Catchment					Country of Origin		Total
	Loenen NL	Bodegraven NL	Oosterhout NL	Cedervænget DK	V.Parad. DK	The Netherlands	Denmark	All events
Clus. 1	28 (60)	16 (64)	17 (61)	14 (67)	14 (70)	61 (61)	28 (68)	89 (63)
Clus. 2	11 (23)	4 (16)	6 (21)	4 (19)	4 (20)	21 (21)	8 (19)	29 (21)
Clus. 3	8 (17)	5 (20)	5 (18)	3 (14)	2 (10)	18 (18)	5 (12)	23 (16)
Total	47 (100)	25 (100)	28 (100)	21 (100)	20 (100)	100 (100)	41 (100)	141 (100)

them. In the present study, cluster analysis was applied to determine whether events could objectively be divided into groups and if so, to characterise the groups. The clustering method used here was Ward's method which is described in SAS (1990).

Factor analysis is in contrast to the cluster analysis designed to make an objective grouping of the variables rather than the observations. The interrelationship between a large number of variables is used to find a smaller number of new uncorrelated variables called factors that explain as much as possible of the variation found in the original data set. The role of factor analysis in the present study is to summarise the variations of the eleven interrelated water quality components thus portraying only the essence of the problem at hand. The essential mathematics behind factor analysis is summarised in Cattell (1965) and an extensive introduction to the techniques is given by Harman (1968). In the present study the factors have been extracted as principal factors and thereafter rotated using varimax rotation.

Distribution analysis has here encompassed a visual study of the factors' histograms with particular emphasis on the nature of the distributions right-hand tail and on the presence of mixture distributions as suggested in Hall et al. (1990). The advantage of looking at the factor's distributions, rather than those of the original observed variables, is that the dominant variations are contained in fewer variables and thus particular features of the distributions will be more distinct.

Correlation analysis, which is a measure of the linear relationship between two variables, has in the present study been used to evaluate the relationships between the pollutant factor variables and the rainfall, hydraulic and seasonal variables from the same runoff events.

From an engineering point of view the rainfall, hydraulic and seasonal variables would be considered as potential explanatory variables and methods such as canonical correlation could have been used to study this relationship. However, in the present study a factor and cluster analysis has been preferred because the aim of the present exploratory study is more that of proposing hypotheses rather than the formulation of an empirical model for the system. It is furthermore the authors' impression that the employed approach will facilitate interpretation and communication of results. Similar multivariate methods have earlier been applied in the context of urban water quality by Mulliss et al. (1994).

RESULTS AND DISCUSSION

The results of the factor analysis on the eleven pollutant variables show that their variation could be well described by three factors. The correlations between the three factors and the original EMC variables (the factor loadings) are given in Table 3.1. The three factors, which are independent linear combinations of the original variables, can be characterised verbally as in the bottom row of Table 3.1. Considering the processes of pollutant storage and transport in a combined sewer system, an objective separation of the variables into particulate and non-particulate is to be expected. The further separation of dissolved non-biological oxygen demand from the other dissolved pollutants could be related to a different source and/or to the processes taking place in the sewer system.

The far right column of Table 3.1 contains a measure of the proportion of the variables' variation that can be described by the three factors. Note that only 60% of the variations in dissolved BOD and only 38% in conductivity can be described by the three factors.

This indicates that variations of these two variables are dissimilar both to the other EMC variables and to each other. The entire set of rainfall-runoff pollution events was by the cluster analysis separated into three major types of events containing respectively 63%, 21% and 16% of the events. From the distributions over clusters (given in Table 3.2) it is evident that the three event types are equally common to all five catchments and to both the countries of origin.

Table 3.1. The resultant rotated factor pattern, the variance explained by each factor and a summary characterisation of each of the three factors (coefficient of absolute value less than 0.4 have been marked "-").

		Factors			Communalities, h_i^2 (%)
Variable		I	II	III	
Dissolved	Non-biological oxygen demand (COD-BOD)	-	0.590	0.466	72
(Not settled	Biological oxygen demand (BOD)	-	0.670	-	60
in 1 hour)	Kjeldahl Nitrogen	-	0.850	-	74
	Total Phosphorus	-	0.775	-	73
Particulate	Non-biological oxygen demand (COD-BOD)	0.939	-	-	89
(Settled	Biological oxygen demand (BOD)	0.849	-	-	74
within 1	Kjeldahl Nitrogen	0.906	-	-	85
hour)	Total Phosphorus	0.719	-	-	58
Total Solids		0.819	-	-	69
Dry Weight		0.823	-	-	81
Conductivity		-	0.580	-	38
Explained Variance (Cumulative %)		43%	66%	70%	
Summary Characteristics of the Three Factors		Particulate pollutants level	Dissolved/ fine particulate pollutants level	Non biologically degradable dissolved pollutants level	

The results of the factor and cluster analyses have been combined in the plot of Factor II against Factor I with the symbol indicating the event type (Figure 3.1). Similar plots with Factor III showed that the event types (clusters) had no relationship to this third factor. Summary characterizations of the event types can be formulated as follows using Figure 3.1a:

Cluster 1 : events with low particulate and dissolved pollutants concentrations,
Cluster 2 : events with high levels of dissolved pollutants but low particulate concentrations,
Cluster 3 : events with high particulate concentrations and medium ranged dissolved concentrations.

Though less pronounced, the plot of non-settleable BOD against settleable BOD (Figure 3.1b) exhibits the same spread of the three event types. The fact that the pattern of spreading is less pronounced relates to BOD having a variation pattern somewhat different to the other included EMC variables.

Figures 3.2 and 3.3 contain histograms of the particulate pollutants (Factor I and settleable BOD) and dissolved pollutants (Factor II and non-settleable BOD) respectively. Being essentially a different representation of the same information contained in Figures 3.1, 3.2 and 3.3 nevertheless illustrate how the different event types or clusters result in probability distributions that do not readily fall within the classes of frequently encountered two parameter normal, lognormal and extreme value distributions.

Considering the histogram of Factor I in Figure 3.2c, it is seen that the distribution of events in Cluster 1 and 2, and that of the events in Cluster 3 barely overlap. Due to the relatively large variance of Factor I for events in Cluster 3 and the small number of events, it is difficult to say whether or not this in fact could be well described by two parameters distributions. In the case of settleable BOD the separation of the events in Cluster 1 and 2 from the events in Cluster 3 is less pronounced. Using purely theoretical considerations, Song (1994) has shown how bimodality can be associated with turbulent flow. This would suggest that the events in Cluster 3 are events dominated by resuspension of sediments in the sewage system. This agrees well with the fact that the average maximum rainfall intensity of the events in Cluster 3 was found to be 6.3 mm/hr. as opposed to the average values found for the Clusters 1 and 2 which were 1.9 mm/hr. and 2.0 mm/hr. respectively.

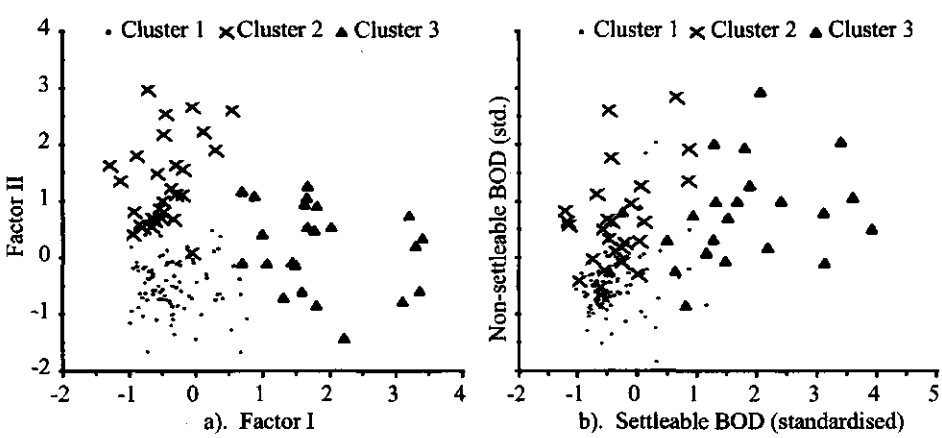


Figure 3.1. Plots illustrating the distribution of the three clusters in the Factor I - Factor II plane and in the settleable BOD - non-settleable BOD plane.

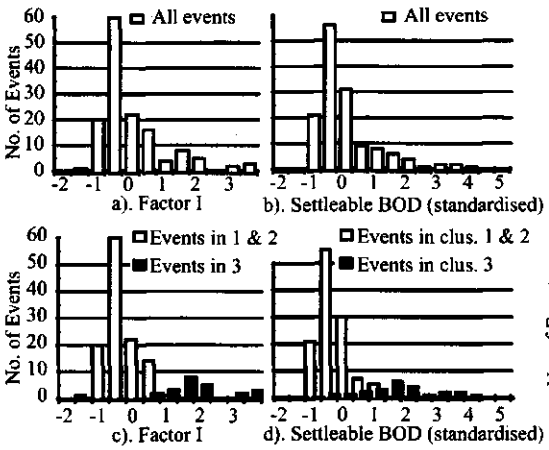


Figure 3.2. Histograms showing the distribution of events in clusters 1 and 2, and cluster 3 with respect to Factor I and settleable BOD.

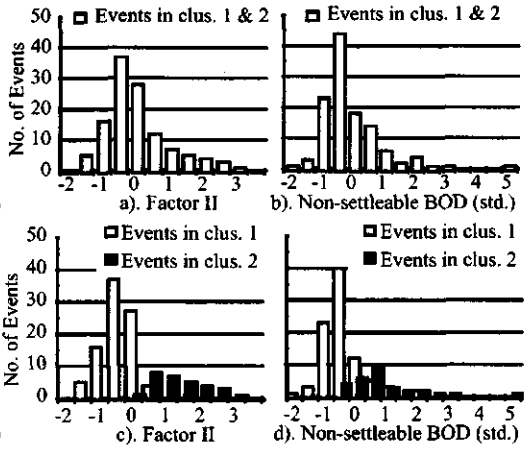


Figure 3.3. Histograms showing the distribution of events in cluster 1 and cluster 2 with respect to Factor II and non-settleable BOD.

Factor II has earlier in the chapter been characterised as describing the dissolved pollutants. The histograms in Figure 3.3 show the empirical distributions of Factor II and non-settleable BOD for events in Clusters 1 and 2. Though less pronounced, the same phenomena are apparent as in the case of the distributions of the particulate pollutants (i.e. Figure 3.2).

CONCLUSIONS

Through a factor analysis it has been possible to summarize the event mean pollutant concentration variables into three independent linear combinations of the eleven studied pollutant variables. Through the cluster analysis, CSO events could objectively be grouped into three different types of events. A summary characterization of the event types was made by relating these three groups to the three independent factors.

Though the observation data originated from three combined sewer catchments in the Netherlands and two in Denmark, it was found that events from all five catchments were similarly distributed in the three event type groups. The distribution analysis that followed the combined cluster and factor analysis has confirmed earlier findings that event mean pollutant concentrations are often best described as originating from mixture distributions. The results suggest that event based combined sewer overflow modelling could be improved either by modelling the different event types separately or by including a class variable indicating the event type and using a discriminant analysis to decide on the event type in question.

REFERENCES

Cattell, R. (1965). Factor Analysis: An Introduction to the Essentials. I. The purpose and underlying models. II. The role of factor analysis in reaserch. *Biometrics*, **21**.

Driver, N. E. and Troutman, B. M. (1989). Regression Models for Estimating Urban Storm-Runoff Quality and Quantity in the United States. *J. Hydrol.*, **109**(3/4), 221-236.

Hall, M. J., Ellis, J. B., and Brizio, M. G. (1990). On the Statistical Distribution of Event Mean Concentrations of Pollutants in Stormwater Runoff. Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Harman, H. H. (1968). Modern Factor Analysis. The University of Chicago Press.

House, M. A., Ellis, J. B., Herricks, E. E., Hvitved-Jacobsen, T., Seager, J., Lijklema, L., Aalderink, H. and Clifforde, I. T. (1993). Urban Drainage - Impact on Receiving Water Quality. *Wat. Sci. Tech.*, **27**(12), 117-158.

Mulliss, R., Ellis, J. B., Revitt, D. M. and Shutes, R. B. E. (1994). An Evaluation of Toxic Influences on Asellus Aquaticus (L) in an Urban Stream Environment. *Wat. Sci. Tech.*, **29**(1-2), pp. 199-207.

SAS (1990). SAS Institute Inc., SAS/STAT User's Guide, Version 6, Fourth Edition, Cary, NC: SAS Inst.Inc., 943pp.

Song, Q. (1994). *On the Dynamical Stochasticity of Event Mean Concentrations of Pollutants in Stormwater Runoff* (Report), Urban Pollution Research Centre, Middlesex University, Bounds Green Road, N11 2NQ, United Kingdom.

PART II

MODELLING

CHAPTER 4

EVENT LUMPED MODELLING OF COMBINED SEWER OVERFLOW

ABSTRACT

The complexity of models chosen for integrated urban water quality management may depend on the engineering alternatives to be examined, the current project phase and the available resources. In this chapter focus is on combined sewer overflow (CSO) models lumped both in space and in time as each rainfall-runoff occurrence is considered to be a separate single event. The resulting short calculation times make them suitable for return period and uncertainty analysis of CSO effects. Two event lumped non-linear regression models have been developed; a sewer overflow volume model and an event mean concentrations (EMC) model. Selection of model structure was based primarily on the nature of the resulting residuals and the identifiability of the estimated parameters. The rainfall, water quantity, chemical oxygen demand, Kjeldalh nitrogen and suspended solids data originate from a Dutch urban catchment. In the overflow volume model the runoff coefficient is expressed as a function of a wetness index which in turn is identified as a function of an estimated cut-off sinusoidal drying rate. The thus estimated drying rate was found to coincide well with mean monthly open water evaporation which is consequently used as input in the final model. Both total and slow settling fractions were included in the EMC model. Focus was here on achieving a description of the joint distribution.

This chapter is based on Grum, M. and Aalderink, R. H. (1999). Event lumped modelling of combined sewer overflow pollution. In preparation.

INTRODUCTION

Extreme statistics of CSO loads and effects are often used as design criteria in urban storm water management. The extreme statistics often take the form of return periods calculated using models having rainfall as major input. This is done both because rainfall has often been observed over a longer period and as rainfall is also the primary driving force this permits the analysis of proposed engineering scenarios.

Models may be dynamic describing the evolution of runoff, pollutant transport and storage over time or they may be event lumped describing only variables such as the total overflow volume and event mean concentration of pollutants as a function of the total depth, duration or mean rainfall intensity. This study is concerned with event lumped models for predicting CSO volumes and pollutant concentrations.

CATCHMENT AND DATA

In this study an extract of the Loenen data set described in Chapter 2 is used. See also Chapter 2 for a definition of a pollutants "slow settling fraction".

APPROACH

The following three components are characteristic of the applied model building approach:

- *a priori* conception of the processes taking place,
- principals of parsimony,
- graphical examination of the difference between observed and modelled values.

Our *a priori* conception of the processes taking place may range from the knowledge that water is pumped to the treatment plant to the idea that the amount of swirled up pollutant material depends on rainfall intensity. Inclusion of this knowledge and understanding both increases the engineering value of the model and may sometimes lead to more parsimonious model designs.

Through care in allotting of parameters the principals of parsimony aim at reducing the non-uniqueness of the estimated parameter set. This can help identify the most relevant terms and reduce the chance that the structure becomes specific to the given data set. However, parsimony alone does not necessarily lead to improved estimates on extrapolation beyond the ranges of the given data set (Reichert and Omlin, 1997).

Graphical examination of the residuals helps to reveal breaches in the underlying assumptions and may suggest modifications leading to an improved model structure. The most important plots include those of residuals against independent variables, candidate independent variables and against candidate independent class variables such as weekday-weekend.

In the approach, the three components are put to use simultaneously. A pattern in a residual plot could, in combination with our conception of the processes taking place and a high correlation found between two parameters, suggest certain alterations in the model structure. A comprehensive treatment of model building

in non-linear regression analysis can be found in Draper and Smith (1989). No outliers have been removed during the regression analysis. With rainfall as driving force, it is likely that model modifications would demonstrate that presumed outliers are neither observations of a distinctly different event type nor the result of typing errors, but simply that they are extreme cases of the occurrences being modelled. In the context of return period analysis it is important that the model is also able to model the extremes.

MODELS

Overflow Volume Modelling

The core of the final overflow model is based on the following balance: runoff will either be pumped to the treatment plant, be stored behind the weir or overflow into the surface water. This results in the following expression for the overflow volume, $V_{CSO,i}$:

$$V_{CSO,i} = \begin{cases} V^*_{CSO,i} & \text{for } V^*_{CSO,i} > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{where} \tag{4.1}$$

$$\underset{[L^3]}{V^*_{CSO,i}} = \underset{[L^2]\,[-]\,[L]}{A \cdot \alpha_i \cdot H_{RAIN,i}} - \underset{[L^2]\,[L/T]\,[T]}{A \cdot Q_{PUMP} \cdot D_{RAIN,i}} - \underset{[L^3]}{V_{STORE}} - \underset{[L^3]}{\varepsilon_{V,i}}$$

where A is the impervious catchment area, $H_{RAIN,i}$ and $D_{RAIN,i}$ are rainfall depth and duration respectively, α_i is the estimated runoff coefficient, Q_{PUMP} is an estimated pump capacity (relative to impervious surface area) and V_{STORE} is the sewer systems static storage capacity behind the weir. This static storage was calculated on the basis of the systems pipe geometry and was assumed to be a known fixed parameter during the parameter estimation. The errors $\varepsilon_{V,i}$ are assumed to be independent and normally distributed with mean zero and variance σ_ε^2. The runoff coefficient was assumed to be a linear function of the wetness index W_i value at the given event. The dimensions shown with L for length and T for time.

$$\alpha_i = a + b \cdot W_i \qquad \text{where} \qquad W_i = W_{i-1} \cdot e^{-r_{DRYING,i} \cdot (t_i - t_{i-1})} + H_{RAIN,i} \tag{4.2}$$

where a and b are estimated parameters representing the minimum runoff coefficient and the wetness dependency respectively. As shown in (4.2) the wetness index W_i, is itself a function of it's value at the previous rainfall event W_{i-1}, the preceding number of dry days $t_i - t_{i-1}$, a first order drying rate $r_{DRYING,i}$ and the depth of the current rainfall event. The annual variation in the drying rate was described by a cut-off sinusoidal expression (4.3) where t_i is the day of the year and where the constant parameters r_O, r_S and r_C were estimated from the rainfall-overflow data along with other model parameters.

$$r_{DRYING,i} = \begin{cases} r_i^* & \text{for } r_i^* > 0 \\ 0 & \text{for } r_i^* \le 0 \end{cases} \qquad \text{where} \quad r_i^* = r_0 + r_S \cdot \sin\left(\frac{2\pi \cdot t_i}{365}\right) + r_C \cdot \cos\left(\frac{2\pi \cdot t_i}{365}\right) \tag{4.3}$$

In the final model the expressions in (4.3) are replaced by the dependency to Dutch mean open water evaporation shown in (4.4).

$$r_{\text{DRYING},\,i} = \mathrm{r}_0 \cdot E(t_i) \tag{4.4}$$

where $E(t)$ is the Dutch long term mean open water evaporation rate for the t^{th} day of the year in [mm/day] and where the estimated constant coefficient r_0 has units [mm^{-1}]. The input evaporation data series used had been linearly interpolated from a series of mean monthly values. Estimation of parameters in the non-linear overflow volume modelling was done using the downhill simplex optimisation on a least square estimation criteria.

Event Mean Concentrations Model

The final form of the event mean concentrations model is shown in (4.5) where i = 1, 2, 3 for Kjeldahl Nitrogen, COD and SS respectively.

$$\ln(C_i) = a_i + b_i \cdot I_{\text{RAIN}} + \varepsilon_i \tag{4.5}$$

where I_{RAIN} is the mean rainfall intensity, a_i and b_i are estimated constant coefficients and ε_i is a random error with zero mean. The inverse of the slow-settling fraction of the pollutants was found to be best described by a straight line using the log of the total concentration as independent variable. This can be written as in (4.6) with i = 4, 5, 6 for that of Kjeldahl Nitrogen, COD and SS respectively.

$$\ln\left(\frac{1}{f_i}\right) = a_i + b_i \cdot \ln(C_i) + \varepsilon_i \tag{4.6}$$

Parameter estimation in these log-linear models was done using a least square criteria on each component. Simultaneous estimation using the simplex optimisation routine on a maximum likelihood criteria was used when examining more complex non-linear model formulations. The two methods yielded very similar results for the presented log-linear model.

RESULTS AND DISCUSSION

Overflow Volume Model

The results of selected overflow volume models are presented in Table 4.2 with the final model presented as Model 4. The results of the three other versions have been included in order to give an impression of the model building process and to illustrate the importance of the included terms. Results presented in Table 4.2 are for the following models.

Model 1. Constant runoff coefficient.
The runoff coefficient, α, in (4.1), is assumed to be constant over time and does therefore not vary over the year.

Model 2. Wetness dependent runoff coefficient with constant drying rate.
Here the runoff coefficient is assumed to be a function of a wetness index as defined by (4.2). The intercept a and gradient b are estimated along with the other parameters. Also estimated from the data is the constant drying rate r_{DRYING}.

Model 3. Wetness dependent runoff coefficient with a cut-off sinusoidal drying rate.
This is the model described by (4.1) to (4.3). The wetness index is here a function of a drying rate, which in turn is a function of the cut-off sinusoidal expression in (4.3). In Table 4.2, row D of model 3, the resulting drying rate, r^*, has been plotted. This has been calculated using the expression in (4.3) with the estimated values of the three parameters.

Model 4. Wetness dependent runoff coefficient with an evaporation dependent drying rate.
This is the model described by (4.1), (4.2) and (4.4). The wetness index is expressed as a function of the drying rate, which in turn is expressed as a function of the mean open water evaporation rate as shown in (4.4). In Table 4.2, row D of Model 4, the evaporation input series, $E(t)$ has been plotted.

The results of Model 4 (far right column in Table 4.2) would suggest that the effective impervious surface area is 9.2 ha (16%) rather than the documented/reported 15.8 ha (28%) and that the sewer receives excess water from an additional pervious area of up to 17.8 ha (31%). The suggested effective impervious area of 9.2 ha can be seen in the far right column of Table 4.2 where the estimated value for the runoff coefficient interceptor, a, is 0.58 rather than 1.0 which would correspond to the documented impervious 15.8 ha. The suggested additional pervious area, which contributes after very wet periods, is seen from the maximum value of the effective time varying runoff coefficient shown in row C of the far right column in Table 4.2. This maximum is 1.7 where 1.0 again corresponds to an area of 15.8 ha. After subtracting the impervious area (9.2 ha) above this gives a value of 17.8 ha. After long wet periods this results in an effective runoff surface of about 27 ha equivalent to 47% of the total catchment area. Characteristic of the gardens in the Loenen catchment are their convex shape and their slight elevation with respect to the impervious roads and pavements. It is likely that excess water from these so called pervious areas enters the sewer system.

The apparent seasonality in the runoff coefficient has been subject to much debate. In the initial studies (Bakker et al., 1989) a runoff coefficient was estimated for each event using a dynamic model and time series of rainfall and water level at the combined sewer overflow situated in the sewers lowest point. Seasonality in these runoff coefficients was evident and contributions from the impervious area was also then given as a possible explanation. No attempts were then made to relate this seasonal variation, through modelling or

otherwise, to surface wetness or moisture. Later studies, also based on dynamic modelling of individual events, revealed that the apparent seasonality could result from the combined effect of more intense thunder showers in the summer months with long drizzly rainfall events in the winter and questionable assumptions concerning the actual pump operation procedure.

The current results also suggest seasonal variation in the runoff coefficient but relates it through the wetness index, to surface moisture. In spite of this highly plausible physical explanation and the good fit between the thus empirically estimated drying rate and mean monthly open-water evaporation rates (see Table 4.2, row D), it is important to realise that because the long less intense rainfall events occur generally during the wet winters and the short intense storms during the summers, it is not possible to completely exclude the theory that the apparent seasonality results from a faulty description of the pump operation procedures. The pump description in this event lumped model assumes that the pump operates at its estimated capacity all through the rainfall-runoff event. In Loenen there are in fact two screw pumps, which are turned on and off depending on the local water level. In the event lumped description an average is assumed and estimated.

The location and the cut-off level of the estimated drying rate curve of Model 3 is seen to coincide well with the Dutch mean monthly open-water evaporation (Table 4.2, row D). In Model 4 the expression (4.4) was therefore used as in the model in place of the cut-off sinusoidal expression in (4.3). The number of model parameters was thus reduced from 6 to 4 without loss of fit. This further removes the somewhat rigid sinusoidal structure from the model. Though wetness indices are widely used in rural rainfall-runoff modelling, references to expressions with first order drying rates were not found.

Parameter correlation coefficients between the Model 3 parameters were below 0.4 (absolute). An exception to this was the estimated cosine coefficient with a correlation of -0.7 with each of the two other parameters in the drying rate expression. The estimated pump capacity of 0.47 mm/hr is low compared to the reported 0.7 to 1.1 mm/hr. The pump capacity was found to correlate somewhat (0.4) with the constants relating the runoff coefficient to the wetness index. This is an example of a case in which our rough *a priori* knowledge of the pump capacity could beneficially have been incorporated quantitatively into the parameter estimation criteria (Reichert, 1997 and Chapter 5).

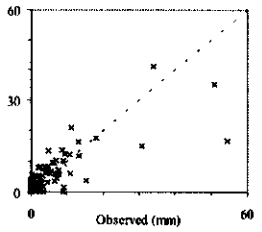
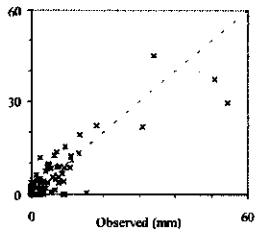
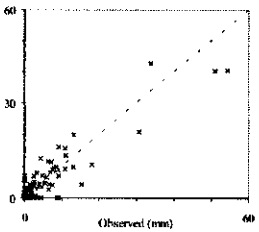
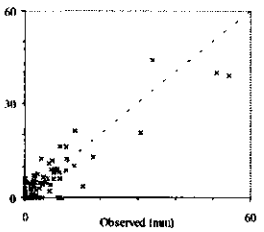
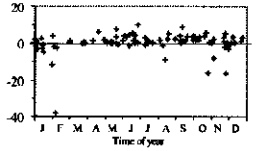
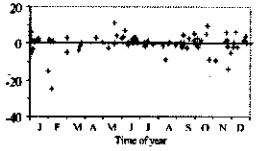
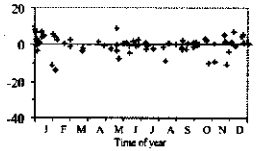
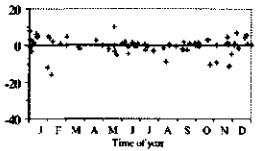
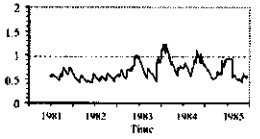
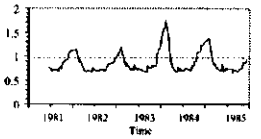
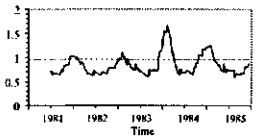
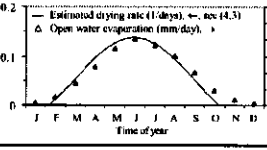
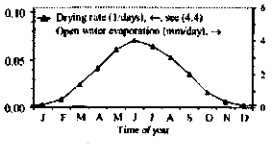
Within the framework of identifiable event lumped rainfall-runoff models future efforts should aim at incorporating mean open water evaporation as a model input, examining the generality and possible weaknesses of using a wetness index with a first order drying rate and comparison with overflow predictions made using dynamic models.

Event Mean Concentrations Model

Results of the EMC and slow-settling fraction modelling are presented in Table 4.3 and Table 4.4. Each column in Table 4.3 represents one of the three pollutant components N-KJ, COD and SS with the total event mean concentration models in the top half and the slow-settling fraction in the bottom half.

For the total EMCs the coefficients of determination, R^2, of 0.24 to 0.35, are much lower than those achieved for the overflow model. A great variety of independent variables and model structures were investigated in an attempt to obtain better predictability. Candidate independent variables included rainfall depth, duration, maximum intensity, functions of time of year, functions of time of day, weekday-weekend, preceding dry weather period, wetness index from the overflow volume model above and various combinations of these. It was possible to obtain better prediction using observed overflow volume and observed maximum discharge intensities as explanatory variables. However, once the predicted values of these variables were introduced the coefficients of determination dropped to about 0.2-0.3.

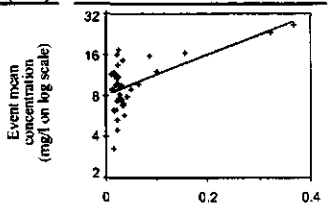
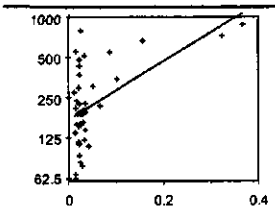
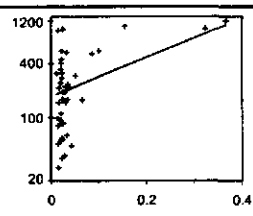
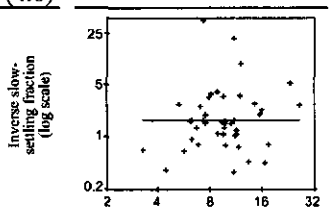
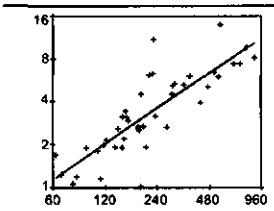
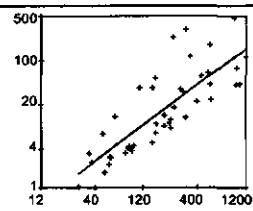
Table 4.2. An overview of the overflow volume models. R^2 is the coefficient of determination and adjusted-R^2 has been adjusted for the number of estimated parameters.

Results of Overflow Volume Modelling				**MODEL 1.** Constant runoff coefficient	**MODEL 2.** Wetness dependent runoff coefficient with constant drying rate	**MODEL 3.** Wetness dependent runoff coefficient with seasonal drying rate	**MODEL 4.** Wetness dependent runoff coefficient with evaporation dependent drying rate
Equations				(4.1)	(4.1)-(4.2)	(4.1)-(4.3)	(4.1)-(4.2),(4.4)
Number of parameters, R^2 (adjusted R^2)				2 parameters, R^2 = 0.68 (0.66)	4 parameters, R^2 = 0.78 (0.77)	6 parameters, R^2 = 0.85(0.84)	4 parameters, R^2 = 0.84 (0.83)
Runoff coefficient: Intercept	[-]	a	(4.1)	0.82 (0.036)	0.20 (0.15)	0.66 (0.056)	0.58 (0.032)
Slope	[1/mm]	b	(4.2)	-	0.0049 (0.0011)	0.0028 (0.00038)	0.0033 (0.00022)
Pump capacity	[mm/hr.]	Q_{PUMP}	(4.1)	0.22 (0.14)	0.24 (0.088)	0.47 (0.090)	0.43 (0.12)
Drying rate	[days^{-1}], [mm^{-1}]	r_0	(4.3)	-	0.015 (0.0026)	0.053 (0.0096)	0.018 (0.0032)
Sine coefficient	[days^{-1}]	r_S	(4.3)	-	-	0.023 (0.012)	-
Cosine coefficient	[days^{-1}]	r_C	(4.3)	-	-	-0.081 (0.014)	-
Error variance	[(mm)^2]	σ_ε^2	(4.1)	5.7^2	4.9^2	4.2^2	4.3^2
A Predicted against observed overflow depth (mm)							
B Error (mm) against time of year (zeros not plotted)							
C Runoff coefficient against time				(constant runoff coefficient)			
D Estimated drying rate (1/day) and mean open water evaporation (Δ) against time of year				-	(constant drying rate)		

Apparent in all three plots of predicted and observed total EMCs (Table 4.3, graph row A) is that the error variance does not appear to be independent of the mean rainfall intensity. Attempts to estimate this dependency using a maximum likelihood criterion were unsuccessful when equal weight was given to observations from events with low and high rainfall intensity. Future efforts should aim at solving this breach of assumptions.

For the slow-settling fractions the coefficients of determination are with total EMCs as independent variables (i.e. their observed values have been used during estimation). On application these are only available as predictions and the actual coefficients of determination with respect to the mean rainfall intensity will be much lower.

Table 4.3. Results of the EMC (top half) and slow-settling fraction modelling (bottom half).

Results of EMC Modelling			**Kjeldahl Nitrogen (N-KJ)**	**Chemical Oxygen Demand (COD)**	**Suspended Solids (SS)**
			Event Mean Concentrations		
No. of parameters			2	2	2
R^2			0.35	0.28	0.24
Adjusted R^2 §			0.33	0.27	0.23
Units*	Symb.	Equat.			
Intercept	a	(4.5)	2.10 (0.06)	5.15 (0.11)	4.91 (0.16)
Gradient	b	(4.5)	3.35 (0.72)	4.99 (1.24)	6.94 (1.91)
Error var.	σ_ε^2	(4.5)	0.33^2	0.58^2	0.89^2
A EMC against event mean rainfall intensity, I_{RAIN}					
			Slow Settling Fractions		
No. of parameters			2	2	2
R^2 †			0.0	0.69	0.65
Adjusted R^2 † §			0.0	0.69	0.64
Units**	Symb.	Equat.			
Intercept	a	(4.6)	0.54 (0.15)	-3.24 (0.46)	-3.48 (0.74)
Gradient	b	(4.6)	0.0	0.82 (0.09)	1.20 (0.14)
Error var.	σ_ε^2	(4.6)	0.95^2	0.37^2	0.90^2
B Inverse of slow-settling fraction against Event Mean Concentration (EMC)					

† Note that these coefficients of determination are with the total event mean concentrations as independent variables. On application these would in turn be a function of the mean rainfall intensity.

§ Adjusted R^2 is the coefficient of determination compensated for the number of estimated parameter in the model.

* Units of a, b and σ_ε^2 are [ln(mg/l)], [ln(mg/l)·min/mm] and [(ln(mg/l))^2] respectively.

** Units of a, b and σ_ε^2 are [-], [1/ln(mg/l)] and [-] respectively.

Table 4.4. Correlation matrix defining the joint probability distribution of the error terms in the EMC and slow-settling fraction model (see (4.5) and (4.6)).

			1 $C_{N\text{-}KJ}$	2 C_{COD}	3 C_{SS}	4 $f_{N\text{-}KJ}$	5 f_{COD}	6 f_{SS}
1	Kjeldahl Nitrogen (N-KJ)	$C_{N\text{-}KJ}$	1.00	0.69	0.54	0.06	-0.51	0.04
2	Chemical Oxygen Demand (COD)	C_{COD}	0.69	1.00	0.94	-0.38	0.06	0.11
3	Suspended Solids (SS)	C_{SS}	0.54	0.94	1.00	-0.47	0.20	0.12
4	Slow-settling fraction N-KJ	$f_{N\text{-}KJ}$	0.06	-0.38	-0.47	1.00	-0.25	-0.04
5	Slow-settling fraction COD	f_{COD}	-0.51	0.06	0.20	-0.25	1.00	0.23
6	Slow-settling fraction SS	f_{SS}	0.04	0.11	0.12	-0.04	0.23	1.00

No explanatory variable was found for the slow-settling fraction of Kjeldalh Nitrogen and the resulting R^2 is therefore zero. As suggested by the cloud of points in the plot of inverse slow-settling fraction of Kjeldahl Nitrogen against its EMC (Table 4.3, N-KJ column graph row B), the value of b in (4.6) was not significantly different from zero and the term was consequently removed from the model. The consequence is that the inverse slow-settling fraction of Kjeldahl Nitrogen is in fact described as a log normally distributed random variable. For many practical applications such as return period analysis of detrimental CSO effects a distribution is much better than no model at all and it is important to realise that a coefficient of determination is zero does not render the model useless.

A linear relationship was found between the logarithms of the inverse of the slow settling fraction and the total event mean concentration for both COD and SS. As shown in the two figures to the right in Figure 4.3 row B, high event mean concentrations were associated with high values of the inverse slow settling fraction. This means that high concentrations are thus predominately associated with increases in material that settles within one hour.

Both in event lumped and dynamic modelling a much higher predictability is found for water quantity than water quality variables. In practice this means that the value of using the relatively well characterised rainfall in evaluating return periods of detrimental CSO effects under various structural or control scenarios is low. The uncertainty in the return periods would be expected to depend to a large extent on the size of the CSO water quality data set itself. The parameter correlation matrix for the water quality model is not presented in full but some main aspects of this 11 by 11 matrix will be shortly discussed. Although the correlation between the a's and b's in (4.5) were generally low the correlation amongst the a's and amongst the b's was generally high (0.5 to 0.9). This would suggest that the number of estimated parameters could be reduced without significant loss of fit by restructuring the model to have a common dependency on the mean rainfall intensity with a scaling factor for each of the three components.

The presented models have been used to evaluate uncertainty in return period analysis of combined sewer overflow effects (Grum et al., 1998 and Chapter 9). In such applications event lumped models have an advantage over dynamic models as they have short simulation times and have reasonably satisfied assumptions concerning errors and uncertainty. The water quality model defined by (4.5) and (4.6) is a single model whose parameters where estimated simultaneously and for which correlation matrix corresponding to the joint distribution of the error terms is given in Table 4.4. Note that particularly the errors of the event mean concentrations of COD and SS are highly interdependent. Chapter 9 introduces a methodology for the analysis of uncertainty in return periods where a distinction is made between inherent variation and uncertainty. The joint distribution of the error terms in the event mean concentrations model describes inherent variation and is in Chapter 9 treated as such.

CONCLUSION

It has been possible to identify and estimate parameters in event lumped models of combined sewer overflow (CSO) volume, event mean concentration and slow-settling fractions of three pollutants; Kjeldahl Nitrogen, Chemical Oxygen Demand (COD) and Suspended Solids (SS). With rainfall as input, high coefficients of determination were found for water quantity but not for the concentration variables.

Using only event lumped rainfall and overflow volume data, the catchment drying rate was identified and found to coincide surprisingly well with mean open water evaporation data. Evaporation was subsequently introduced into the model as an explanatory variable. Future efforts should be made to examine the generality and value of the presented wetness index particularly in urban drainage systems receiving runoff from pervious and semi-pervious surfaces.

REFERENCES

Bakker, K., Timmer, J. L. and Wensveen, L. D. M. (1989). De vuiluitworp van gemengde rioolstelsels. NWRW/STORA Report no. 5.2, Min. Of Housing, Phys. Planning and Env., SDU Publs., P.O. Box 20014, 2500EA 's-Gravenhage.

Draper, N. R. and Smith, H. (1989). Applied regression analysis. Wiley series in probability and mathematical statistics. ISBN: 0-471-02995-5.

Grum, M. and Aalderink, R. H. (1998). Uncertainty in Return Period Analysis of Combined Sewer Overflow Effects Using Embedded Monte Carlo Simulations. *Wat. Sci. Tech.*, 39 (4), pp. 233-240. Published after oral presentation at the AQUATECH conference in Amsterdam, September 1998.

Reichert, P. (1997). On the necessity of using imprecise probabilities for modelling environmental systems. *Wat. Sci. Tech.*, **36**(5), 149-156.

Reichert, P. and Omlin, M. (1997). On the usefulness of overparameterized ecological models. *Ecological Modelling*, 95(2-3), pp.289-299.

CHAPTER 5

DETERMINISTIC AND STOCHASTIC MODELLING

ABSTRACT

In this chapter the underlying assumptions of deterministic and stochastic modelling are examined in the context of water quality modelling. Particular attention is given to assumptions concerning the source of the deviation between modelled and observed values and how this results in two very different approaches to parameter estimation. Potentials and limitations of including physical, chemical and biological theory into stochastic models are evaluated. Methods and approaches generally associated with stochastic modelling are discussed and applied in the context of both approaches. These include the virtues of the linear and non-linear state space formulations, common estimation criteria, exploring parameter space and aspects of identifiability including the quantitative incorporation of a priori knowledge into the parameter estimation process.

A case study is presented in support of the theoretical outline. A rainfall-runoff model for a combined sewer system is estimated both as a deterministic and stochastic model. The physical description used in the two models is identical; they differ only in their error assumptions. The modelling tool used was developed specifically to estimate parameters of a model as both deterministic and stochastic. The case study also highlights some of the practical problems that need to be addressed when incorporating water and water quality theory into stochastic modelling.

INTRODUCTION

Models are often used to support decisions in urban water quality management. There exist many different approaches to water modelling. Two main lines are the traditional deterministic and the stochastic modelling approaches. In both practice and research the choice of approach appears to depend more on the background of the modeller than on the problem being addressed and the available economic resources.

With the aim of contributing to bridging this gap between traditional deterministic and stochastic modelling approaches, this chapter firstly addresses the central differences between the two approaches. As argued, via the parameter estimation problem, it is important not only to realise but also to keep in mind that the difference between stochastic and deterministic modelling lies in the implicit error assumptions.

Secondly, the chapter outlines and discusses some principles and methods relevant to both approaches but which, for historic reasons, are generally only associated with stochastic modelling. These include the state space formulation, various estimation criteria, empirical and mechanistic model structure, identifiability and parsimony. A case study is presented in which both methodologies and a number of the associated methods are applied to a water quantity model of the Loenen sewer system.

Out of this chapter spring some key questions: When should which modelling approach be used? How much detail should be included in the model? These and other questions are discussed in Chapter 8, which concludes the modelling part of this thesis.

DETERMINISTIC AND STOCHASTIC MODELLING

Parameter Estimation in Deterministic Modelling: NO UP-DATE

Traditionally deterministic water quality models are calibrated against observed data. The simulation results are made to fit the observed data as best possible by adjusting the model parameters to values that the modeller's engineering experience tells him are reasonable values. In other cases deterministic models have their parameters estimated by using search algorithms that minimise the sum of squared errors. What ever the case the following points are in principal equally valid and "parameter estimation" will here be used to describe either of the two methods.

The parameters of a deterministic model are estimated by sequentially running a full simulation, comparing the results with the observations and evaluating the fit. This is repeated several times usually as part of some structured optimisation procedure. When no further improvement is obtained the process stops. Figure 5.1 illustrates a single such full simulation (the solid line) and the observations (the crosses). Modelled is the response of a fictitious flow to the fictitious rainfall input plotted above.

What in fact is done during one such simulation can be described in other words. The simulation starts at time $t=0$ and an observation is first encountered at time $t=1$. At this point the deviation between the modelled value is noted but no adjustment is made to the estimated value of the flow. It is not *up-dated*. Total confidence is placed in the modelled value, which is taken as our best estimate of reality in spite of the fact that this didn't coincide with the observation. The implicit assumption is that the deviation between the

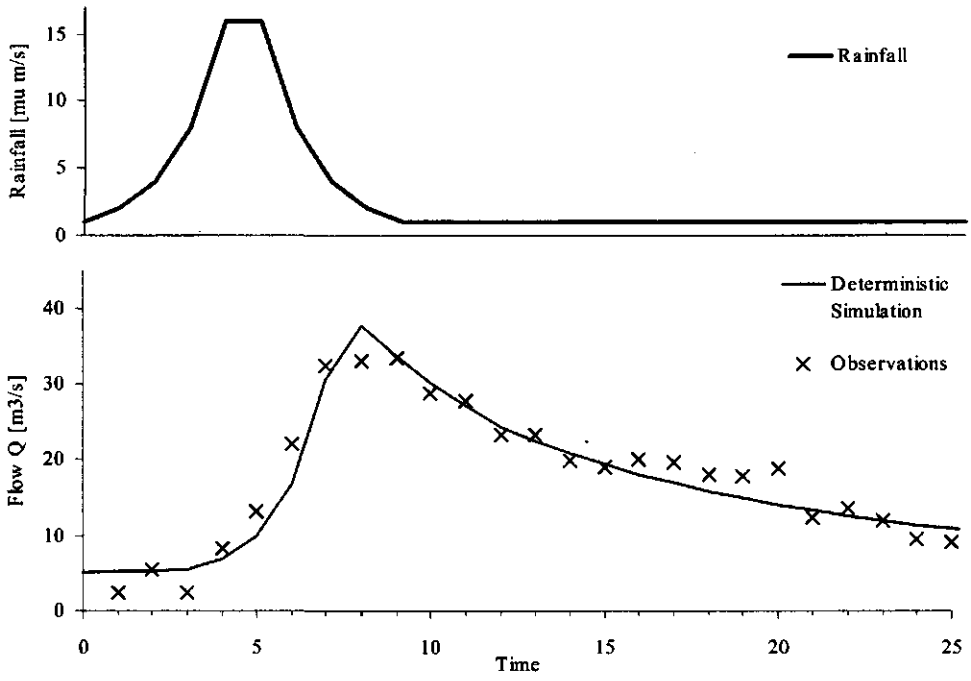


Figure 5.1. During parameter estimation and/or calibration in traditional deterministic modelling deviation between model and observation is implicitly assumed to result from observation error only. Therefore no adjustment is made to the flow on the presence of an observation. (Example shows fictitious data).

modelled and observed values results from observation error alone, no up-dates are made to the state of the system and we continue our simulation of reality from the modelled value.

- *In deterministic modelling the implicit assumption is that our description of reality is perfect and that all deviation between modelled and observed values is the result of observation error.*

Note that the term *observation error* is used rather than *measurement error*. This is because observation error encompasses measurement error, sampling error and error due to inhomogeneity.

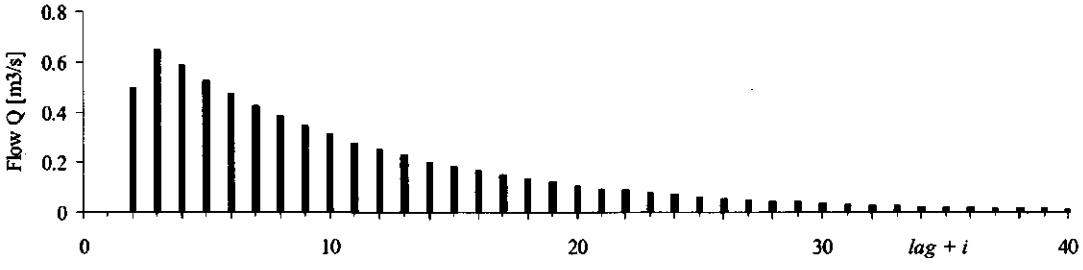


Figure 5.2. A fictitious unit hydrograph (impulse response function) defining the values of a_i in (5.1).

Parameter Estimation in Stochastic Transfer Function Modelling: COMPLETE UP-DATE

Consider the following model in which the flow, Q, is expressed as a function of earlier observations of rainfall, R, with an added error, e. The model is defined in discrete time where for example the subscript *t-2-lag* refers to the variables value *(2+lag)* time steps earlier than the time t in which the variable is defined.

$$Q_t = a_0 \cdot R_{t-lag} + a_1 \cdot R_{t-1-lag} + a_2 \cdot R_{t-2-lag} + a_3 \cdot R_{t-3-lag} + \ldots + a_n \cdot R_{t-n-lag} + e_t \qquad (5.1)$$

This could for example be a simple unit hydrograph consisting of a rise followed by an exponentially decay with the values of a_0, a_1, a_2, ..., a_n as shown in Figure 5.2.

Ljung (1987) uses a similar model as an initial attempt in an example problem where river flow is modelled as a function of rainfall. Ljung then proceeds to test a model of the following form (5.2) and ascertains that the error variance is significantly reduced by including the flow's value in the previous time step as an explanatory variable.

$$Q_t = \delta \cdot Q_{t-1} + \omega_0 \cdot R_{t-lag} + \omega_1 \cdot R_{t-1-lag} + e_t \qquad (5.2)$$

True enough, the flow is now expressed as a function of its value in the previous time step but in the paragraphs below it will be shown that this is a mere technicality when comparing it to the model defined by equation (5.1). The error variance found for (5.1) is undoubtedly much smaller then that found for (5.2). But

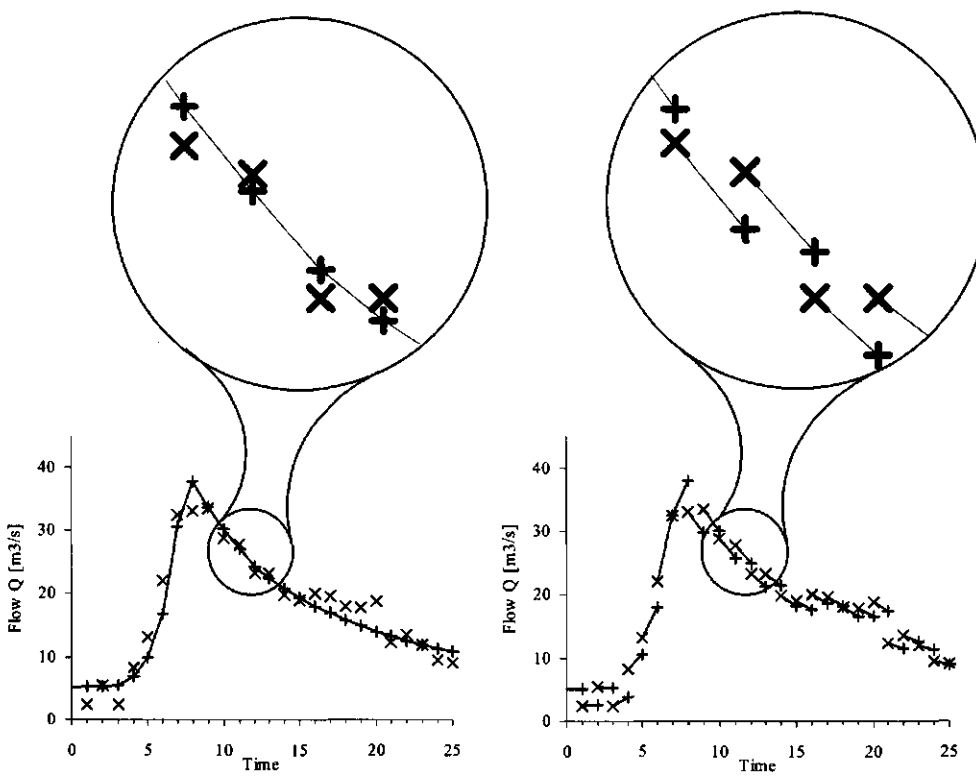


Figure 5.3. Left: Parameter estimation with no up-date implying no model error and thus a deterministic model (5.6). Right: Parameter estimation with complete up-date. Up-date implies model error is present thus a stochastic model (5.2). Observation = cross and prediction = plus.

do these two errors describe the same thing? Are they at all comparable?

The prediction of the flow, Q, as defined by the model in the expression (5.1) can be written as

$$\hat{Q}_t = a_0 \cdot R_{t-lag} + a_1 \cdot R_{t-1-lag} + a_2 \cdot R_{t-2-lag} + a_3 \cdot R_{t-3-lag} + \ldots + a_n \cdot R_{t-n-lag} \tag{5.3}$$

where the hat symbol, ^ , is used to indicate that this is a prediction of the variable. A response of the form shown in Figure 5.2 with an initial rise followed by an exponential decay can be expressed using fewer parameters (5.4).

$$\hat{Q}_t = d \cdot \hat{Q}_{t-1} + w_0 \cdot R_{t-lag} + w_1 \cdot R_{t-1-lag} \tag{5.4}$$

where the model parameters d, w_1 and w_2 could be found from the following relation to the response function defined by the a's in (5.3) as follows

$$\begin{aligned} a_0 &= w_0 \\ a_1 &= d \cdot w_0 + w_1 \\ a_2 &= d^2 \cdot w_0 + d \cdot w_1 \end{aligned} \tag{5.5}$$

Thus, the complete model in (5.1) may be rewritten with fewer parameters as follows (5.6).

$$Q_t = d \cdot \hat{Q}_{t-1} + w_0 \cdot R_{t-lag} + w_1 \cdot R_{t-1-lag} + e_t \tag{5.6}$$

At first sight the two models defined be the equations (5.6) and (5.2) now look very much alike. There is, however, a world of difference.

At each time step during parameter estimation ***in the first model*** (5.6) (a reformulation of (5.1)),

- prediction departs from the ***predicted*** value in the previous time step,
- observed values are used only to calculate the error, e_t, and subsequently to find the fit,
- no adjustment is made to the current flow estimate, that is, NO UP-DATE IS PERFORMED,

thus

- deviation between modelled and observed values is assumed to result from observation error alone,
- the model's system description is assumed to give a perfect description of the system behaviour,
- the model is a DETERMINISTIC model.

At each time step during parameter estimation ***in the second model*** (5.2),

- prediction departs from the ***observed*** value in the previous time step,
- the current flow estimate is adjusted to be equal to the observed flow, that is, a COMPLETE UP-DATE IS PERFORMED,

thus

- deviation between modelled and observed values is assumed to result from model error alone,
- the model's system description is assumed only partly to describe system behaviour and the remaining behaviour is assumed to be random compared to given explanatory variable (in this case rainfall R_t),
- the model is a STOCHASTIC model.

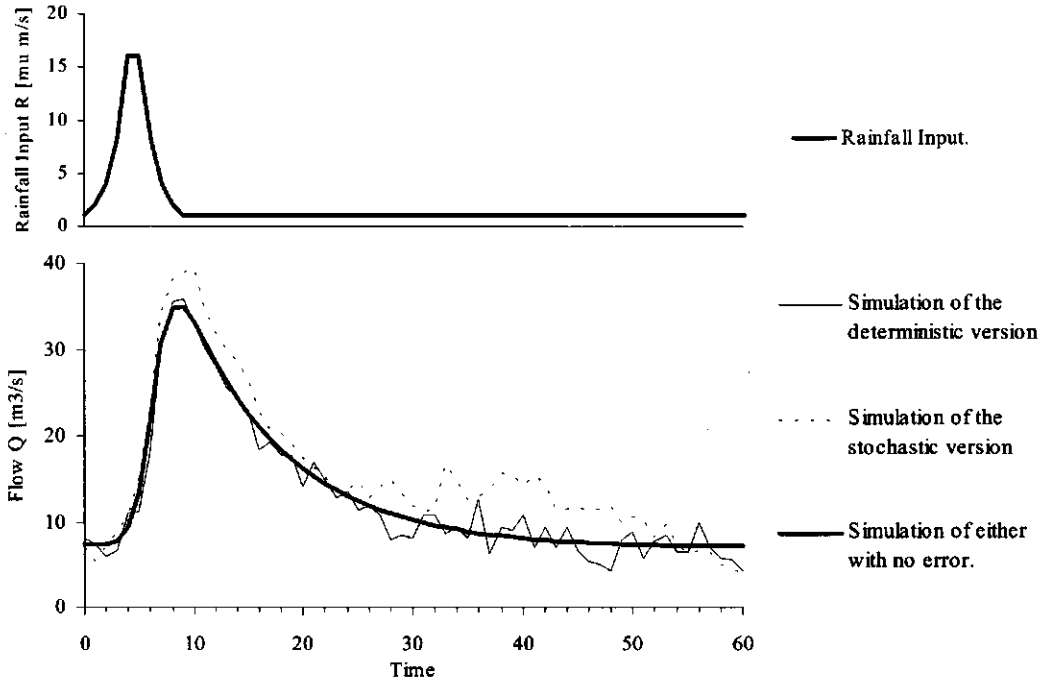


Figure 5.4. Comparison of simulations of the two models (5.6) (solid line) and (5.2) (dashed line). The simulations have been performed with the same model parameters ($d = \delta = 0.9$, $w_0 = \omega_0 = 0.5$, $w_1 = \omega_1 = 0.2$ and $lag = 2$) and the same error variance σ_e^2 of 1.5^2.

The error term in (5.6) is an observation error whereas the error term in (5.2) is model error describing random behaviour of the system. The two error terms represent very different things and a direct comparison is therefore uninteresting. What generally would be interesting after parameter estimation with each of the models would be to evaluate whether the resulting error series, the residual series, in fact was made up of *independent* realisations of a random variable. This is discussed in more detail later in the chapter.

In the above discussion the difference between the two simple models is highlighted in view of the parameter estimation or calibration. This may be supported further visually by model simulations. Simulations of the two models (5.6) and (5.2) have been plotted in the same graph in Figure 5.4. The simulations have been performed with the same model parameters ($d = \delta = 0.9$, $w_0 = \omega_0 = 0.5$, $w_1 = \omega_1 = 0.2$ and $lag = 2$) and the same error variance σ^2 of 1.5^2. The visual difference between the two models is clear and unmistakable. The deterministic version of the model is a smooth curve with an overlaid fuzz whereas the stochastic version has more the look of a shiver. This is because in the first deterministic version of the model (5.6) the flow at time t is a function of an error term only of time t (the observation error). In the stochastic version of the model (5.2) the flow is not a function of its predicted value in the previous time step but of its simulated value and therefore a function of the error terms of all previous time steps (the model error).

In the deterministic version (5.6) $$Q_t = f\left(R_{t-lag}, R_{t-1-lag}, R_{t-2-lag}, \ldots, e_t\right) \quad (5.7)$$

In the stochastic version (5.2) $$Q_t = f\left(R_{t-lag}, R_{t-1-lag}, R_{t-2-lag}, \ldots, e_t, e_{t-1}, e_{t-2}, e_{t-3}, \ldots\right) \quad (5.8)$$

Generally the following can be said about the stochastic models traditionally applied in statistical time series analysis such as the transfer function models in Box and Jenkins (1976) and Chatfield (1996).

- *In stochastic modelling with complete up-date the implicit assumption is that the observations of reality are perfect and that all deviation between modelled and observed values is the result of model error describing random behaviour of the system.*

A complete up-date of the state variable(s) is performed as the modelled variable is adjusted to be equal to the observed value whenever an observation is available.

Parameter Estimation with Separation of Variance: KALMAN UP-DATE

Models are always a simplification of reality and assuming that all deviation results from observation error alone would generally be incorrect. On the other hand observations are also often uncertain due to inhomogeneity, variations within sampling and analysis procedure. A third possibility is to accept the presence of both observation and model error.

When an observation is encountered, whilst modelling in time, an improved estimate of the variables true value may be obtained as a weighted average between the model's predicted value and the observed value. This weighted average is called the up-dated value. Modelling would then proceed from this new up-dated value.

In the deterministic version of the simple unit hydrograph model (5.6) there was no up-dating during parameter estimation. In the stochastic version presented in (5.2) there was complete up-date during parameter estimation. A model that up-dates to a weighted average between the predicted and observed values could be formulated by simply merging the two models (5.6) and (5.2) to the form (5.9).

$$Q_t = \delta \cdot \left(k \cdot Q_{t-1} + (1-k) \cdot \hat{Q}_{t-1} \right) + \omega_0 \cdot R_{t-lag} + \omega_1 \cdot R_{t-1-lag} + e_t \qquad (5.9)$$

The weighting factor k is called the Kalman gain and the whole methodology for separating the variation into observation and model errors is called the Kalman filter after the author who first presented it in the 1960s. Note that Q with the hat is the predicted flow whereas Q without the hat is the observed flow.

In stochastic modelling with Kalman up-date,

- *In stochastic modelling with Kalman up-date the implicit assumption is that deviation between modelled and observed values are the result of both observation errors and model noise.*

Figure 5.5 illustrates the how prediction in the Kalman filter proceeds from a weighted average of the model's predicted value and the observed value.

The value of the Kalman gain, k, depends on the uncertainty associated with the model prediction relative to the observation's uncertainty. For a one dimensional model such as (5.9) the Kalman gain can be expressed as follows.

$$k = \frac{\sigma^2_{MODELLED,\,t}}{\sigma^2_{MODELLED,\,t} + \sigma^2_{OBSERVED}} \tag{5.10}$$

where $\sigma^2_{MODELLED,\,t}$ is the variance of the model's predicted value (at time t) and $\sigma^2_{OBSERVED}$ is the observation variance.

Figure 5.6 graphically illustrates how the Kalman up-date works. In Figure 5.6A the uncertainty of the model's prediction is illustrated by its broad and flat probability density function. In such a case the up-dated value will be close to the relatively more certain observation. In Figure 5.6B it is the model prediction that is the most certain and consequently the up-dated value will lie closer to the predicted value.

Modelling the Variances

Above it is outlined how the Kalman gain is an expression of the uncertainty in the model prediction relative to the combined uncertainty. But where do these variances come from.

The observation error variance is most often assumed to be the same for all observations of a given variable. Thus not changing in time. If a maximum likelihood estimation criterion is used (see section on estimation criteria below) for an off-line parameter estimation then the observation error variance can be estimated from the data along with the other model parameters.

However, the variance of the model prediction will generally vary in time. The prediction variance will increase with time as the duration since the last observation increases. The variance will also increase or decrease depending on the rigidness of the system itself in its current state. For example, in the case study that follows the variance of the predicted water level abruptly decreases once overflow over the weir begins.

In using the Kalman filter also the variances of the system states have to be modelled in time. The rate at which the variance increases in time may be assumed constant and this variance may be estimated along with other model parameters. On the presence of an observation the variance of the prediction is up-dated in the same way as the prediction itself. If the modelled value is relatively uncertain then the presence of a new

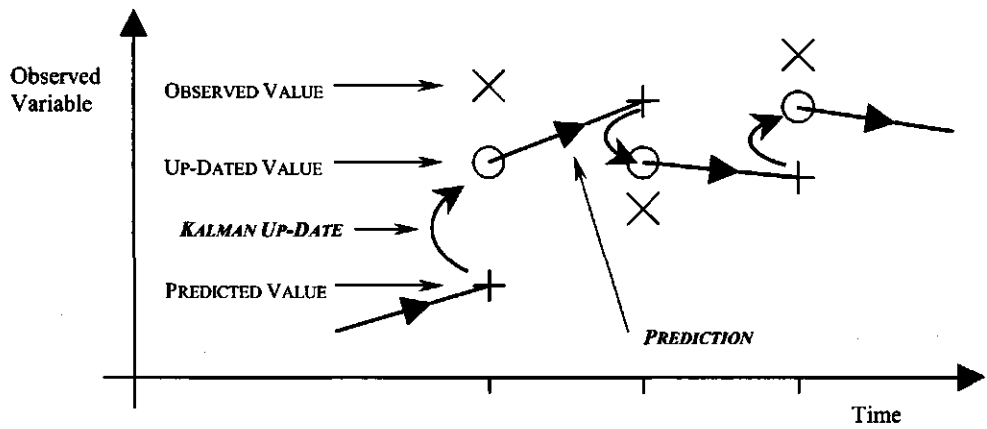


Figure 5.5. Sketch of the principal elements of the Kalman filtering. Both model and observation errors are assumed to be present and prediction therefore proceeds from a weighted average between the observed and the modelled values: the Kalman up-date.

observation will result in a major reduction of the variance. Conversely, if the modelled value is relatively certain then the new observation only results in a minor reduction of the variance.

In Appendix A and Appendix B are the equations for the linear and non-linear Kalman filter respectively. Included are the equations for up-dating the state variances and for calculating their propagation in time.

The Kalman filter as Convergence Routine and with Off-Line Estimation

The Kalman filter has often been used as a convergence routine in parameter estimation. This is done by placing the essentially "constant" parameters whose values are to be estimated as state variables in the state vector and assigning a simple empirical model such as a first order autoregressive model and letting the parameter converge to its "true" value in time. For this to work successfully the system variances and observation variances have to be fairly well known.

In the case study that follows and in Chapter 6, the Kalman filter has *not* been used in its capacity of parameter estimator. The Kalman filter has been used uniquely for obtaining an improved estimate of the observed and unobserved system state variables on the presence of an observation. This is then combined with the optimisation of a likelihood criterion that is calculated once for each complete run-through of the time series used for parameter estimation. The model parameters, including the error variances, are estimated

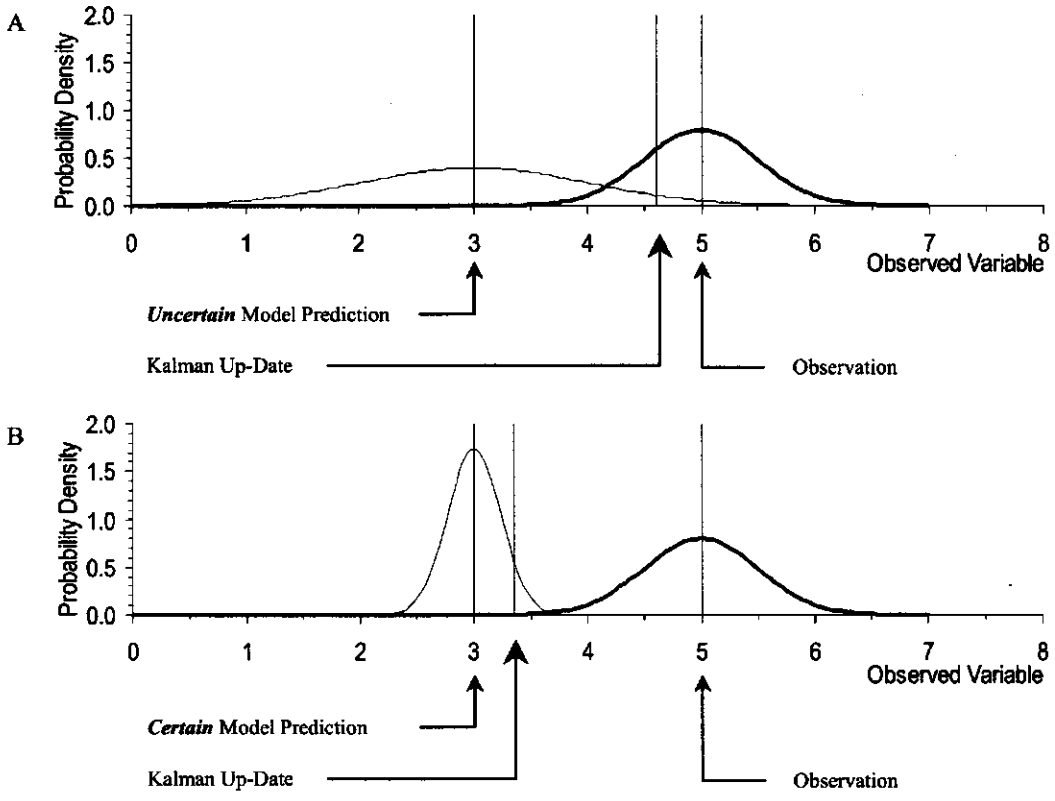


Figure 5.6. Illustration of the Kalman up-date in the cases where the model's prediction is relatively uncertain and relatively certain (A and B respectively). Each graph depicts the probability density functions of observed variable at an instance in time.

by maximising the logarithm for the likelihood criterion. Although essentially the same thing, these two uses of the Kalman filter are in practical terms very different and should not be confused. In certain practical applications such as model based real time control it may be advantageous to combine the two different uses of the Kalman filter.

Error Assumptions

The main distinction between deterministic and stochastic modelling lies in the implicit error assumptions. In deterministic modelling our parametric description of reality is assumed to be perfect and all deviation between modelled and observed values is assumed to be the result of observation error. In stochastic modelling our parametric description is assumed only to be a partial replication of reality and at least part of the deviation between modelled and observed values is assumed to be the result of model error. Figure 5.7 summarises the error assumptions into the two extremes of no up-date and complete up-date, and with the Kalman up-date in-between.

Characterisation of dynamic models should in first place relate to their underlying assumptions. The underlying assumptions distinguishing deterministic models from stochastic models are the error assumptions that on the presence of observations lead to no up-date or to an up-date respectively.

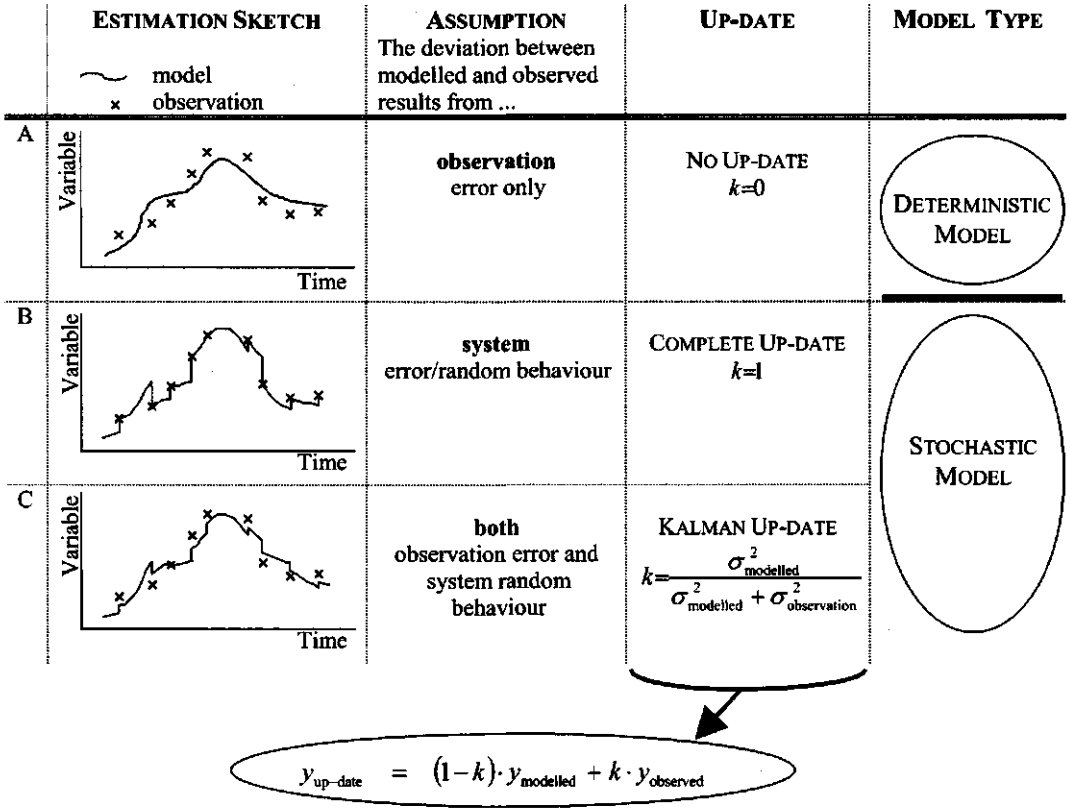


Figure 5.7. Summary for the comparison of assumptions made during deterministic and stochastic modelling. The up-date is here illustrated for a one state variable model.

STATE SPACE FORMULATION

Linear

The model in (5.9) would generally not be written in that form. A more general and more convenient form is the state-space representation with a state equation and an observation equation. In (5.11)-(5.12) below the unit hydrograph model of equation (5.9) has been rewritten into a state-space form.

The system or state equation:

$$Q_t = \delta \cdot Q_{t-1} + \omega_0 \cdot R_{t-lag} + \omega_1 \cdot R_{t-1-lag} + e_{1,t} \tag{5.11}$$

The observation equation:

$$Q_{OBS,t} = Q_t + e_{2,t} \tag{5.12}$$

Where $Q_{OBS,t}$ is now the observed flow. The state variable, Q_t, defines the current state of the system just as the state equation defines the system dynamics. The observation equation describes how the system is observed. In (5.12) the flow variable is observed directly and only an observation error term is added.

At first glance it may appear as if one would never be able to distinguish between the model error, $e_{1,t}$ and the observation error, $e_{2,t}$. Looking carefully, however, it is evident that the observed flow, $Q_{OBS,t}$, is a function of all past values of the model error whereas it is only a function of the observation error at that particular moment in time.

On simulation the flow, Q_t, would be calculated using (5.11) and its observed value, $Q_{OBS,t}$, would be obtained by adding the observation error to Q_t. During parameter estimation the Kalman filter would be used. Flow prediction would be done using (5.11) without its error term and the equivalent second moment expression would be used to predict the prediction variance at time t as a function of the prediction variance at time $t-1$. The Kalman gain would be calculated using (5.10) and this in turn used to calculated up-dates of the flow and flow variances.

Notice that if $\underline{e}_{1,t}$ in (5.11) had a variance equal to zero then the equations (5.11)-(5.12) would be identical to the deterministic model (5.6). If instead $\underline{e}_{2,t}$ in (5.12) had a variance equal to zero then (5.11)-(5.12) would be the same as the stochastic model with complete up-date in (5.2).

In the above the deterministic modelling with no up-date and stochastic modelling with both complete and Kalman up-dates have been presented using a simple flow model with one state variable and one observed variable. Often, however, it is not sufficient to use only one state variable in defining the state of a system. When working with more than one state variable it is helpful to place them in a state vector that is generally written with the symbol $\underline{X}_t$. The generalised linear state-space model can then be written as,

System Equations:

$$\underline{X}_t = \underline{\underline{A}}\,\underline{X}_{t-1} + \underline{\underline{B}}\,\underline{u}_{t-1} + \underline{e}_{1,t} \tag{5.13}$$

Observation Equations:

$$\underline{Y}_t = \underline{\underline{C}}\,\underline{X}_t + \underline{e}_{2,t} \tag{5.14}$$

Where $\underline{X}_t$ is the state vector, $\underline{\underline{A}}$ is the system matrix containing coefficients to the state vector of the previous time step, $\underline{u}_{t-1}$ is the vector of inputs, $\underline{\underline{B}}$ is inputs matrix containing coefficients to the input variables, $\underline{e}_{1,t}$ is the model error, $\underline{Y}_t$ is the vector of observations, $\underline{\underline{C}}$ is the observation matrix with coefficients defining which state variables or linear combination of state variables is observed, and $\underline{e}_{2,t}$ is the observation error. Refer to see Harvey (1993) and Madsen (1995) for more details on the linear state space formulation and Kalman filtering.

The equations of the linear Kalman filter, which may be used for estimation in linear state-space models in the form of (5.13)-(5.14), are listed in Appendix A.

Non-linear

For much practical water quality engineering simplification to linear models would be unreasonable or even impossible. In such cases the extended Kalman filter may be used with the non-linear state-space formulation. The non-linear state-space formulation in discrete time may be expressed as follows,

System Equations:

$$\underline{X}_t = f\left(\underline{X}_{t-1},\ \underline{u}_{t-1}, \underline{\theta}\right) + \underline{e}_{1,t} \tag{5.15}$$

Observation Equations:

$$\underline{Y}_t = h\left(\underline{X}_t, \underline{\theta}\right) + \underline{e}_{2,t} \tag{5.16}$$

where $\underline{X}_t$ is the state vector, $f()$ is the non-linear system function, $\underline{u}_{t-1}$ is the vector of inputs, $\underline{\theta}$ are the constant parameters, $\underline{e}_{1,t}$ is the model error, $\underline{Y}_t$ is the vector of the observed variables, $h()$ is the non-linear observation function defining how the state variables are observed, and $\underline{e}_{2,t}$ is the observation error.

Some times none of the state variables may in fact be observed. Water volumes or flows may, for example, be the modelled state variables but in the observation equation these would often be converted to water levels which have been or are being monitored. Modelling suspended solids and dissolved organic matter as state variables the observed variables may be turbidity and UV-absorption.

Measurements such as turbidity and UV-absorption are often supplemented by less frequently sampled and analysed quantities such as suspended solids, COD, BOD or others. These would generally be included in the observation vector $\underline{Y}_t$ and may be a function of the same state variables in the state vector $\underline{X}_t$ as the more frequently observed variables. In periods where only the frequently observed variables are available the less frequently observed variables would be treated as missing. Single variables in the observation vector can be treated as missing by momentarily giving them any arbitrary value and letting their observation

variance be equal to an extremely high value. Extremely uncertain observations have no effect on the up-date of the states. Stochastic time series modelling with *complete up-date* in discrete time requires equidistant sampling and missing values is generally problematic. Stochastic modelling with *Kalman up-date* has the distinct advantage of accepting irregular sampling, not requiring the same sampling schedule for all variables and handling missing values without violating the underlying assumptions.

Much traditional deterministic water quality modelling is in fact implemented in a form that differs only in format and interpretation from the non-linear state-space formulation in (5.15) and (5.16). The main difference is that in traditional deterministic modelling the model error $\underline{e}_{1,t}$ is assumed always to be equal to zero. The observation function $h(\,)$ would then generally pick out the state variables that have been measured. During calibration or parameter estimation the sum of squares of the observation errors, $\underline{e}_{2,t}$, which in this deterministic case is the deviation between the observed and the modelled, would be minimised either by visual evaluation of the fit or by using an optimisation algorithm.

When rewriting or reinterpreting traditional deterministic forms into the non-linear state-space form it is important to realise that all quantities which define the state of a system at any given moment must be state variables. If one is modelling an object falling through space then a definition of it's state requires information on both its position and its velocity. Similarly, a state-space representation of a one dimensional hydrodynamic dynamic pipe flow model involving solving the Saint Venant equations would require that both the water levels and water velocities at every calculation point in the system were state variables.

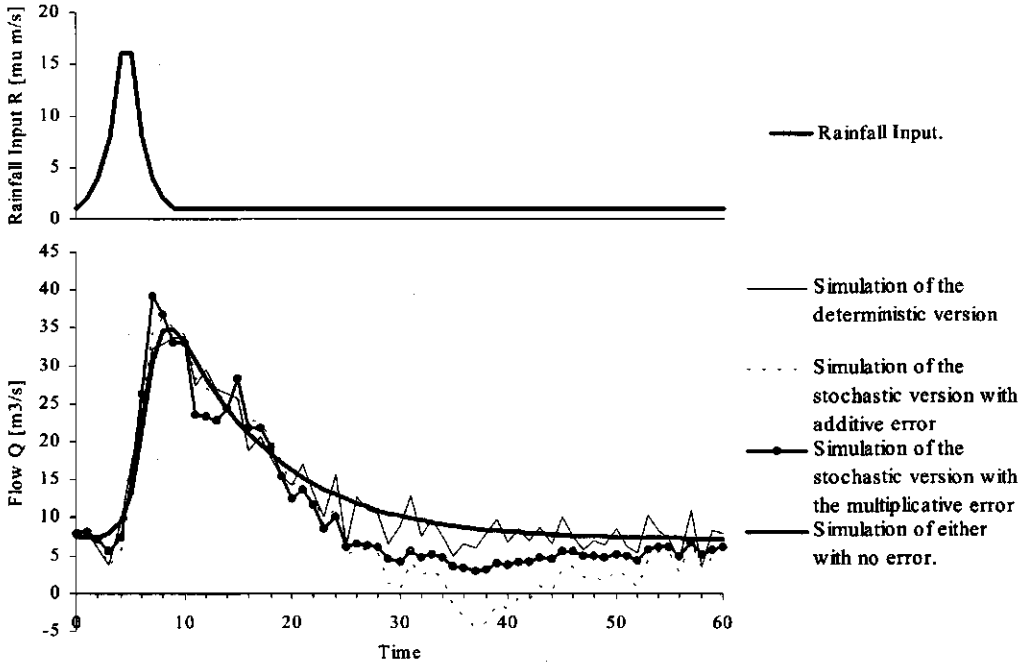


Figure 5.8. Comparison of simulation of the three models: deterministic model (5.6), stochastic model with additive error (5.2) /(5.11) and the stochastic model with multiplicative error (5.17). The simulations have been performed with the same model parameters as in Figure 4 and with a dimensionless error for multiplicative model of 0.15.

Up-Dating Unobserved State

In most of the above discussion the up-dated state variable was also an observed variable. A characteristic of the Kalman filter is that although all state variables are up-dated, only a few of these actually need to be observed. Through the interrelationship between state variables represented by partial derivatives of the system function with respect to the state variables at the previous time step and through the covariance of the state predictions, observations of one of the state variables will contain information about other state variables. Thus direct or indirect measurement of only one of the state variables will generally result in an up-date of all state variables. In this sense the Kalman filter is well suited in the design of software sensors (e.g. Carstensen et al., 1996).

State Dependent Errors

One of the assumptions of the state space formulation in (5.13)-(5.14) and (5.15)-(5.16) is that the errors are normally distributed. With, for example, a component concentration as state variable the situation often arises where the added error is negative but in absolute value greater than the predicted concentration. A simulation of the system would result in negative concentrations. For water quantity the same problem arises as flow or water volumes drop to zero. In such situations it is simply not reasonable to assume that the errors are normally distributed with the same distribution irrespective of the state variable's current value.

One way of handling state dependent model errors is by introducing an error state for each of the state variables whose randomness should depend on the value of the state variable. This method is outlined below.

Written with the applied multiplicative state dependent error the system equation for the little unit hydrograph example (5.6) becomes

$$Q_t = \left(\delta \cdot Q_{t-1} + \omega_0 \cdot R_{t-lag} + \omega_1 \cdot R_{t-1-lag}\right) \cdot e^{\sigma \cdot \varepsilon_{t-1} - \frac{1}{2} \cdot \sigma^2} \tag{5.17}$$

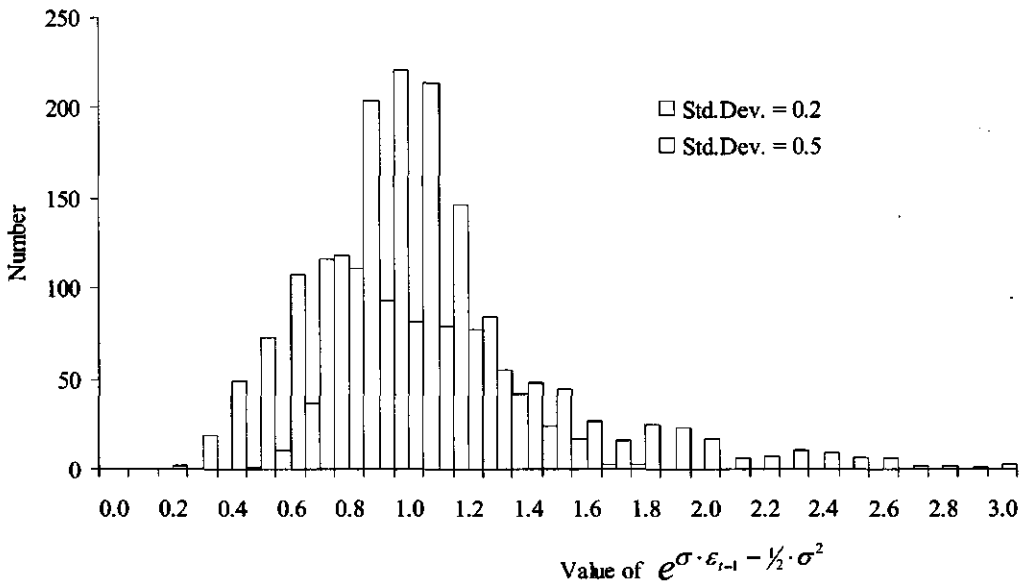


Figure 5.9. Histogram showing the distribution of the random factor by which the state variable is multiplied. The error ε in the exponential expression is a normally distributed random variable with mean 0 and variance 1. Histograms are shown for standard deviations, σ = 0.2 and 0.5.

where ε is a normally distributed random variable with mean 0 and variance 1, and where σ is the dimensionless standard deviation of the multiplicative factor (i.e. the exponential term in (5.17)). The random variable ε is in fact placed in the state vector with a constant expectancy of zero. In this way the Extended Kalman filter (non-linear) can be used without violating the condition that the errors must be independent of the state variables. A stochastic simulation of the system defined by (5.17) has been plotted in Figure 5.8. In the same figure also a simulation of (5.2) is depicted in which the flow variable becomes negative, which is in this case considered physically impossible. Note that the model with the multiplicative error term remains positive by definition because the multiplicative exponential error term is always greater than zero.

It is important that the exponential factor in (5.17) has to have a mean of 1 which it in this case does so long as the error ε is a normally distributed random variable with mean 0 and variance 1 and the dimensionless standard deviation σ is positive. Figure 5.9 shows the distribution of the factor for standard deviations σ of 0.2 and 0.5.

ESTIMATION CRITERIA

Least Squares

The criterion for calibration of traditional deterministic models is often a visual evaluation of the simulated and observed values plotted in the same graph. Once an automatic search routine is coupled the most frequently used criterion is the least square (LS). Using the LS criterion one minimises the sum of the squared errors.

One of the assumptions associated with least square estimation is that the errors are normally distributed with the same variance irrespective of the value of the observed variable. There are many cases in water quality modelling where this assumption is not reasonable. For both water quantity and water quality modelling the error is often highly dependent on the variable value. At small concentrations the modelling errors may be much lower than at high concentrations.

Using LS estimation on the errors divided by the predicted value often leads to an exaggerated importance given to the prediction with low values. In some cases a weighting scheme is devised to emphasise fit of peak values or low values but it often turns out that the parameter estimates are very sensitive to the selected weights.

Maximum Likelihood

A criterion very rarely seen applied in the context of deterministic models is the maximum likelihood (ML) criterion. During ML estimation one seeks the parameter set that results in the most probable set of errors assuming a given distribution. Generally the logarithm to the likelihood is maximised. Where the assumed distribution is a normal distribution with mean zero then its variance would generally be unknown and have to be estimated along with the other model parameters.

Using linear deterministic models with the LS and the ML criteria applied directly to the deviation between the predicted and observed lead to the same set of estimated parameters. With non-linear models the

estimates will generally not be the same. The advantage of the ML criterion is that it provides more flexibility in terms of assumptions relating to the distribution of the errors.

An additional advantage of ML estimation is that multiple criteria estimation problems can readily be combined into a single criterion by minimising the product of the individual likelihoods. If, for example, observations of both dissolved oxygen and biological oxygen demand (BOD) were to be used in calibrating a surface water model, it would often not be possible to estimate the parameters for each model sequentially. This is because oxygen removal depends on the degradation of organic matter and the degradation of organic matter also depends on dissolved oxygen in the water. Using a maximum likelihood criterion parameter estimation in the whole model would be done by minimising the sum of the logarithms to the individual likelihoods. In LS estimation weights would have to be selected for each component and the result turns out to be very sensitive to the selected weights.

Using the maximum likelihood criterion it is possible to construct a parametric model for the error variance and have these estimated along with other model parameters. This is done for example in the non-linear regression models used in relating measurements of turbidity in a combined sewer to variables such as suspended solids and COD in Moens et. al. (1999).

Robust estimation can readily be applied when using the maximum likelihood criterion. This works as a sort of on-line suppression of outliers. There exist several different Robust estimators with the shared feature that extremely unlikely observations end up having little or no influence on the likelihood value. It is generally not a good idea to start ones parameter estimation using a robustness criterion because the parameter values may be so far off that a very large proportion of the observations become categorised as outliers.

In stochastic modelling with the Kalman filter as state estimator LS estimation only works in cases where the error variances are known. This is however rarely the case and therefore the ML criterion usually has to be used.

A further advantage of the maximum likelihood criterion is the relatively simple manner in which *a priori* knowledge on likely parameter values can be incorporated into the estimation procedure. This is discussed in more detail in the section on identifiability below.

A disadvantage of ML estimation is that at least one extra parameter, namely the error variance, has to be estimated for each error series. In some cases it may therefore be a good idea to first obtain rough parameter estimates from a LS estimation and to use these as starting values for a ML estimation.

IDENTIFIABILITY

Identifiability deals with the uniqueness of an estimated parameter set in shedding light on questions such as: Do some of the parameters have little or no influence on the goodness of fit? Could an increase in one parameter be compensated by a decrease in another parameter to give an equally good fit? Are certain combinations of parameters interchangeable with other combinations of parameter? How much uncertainty is associated with the estimated parameters and what is the uncertainty once parameter interchangeability is considered? Within the world of time series analysis and stochastic modelling analytical methods exist to examine the identifiability of the parameters in a given model structure on the basis of a given data set and considering *a priori* knowledge of the parameter values. This section outlines some essential aspects of these methods and how they can be applied in the field of water quality modelling. In the case study later in the chapter the methods have been applied to combined sewer modelling problems.

Identifiability as it is used here deals with examining the uniqueness of an estimated parameter set. However it should be realised that the field of model identification is much broader covering also the revealing or disclosure of model structure in empirical time series modelling or in mechanistic models such as the disclosure of the evaporation dependency in Chapter 4.

Identifiability is a combination of two elements: the model whose structure has to be identifiable and the given data that has to be informative (Madsen and Holst, 1998). An outline of these two distinctly different aspects of identifiability is followed by an introduction to how an exploration of the parameter space can be used to examine their combined effects.

Structural Identifiability

Sometimes the structure of a model is such that no distinction can be made between two or more model parameters. Consider for example the model in (5.18) where the flow, Q, is the flow in a sewer receiving rainfall-runoff from two different sub-catchments, A and B.

$$Q_t = \delta \cdot Q_{t-1} + a_0 \cdot R_{t-lag} + a_1 \cdot R_{t-1-lag} + b_0 \cdot R_{t-lag} + b_1 \cdot R_{t-1-lag} + e_{1,t} \tag{5.18}$$

where Q_t is the modelled flow, R_t is the rainfall input, e_t is the error term and δ, a_0, a_1, b_0 and b_1 are model parameters to be estimated from observations of rainfall and flow. Although the parameters a_0 and a_1, and b_0 and b_1 are physically associated with the sub-catchments A and B, it is not possible to distinguish these parameter pairs from one and other. An increase in a_0 could be directly compensated by a decrease in b_0.

In most cases problems of structural identifiability will be less obvious than in the example above. Often the problems will only be present for certain ranges of the input and output data, and they will be difficult to distinguish from problems of poor excitation in the data set. Consider for example the Monod expression (5.19) for the growth rate, μ, as a function of substrate concentration, S, with the two parameters, the maximum growth rate, μ_{max}, and the half saturation constant, K_S, which have to be estimated from a given data set.

$$\mu = \mu_{max} \cdot \frac{S}{K_S + S} \tag{5.19}$$

There is no immediate sign of structural identifiability problems. However as the substrate concentration, S, decreases and becomes smaller than the value of the half saturation constant, K_S, one sees that the maximum growth rate and the half saturation constant, K_S, become interchangeable.

Re-parameterisation of the model should be considered in such cases with structural identifiability problems within a limited range. This will avoid unnecessary complications during parameter statistics. The gradient of the curve in the first part of the Monod expression can be expressed as,

$$grad = \frac{\mu_{max}}{2 \cdot K_S} \tag{5.20}$$

The half saturation constant, K_S, can then be substituted out of the Monod expression (5.19) to give,

$$\mu = \frac{\mu_{max} \cdot 2 \cdot grad \cdot S}{\mu_{max} + 2 \cdot grad \cdot S} \tag{5.21}$$

and the maximum growth rate, μ_{max}, and gradient, *grad*, would now be the new parameter set to be estimated from the data. This algebraically cosmetic re-parameterisation is likely to have a positive effect on the performance of the non-linear search algorithm and undoubtedly on examining the model identifiability after parameter estimation. The structurally induced correlation between the two parameters is avoided by the simple re-parameterisation. Within the field of water quality modelling, the identifiability of the parameters in the Monod model and other growth-degradation models has been treated in Reichert and Omlin (1997), Vanrolleghem and Keesman (1996), Vanrolleghem et. al. (1995) and Dochain et. al. (1995).

Excitation

In order to calibrate a rainfall-runoff model during rainfall it is necessary to have data from rainfall periods. That is obvious. It is however also important that the data set does not consist of simply a single constant drizzle over a long period. A rainfall-runoff model describes the relationship between input rainfall and the subsequent runoff. For a good calibration or parameter estimation it is important to have a dynamic input in order to be able to distinguish between parameters.

A dynamic input is however not always enough. It is important that the frequencies of the variations in the inputs coincide fairly well with the characteristic time constants of the system whose parameters are being estimated (Sadegh et al., 1995). If the variations in the input are much faster or much slower than the time constants of system then the input will have an effect on the system similar to that of a constant input signal. This is illustrated for a unit hydrograph model in Figure 5.10.

For a simple linear rainfall-runoff model essentially characterised by a unit hydrograph an impulse input produces a response from which the transfer function of the system can be obtained. For more complex systems with several simulations process such as rainfall-runoff, sediment resuspension, biological degradation of organic matter and oxygen depletion a single impulse or step input will often result in limited information about the system parameters. Sadegh (1996) presents the theoretical and practical considerations concerning the experimental design of linear and non-linear complex systems built on physical and chemical theory.

Although both deterministic models and stochastic models with Kalman up-date do not require that observations be equally spaced in time it is important that observation be made when "things are happening". A sewer suspended solids model, such as that estimated in Chapter 7, may, for example, involve a threshold point for which higher flow rates or velocities result in an initialisation of sediment resuspension. In order to identify this point and to distinguish its value well from other parameters it is necessary that measurements of the suspended solids have been made both before and after the threshold has been crossed. See Chapter 7 on Random Coefficient Modelling for more discussion of the particular case of resuspension modelling.

Exploring Parameter Space

Once the estimation criterion has been optimised one is interested in knowing how unique the parameter set is. This is done by examining the sensitivity of the criterion value to changes in the parameters. This is in the present thesis broadly described as "parameter statistics".

Parameter statistics involves the evaluation of

- parameter variances and confidence limits,
- covariance and correlation matrix,
- confidence contour plots.
- eigen values and eigen vectors of the Fischer's information matrix,
- eigen values and eigen vectors of the parameter correlation matrix,
- experimental design criteria,
- projections of eigen vectors,
- parameter confidence limits under consideration of covariation.

The parameter statistics are carried out under the assumption that the model is locally linear. In a large number of cases this will in fact be a reasonable assumption however in some cases parameter transformations will be required in order to be able to make this assumption of local linearity. Parameter transformation is discussed below and applied in the case study that follows.

Parameter Variance. With maximum likelihood (ML) estimation in linear models the estimated parameters are normally distributed. The likelihood, L, that the parameter θ is equal to θ rather than its estimated value $\hat{\theta}$ is defined by the probability density function of the normal distribution. In such a case where *only one parameter* has been estimated the likelihood takes the form of (5.22).

$$L = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{(\theta-\hat{\theta})^2}{2\cdot\sigma^2}} \tag{5.22}$$

where

L	is the likelihood (i.e. the probability density),
$\hat{\theta}$	is the estimated parameter value, i.e. its value at the point of maximum likelihood,
θ	is the parameter value for which the likelihood is expressed,
σ	is the parameter standard deviation.

Taking the log of both sides (5.22) can be rewritten as,

$$-\log L = \frac{(\theta-\hat{\theta})^2}{2\cdot\sigma^2} - \log\frac{1}{\sqrt{2\pi}\cdot\sigma} \tag{5.23}$$

Notice that this results in a quadratic equation with respect to the negative logarithm to the likelihood. During maximum likelihood estimation it is generally the $-\log L$, that is a quadratic function, which is minimised.

The gradient of the negative log likelihood function is then the derivative of (5.23) with respect to the parameter value.

$$\frac{d(-\log L)}{d\theta} = \frac{\theta-\hat{\theta}}{\sigma^2} \tag{5.24}$$

Notice here that as we would expect in the optimum the gradient of the likelihood function is equal to zero when $\theta = \hat{\theta}$. The gradient or slope of the negative log likelihood in (5.24) can further be differentiated giving the curvature of the likelihood function (5.25).

$$\frac{d^2(-\log L)}{d\theta^2} = \frac{1}{\sigma^2} \tag{5.25}$$

Thus in the one-dimensional case, where only one parameter has been estimated, the parameter variance turns out to be the reciprocal of the curvature of the negative log likelihood function (5.25).

$$\sigma^2 = \left(\frac{d^2(-\log L)}{d\theta^2} \right)^{-1} \tag{5.26}$$

The above outline of how parameter variances are defined is based on the estimation using the ML criterion. Refer to Harvey (1993) or Madsen (1995) for further details on ML estimation criteria. It turns out that with LS estimation the covariance matrix can be expressed as the product of twice the mean squared error and the inverse of the curvature of the sum of squared errors. More on parameter statistics with LS estimation can be found in several standard statistical textbooks such as Draper and Smith (1981).

Parameter Covariation. A parameter's standard deviation represents its uncertainty under the condition that all the other parameters in the model actually have the value that they have been estimated to have. Most often the other estimated parameters are also uncertain and interchangability between parameters may mean that the uncertainty of a given parameter is much larger once the uncertainty of all the other parameters is taken into account.

Like the variance of a single parameter (5.26) the covariance between parameters is the inverse of the curvature in their combined direction. For several parameter covariances are generally arranged in a covariance matrix with the variances in the diagonal. Using the ML criterion the parameter covariance matrix can be found by examining the curvature of the negative log likelihood function. The covariance matrix can be expressed as follows,

$$\underline{\underline{V}} = \left[\frac{\delta^2(-\log L)}{\delta\theta_i \cdot \delta\theta_j} \right]^{-1} \tag{5.27}$$

where

$\underline{\underline{V}}$ is the parameter covariance matrix,

i and j indicate the i'th and the j'th parameters respectively.

The parameter covariance matrix is a symetrical matrix with the parameter variances in the diagonal and their covariances on the off-diagonals.

In practical engineering problems the covariance matrix will often be difficult to intepret directly because the different parameters have different units and the covariances will each have units of two parameters. The covariace is however an important stepping-stone to calculating other quantities which may be more usefull to the modeller. These are the parameter correltation matrix, eigen values and the eigen vectors.

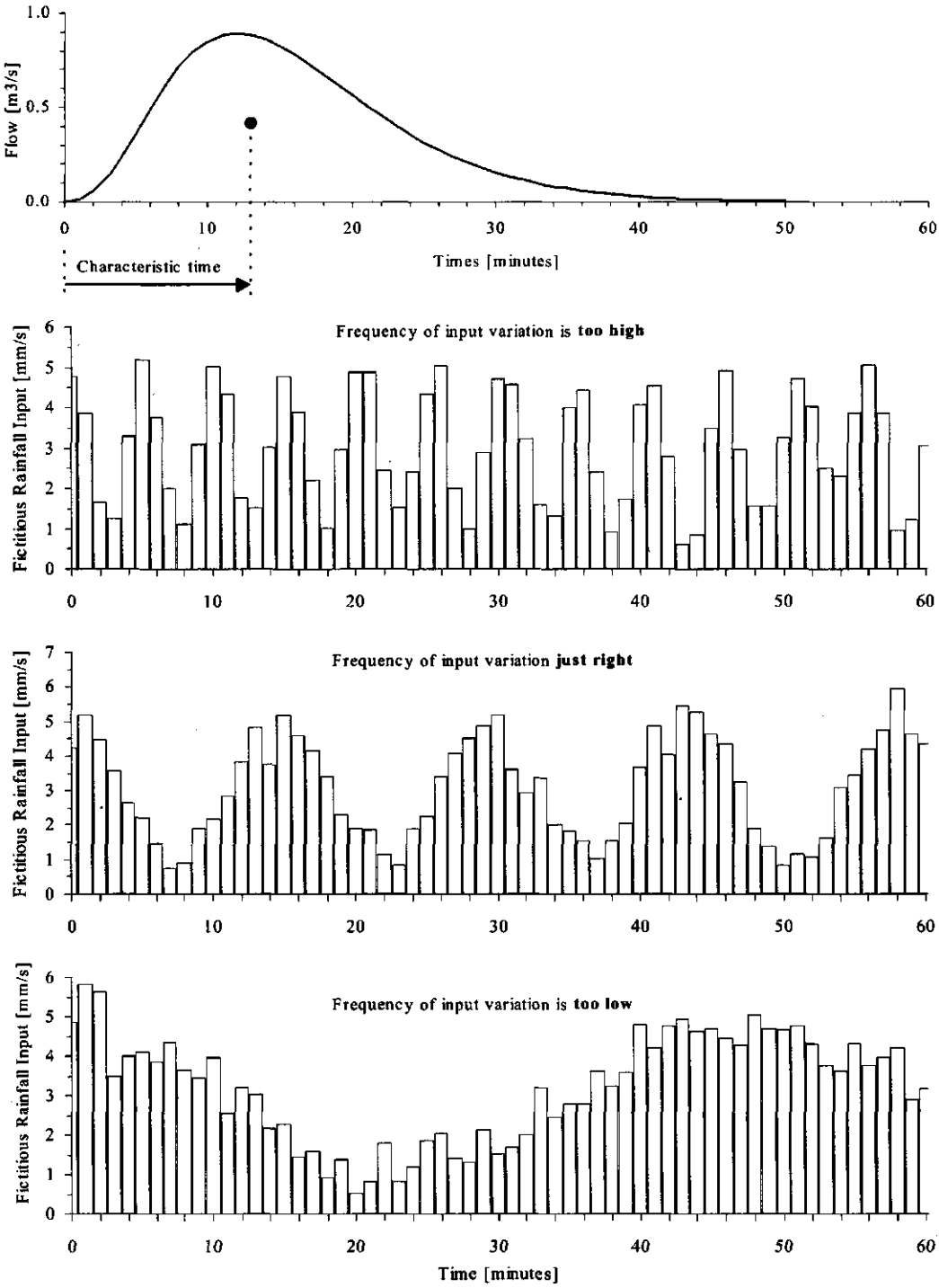


Figure 5.10. Characteristics of data suitable for parameter estimation in a runoff model with the given unit hydrograph or impulse response (top). The three fictitious rainfall input series having frequencies of variation that are (from up to down) too high, "just right" and too low (see discussion in text).

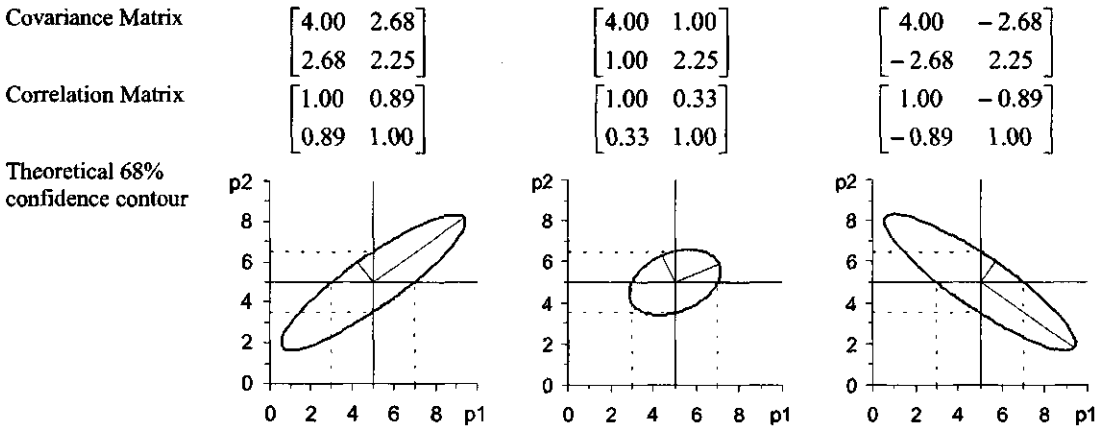


Figure 5.11. Illustration of the effect of covariance between the parameters p_1 and p_2 for three cases with high positive, low and high negative correlation respectively. In all three cases the parameter estimates are (5,5) and their variances (4.00, 2.25).

Parameter correlation is the covariation considered relative to the standard deviation of the respective parameters. The correlation between two parameters can be calculated from covariances by,

$$\rho_{i,j} = \frac{v_{i,j}}{\sqrt{v_{i,i} \cdot v_{j,j}}} \tag{5.28}$$

where

$\rho_{i,j}$ is the correlation between the i'th and the j'th parameter,

$v_{i,j}$ is the covariance between the i'th and the j'th parameter.

Consider a model with two estimated parameters. Given that the parameter covariance matrix was found to be as in the left of (5.29) then the corresponding correlation matrix is as on the right.

$$\underline{\underline{V}} = \begin{bmatrix} 4.00 & 2.68 \\ 2.68 & 2.25 \end{bmatrix} \qquad \underline{\underline{P}} = \begin{bmatrix} 1.00 & 0.89 \\ 0.89 & 1.00 \end{bmatrix} \tag{5.29}$$

The parameter correlation is dimensionless with a value between −1 and +1. A correlation close to −1 is described as a high negative correlation and indicates that an increase in the one parameter may be compensated by a decrease in the other parameter to give the same fit (i.e. the same value for the estimation criterion). Similarly a correlation close to +1 indicates that an increase in the on parameter may be compensated by an increase in the other.

When is parameter correlation too high? This is a question that will soon arise once models are applied to practical problems. Ideally one would prefer to have no parameter correlation. This would make things easy in that the parameter uncertainty under consideration of the uncertainty of the other parameter would be the same as the parameters uncertainty under the assumption that all other parameters were fixed. Unfortunately this is most often not the case. Parameters are often highly correlated and the consequences have to be considered.

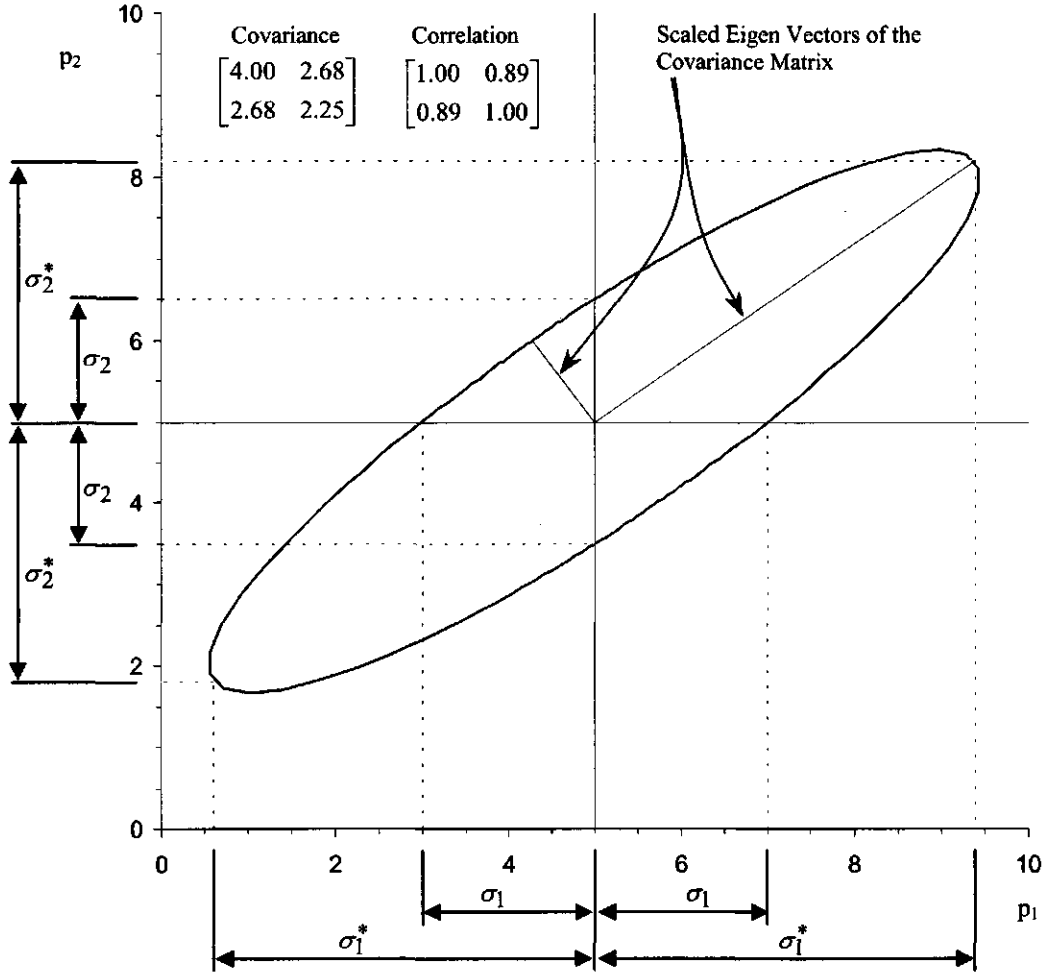


Figure 5.12. The projections of the covariance matrix eigen vectors, σ*, can give an indication of the parameter uncertainty when parameter covariation or interchangeability is considered. This is illustrated here for the two parameters.

Generally speaking a parameter correlation of 0.89 as seen in (5.29) would be considered high. However, as we shall see in the following, this depends very much upon the parameters that have been estimated and on their variances.

An adequate evaluation of the parameter uncertainty after parameter estimation involves calculating the eigen vectors of the covariance matrix, scaling them and projecting these onto the parameter axes. The eigen vectors of covariance matrix form a set of orthogonal vectors. An eigen vector therefore defines a combination of parameters which are independent of the combinations defined by the other eigen vectors. A covariance matrix has as many eigen vectors as there are parameters. The eigen vectors are often scale to have a length of one and are then called normalised eigen vectors. Returning to the two dimensional example of (5.29) the normalised eigen vectors are,

$$\underline{\varepsilon}_1 = \begin{bmatrix} 0.81 \\ 0.59 \end{bmatrix} \qquad \underline{\varepsilon}_2 = \begin{bmatrix} -0.59 \\ 0.81 \end{bmatrix} \tag{5.30}$$

These normalised eigen vectors can be scaled so that they define the axes of an ellipse whose intercepts with the x-axes and y-axes are equal to the standard deviations of the first and the second parameter respectively. The scaled eigen vectors and the thereby defined ellipse are shown in Figure 5.11 (left) for a two parameter problem. Notice that the intercepts with the x-axis and the y-axis are 2 and 1.5 corresponding to the square roots of the diagonal elements in the covariance matrix in (5.29) left. If the ellipse is drawn proportional to the standard deviations it will enclose the 68% confidence region. Had it been drawn proportional to roughly twice the standard deviation it would enclose the 95% confidence region. Once again these consideration assume that the model is linear in the region of the criterion optimum. In Figure 5.11 are also plots of the ellipses corresponding to parameters with almost no correlation and parameters with negative correlation respectively.

It is evident from the above that the covariance matrix eigen vectors plotted in Figure 5.11 give information on how parameter interchangeability affects the parameter standard deviations. In many modelling problems it is of greater interest to have the "effective standard deviation" or "unconditional standard deviation", σ^*, based on the longest projection of the tip of the scaled eigen vectors than the actual standard deviation. For the two- parameter problem in (5.29) this has been illustrated in Figure 5.12. Given this "unconditional standard deviation" the engineer is in a better position to evaluate the actual range in which the parameter value may lie.

Note that projecting the tip of the eigen vector is in fact a compromise. The projected point ought to be the furthest left point at which the tangent (or hyperplane) to the ellipse is perpendicular to the axis onto which it is being projected. When the correlation is large the tip of the eigen vector is a fair approximation.

Parameter Line Plots and Confidence Contours. The above parameter statistics rest on the assumption that the model is locally linear in the region of the parameter estimates. This assumption has to be reasonably fulfilled for the parameter statistics to be reliable and meaningful.

An initial way to verify that this condition of local linearity is reasonably satisfied is to evaluate and plot the value of the estimation criterion for parameter values just above and below the estimated value. During this exercise all other parameters should be kept fixed to their estimated values. In a region around the optimum

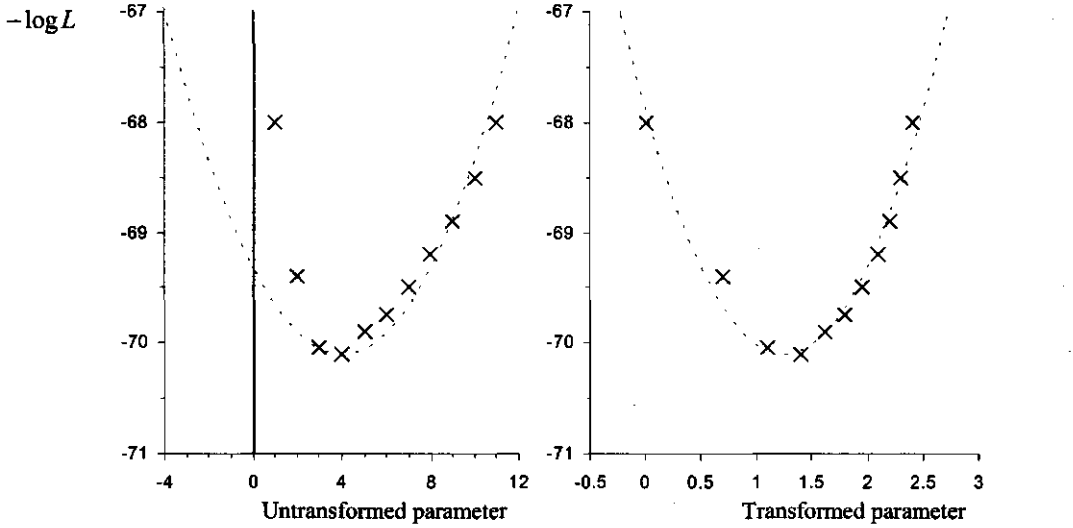


Figure 5.13. Parameter may have to be transformed in order to satisfy the condition of local linearity. The dashed line represents the estimated curvature based on the maximum likelihood evaluations (crosses).

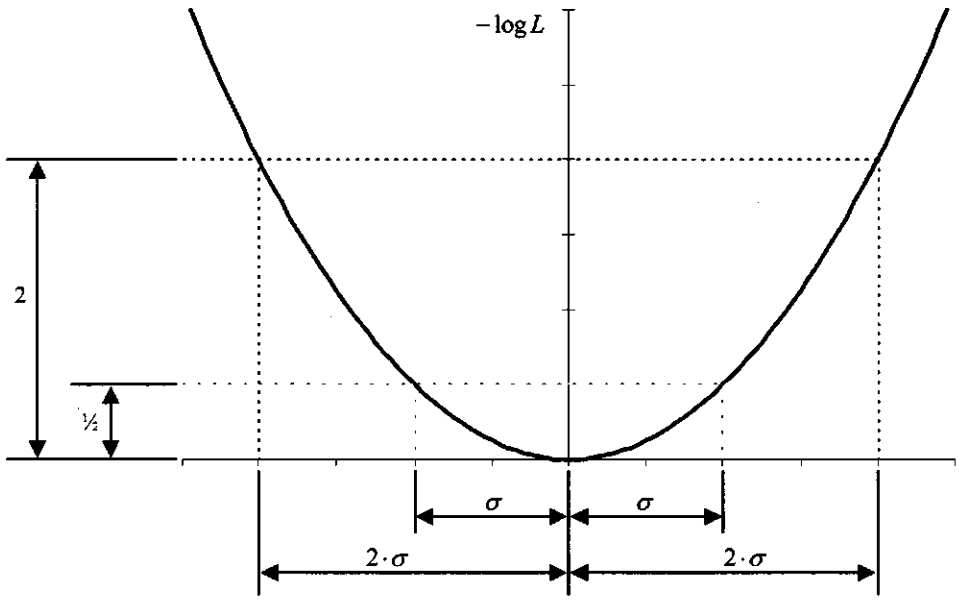


Figure 5.14. The rise in the negative log likelihood function corresponding to the standard deviation and twice the standard deviation away from the estimated parameter value.

the plotted curve should have the shape of a quadratic function. In particular it should be symmetrical about the estimated parameter values.

Local non-linearity can often be tackled by transforming model parameters. A parameter transformation does not change the estimated value of the parameter but only the type of distribution implicitly assumed for its uncertainty. Parameter transformation is in fact a re-parameterisation much like that presented earlier in going from (5.19) to (5.21). If the right transformation is selected then the parameter uncertainty of the new transformed parameter is normally distributed just as is the uncertainty of parameters in linear models.

Figure 5.13 contains such a plot of the maximum likelihood criteria against the parameter values before and after transformation (left and right respectively). In this illustrative example a log transformation was found suitable. Notice that the if the transformation had been neglected then the lower 95% confidence limit (corresponding to a negative log likelihood of –68.0) would have been negative. However, the physical interpretation given to this parameter impels it to be positive and the likelihood is not even defined for negative values of the parameter. By performing statistics on the transformed parameter the more realistic 95% confidence limit is found at about 0. Converted back to the untransformed parameter this gives a value of 1. Once the required transformations have been found and applied, the covariance matrix may be estimated and parameter statistics evaluated.

Another way of verifying that the condition of local linearity is reasonably satisfied is to create confidence contour plots based on evaluations of the estimation criterion at grid points of the parameter planes. Local linearity will exhibit itself by having elliptically shaped confidence contours with clear, though possibly tilted, axes of symmetry. Non-linearity show itself by having skewed or even banana shaped contour lines. Theoretically, however, even "nice" elliptical confidence contours for each parameter pair gives no guarantee of local linearity. The "covariation" in the third or higher dimension may still be skewed. In many practical cases it is sufficient to evaluate the confidence contours in two parameter planes but a stronger test/check would be to evaluate the confidence contours in the planes of all pairs of eigen vectors of the parameter correlation matrix. As the eigen vectors are orthogonal by definition the confidence contours

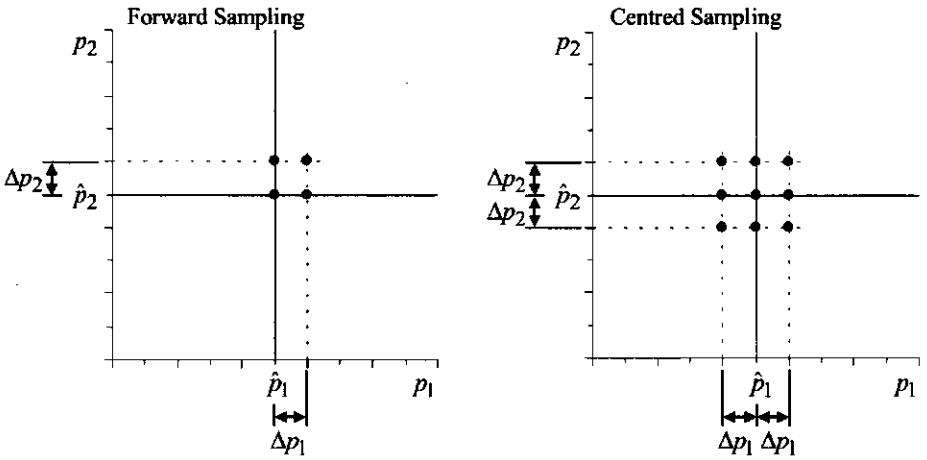


Figure 5.15. Forward and centred designs for estimating the parameter covariance matrix in non-linear models. Illustrated here for a two parameter model.

should be perfect circles and inconsistencies with the assumptions of local linearity would be readily disclosed.

What is then the region within which the model should be linear? This would depend on the required confidence level with which one wishes to evaluate the parameter value. As a rule of thumb the 95% confidence region can be used. The rise in the negative log likelihood function corresponding to the 95% confidence contour can be found from the relationship in equation (5.23) which may be rewritten to,

$$-\log L(\theta) = \frac{1}{2 \cdot \sigma^2} \cdot \theta^2 \tag{5.31}$$

where the optimum for an instant has been translated to the origin (i.e. $-\log L(0) = 0$). For the standard deviation we find $-\log L(\sigma) = \frac{1}{2}$ and for twice the standard deviation corresponding roughly to the 95% confidence limit that $-\log L(2\sigma) = 2$. The contours with negative log-likelihood values of ½ and 2 higher than the negative log-likelihood value in the optimum correspond to the 68% and the 95% confidence contours respectively. This is illustrated in Figure 5.14.

With many model parameters one ought to draw confidence contours in the planes of the eigen vector pairs in order to inspect whether assumptions of linearity were satisfied. If these turn out to be circles up until around the 95% confidence contour then the assumption of local linearity is satisfied. Banana shaped contours would suggest that some non-linear parameter transformation may be needed in order to perform the parameter statistics. Though drawing confidence contours in the planes of the eigen vector pairs is an ultimate check for local linearity it may often be difficult to decide on the required parameter transformations due to the required multi-dimensional hyperspace abstraction.

Estimating the Parameter Covariance Matrix. For linear models the parameter covariance matrix can be calculated directly from the parameter estimates and the observed values. For non-linear models however the curvature of the criterion surface has to be found by evaluating the criterion at different points in parameter space. This can be done in many different ways. One way is to evaluate the estimation criterion at a number

of points and then to fit these points to a quadratic function. In the present research work this approach has been used in combination with a variety of different sampling designs.

In general the more sampling points taken the better the estimate of covariance matrix. Two standard designs are forward and centred sampling, which have been shown for a two-parameter model in Figure 5.15. The covariance matrix obtained should theoretically be independent of the size of the small step, Δp_i, away from the estimated value. In practice however the criterion surface of non-linear models is often uneven and it is therefore a good idea to verify that the same covariance matrix is obtained for different delta steps.

Using a centred sampling design is more costly in terms of criterion evaluation (and therefore model simulations) but gives a better robustness against unevenness in the criterion surface. Fitting the quadratic function to the sampled points can be done using the explicit least square solution to the general linear model. The standard deviations and the coefficients of variation of the estimated quadratic coefficients can be computed at the same time and used to give an indication of whether there is agreement with the local linearity assumption. At the same time the coefficients of variation disclose where problems may lie and which parameters or parameter pairs ought to be looked at in more detail. Lack of symmetry and poor agreement with the local linearity assumption in general cannot be detected using forward sampling.

With non-linear least square estimation it is possible to estimate the covariance matrix from changes in residual series with respect to changes in the each of the model parameters. This matrix with as many columns as parameters in the model and as many rows as observations in the data is sometimes called the Jacobian matrix. Its cross-product gives an estimate of the Fischer matrix which is proportional to the inverse Covariance matrix.

A basic aspect of experimental design is that the more we know what we are looking for, the better we are able to design an experiment to find it. In the same way, if the covariance matrix is known, then it is easier to create a good sampling design for its estimation. During the course of this work a two step procedure for the estimation of the covariance matrix was developed. In the first step a rough estimate is obtained using a central design. This first estimate is then used to design a sampling strategy for the second step in which the central design is rotated and stretched so that it coincides with the tips of the covariance matrix eigen vectors. Especially when computing the covariance matrix of a model with many parameters, this improves the estimate of the covariance matrix considerably.

Sometimes problems arise where the inverse to the covariance matrix is non-positive definite and can not be inverted. This may be due to the fact that the optimum has not yet been reached but it may also result from the combined effect of the limited numerical precision, a rough criterion surface and large differences between the longest and the shortest eigen vectors. The two step procedure provides an excellent solution to this problem. During the first step the value of the curvature of the criterion surface is slightly exaggerated by subtracting a small value from the criterion in the optimum before fitting to the quadratic equation. This first under-estimated covariance matrix then forms the basis for the improved design in the second step. As sampling in the second step takes place at the tip of the eigen vectors all points would be expected to have a negative log likelihood value of at least ½ higher than that of the optimum and the inteference of surface roughness and numerical precision is less probable.

A Priori Information on Parameter Values

On calibration of traditional deterministic sewer or other water models the engineers has a feeling for the realistic ranges for the various model parameters and a rough idea of what value to expect. The engineer's prior knowledge about the parameter value is combined with the available data to find the best set of parameters for the specific case.

Attempts to automate the process of deterministic model calibration have often failed because only the available data was used. The engineer's prior knowledge was ignored and the information content of the observed data was not enough to identify all model parameters. Stochastic modellers wishing to include only parameters and structure that could be identified from the available data ended up reducing the models to such simple structure that the models were unsuitable in terms of the engineer's needs for extrapolation.

Maximum likelihood estimation makes it possible to take a Bayesian approach to the parameter estimation problem (Madsen and Holst, 1998). Model parameters are considered as random variables with a mean and standard deviation both before and after including the information contained in the available data. Before parameter estimation we have the a priori parameter distributions, which represent the "engineer's prior knowledge about the parameter value". This information is combined with the observed data to give the new *a posteriori* distributions.

In its original form the likelihood estimation criterion is the conditional probability density of the observations for a given set of model parameters. This likelihood can further be multiplied by the probability density of these parameters given their *a priori* distributions to form the maximum *a posteriori* estimation criterion. The maximum *a posterioiri* estimation criterion is popular in grey-box modelling, that is in stochastic modelling with some physical, chemical or biological theory in its structure. Therefore the a posteriori estimation criterion is also called the grey-box estimator (Tulleken, 1993).

As discussed earlier in the section on persistence of excitation, data sets may be more or less informative. Using the maximum *a posteriori* estimation criterion one obtains information on how informative the data is on specific parameters. If there is a large difference between the *a priori* distribution and the *a posteriori* distribution then the data has been rich on information about this particular parameter. If the two distributions are practically the same then the data set has had little to contribute in terms of determining the value of this parameter. If the means of the two distributions are the same but the variance of the *a posteriori* estimate much smaller than the *a priori* estimate, then the data has been rich on information concerning that parameter and the *a priori* parameter estimate was already a good estimate. As the parameter becomes more and more certain then even data sets that are information rich will not change the mean and variances of a good *a priori* estimate. The main advantage of using an *a priori* distribution is that though the data has little information concerning a given parameter, the parameter value does not wander off and possibly corrupt the estimation of other parameters.

CASE STUDY: MODELLING IN THE LOENEN SEWER SYSTEM

AIM

The aim of this case study has been to illustrate and compare concepts and approaches presented and discussed earlier in this chapter. A rainfall-runoff model of a sewer system has been estimated both as a deterministic and a stochastic model. The non-linear state space formulations are described including the construction of the state dependent error for the stochastic model. The quantitative incorporation of *a priori* knowledge into the estimation criteria is illustrated for both formulations. Parameter space is explored in view of identifiably and appropriateness of the assumptions made.

The rainfall-runoff models have been estimated using data from the Loenen catchment, which is described in Chapter 2. In this case study only water quantity has been considered. Attempts to estimate stochastic model for water quality variables such as suspended solids and chemical oxygen demand were unsuccessful because there were too few data points and because all sampling had taken place during actual combined sewer overflow only. In such cases, random coefficient modelling studied in Chapter 7 are a possible alternative.

During this study a modelling tool was developed with the primary objective of being able to estimate the same physical model description as a deterministic model and as a stochastic model. The tool was thus specially designed to estimate parameters in non-linear state space models.

MODEL

Structural Equations

The structural description of the rainfall-runoff system consisted of three linear reservoirs in series followed by a single non-linear reservoir (5.32) simulating the combined sewers static storage below the edge of the overflow weir. This is illustrated in Figure 5.16.

The continuity equation for the linear reservoirs is given as

$$\frac{dV_i}{dt} = Q_{\text{in},i} - Q_{\text{out},i} \tag{5.32}$$

where V_i is the volume of water in the i'th reservoir. The proportionality between the flow out, $Q_{\text{out},i}$, and the volume stored in the linear reservoir, V_i, is defined by the reservoir emptying rate or inverse time constant, k, with the dimensions [T^{-1}].

$$Q_{\text{out},i} = k \cdot V_i \tag{5.33}$$

Tests in which the volume in (5.33) was raised to an estimated power showed no significant improvement compared to the linear expression and the estimated power was not significantly different from 1.

Flow into the reservoir is equal to the flow out of the preceding reservoir so that $Q_{\text{in},i} = Q_{\text{out},i-1}$ for $i > 1$ and for $i = 1$, that is the flow into the very first reservoir,

$$Q_{\text{in},1,t} = (\text{a} + \text{b} \cdot W_j) \cdot \text{A}_{\text{imp}} \cdot R_t \tag{5.34}$$

where a and b are constant parameters estimated from the given data,

W_j is a wetness index value of the j'th rainfall-runoff event as defined and calculated in Chapter 4 using event lumped models,

$(\text{a} + \text{b} \cdot W_j)$ together form a wetness dependent runoff coefficient,

A_{imp} is the impervious area,

R_t is the rainfall a time t.

In the text that follows "a" is referred to simply as the runoff coefficient and "b" is called the wetness dependency.

The continuity equation governing the final reservoir was defined as follows.

$$\frac{dV_{\text{Final}}}{dt} = Q_{\text{in, Final}} - Q_{\text{pump}} - Q_{\text{overflow}} \tag{5.35}$$

where $Q_{\text{in, Final}}$ is the flow out of the linear reservoir which empties into final reservoir, Q_{pump} is the pump flow and Q_{overflow} is the flow discharged over the overflow weir into a surface water pond.

The expression describing the behaviour of the pumps has two estimated parameters: the level at which the pumps were switched on and the total pump capacity. The total pump flow at any given time, Q_{pump}, is defined by the following expression,

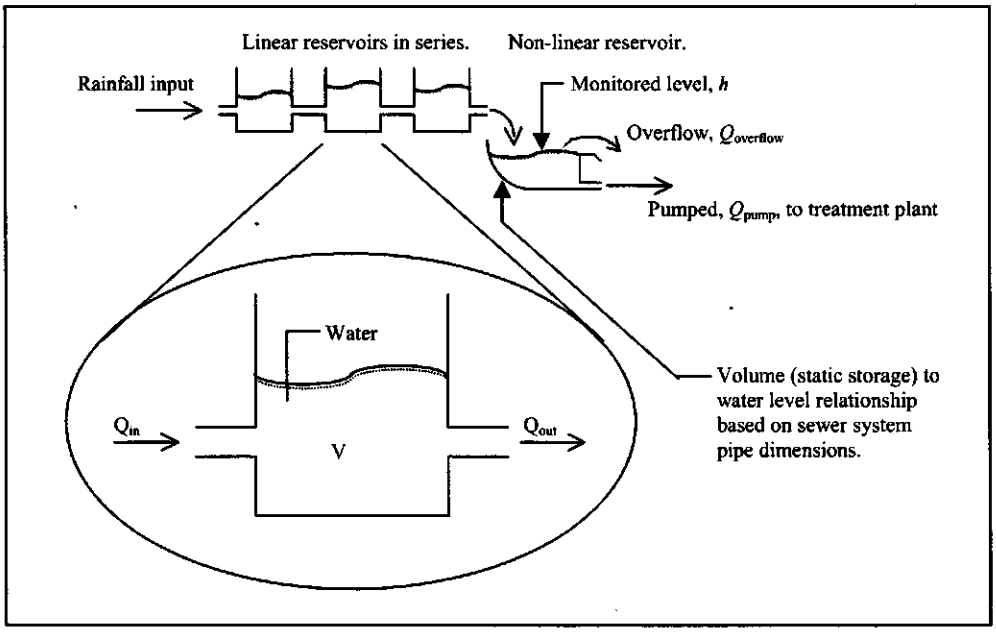


Figure 5.16. Schematic diagram of the rainfall-runoff model consisting of three linear reservoirs in series with a final non-linear reservoir incorporating the geometric level to volume relationship.

$$Q_{pump} = \begin{cases} 0 & \text{for } h < h_{pump-on} \\ Q_{pump-cap.} & \text{for } h \geq h_{pump-on} \end{cases} \quad (5.36)$$

where $Q_{pump-cap.}$ is the pump capacity and $h_{pump-on}$ is the level at which the pump turns on. Both these parameters are estimated from the available data.

Combined sewer overflow takes place only when the water level, h, is above the overflow weir, h_{weir}, and can be expressed as follows.

$$Q_{overflow} = \begin{cases} 0 & \text{for } h < h_{weir} \\ \alpha \cdot (h - h_{weir})^{\beta} & \text{for } h \geq h_{weir} \end{cases} \quad (5.37)$$

where α and β are constant parameters estimated from the available data. The weir level, h_{weir}, is reported to be at 18 meters (+ NAP). This is assumed to be so correct that the position of this concrete edge is not estimated from the data.

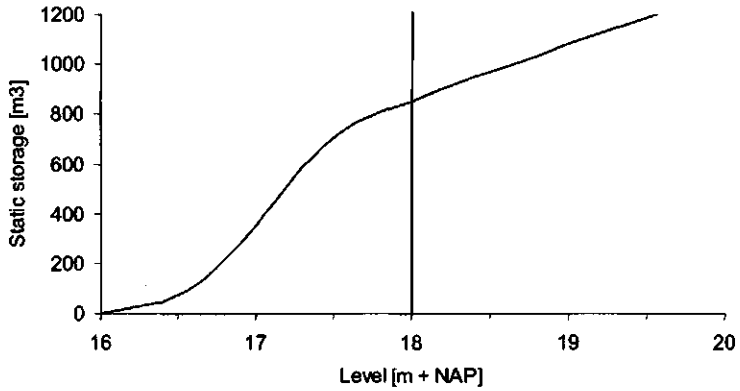


Figure 5.17. Curve of the geometrically determined static storage volume behind the overflow weir as a function of the water level at the weir.

The models final reservoir is a non-linear reservoir representing the sewers static storage behind the overflow weir. Dimensions of the sewer pipes and manholes have been used to calculate a fixed relationship between the water level and the static storage in this final reservoir. This relationship is shown in Figure 5.17.

Solving the Differential Equations

The rainfall-runoff model thus consists of linear and non-linear differential equations. The volume of water in each of the reservoirs is considered as a state variable. One way to solve the stochastic system is to consider it to be a set of stochastic differential equations. This is done in Chapter 6. Another option is to solve the differential equations using a standard ordinary differential equation (ODE) solver, in this case the Runge-Kutta method, and then to interpret this solution as being the predicted state corresponding to the function f( ) in (5.15). This is the approach adopted in the present study. For the linear reservoirs there is of course no need for an ODE solver as the analytical solution exists.

The Observation Equation

In the present case study the observed variable was the water level at the combined sewer systems only overflow weir. This level is an indirect observation of the volume of water in the final reservoir. The observation equation consisted of the volume-to-level relationship defined by the systems pipe and manhole structure. This relationship is shown in Figure 5.17. This relationship essentially defines the function *h( )* of the observation equation (5.16).

The Partial Derivatives

The partial derivatives of the state predictions with respect to the states' values at the previous time step are then computed by recalculating the predicted states for small changes in each of the state variables respectively. These are the partial derivatives that have been used in the Kalman filter in this case study.

State Dependent Errors

It would be unreasonable to assume that the size of the system error should be the same irrespective of the volume of water in the reservoirs. Indeed if this were the case a simulation of the system would during periods with low flow be likely to yield negative volumes. One solution to this problem is to model the system of volumes with state dependent error terms. The state dependency can be introduced by inserting an *error state* into the state equation for each of the volume states. The dependency was then defined as in equation (5.38) where the state index *i* only counts even numbers corresponding to the reservoir states (the odd states are now error states).

$$V_{i,t} = V_{i,t}^{*} \cdot \left(1 + \varepsilon_{i-1,t-1}\right) + \varepsilon_{i,t} \tag{5.38}$$

where $V_{i,t}$ is the volume in reservoir *(i/2)* at time *t*, $V_{i,t}^{*}$ is the volume in reservoir *(i/2)* at time *t* as calculated based on the solution of the defining differential equations. The first error term, $\varepsilon_{i-1,t-1}$, is in fact the value of the associated error state variable at time *t-1* and the second error term, $\varepsilon_{i,t}$, is the usual additive error defined in the generalised state space model. Inserting an error states in this way does have the disadvantage of doubling the size of the state space system being modelled.

The error term of interest in (5.15) is thus not the states own additive error, $\varepsilon_{i,t}$, but that of the preceding state. The expected value of the error states, corresponding to *f(...)* in (5.15), is, by definition, always equal to zero and the variance of their added errors was estimated as a parameter in the model. The variances of the added errors of the volume states themselves were fixed to zero. Tests carried out to evaluate added error terms different from zero showed that this did not improve the model at all.

A priori Parameter Estimates

Our engineering knowledge and understanding of the system suggests certain expected ranges for some of the parameters whose values will be estimated from the data. This knowledge was expressed as probability distributions for four of the model parameters: the runoff coefficient, the wetness dependency, weir overflow power and the pump capacity. These *a priori* distributions have been included in the results Table 5.1.

RESULTS

Parameter Estimates

Table 5.1 shows the results of estimating parameters as

- **a deterministic model** without any up-date of on the presence of observations,
- **a stochastic model** up-dating using the extended Kalman filter to up-date on the presence of an observations.

The standard errors and the corresponding confidence limits listed in Table 5.1 are the standard errors of the parameters when considering covariation between the estimated parameters as outlined earlier in the chapter. In the column on the far right parameter plots show the estimated parameters and their 95% confidence range. The *a priori* distributions have been included in the table and plots for the four parameters for which *a priori* distributions were used.

The parameters marked with two asterisks were transformed using the natural logarithm and the given standard error in Table 5.1 corresponds to the transformed variable. The transformation of these parameters was found necessary in order to reasonably be able to make the assumption of local linearity. This assumption is required to be able to estimate parameter variances and covariances.

The two sets of parameter estimates are generally very similar. An exception is the reservoir emptying rate for the linear reservoirs which is higher for the deterministic model, 0.33 per minute, compared to 0.24 for the stochastic model. Converted to reservoir time constants this corresponds to 3.0 and 4.2 minutes for the deterministic and stochastic models respectively. Three linear reservoirs were used and this rate thus corresponds to a mean runoff time of 9.0 and 12.6 minutes respectively. Another exception is the estimated value for the pump starting level, which was estimated to be 16.35 and 16.06 meters above the Dutch national reference NAP. This parameter is a threshold value that is only crossed at the beginning of each event. Both its value and its associated statistics should be considered with some caution.

Finally the observation error estimated for each of the models differs two orders of magnitude. This is not surprising as the underlying assumption in the case of the deterministic model is that all deviation between modelled and observed is observation error as opposed to the stochastic model were both system error and observation error are assumed to be present. It is here not possible to compare the variances of the system errors and the observation error because the variances are in different units. The system errors are multiplicative to water volumes whereas the observation errors are additive to a water level. Parameter uncertainty in the stochastic model appears to be smaller than that in the deterministic model. This could be due to more appropriate assumptions. However this is difficult to judge from the given case alone.

The *a priori* probability distributions do not appear to have had much influence on the final set of estimated parameters. If the observed dataset had been low on information about the value of one of the parameters for which an *a priori* distribution had been given then the *a posteriori* distribution would have been similar to the given *a priori* distribution. This is not the case for any of the four parameters. The variances of the *a posteriori* parameter distributions are much smaller than the given *a priori* distributions. Interesting is of course the fact that the estimated wetness dependency appears to be lower than what had been estimated in Chapter 4 (which had formed the basis for the rough *a priori* estimate of the runoff coefficient and wetness indices).

Parameter Correlation Matrix

The parameter correlation matrix for both the deterministic and stochastic models are given in Table 5.2. Several different sampling methods and step sizes were used to estimate the correlation matrix to ensure that the values found gave a fair representation of the curvature of the objective function surface. The parameter correlations are generally low for both models. All values larger than 0.4 have been highlighted in bold type.

Closer examination of the correlation matrix reveals that it is not always the same parameters that are correlated in each of the two models. The largest discrepancy turned out to be the correlation between the "Reservoir time constant" and the "Weir overflow coefficient" which was found to be 0.26 and –0.31 for the two models respectively. The models also had very different correlation between the "Weir overflow coefficient" and the "Weir power" (0.8 for the deterministic model and 0.34 for the stochastic model). In the latter case some correlation would be expected but no immediate reason has been found for the large differences between the models. The confidence contours in the plane spanned by this parameter pair is plotted and discussed later.

Trajectories

Figure 5.18 shows the predicted and observed values for one event selected from the 19 rainfall-runoff events used for the parameter estimation. The central plot shows the deterministic estimation and the bottom plot shows the stochastic estimation. The rainfall for the given runoff event is shown in the top plot. Figure 5.19 is an extract of Figure 5.18 zoomed in on the time steps 180 to 480 minutes after the start of the rainfall event.

In the deterministic plot (Figure 5.18 centre and Figure 5.19 centre) the up-dated value is equal to the predicted value. This it does by definition as no up-date is performed during deterministic parameter estimation. In the plot of the stochastic estimation (Figure 5.18 bottom and Figure 5.19 bottom) shows that the observation error is very small compared to the system error and the up-dated values are practically equal to the observed values. One can also see how the uncertainty generally increases when the time-since-last-observation increases. As expected the uncertainty tends towards a maximum value which depends on how rigid the system is. In this case the system is more rigid once overflow starts. This is because once overflow starts a given change or uncertainty in the upstream parts of the system will have a smaller effect on the water level at the weir.

In both models, but particularly for the deterministic estimation, it is apparent that the modelled values lie generally below the observed values during this particular event. This is because the parameters have been estimated from a large number of events. This leads to questions as to whether or not the runoff coefficients dependency on the wetness index is an appropriate model and questions concerning the assumption that the parameter values are constant from event to event. The latter question leads directly on to random coefficient modelling which is studied in Chapter 7 for a deterministic pollutant resuspension model.

A topic, which has not been treated in much depth in this thesis, is the analysis of residuals. Residuals are the deviations between modelled and observed values. For time series with equidistant observations the autocorrelation function of the residual series (i.e. linear correlation between the series' value and its own values at other time steps) is an important tool in finding out whether a model describes the observed process well and whether the implicit assumptions are reasonably satisfied. A low autocorrelation means that the residuals contain little or no further systematic variation that could be modelled. In the present study the data are very irregularly spaced in time and it was not possible to calculate the residual autocorrelation function directly. Some effort was given to developing a non-parametric autocorrelation function that would be generally applicable irrespectively of the spacing of the observations.

Table 5.1. Results of the parameter estimation for both the deterministic and the stochastic model. *A priori* distributions are given where used.
* Parameter statistics with collinearity considered. **Marked parameters where log transformed (thus the non-symetrical confidence ranges).

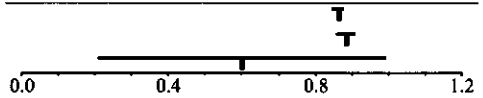
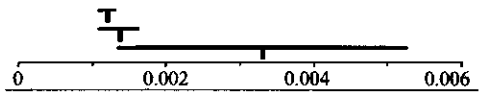
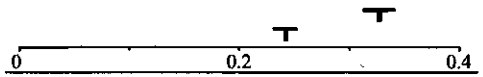
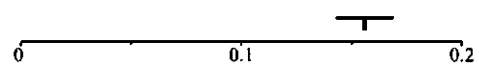
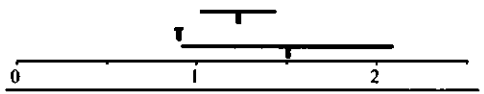
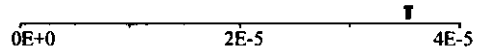
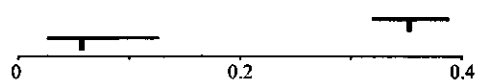
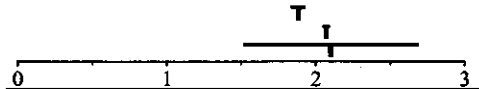
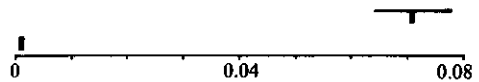
Nr.	Name	Symbol	Unit	Model	Estimate	Std. Error*	Lower 95%*	Upper 95%*	Parameter Confidence Plots
1	Runoff coefficient	a	[-]	Deterministic	0.86	0.0084	0.85	0.88	
				Stochastic	0.88	0.0133	0.86	0.91	
				a priori dist.	0.60	0.2000	0.21	0.99	
2	Wetness dependency	b	[mm^{-1}]	Deterministic	0.0012	5.93E-05	0.0011	0.0013	
				Stochastic	0.0014	1.43E-04	0.0011	0.0016	
				a priori dist.	0.0033	1.00E-03	0.0013	0.0053	
3	Reservoir emptying rate (inv. time constant)	k	[min^{-1}]	Deterministic	0.33	0.023 **	0.31	0.34	
				Stochastic	0.24	0.023 **	0.23	0.25	
4	Variance of reservoir error	σ_1^2	[-]	Deterministic					
				Stochastic	0.16	0.043 **	0.14	0.17	
5	Weir overflow coefficient	α	[m$^{3-\beta}$·min^{-1}]	Deterministic	3652	0.590 **	1148	11614	
				Stochastic	1143	0.028 **	1082	1208	
6	Weir power	β	[-]	Deterministic	1.23	0.107	1.02	1.44	
				Stochastic	0.90	0.009	0.88	0.92	
				a priori dist.	1.50	0.300	0.91	2.09	
7	Variance overflow error	σ_2^2	[-]	Deterministic					
				Stochastic	3.53E-05	0.0048 **	3.49E-05	3.56E-05	
8	Pump start level (+16m + NAP)	$h_{pump-on}$	[m]	Deterministic	0.352	0.051 **	0.319	0.388	
				Stochastic	0.057	0.408 **	0.026	0.126	
9	Pump capacity	$Q_{pump-cap}$	[m^{3}min^{-1}]	Deterministic	1.88	0.0262	1.83	1.93	
				Stochastic	2.07	0.0003	2.07	2.07	
				a priori dist.	2.10	0.3000	1.51	2.69	
10	Variance of observation error	σ_{obs}^2	[(m)^2]	Deterministic	7.06E-02	4.95E-02 **	6.40E-02	7.78E-02	
				Stochastic	8.77E-04	2.38E-03 **	8.73E-04	8.82E-04	

Table 5.2. Parameter correlation matrix for both the deterministic (before of the "/") and the stochastic model (after the "/").

	Parameter name	Symbol	Unit	1	2	3	4	5	6	7	8	9	10
1	Runoff coefficient	a	[-]	1.00/ 1.00	**-0.55/ -0.53**	-0.04/ 0.25	--- / 0.16	0.08/ 0.07	-0.01/ -0.24	--- / 0.18	0.08/ 0.02	0.27/ -0.07	0.08/ -0.04
2	Wetness dependency	b	[mm^{-1}]	**-0.55/ -0.53**	1.00/ 1.00	-0.08/ -0.36	--- / -0.40	-0.13/ 0.08	-0.05/ 0.05	--- / -0.10	-0.03/ -0.04	0.32/ 0.03	-0.01/ 0.03
3	Reservoir time constant	k	[min^{-1}]	-0.04/ 0.25	-0.08/ -0.36	1.00/ 1.00	--- / 0.06	0.26/ -0.31	0.25/ 0.27	--- / 0.05	0.07/ 0.07	0.09/ -0.02	0.03/ -0.01
4	Variance of reservoir err.	σ_1^2	[-]	--- / 0.16	--- / -0.40	--- / 0.06	--- / 1.00	--- / -0.07	--- / -0.14	--- / 0.03	--- / 0.02	--- / -0.01	--- / -0.01
5	Weir overflow coefficient	α	[$m^{3-\beta} \cdot min^{-1}$]	0.08/ 0.07	-0.13/ 0.08	0.26/ -0.31	--- / -0.07	1.00/ 1.00	**0.80**/ 0.34	--- / 0.01	0.03/ -0.03	0.05/ 0.00	0.12/ 0.00
6	Weir power	β	[-]	-0.01/ -0.24	-0.05/ 0.05	0.25/ 0.27	--- / -0.14	**0.80**/ 0.34	1.00/ 1.00	--- / -0.04	0.02/ 0.02	0.07/ 0.01	0.11/ 0.01
7	Variance of overflow err.	σ_2^2	[-]	--- / 0.18	--- / -0.10	--- / 0.05	--- / 0.03	--- / 0.01	--- / -0.04	--- / 1.00	--- / 0.02	--- / -0.30	--- / -0.30
8	Pump start level	$h_{pump\text{-}on}$	[m]	0.08/ 0.02	-0.03/ -0.04	0.07/ 0.07	--- / 0.02	0.03/ -0.03	0.02/ 0.02	--- / 0.02	1.00/ 1.00	0.19/ 0.02	-0.03/ -0.04
9	Pump capacity	$Q_{pump\text{-}cap}$	[$m^3 min^{-1}$]	0.27/ -0.07	0.32/ 0.03	0.09/ -0.02	--- / -0.01	0.05/ 0.00	0.07/ 0.01	--- / -0.30	0.19/ 0.02	1.00/ 1.00	0.08/ -0.33
10	Observation err. variance	σ_{OBS}^2	[$(m)^2$]	0.08/ -0.04	-0.01/ 0.03	0.03/ -0.01	--- / -0.01	0.12/ 0.00	0.11/ 0.01	--- / -0.30	-0.03/ -0.04	0.08/ -0.33	1.00/ 1.00

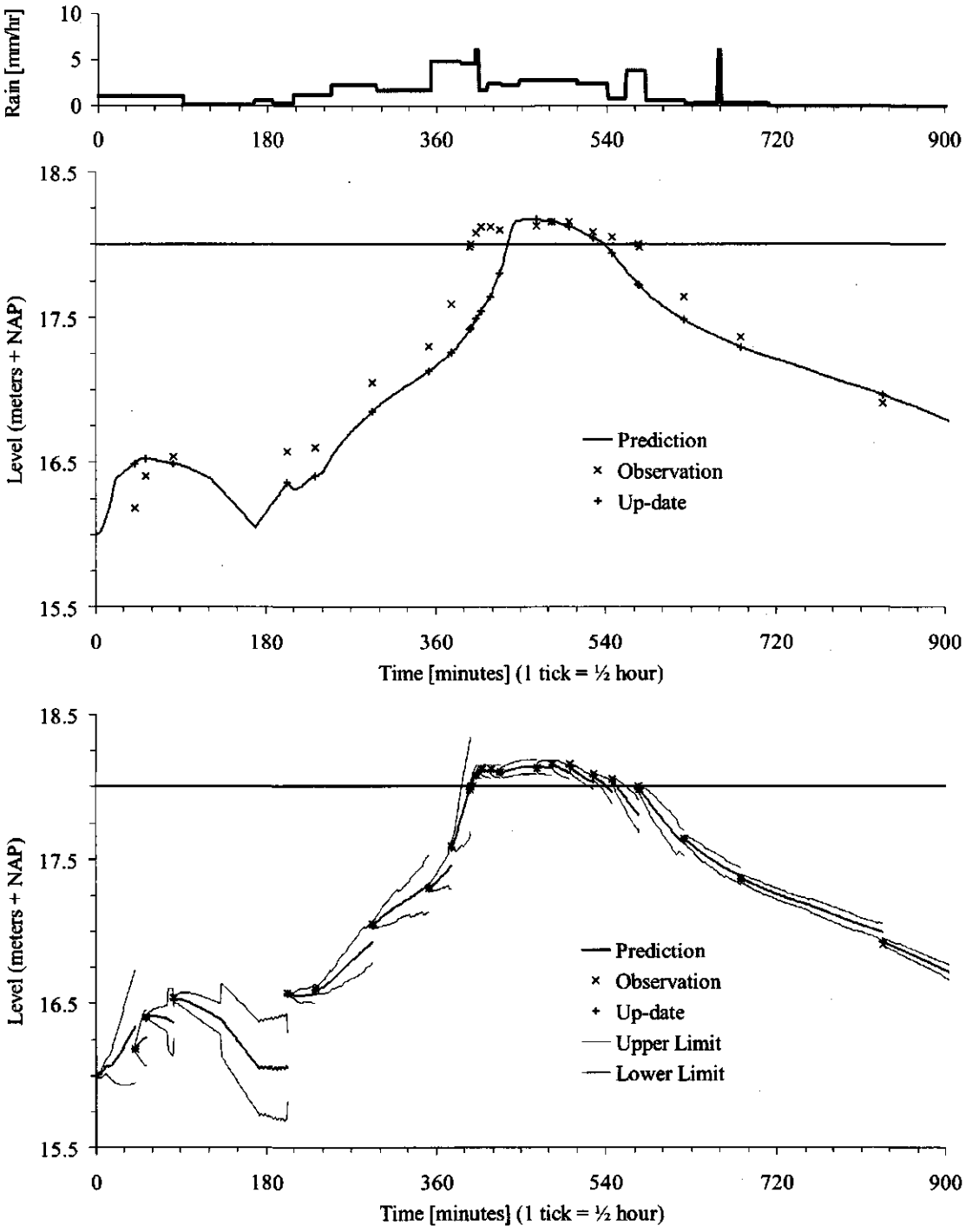


Figure 5.18. The observed and modelled values of water level at the overflow weir for both the deterministic (centre) and the stochastic estimation (bottom). Plotted is one event selected from the 19 rainfall-runoff used during parameter estimation. The top plot shows the associated rainfall series. Limits are 95% confidence limits.

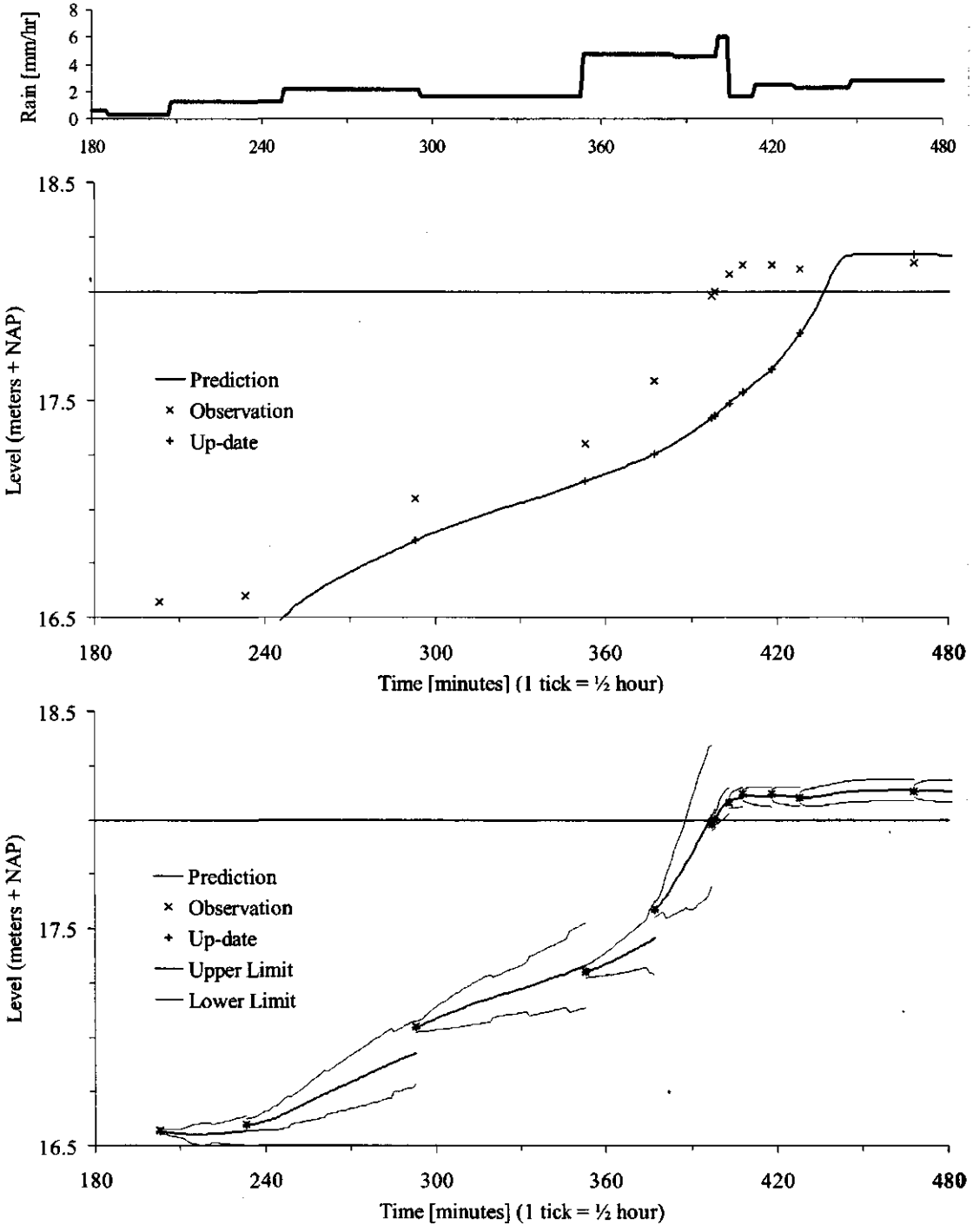


Figure 5.19. The observed and modelled values of water level at the overflow weir for both the deterministic (centre) and the stochastic estimation (bottom). Data show is the same as in the previous figure but zoomed in on time steps 180 to 480 minutes. Limits are 95% confidence limits.

The developed non-parametric auto- and cross-correlation function proved for the present data to be extremely sensitive to the selection of bandwidth, which was required. Although non-parametric autocorrelation functions for irregularly spaced data would be of immense value to practical water quality engineering, further developments in this direction lie beyond the scope of this study. Viewing the plot of the deterministic estimation in Figure 5.18 (centre) and Figure 5.19 (centre) it is evident that the deviation between the observed and modelled is not random. This suggests that the assumptions are not completely satisfied. This is seen to be better for the stochastic model.

Confidence Contours

Figure 5.20 shows the confidence contours for the plane defined by the parameters "Weir overflow coeffiecient" and "Weir power". The plots have been drawn both for the deterministic estimate (left) and for the stochastic estimate (right). In each case the found optimum is shown by a the black dot and the distance between the confidence contours is equivalent to a change of 2.0 in the log-likelihood value. Only the first 15 lines have been shown. The inner most line corresponds roughly to twice the standard deviation and therefore represents the 95% confidence region for the parameter pair.

This pair of parameters is shown here because a large discrepancy between the pairs parameter correlation in the two models (see Table 5.2). A correlation of 0.80 was found for the deterministic model whereas that for the stochastic model was only 0.34. It is apparent from the contour plots that the deterministic estimate of these parameters is more uncertain than that of the stochastic estimate. However, it also would appear that the ellipses are equally elongated and therefore we would expect the correlation coefficients to be roughly the same for the two models. From the confidence plots it would appear that the pair's correlation in the stochastic model (i.e. 0.34 in Table 5.2) has been somewhat under estimated and is in fact probably closer to the value found for the deterministic model (i.e. 0.80 in Table 5.2). The reason for this is likely to be unevenness of the objective function surface in the region of the optimum.

Estimation of the covariance matrix (used for the correlation matrix Table 5.2) is based on the assumption of local linearity which in the confidence contour plots should exhibit itself as "nice" ellipses with the optimum at their centre. Indeed the confidence contours do appear to be fairly elliptical but the optimum for the deterministic model does not appear to lie in the centre of the elliptical shapes. This is because the objective function surface is very flat along the bottom of the elliptical valley and possibly a bit rough so that the

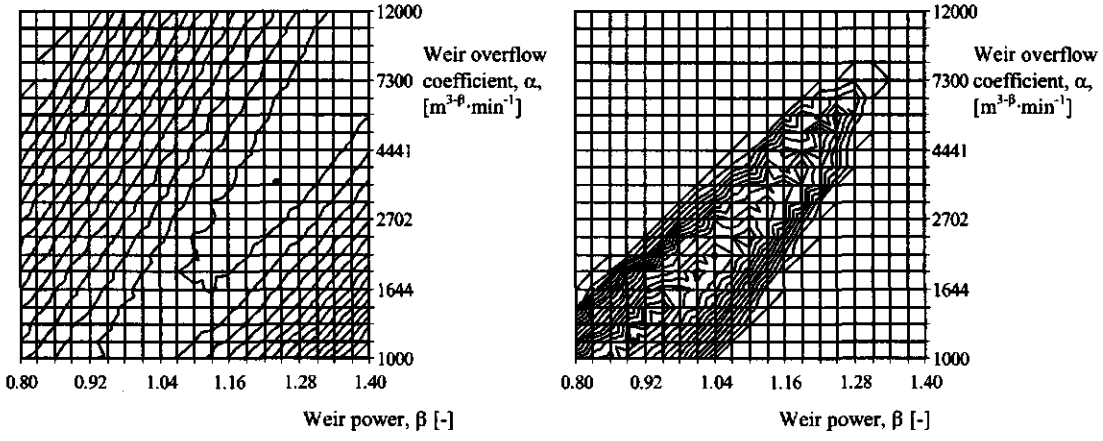


Figure 5.20. Confidence contour plots for the plane defined by the parameters "Weir overflow coeffiecient" and the "Weir power". Plot shown for the deterministic estimation (left) and for the stochastic estimation (right). Each contour represents a change of 2.0 in the log-likelihood value.

optimisation routine has been unable to find the very optimum. The point found is probably not a distinct "local minimum" because such a point would manifest itself in the parameter contour plot. It is more probably a result of the fact that the roughness of the objective function surface is large compared to the curvature. This would generally introduce some error on the estimated confidence limits purely due to the fact that the surface is not symmetrical around the optimum. Although the position of the optimum in the two models is not the same, it is worth noting that the optima lie in line with the tips of the each others confidence contours and are therefore in fact not that different.

Error Modelling

Introducing a multiplicative error on the reservoir water volumes by adding one error state for each of the physically interpretable states does indeed result in a more complex and demanding set of equations. An alternative to this is to transform the state variables. This approach was originally discarded because initial attempts using the natural logarithm to transform the volume gave poor results and much instability in the Kalman filter. Experience from similar modelling work carried out after this thesis work does however suggest that state transformation is indeed a viable and interesting option if only a suitable transformation can be found. The transformed variable(s) should be such that the assumptions of normally distributed additive error are reasonably satisfied. Promising results have been obtained using the transformation below for non-negative water quality variables such as volumes, flows, masses and concentrations.

$$f(x) = \frac{x^2 - c^2}{x} \tag{5.39}$$

where x is the untransformed variable, $f(x)$ is the transformed variable and where c is the intercept on the untransformed variable axis at which point transformed variable is equal to zero. The constant c can be estimated as a parameter in the model or otherwise appropriately chosen. The values of the models estimated parameters are not likely to be sensitive to the value of this transformation parameter. The advantage of using this transformation for quantities such as water volumes and component concentrations is that there is practically no transformation of the variable when it is much larger than the intercept value c. Stochastic water quality modelling would be strengthened by more research directed to finding suitable transformations for types of state variables typical within the field.

CONCLUSION

In this chapter deterministic and stochastic modelling approaches have been contrasted and compared. The underlying assumptions about the source of the deviation between modelled and observed values have been highlighted as the core difference between the two approaches. The discussion has been supported by the case study of a rainfall-runoff model for a combined sewer system. In Chapter 8 a broader discussion on deterministic and stochastic modelling in water quality management will be presented.

Whether the assumptions made are reasonable or not, will not only depend on how good a description one has of the mechanisms and processes involved, but also on the precision required and the available resources. In practical terms stochastic modelling with physical, chemical or biological descriptions means that also the uncertainty of a systems state has to be modelled in time. This is an increased effort which inevitably leads the practitioner to two questions: Does the improvement in modelling and subsequently design and decision making stand in proportion to the increased effort compared to traditional deterministic

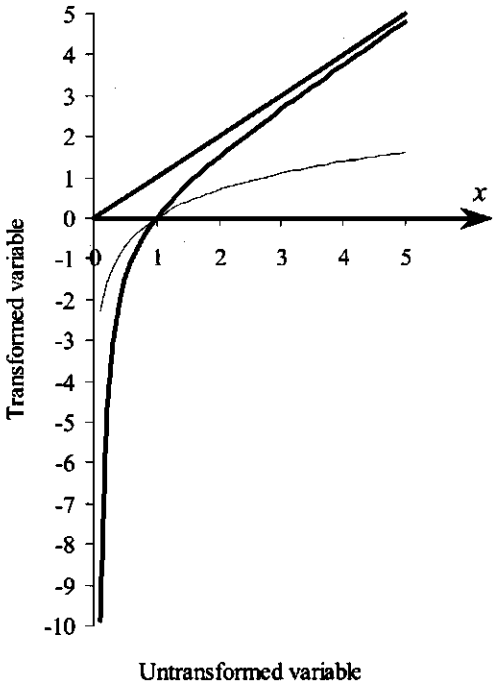


Figure 5.21. The state transformation (5.39) proposed (bold) as a suitable alternative to the log transformation (fine) for transforming environmental water quality state variables in stochastic modelling. Here c is chosen as 1 to make least difference to the natural logarithm transformation.

modelling? And if so, can the customer be convinced that this is the case? These questions lead beyond the scope of this thesis work but point to potential problems in bridging the gap between science and practice within the field of water quality management.

There are still a number of open questions when it comes to applying stochastic modelling to integrated urban water management and water quality modelling in general. Efforts should be made to find suitable transformations for typical water quality parameters to improve parameter statistics through a better adherence to the condition of local linearity. Studies should be carried out to identify state transformations that could be used to transform typical water quality state variables in such a way that random behaviour can be described with the normally distributed additive errors. Other aspects of applying stochastic modelling to water quality that would merit further study and experience are numerical methods for computing the partial derivatives needed in the Extended Kalman filter and for which analytical expressions are not available, estimation and parameter statistics of threshold values and the use and sensitivity of a priori parameter distributions. Generally speaking there is a need for stochastic tools aimed specifically at working with the sub-class of stochastic water quality and water engineering models.

Once the essential differences between stochastic and deterministic modelling have been pinned down it is possible to spot methods and approaches which may be beneficially transferred from the one to the other. As illustrated stochastic models, which have traditionally been empirical, would benefit from the inclusion of mechanistic descriptions. This would improve the possibilities of extrapolation to proposed scenarios and

may lead to a more parsimonious description. Deterministic modelling can, without taking the leap all the way to stochastic modelling, beneficially adopt several methods used in stochastic modelling. These include automated calibration/parameter estimation, parameter statistics, experimental design and systematic inclusion of *a priori* parameter knowledge.

REFERENCES

Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis: Forecasting & Control. Holden-Day Series in Time Series Analysis and Digital Processing. Holden-Day, Oakland, California, USA.

Carstensen, J., Harremoës, P. and Strube, R. (1996). Software sensors based on the grey-box modelling approach. *Wat. Sci. Tech.*, 33(1), pp. 117-126.

Chatfield, C. (1996). The analysis of time series. An introduction. 5. ed. Champman Hall. London.

Dochain, D., Vanrolleghem, P. A. and Van Daele, M. (1995). Structural identifiability of biokinetic models of activated sludge respiration. *Water Research*, 29(11), pp. 2571-2578.

Draper, N. R. and Smith, H. (1981). Applied regression analysis. Wiley series in probability and mathematical statistics. ISBN: 0-471-02995-5.

Harvey, A. C. (1993). Time Series Models. 310 pp.

Madsen, H and Holst, J (1998). Modelling Non-Linear and Non-Stationary Time Series. Lecture notes for the PhD. cource DTU-0417 Advanced Time Series Analysis. IMM-Institute of Mathematical Modelling, building 321, Technical University of Denmark, DK 2800 Lyngby. www.imm.dtu.dk.

Madsen, H (1995). Tidsrækkeanalyse (Time Series Analysis). IMM, DTU. Lecture notes used at IMM/DTU, University of Copenhagen (KUIMS), Copenhagen Business School, and the University of Iceland. 284 pp., www.imm.dtu.dk

Moens, M., Grum, M. and Aalderink, R. H. (1999). Inzet van troebelheidsmeters ter bepaling van CZV en drogestof in gemengde rioolstelsel. H 2 O, 32(8), pp. 22-25.

Ljung, L. (1987). System identification. Theory for the user. Prentice-Hall. New York.

Reichert, P. and Omlin, M. (1997). On the usefulness of overparameterized ecological models. *Ecological Modelling*, 95(2-3), pp. 289-299.

Sadegh, P., Holst, J., Madsen, H. and Melgaard, H. (1995). Experiment Design for Gray-box Identification. Int. Journal of Adaptive Control and Signal Processing, Vol.9, pp. 491-507.

Sadegh, P. (1996). Experiment Design and Optimization in Complex Systems. PhD. Thesis (IMM-PHD-1996-23). Institute of Mathematical Modelling, Technical University of Denmark, DK-2800 Kngs. Lyngby, Denmark, www.imm.dtu.dk.

Tulleken, H. (1993). Grey-box modelling and identification using physical knowledge and bayesian techniques. *Automatica* 29(2), 285-308.

Vanrolleghem, P. A. and Keesman, Karel J. (1996). Identification of Biodegradation models under model and data uncertainty. *Wat. Sci. Tech.*, 33(2), pp. 91-105.

Vanrolleghem, P. A., Van Daele, M. and Dochain, D. (1995). Practical identifiability of a biokinetic model of activated sludge respiration. *Water Research*, 29(11), pp. 2561-2570.

CHAPTER 6

MODELLING USING STOCHASTIC DIFFERENTIAL EQUATIONS

ABSTRACT

The model studied in this chapter incorporates notions of physical theory in a stochastic model of water level and particulate chemical oxygen demand (COD) at the overflow point of a Dutch combined sewer system. A stochastic model based on physical mechanisms has been formulated in continuous time. A tool dedicated to modelling using stochastic differential equations with the extended Kalman filter has been used. Estimation is based on a maximum likelihood criterion in conjunction with a non-linear state space formulation decomposing the error term into system noise terms and observation errors. The bias generally obtained in deterministic modelling, by invariably and often inappropriately assuming all error to observation error, is thus avoided. Continuous time stochastic modelling incorporating physical, chemical and biological theory presents a possible modelling alternative. These preliminary results suggest that further work is needed in order to fully appreciate the method's potential and limitations in the field of urban runoff pollution modelling.

This chapter is based on Grum, M. (1998). Incorporating Concepts From Physical Theory into Stochastic Modelling of Urban Runoff Pollution. Wat. Sci. Tech., 37(1), 179-185. Published after presentation at 2th International Conference on The Sewer as a Physical, Chemical and Biological Reactor, Aalborg, May 1997.

INTRODUCTION

In the preceding chapter the differences between traditional deterministic and stochastic modelling have been outlined in terms of their implicit assumptions with respect to the origin of the deviation between the modelled values and the observations. The methods have been compared and studied using a non-linear version of the Kalman filter implemented in discrete time.

This chapter looks at the possibilities and limitations of using *stochastic differential equations* for sewer system modelling. Stochastic differential equations distinguish themselves from the discrete time stochastic models presented in the two preceding chapters by being defined in continuous time. This has a number of advantages such as reduced sensitivity to the discrete time step that inevitably has to be chosen, as observations are generally always discrete. Stochastic differential equations have the further advantage that physical, chemical and biological process theory is often expressed in the form of ordinary differential equations and can in some cases be applied with only minor adjustments.

Like all stochastic modelling methods, modelling using stochastic differential equations is associated with techniques to assure model identifiability. Initially this involves studying parameter standard errors and correlation matrix. A co-linearity analysis would give a more complete picture of the parameters space and model identifiability. All these parameter statistics are based on the assumption the model is linear with respect to the model parameters. In the case of non-linear models the local validity of this assumption may be examined. This may be done using contour plots. Parameter transformations may be introduced before re-estimating the parameter statistics in order to live up to the assumption of local linearity.

DATA

This study has been carried out using the Loenen data set as described in Chapter 2. The observed variables that have been modelled are the water level at the overflow structure and the suspended or fast settling COD concentration. Table 6.1 gives a summary of the data used for this particular study.

METHODS

Modelling using Stochastic Differential Equations

Modelling using stochastic differential equations is a good tool for combining information from physical, chemical and biological theory with information from data and is therefore often also called grey box modelling (Madsen and Holst, 1996). In this section is a brief outline of a few essential elements of modelling using stochastic differential equations. Refer to Madsen and Holst (1998) and Madsen and Melgaard (1991) for further details on the mathematics behind modelling with stochastic differential equations. In the field of water resources research these methods have been applied to practical problems by several authors including Carstensen and Harremoës (1997), Jacobsen et al. (1996) and Carstensen et. al. (1996).

Table 6.1. Summary of the data studied.

	Level at Overflow	Suspended COD conc.
Mean	17.5 m	150.0 mg/l
Minimum	16.3 m	0 mg/l
Maximum	18.5 m	997 mg/l
No. events	22	19
No. observations	623	253

Consider, as an example, the usual (deterministic) continuity equation for a reservoir which could be written as:

$$\frac{dV}{dt} = Q_{in} - Q_{out} \tag{6.1}$$

V = the volume stored in the reservoir,
Q_{in} and Q_{out} = the flow into and out of the reservoir respectively.

The flows would often through some linear or non-linear storage function be connected to the reservoir volume and input flows. In accepting that the above differential equation does not cover "the whole truth", a stochastic term is added to (6.1) to give:

$$\frac{dV}{dt} = Q_{in} - Q_{out} + \frac{d\omega}{dt} \tag{6.2}$$

ω = a stochastic process assumed to have independent increments (a Weiner process).

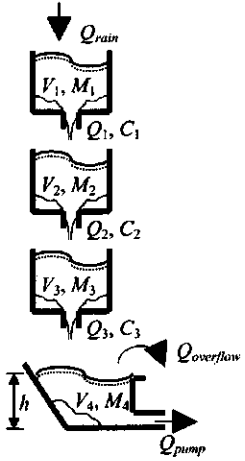
Consider the case in which the water level in the reservoir has been monitored. Assuming the reservoir to be such that the level can be calculated as a function of the volume, $f(\ )$, the following observation equation can be formulated:

$$h_t = f(V(t)) + \varepsilon_t \tag{6.3}$$

h_t = the monitored water level and $V(t)$ = the modelled volume in the reservoir,
ε_t = observation error (independent of the stochastic process ω).

The system and observation errors can not be found directly. However, given their variances the Extended Kalman filter can on the presence of a new observation be used to calculate the best estimates of the system's current state. Thus, at every available observation, the Kalman filter makes a weighting between "what we calculate" and "what we see" (see Chapter 5 and Chapter 8). This up-date affects all state variables and not only those which have been observed or which relate directly to the observed variables. Simulation or prediction then proceeds from these new estimates of the system state variables. The variances of the system and observation errors, and the other constant model parameters are estimated by an off-line optimisation of a maximum likelihood criteria. In the present study a modified version of the program CTLSM (Continuous Time Linear Stochastic Modelling), due to Melgaard and Madsen (1993), was used. This program, which is specifically designed for parameter estimation in non-linear stochastic differential equations, uses a modified quasi-Newton off-line optimisation of the parameter estimation.

Table 6.2. System model overview.

System sketch	Reservoir Storage Relationship	Stochastic Differential Equation: WATER QUANTITY (Level at overflow structure)	Stochastic Differential Equation: WATER QUALITY (Suspended COD concentration)
	$Q_{rain} = \phi A_{impervious\_area}\, r(t)$		
	$Q_1 = kV_1$	$\frac{dV_1}{dt} = Q_{rain} - Q_1$	$\frac{dM_1}{dt} = -Q_1 \cdot C_1 + g\left((Q_{rain} + Q_1)/2\right) - \beta_1 \cdot M_1 + \frac{d\varpi_2}{dt}$
	$Q_2 = kV_2$	$\frac{dV_2}{dt} = Q_1 - Q_2$	$\frac{dM_2}{dt} = Q_1 \cdot C_1 - Q_2 \cdot C_2 + g\left((Q_1 + Q_2)/2\right) - \beta_1 \cdot M_2 + \frac{d\varpi_2}{dt}$
	$Q_3 = kV_3$	$\frac{dV_3}{dt} = Q_2 - Q_3$	$\frac{dM_3}{dt} = Q_2 \cdot C_2 - Q_3 \cdot C_3 + g\left((Q_2 + Q_3)/2\right) - \beta_1 \cdot M_3 + \frac{d\varpi_2}{dt}$
	$Q_{overflow} = w_{coef} \cdot (h - h_{weir})^{\frac{3}{2}}$ and $h = f(V_4)$	$\frac{dV_4}{dt} = Q_3 - Q_{overflow} - Q_{pump} + \frac{d\varpi_1}{dt}$	$\frac{dM_4}{dt} = Q_3 \cdot C_3 - (Q_{overflow} + Q_{pump}) \cdot C_4 + g\left((Q_3 + Q_{overflow} + Q_{pump})/2\right) - \beta_1 \cdot M_4 + \frac{d\varpi_2}{dt}$
	Observation Equations:	$h_{observed,t} = f(V_4(t)) + \varepsilon_{1,t}$	$C_{observed,t} = M_4(t)/V_4(t) + \varepsilon_{2,t}$

Where

- The stochastic terms $d\varpi_1/dt$ and $d\varpi_2/dt$ are the derivatives of additive random processes (Wiener processes) for the quantity and quality equations respectively.
- The discrete time observation error terms ε_1 and ε_2 are assumed to be normally distributed random variables with mean zero and variance $\sigma^2_{\varepsilon 1}$ and $\sigma^2_{\varepsilon 2}$ respectively.
- $A_{impervious\ area}$ is the impervious area (fixed), ϕ is a runoff coefficient, $r(t)$ is the rainfall series and Q_{rain} is the flow into the first reservoir.
- V_i, Q_i, M_i and C_i are the volume, flow out, suspended COD mass and concentrations respectively (of the i'th reservoir).
- k is the proportionality constant (equal to inverse mean retention time) of the first three linear reservoirs, h_{weir} the weir level (fixed) and w_{coef} the weir coefficient.
- $g(\,)$ is resuspension of sewer sediments in each reservoir given as a function of the flows in the reservoir and with units [g/s]:

$$g(Q) = \begin{cases} 0 & \text{for } Q \le Q_{threshold} \\ \beta_2 \cdot (Q - Q_{threshold}) & \text{for } Q > Q_{threshold} \end{cases}$$

where β_2 has units [g/m^3] or [mg/l] and β_1 is the settling rate of the suspended COD with units [1/s].

$f(\,)$ is a function giving the water level at the overflow point as a function of the static storage in the sewer system. This relationship is based on the pipe dimensions of the actual sewer system

Table 6.3. Results of the parameter estimation in the final water quantity model.

No.	Parameter name	Units	Parameter symbol	Estimated value	95% confidence limits lower	95% confidence limits upper	Reported value
1	Runoff coefficient	[-]	ϕ	0.71	0.62	0.80	0.7-1.7
2	Proportionality constant	[1/min]	k	0.18	0.16	0.20	-
3	Overflow weir coefficient	[m^2/s]	w_{coef}	3.6	1.9	5.3	2.79
4	System noise variance	[$(m^3/s)^2$]	σ^2_ω	0.22^2	0.15^2	0.27^2	-
5	Observation error variance	[$(m)^2$]	σ^2_ε	0.097^2	-0.0038	0.15^2	-

Table 6.4. Parameter correlation matrix (for the final water quantity model).

No.	Parameter name	Unit	1	2	3	4	5
1	Runoff coefficient	[-]	1.00	-0.65	-0.51	0.87	-0.96
2	Proportionality constant	[1/min]	-0.65	1.00	-0.03	-0.65	0.70
3	Overflow weir coefficient	[m^2/s]	-0.51	-0.03	1.00	-0.08	0.26
4	System noise variance	[$(m^3/s)^2$]	0.87	-0.65	-0.08	1.00	-0.92
5	Observation error variance	[$(m)^2$]	-0.96	0.70	0.26	-0.92	1.00

Water Quantity Modelling (Level)

Water quantity has been modelled as three linear reservoirs in series followed by a final reservoir representing the static storage volume in the pipe system immediately before the overflow weir. Table 6.2 presents a combined representation of the reservoirs as sketches with the corresponding storage equations and stochastic differential equations. The flow out of the first three reservoirs is assumed to be proportional to the volume of water stored in each reservoir. The flow out of the final reservoir includes the flow pumped to the treatment plant (here assumed equal to the pump capacity) and the flow out of the system through the CSO structure which is calculated as a function of the water level at the overflow structure. The level is in turn calculated as a function of the volume stored in this final reservoir based on the pipe dimensions of the sewer system itself. If the water level is above the assumed known "pump on level" then the flow to the treatment plant is set equal to the pump capacity and otherwise to zero. In order to maintain a low number of parameters during these preliminary studies the rainfall is multiplied by a runoff coefficient and there are therefore assumed to be no initial losses. Five model parameters: runoff coefficient, reservoir storage constant, overflow weir coefficient, variance of reservoir noise and the variance of the observation error are estimated.

Water Quality Modelling (Suspended COD)

The suspended COD model is based on the premise that the main source of particulate pollution in the overflow is as a result of resuspended sewer sediments. The input to the first reservoir is thus assumed to contain no suspended material. It is further assumed that each reservoir has an infinite amount of available deposited sediment (see Results and Discussion for comments on this assumption). The water quantity model provides no flow velocities to which resuspension rate could be related. The resuspension rate is therefore calculated as a function, g( ), of the flow rates in and out of the reservoirs. Above an estimated threshold value, resuspension is assumed to be proportional to the flow rate. The stochastic differential equations for suspended mass of COD are shown in the far right column of Table 6.2. In this suspended COD model, three parameters are estimated: settling rate, threshold flow rate and a proportionality constant.

RESULTS AND DISCUSSION

Water Quantity (Level)

The water quantity model presented here contains 5 estimated parameters. These are listed in Table 6.3 where their estimated values and confidence limits for the final model are also given. All parameters except the observation error variance were found to be significantly different from zero. The estimated standard error for the observation error variance is so large that the 95% confidence interval spans down to below zero. This suggests that, in practical terms, there is no real observation error compared to system error, which describes the incompleteness of the models description of system inputs and behaviour.

The values listed in the far right column are those reported in some of the literature mentioned earlier in the data section. It is apparent from the values in Table 6.4 that the parameter correlations were generally rather high. This is particularly the case for correlation between the observation error variances and both the runoff coefficient and the system noise variance.

A number of parameters including pump on level, pump capacity and overflow weir coefficient were fixed to their "known" values (see Chapter 5). Future efforts should aim at having also these parameters estimated from the data as has been done in the equivalent deterministic model in Grum and Aalderink (1997) and Chapter 5.

Attempts were also made to identify a separate noise term for each of the four reservoirs and also to identify a common variance for the first three linear reservoirs with a fourth variance on the final reservoir. In both cases it was not possible to obtain convergence to a single or reproducible likelihood optimum. This could relate both to structural aspects of the model but also to the quality of data, in terms of both excitation and sampling frequency. The identifiability of stochastic terms in reservoir modelling (linear and non-linear) is clearly an area in need of much more experience.

Water Quality (Suspended COD)

Reproducible results proved hard to obtain in the case of the water quality model. This was probably for a large part due to the small number of observed data available, namely only an average of only 13 observations per event for 19 rainfall-runoff events. Table 6.5 and Table 6.6 contain the results of the parameter estimation. The threshold flow rate of the resuspension equation is not significantly different from zero and could therefore be excluded in order to obtain an improved model.

The variance of the system noise term is seen to be significantly different from zero as is the variance of the observation error. These results would suggest that both system noise and observation error should be modelled. The absolute value of $1.58 \cdot 10^{-5}$ mg/l for the observation error standard deviation however appears to be suspiciously small. Suspiciously small were also the correlation coefficients between the observation error variance and the other model parameters. It should be noted that though the results were reproducible from different parameter starting values not *all* starting values converged to the optimum point.

Other resuspension models were examined. These included the estimation of a maximum amount of available deposited sediment and other functions relating the resuspension rate to the flow rate. None of these gave better results than the presented model.

Table 6.5. Results of the parameter estimation in the final water quality model (suspended COD).

No.	Parameter name	Units	Parameter symbol	Estimated value	95% confidence limits lower	upper
1	Flow rate threshold	$[m^3/s]$	$q_{threshold}$	$8.7 \cdot 10^{-10}$	$-2.0 \cdot 10^{-10}$	$2.0 \cdot 10^{-9}$
2	Settling rate	[1/s]	β_1	0.00294	0.00293	0.00295
3	Proportionality constant	$[g/m^3]$	β_2	1749	1705	1791
4	System noise variance	$[(g/s)^2]$	σ^2_ω	1378^2	1364^2	1400^2
5	Observation error variance	$[(mg/l)^2]$	σ^2_ε	$(1.58 \cdot 10^{-5})^2$	$(1.57 \cdot 10^{-5})^2$	$(1.68 \cdot 10^{-5})^2$

Table 6.6. Parameter correlation matrix (for the final water quality model, suspended COD).

No.	Parameter name	Units	1	2	3	4	5
1	Flow rate threshold	$[m^3/s]$	1.00	0.35	0.62	-0.85	-
2	Settling rate	[1/s]	0.35	1.00	0.11	-0.07	-
3	Proportionality constant	$[g/m^3]$	0.62	0.11	1.00	-0.87	-
4	System noise variance	$[(g/s)^2]$	-0.85	-0.07	-0.87	1.00	-
5	Observation error variance	$[(mg/l)^2]$	-	-	-	-	1.00

CONCLUSIONS

It has been possible to identify and estimate the parameters of a sewer system water quantity model. The results of the water quantity model suggest that the generally accepted assumption that all error results from observation error, which is the basis of deterministic model calibration, is not a valid assumption. On the contrary these results suggest that most of the deviation between model prediction and observation are a result of sewer system behaviour that the model does not explain. It is difficult to conclude anything from the results of the water quality model. They suggest that more research efforts should be put into applying existing experimental design and sampling frequency theory to water quality modelling. It also indicates that a new approach to data collection is required when subsequent modelling is amongst the primary purposes for the data collection.

The results from both models and from the study as a whole suggest that more efforts should be put into exploring and gaining experience in stochastic modelling with physically interpretable parameters. This is particularly important with respect to the identifiability of both the mechanistically interpretable model parameters and of the system noise components.

REFERENCES

Carstensen, J. and Harremoës, P. (1997). Time series modelling of overflow structures. *Wat. Sci. Tech.*, 36 (8-9), pp. 45-50.

Carstensen, J., Harremoës, P. and Strube, R. (1996). Software sensors based on the grey-box modelling approach. *Wat. Sci. Tech.*, 33(1), pp. 117-126.

Grum, M and Aalderink, R H (1997). A Statistical Approach To Urban Runoff Pollution Modelling. *Wat. Sci. Tech.*, 36(5), pp. 117-124,

Jacobsen, J. L., Madsen, H. and Harremoës, P. (1996). Modelling the transient impact of rain events on the oxygen content of a small creek. *Wat. Sci. Tech.*, 33(2), pp- 177-185.

Madsen, H and Holst, J (1998). Modelling Non-Linear and Non-Stationary Time Series. Lecture notes for the PhD. cource DTU-0417 Advanced Time Series Analysis. IMM-Institute of Mathematical Modelling, building 321, Technical University of Denmark, DK 2800 Lyngby. www.imm.dtu.dk.

Madsen, H and Melgaard, H (1991). The Mathematical and Numerical Methods used in CTLSM (Continuous Time Linear Stochastic Modelling) - a program for ML-estimation in, continuous time dynamic models. IMM-Institute of Mathematical Modelling, building 321, Technical University of Denmark, DK 2800 Lyngby, www.imm.dtu.dk.

Melgaard, H and Madsen, H (1993). CTLSM version 2.6 - a program for ML-estimation in stochastic differential equations. Tech. rep. No. 1/1993, IMM-Institute of Mathematical Modelling, building 321, Technical University of Denmark, DK 2800 Lyngby.

CHAPTER 7

RANDOM COEFFICIENT MODELLING

ABSTRACT

Combined sewer water quality models are sometimes able to capture the general dynamics of the changes in pollutant concentrations but will for some events over-predict and for other events under-predict. For some events the characteristic dynamic pattern will appear too early and for other again too late. This is because during calibration model parameters have either been considered as universal constants valid for all events or as depending in some deterministic way on explanatory variables such as the time of year or the sewer system's recent history. However, for the majority of water quality variables and parameters describing the state of a combined sewer system at the onset of each rainfall event are only very poorly known. These include quantities such as amount of deposited sediment, the strength of the sediment crust, concentrations of sediment pore water, the size distribution and settling rates of the particulate matter and even to some extent the quantity and quality of the dry weather flow.

The objective of this study has been to evaluate the potential of applying random coefficient modelling in order more accurately describe pollutant concentrations in combined sewers during rain. In random coefficient models certain selected parameters are assumed to be realisations of a random variable at each event. Emphasis is on models that are applicable to return period analysis of extreme pollution events in an integrated approach. A random coefficient model is estimated for suspended chemical oxygen demand (COD) in a combined sewer and its performance is compared to the equivalent constant coefficient model. As well as highlighting the advantages and the potential of the approach, drawbacks and limitations of both the approach and the present study are also discussed.

INTRODUCTION

The parameters of a combined sewer model are generally considered to be constant over time or to follow a yearly pattern. Their values are calibrated or estimated by comparing simulated model responses to measured responses. A frequently encountered problem is that, although the general shapes of the modelled responses are often very similar to the observed responses, the modelled response will for one event be too high, for another too low, for a third too early and a fourth too late. This is illustrated in Figure 7.1 where the variable on the y-axis could for example be suspended solids in a combined sewer system. If one attempts to calibrate the model on each of the single events individually this discrepancy between modelled and observed values generally disappears.

The approach applied in this study is called random coefficient modelling because it assumes that whilst some parameter values may be constant in time, other parameters vary from event to event as realisations of a random variable. The random coefficients are thus quantities that vary from event to event but for which no deterministic relationship can be found to explain these changes quantitatively. In combined sewer water quality modelling one would expect values such as, for example, the amount of deposited material, its content of degradable matter and the strength of the sediment crust, to vary from event to event. Numerous examples of modelling and studies of sewer system have however shown that the number and complexity of phenomena taking place in a combined sewer would require an endless number of parameters to be calibrated. One may well end up including phenomena that are less significant than phenomena that have yet to be revealed.

This problem of model assumptions not considering variations in system characteristics from event to event is visually more evident during deterministic modelling than during stochastic modelling. On stochastic modelling the constant up-date of the system states would often mean that only a detailed analysis of the residuals would reveal the need for random coefficient modelling. It is important to realise that, although this study deals with random coefficient modelling in the context of deterministic modelling, also stochastic models can be formulated and estimated with some random coefficients whose randomness is not described as randomness in time but as random realisations at each occurrence.

The objectives of the study are to evaluate the potential of random coefficient modelling of combined sewer water quality modelling. The advantages and disadvantages are examined through the application of the methodology to a simple deterministic model of suspended chemical oxygen demand (COD) in combined sewer overflow. The application of random coefficient models to integrated return period analysis is part of Chapter 9 of this thesis. There are very few studies in which random coefficient modelling has been applied within the field of water modelling. An example is Gwo-Fong and Yu-Ming (1996) who have studied the use of random coefficient modelling for hydrological runoff models with the storage coefficient treated as a random variable in a dynamically stochastic model applied to a rural watershed.

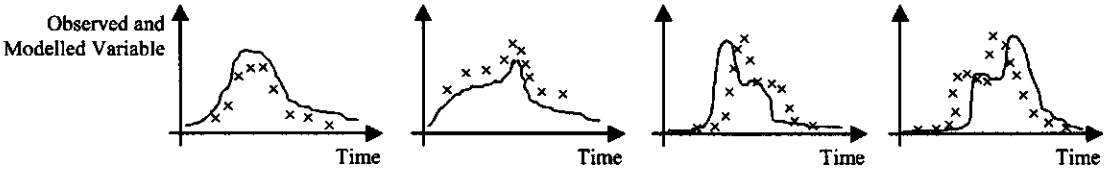


Figure 7.1. With constant coefficient models the principle dynamics may be well explained but for different rainfall-runoff events the simulation may lie too high, too low, too early or too late.

DATA

The data used for this study are part of the Loenen data set outlined in chapter 2. Water quality variables including COD were measured before and after settling for one hour. The difference between the two values represents a fast settling fraction or particulate fraction and is in this chapter called suspended COD. Only events with at least eight observation of suspended COD were selected for the study. This resulted in nineteen rainfall-runoff events with a total of 253 observations of suspended COD from samples taken during overflow.

METHOD

Parameter estimation in non-linear random coefficient modelling involves an outer optimisation routine for estimating the parameters considered to be constant. Each evaluation of the estimation criterion's value for a given set of the constant parameters involves event-level optimisations in which the random coefficient's values are estimated for each event. Thus, as illustrated in Figure 7.2, each evaluation of the objective function, i.e. the estimation criterion, in the outer optimisation involves complete optimisations to find the values of the random coefficients. A least square estimation criterion was used in the present study. Refer to statistical texts such as Bondeson (1989), Johansen (1984) or Nicholls and Quinn (1982) for a comprehensive introduction to random coefficient modelling.

Parameter statistics including covariance matrix was here estimated using a jack-knife technique in which the complete parameter estimation is carried out using all events and then repeated whilst systematically excluding one of the rainfall-runoff events. The jack-knife technique is discussed in more detail in Chapter 9 and a general introduction can be found in statistical text books such as Stuart and Ord (1994).

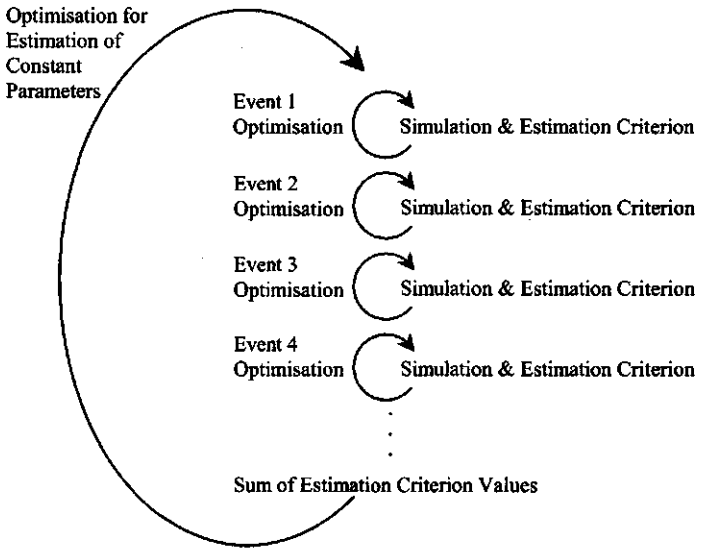


Figure 7.2. The constant parameters and the random parameters/coefficients are estimated simultaneously. Each function evaluation in the outer optimisation involves complete optimisations of the parameters varying from event to event.

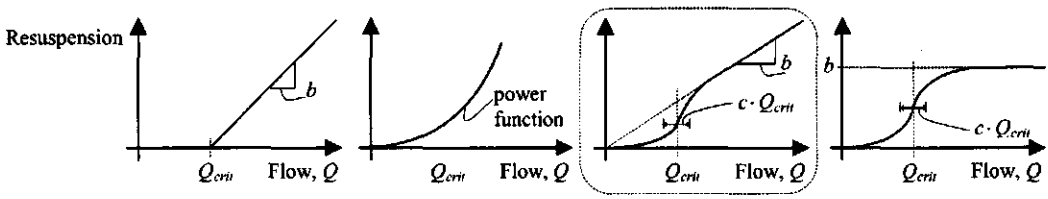


Figure 7.3. Four different relationships between flow and resuspension were initially tested. The encircled relationship was selected for this study and is expressed in equation (7.1).

Suspended Pollutants Models

Suspended COD was modelled as a first order settling rate with a flow dependent resuspension of both surface and in-sewer deposits. Four different expressions for resuspension were initially tested and one was selected for this study. The four expressions relating resuspension to flow are shown in Figure 7.3 where the encircled relationship gave the best fit and was subsequently selected this study. The mathematical expression for this resuspension relationship is expressed in (7.1).

$$\frac{dM_{COD}}{dt} = \overbrace{-a \cdot M_{COD}}^{\text{Settling}} + \overbrace{\frac{b \cdot Q}{2}\{1 + \tanh[c \cdot Q_{crit} \cdot (Q - Q_{crit})]\}}^{\text{Resuspension (Figure 7.3)}} \quad (7.1)$$

where M_{COD} is the suspensed mass of COD in the reservoir,

Q is the flow rate (see section on underlying water quantity model below),

a is a first order settling rate,

b is the gradient in the flow to resuspension relationship (see Fig. 7.3),

c defines (with Q_{crit}) the smoothness of the rise to linear resuspension (see Fig. 7.3),

Q_{crit} is the critical flow defining the position of the rise to linear resuspension (see Fig. 7.3).

In the final version of the random coefficient model the parameters a, b and c were all considered as constant parameters and Q_{crit} as the random parameter who's mean and variance were of interest. The critical flow rate, Q_{crit}, thus takes a different value for each rainfall-runoff event corresponding to variations which would be expected in the critical shear stress at which resuspension would begin. For purposes of comparison the model was also estimated with the critical flow rate, Q_{crit}, as a constant for all 19 rainfall-runoff events.

The sewer system is relatively flat and it was assumed that an infinite amount of sediment was available. Tests were carried out to confirm that setting or estimating a maximum available amount of sediment did not change the fit of suspended COD in the sewer overflow.

The Underlying Water Quantity Model

The suspended COD model was built upon a water quantity model consisting of three linear reservoirs in series discharging into a non-linear reservoir representing the sewer system's static storage volume. This is the deterministic version of the model presented in Chapter 5. The parameters of the water quantity model had been estimated beforehand and were not estimated as random coefficients.

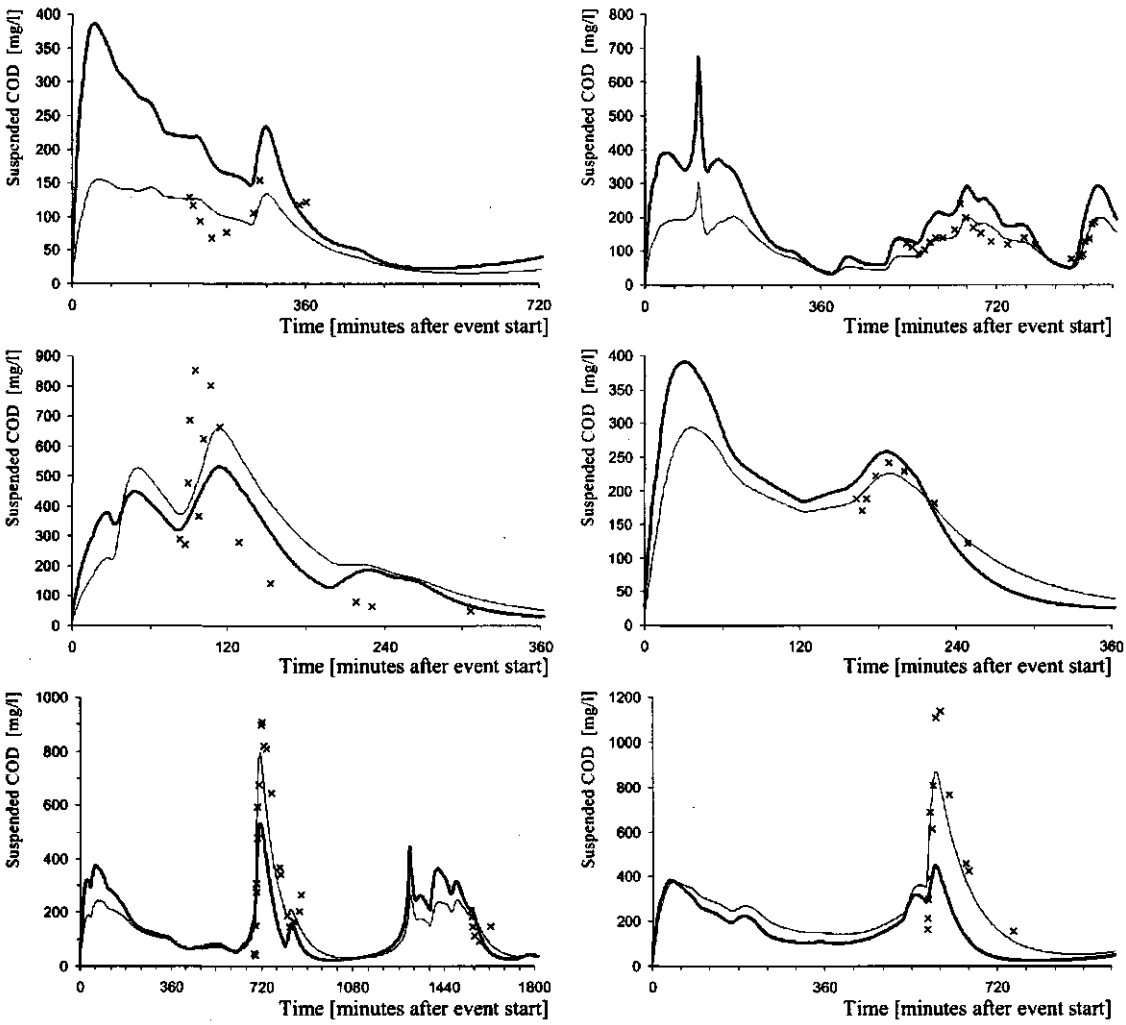


Figure 7.4. Results of the suspended COD modelling showing the observed values (crosses) with the constant coefficient model simulation (thick line) and the random coefficient model (thin line). In the top two event the constant coefficient model simulated too high, in the bottom two event too low and fair for the middle two events. Notice that simulations of the random coefficient model lie neither particularly high nor low for any of the events.

At the beginning of each event the suspended COD concentration in all four reservoirs was considered to be zero. Thereafter the concentration rose and fell according to the expression in (7.1), which was applied to each of the model reservoirs with resuspension depending on the average of inflow and outflow. Suspended COD was then carried downstream by advection. No dispersion was applied to the suspended pollutant.

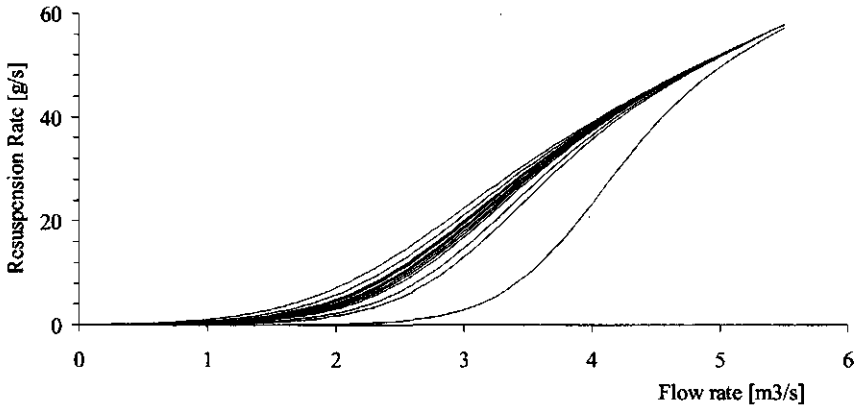


Figure 7.5. Estimated relationship between flow rate and resuspension in the random coefficient model. Each line is the resuspension curve corresponding to one of the 19 rainfall events used.

RESULTS AND DISCUSSION

The results of the random coefficient model and the corresponding constant coefficient model are presented in Table 7.1. Notice that for the random coefficient model the standard deviation of the critical flow rate, Q_{crit}, expresses variation from event to event. This should not be confused with the parameter standard errors, which are the results of the Jack-knife estimation and which represent uncertainty in the value of the estimated constant parameters.

Table 7.1. Results of constant and random coefficient modelling. The parameter Q_{crit} Std.dev. is the standard deviation of the event to event variation in the critical flow Q_{crit}. (**See text on interpretation of R^2).

			Constant Coefficient Model			**Random Coefficient Model**		
	Parameter	Unit	Value	Low 95%	Upp.95%	Value	Low 95%	Upp.95%
Settling rate	a	[1/s]	4.39e-4	3.61e-4	5.35e-4	2.44e-4	2.15e-4	2.78e-4
Resuspension	b	[g/m^3]	5.79e2	3.67e2	7.92e2	1.06e4	0.82e4	1.30e4
	c	[s/m^3]	0.498	0.359	0.660	0.317	0.256	0.377
	Q_{crit}	[m^3/s]	0.833	0.459	1.207	2.85	2.71	2.99
	Q_{crit} Std.dev.*	[m^3/s]	—	—	—	0.306	0.030	0.582
Coef. of Det.	R^2 **	[-]	0.47	—	—	0.67	—	—

The coefficient of determination, R^2, is seen to rise from 0.47 to 0.67 when the critical flow rate is assumed to be a random variable and therefore estimated for each event. For six of the 19 rainfall-runoff events the simulated and observed suspended COD concentrations have been plotted in Figure 7.4. Notice that for the top two events the constant coefficient model was generally too high, for the two events in the centre both models perform equally well and for the bottom two events the constant coefficient model was too low. This is exactly the sort of improvement one wants and expects when applying random coefficient modelling.

In Figure 7.5 are the flow-resuspension relationships found for each of the 19 events. The curves are defined by the last term in (7.1) and all have the same value for b and c but different values for the critical flow rate, Q_{crit}.

There is a significant improvement in the coefficient of determination, R2, from 0.47 to 0.67. However, it is important to keep in mind that the coefficient of determination for the random coefficient model is in a sense conditional to our knowing the values of the critical flow rate, Q_{crit}, for every event. In the random coefficient model these are in fact not known but estimated.

Table 7.1 reveals a poor agreement between particularly the settling rates in the two models. This should be viewed in consideration of the high correlation found between settling rate and resuspension (in all models about 0.9 between a and b). Settling and resuspension are two phenomena working practically directly against each other and almost the same result can therefore be obtained by increasing the two parameters a and b. The result is a high correlation between the parameters and this sets extremely high demands on the quality of the required data. Had the data sets included values for the period before actual overflow started this correlation would probably have been much lower. Both of the presented models have been estimated using the least square criterion and without any quantitative *a priori* information about the parameter values. The inclusion of quantitative *a priori* knowledge about sewer sediment settling rates using *a priori* distributions with a maximum likelihood criterion as outlined in Chapter 5 is likely to have overcome this problem. The *a priori* distributions could originate from various laboratory or field experiments such as reported in Hvitved-Jacobsen et al. (1998) and Gent et al. (1996). It is important that this *a priori* knowledge is incorporated in the form of *a priori* distributions in a maximum likelihood estimate and not just as fixed values in order to permit the assimilation of data information on any of the parameters.

Attempts were made to correlate the critical flow rates found for each event with variables such as time of year, preceding dry weather period and the wetness index presented in Chapter 4. No relationship with these variables could be found. Until such a relationship to an explanatory variable is found, the variation in Q_{crit} has to be considered as random variation. On application of the model in an integrated return period analysis, such as presented in Chapter 9, this variation from event to event is included as inherent variation.

Attempts were made to investigate whether further improvements to the fit could be obtained by having other parameters estimated as random coefficients varying from event to event. Improvements were very small and it was soon apparent that the pairs of parameters estimated for each individual event were highly interchangeable and could have very different values for the same goodness of fit. It is evident that care should be take not to classify more parameters as random coefficients than can be identified by the individual events.

An obvious disadvantage of random coefficient modelling is the long time needed for parameter estimation due to the sub-optimisations which have to be done in order to estimate the values of the random coefficients of each event. In models where there is a linear relationship between the response and the random coefficients, the sub-optimisations are replaced by the equivalent explicit expressions for each set of constant coefficients.

CONCLUSION

A random coefficient model has been formulated for the suspended COD concentration in a combined sewer system. The model comprises a settling term and a resuspension term whose parameters were estimated using data from nineteen rainfall-runoff events. Only the critical flow rate, comparable to a critical shear stress, was assumed to be the realisation of a random variable whilst all other parameters were estimated as constant coefficients. Results show a clear improvement in fit compared to the equivalent constant coefficient model. Rather than an increased degree of determination, this improved fit reflects the more appropriate modelling assumptions and consequently a reduced bias on the estimates of the model parameters in general.

The high correlation found between settling and resuspension parameters confirms not only the need for improved data but also the need for the inclusion of quantitative *a priori* knowledge concerning parameter values of even very simple models with only few parameters. Once established with the inclusion of quantitative *a priori* knowledge in a maximum likelihood estimate, random coefficient modelling is likely to become extremely useful methodology. Given the high degree of randomness within water quality processes and phenomena random coefficient modelling undoubtedly has a high future potential in the context of both deterministic models (as here) and in the context of stochastic models.

REFERENCES

Bondeson, J. (1989). Random coefficient regression models in biostatistics. University of Lund and Lund Institute of Technology. Department of Mathematical Statistics, Lund, Sweden. 151 pp.

Gent, R., Crabtree, B. and Ashley, R. (1996). A review of model development based on sewer sediments research in the UK. *Wat. Sci. Tech.* 33 (9), pp 1-7.

Gwo-Fong, Lin and Yu-Ming, Wang (1996). General stochastic instantaneous unit hydrograph. *J. of Hydrology*, Vol. 182, Issue. 1-4, pp 227-238.

Hvitved-Jacobsen, T., Vollertsen, J. and Tanaka, N. (1998). Wastewater quality changes during transport in sewers - an integrated aerobic and anaerobic model concept for carbon and sulfur microbial transformations. *Wat. Sci. Tech.*, 38(10), pp 257-264.

Johansen, S. (1984). Functional relations, random coefficients and nonlinear regression with application to kinetic data. Series Lecture notes in statistics 22, 126 pp, Springer, New York

Nicholls, D.F. and Quinn, B.G. (1982). Random coefficient autoregressive models : An introduction. Series Lecture notes in statistics ; 11, 154 pp, Springer, New York.

Stuart, A. and Ord, J. K. (1994). Kendall's Advanced Theory of Statistics. Vol. 1. Distribution Theory. Edward Arnold.

CHAPTER 8

DISCUSSION ON MODELLING

ABSTRACT

The objective of this chapter is to discuss, in a broader perspective, the approaches, methodologies and results presented in this part of the thesis. While focusing on deterministic and stochastic modelling and on aspects of identifiability, the discussion goes beyond the combined sewer modelling to water quality modelling in general. Finally, the chapter closes this part of the thesis with recommendations for both practice and research.

A part of the discussion in this chapter is based on Grum, M. (1997). WATERMATEX '97: A Symposium Review. Water Quality International, Sept./Oct. 1997, pp 37-39.

INTRODUCTION

Mathematical models aiming at simulating the system behaviour play an important role in integrated urban water management. Models are used to assess the impacts of combined sewer overflow and to identify optimal amelioration strategies. Optimal water quality management requires an integrated approach. Integrated urban water management may imply integration at many different levels. Integration may refer to the combined consideration of water quantity and water quality. In this chapter and in the thesis as a whole the term integration is primarily used to refer to the combined consideration of various parts of the water system such as the sewer system, the wastewater treatment plant, the recipient surface waters, the groundwater aquifers and the drinking water supply system extracting water from these aquifers. In recent years "an integrated approach to water management" has often been be used to refer to the integration of social aspects into water management and policy formulation.

Discharge from combined sewer systems is often a limiting factor to achieving ecologically healthy urban waters. Mathematical models of relevant parts of the integrated water system can be used to calculate return periods of detrimental effects and to analyse the consequences of proposed amelioration projects.

With an emphasis on combined sewer overflow, the research presented in this part of the thesis has focused on various approaches to modelling water quantity and quality. Common to all approaches is that the engineers a priori perception of the system is combined with information present in monitored data. Model structure and parameters are thus in part estimated from the available data. A primary aim has been to compare and contrast how information in data is used in deterministic and stochastic modelling approaches. This chapter discusses the previously presented conclusions in a broader perspective. The chapter concludes with recommendations directed towards both research and practice.

DETERMINISTIC AND STOCHASTIC MODELLING

When encountering an observation during model calibration, we can basically do three things. First, we could say that "our system description is perfect", note the deviation between modelled and observed and thereafter continue our modelling in time from this same modelled value. Alternatively, we could say "we'd better trust in what we see" (our observed value), note the deviation between modelled and observed, adjust the state of our system to fit the observed value and thereafter continue our modelling in time from this observed value. In the first case, we implicitly assume all deviation between the modelled and observed values to be a result of observation error alone; in the second, we assume that this deviation results only from random behaviour of the system (i.e. an incomplete system description).

A third option would be to say that "the truth" lies in a bit of both. In which case we would use the Kalman filter which makes a weighting between "what we see" and "what we model", and our modelling in time then proceeds from this new up-dated point. If we, on the presence of an observation, use the Kalman filter to estimate the state of our system, we are acknowledging the presence of both observation error and random behaviour of the system (i.e. modelling error). It is precisely this acknowledgement that gives the Kalman filter (and related forms of up-dating) such immense potential value to environmental and water quality modelling where all influential details never can be modelled and where inhomogeneity and measurement imprecision inevitably give rise to observation errors.

The main entry of the Kalman filter into both hydrology and water quality took place some twenty years ago with a conference titled "Application of Kalman Filter to Hydrology, Hydraulics and Water Resources" (Chiu, 1978). During the following ten years there was a rapid increase in the number of published applications of Kalman filtering in water quality modelling. The majority of these applications came from stochastic modellers and as a consequence (at that time), somewhat unfortunately, were empirical with little or no mechanistic interpretation. However, the state-space form, over which the Kalman filter is defined, is ideal for mechanistic system descriptions. Kalman had already emphasised this, but in the field of hydrology and water quality this seems to have been recognised only by a few (e.g. Beck and Young, 1976). Filtering was often done with model parameters varying with the state of the system and using the convergence properties of the filter to estimate their values. Stories of the practical problems of choosing suitable starting values, of difficulties in establishing system noise and observation error variances, of convergence to wrong parameter values and stories of instability abounded. From the middle of the eighties up until recent years the application of the Kalman filter within water quality seems to have been just quietly simmering.

In the proceedings of the above mentioned conference in 1978, Prof. R. E. Kalman, referring to the Kalman filter, wrote "Fortunate is the user, for the less he knows about probability the more successful he is going to be ... ". Ironically, these words straight from the horse's mouth stand quite in contrast to what we are actually seeing today. Filtering in water quality modelling is today almost always linked to a maximum likelihood criterion and often also to a coupled or off-line optimisation routine. At the same time practical problems of starting values, unknown variances and instability seem to have become less problematic.

The following summarises some recent examples of the use of Kalman filtering within water quality modelling. Jacobsen et al. (1997) present a stream model built up of linear reservoirs in series in a continuous state-space form with the Kalman filter for up-dating. Model parameters, including the variance of both system and measurement error terms, are estimated on the basis of a maximum likelihood criterion which is optimised off-line. In three river models with mechanistic descriptions of BOD variations and algal growth, Qian (1997) also uses a modified version of the Kalman filter for up-dating the system state. He avoids the linear approximations of the extended Kalman filter by an interesting evaluation at each time step of the likelihood ratio between a number of parallel simulations, each made with different parameter values. Essentially, the advantage of the method is that it provides more sensitive parameter trajectories for system identification, but the author points further to the examination of trajectories of the likelihood ratios themselves and those of the system variance, which is also assumed to vary. Carstensen et al. (1996) demonstrate how Kalman filtering with simple mechanistic model structure can be used as software sensors to estimate the values of unobserved system states.

In an effort to bridge the gap between the traditional deterministic and the empirical stochastic approach, Chapter 5 draws attention to the most essential distinction between dynamic deterministic and stochastic dynamic modelling: the difference in assumptions made concerning the origin of the deviation between modelled and observed values. Isolating and characterising the core difference between deterministic and stochastic modelling has disclosed related methods traditionally applied only within one of the approaches but which in principal would be of relevance and interest to both. From the field of stochastic modelling these include parameter statistics, experimental design theory, inclusion of a priori knowledge in parameter estimation and the concepts of identifiability as a whole. From deterministic model the most important aspect is the inclusion of a priori knowledge of the physical, biological and chemical processes into the model structure.

In the case study of Chapters 5 parameters of a combined sewer model are estimated under both deterministic and stochastic modelling assumptions. To this end a tool was developed with the main aim of rapidly being able to switch from one way of estimation to the other. The tool was built upon a discrete time

non-linear state space formulation. Modelling of the state's second moments and the application of the Kalman filter was active only during stochastic estimation. It was also possible to apply both the maximum likelihood and the least square estimation criteria to either of the two approaches.

That water quality modellers of the traditionally deterministic approach and those of the traditionally stochastic approach have moved closer to one another is certain. The traditional deterministic modellers have dropped ideas of incorporating all conceivable processes into their models. Modellers from the traditionally stochastic approach have recognised that if an engineer should have any use for a model, then the purely empirical formulation often must be replaced by some degree of mechanistic description relating variables via the physical, biological or chemical processes in question.

However, on one particular aspect a clear difference persists. On model calibration there are those who assume that all deviation between observed and simulated results only from observation error and then there are those who assume that this deviation also results from an incomplete system description. Unfortunately, this difference in assumptions relates not as one might logically expect to the nature of the systems, the nature of the available data, the available knowledge and understanding of the system or the available resources, but simply to the background of the modeller.

Perhaps the time has come where the use of stochastic state filtering, such as Kalman filtering, in water quality modelling will no longer be merely simmering but will take the central role which it ought to have. Deviation between "what we model" and "what we see" is a result of both modelling and observation error.

IDENTIFIABILITY IN WATER QUALITY MODELLING

There has been a tendency to persistently add new processes to water quality models in an attempt to improve their ability to simulate system behaviour. This often has two unfortunate consequences. Firstly, calibration or parameter estimation leads to non-unique sets of parameters. Secondly, one may often include processes that are less important than processes which have not been included because they are not even known.

Parsimony is a concept well developed within the field of statistics and can briefly be summarised as "care in the allotting of model parameters". Working with empirical models one would simply exclude all model structure that does not significantly improve a models fit with the available observations. In models where physical, chemical or biological interpretations are attached to the parameters exclusion of model structure may render a model unsuitable for its intended purpose.

Reichert and Omlin (1997) point out that we should be careful not to take identifiability a bit too far. There are many cases where the inclusion of unidentifiable model structure from an engineering point of view is indispensable. Consider a sewer system designed to run with full pipes every five years. Even a very long monitoring period is most likely to fail in obtaining data which will permit the identification of model structure related to processes which occur only during full pipe flow. One's engineering knowledge may however strongly suggest that the sewer behaves quite differently in its full state; one would perhaps be better off extrapolating with unidentified model structure than charging on with something identified under quite different conditions. The example in Reichert and Omlin (1997) relates to the indentification of Monod growth kinetics, but many other water quality examples could be given.

When accused of using over-parameterised models which could be calibrated to fit almost any signal, the water quality engineer would often answer something in the lines of "yes, but my engineering intuition and experience tells me what the realistic parameter values are and the calibration is only a matter of making the case-specific adjustments". This does of course not quite correspond to our formal concepts of identifiability, yet there is some truth about it. Reichert (1997) outlines how such diffuse a priori knowledge on our parameter values can and should be handled in a structured manner. He emphasises the fact that defining probability distributions for our knowledge is not enough. Our prior knowledge is diffuse and must therefore be formulated as imprecise probabilities. Five methods of characterising imprecision in probability distributions are outlined and discussed, and with a didactic example the author points to the advantages of a Bayesian approach in incorporating our prior knowledge into the parameter estimation phases of modelling. The Bayesian approach relies on conditional probabilities expressions in characterising the uncertainties of variables that are causally dependent on the value other uncertain variables or parameters. The near future will undoubtedly present some interesting applied examples from various fields of water quality. This point is not without bearing on discussion above on the usefulness of including unidentifiable model structure and will undoubtedly change the practical implications of systems identification.

In Chapter 5 of the present thesis *a priori* parameter knowledge was quantitatively included by using a Bayesian approximation to reformulate the maximum likelihood criterion into a maximum *a posteriori* likelihood as outlined in Madsen and Holst (1996). The possibility of quantitatively including a priori parameter information gives rise to a new situation with respect to the identifiability of model structure with physical, chemical or biological interpretation attached to the model parameters. The maximum *a posteriori* criterion can be used with both deterministic and stochastic parameter estimation. More structure will be acceptable once we include our knowledge and experience concerning likely parameter values in the form of a priori distribution. Though a comprehensive investigation of the advantages, interpretation and possible pitfalls of such inclusion of a priori knowledge lies beyond the scope of this thesis, the results of Chapter 5 did confirm the expectations.

The combined sewer water quality modelling of the preceding chapters has shown that aspects of experimental design theory have to be considered in the planning phase of a monitoring campaign aimed at system identification and parameter estimation or calibration. The principal input to a combined sewer is rainfall, which cannot be designed as such. However, experimental design theory can assist in identifying the best sampling points within the sewer network and indicate how long before and after an overflow it is necessary to extract samples from the water. The suspended COD models of Chapter 6 and Chapter 7 would have been better identified and the resulting parameter sets less ambiguous had water quality samples been taken before the critical flow at which resuspension started had been passed and if further samples had been taken after high flows had subsided. For water quality models used for studying the response of surface waters to combined sewer overflows controlled overflows may be designed with a given dynamic characteristics in order best to identify the model parameters.

Other aspects of experimental design with bearing upon subsequent model identification are sampling frequency and regularity. They become particularly relevant in relation to the identification of modelling error terms in dynamic systems during stochastic modelling. This relates to complexities such as Shannon's sampling theory, which, in simple terms, states that you will have difficulties observing Brownian motion of a pollen grain on a water surface if you only open your eyes every second minute. In contrast, the identifiability of deterministic structure (empirical or mechanistic) relates more directly to the amount of available data and the persistence of excitation.

CONCLUSIONS

The overall objective of this part of the thesis was to examine modelling approaches and methodologies in an attempt to improve combined sewer water quality modelling. The following have been identified as methods whose implementation would contribute to such improvement.

- quantitative inclusion of a priori knowledge in automated parameter estimation,
- application of parameter statistics and experimental design theory,
- random coefficient modelling,
- stochastic modelling with physical, chemical and biological model structure.

Incorporation of the first two points into traditional deterministic combined sewer modelling is a fairly straightforward task. The maximum likelihood criterion should be applied to the existing deterministic modelling approach and a rapid but robust optimisation routine should be adopted. The main hindrance to immediate application of parameter statistics and experimental design criteria is probably the parameter transformations which would be required in order to be able to make assumptions of local linearity in the often non-linear combined sewer water quality models.

Modellers should recognise the fact that explanatory variables are not always available for characterising the state of a sewer system at the start of a rainfall event. No explanatory variables may be known or no observations may have been made of their values. In recognising this, random coefficient modelling should be applied so that parameters, such as the strength of sediment crusts and pore water concentrations, may be considered as random variables varying from event to event. Random coefficient modelling requires more computational resources and would today perhaps only be realistically applied to simpler models with short simulation times and few estimated parameters.

Finally, stochastic modelling would bring a major improvement to combined sewer water quality modelling as it would avoid the bias on parameter estimates which is introduced due the poor agreement present in deterministic modelling between the implicit assumptions on the errors and the actual situation. In certain situations a change from a deterministic to a stochastic formulation requires little effort but in other situations simplifications have to be made to the model in order to make it appropriate for stochastic modelling. A stochastic state space formulation of a hydrodynamic sewer model would involve state variables for both flow velocity and water depth at each calculation point in the system and is therefore not likely to be just around the corner. Where exactly the limit lies with respect to the number and remoteness of unobserved system states is one of the questions concerning the applicability of stochastic modelling that have yet to be answered. Generally speaking, the widespread use of stochastic modelling in water quality engineering would require professionally developed tools specifically directed towards stochastic modelling for water quality management.

The modelling work in this part of the thesis has revealed that water quality models have a number of common characteristics for which generalisations in terms of statistical and stochastic approaches ought be sought. Questions that should be addressed are:

- Which parameter transformation would be appropriate for which types of parameters (e.g. runoff coefficient, growth rates, settling rates, resuspension parameters, storage constants, yields and available deposits)?
- Which system variables would be suited as state variables under various circumstances?
- What would be appropriate state transformations for application with Kalman filtering?
- How should the system error be structured for various processes?
- What is a realistic number and remoteness of unobserved system states in stochastic models?
- Where do the critical parameter interchangeability problems occur and which measures can be taken to overcome them (re-parameterisation and improved monitoring)?

The results of such generalisation studies would be an asset not only to combined sewer water quality modelling but to water quality modelling and engineering as a whole.

REFERENCES

Beck, M. B. and Young, P. C. (1976). Systematic identification of DO-BOD model structure, *Journal of Environmental Engineering Division*, ASCE, 102, 909-927, 1776.

Carstensen, J., Harremoës, P. and Strube, R. (1996). Software sensors based on the grey-box modelling approach. *Wat Sci. Tech.*, 33(1), pp 117-126

Chiu (Ed.)(1978) Application of Kalman Filter to Hydrology, Hydraulics and Water Resources. Proceedings of AGU Chapman Conference held at University of Pittsburgh, Pennsylvania, USA, May 22-24, 1978, Publ. by Stochastic Hydraulics Program, Dept. of Civil Engineering, Univ. of Pittsburgh, Pennsylvania, USA

Jacobsen, J. L., Madsen, H., Harremoes, P. (1997) A stochastic model for two-station hydraulics exhibiting transient impact. *Wat. Sci. Tech.*, 36(5), pp. 19-26.

Madsen, H. and Holst, J. (1998). Modelling Non-Linear and Non-Stationary Time Series. Department of Mathematical Modelling, Technical University of Denmark, Building 321/305, DK-2800 Kgs. Lyngby, Denmark. www.imm.dtu.dk

Qian, Song S. (1997). Water quality model structure identification using dynamic linear modeling: river cam case study revisited. *Wat. Sci. Tech.*, 36(5), pp. 27-35.

Reichert, P. (1997). On the necessity of using imprecise probabilities for modelling environmental systems. *Wat. Sci. Tech.*, 36(5), pp.149-156.

Reichert, P. and Omlin, M. (1997). On the usefulness of overparameterized ecological models. *Ecological Modelling*, 95(2-3), pp. 289-299.

PART III

UNCERTAINTY IN RETURN PERIOD ANALYSIS

CHAPTER 9

UNCERTAINTY ANALYSIS USING EMBEDDED MONTE CARLO SIMULATIONS

ABSTRACT

The return periods of detrimental effects are often used as design criteria in urban storm water management. Considerable uncertainty is associated with the models used. This is either ignored or pooled with the inherent event to event variation such as rainfall depth. It is here argued that uncertainty and inherent event to event variation should be treated separately, in providing engineers and managers with the distributions of return periods. It is then possible to base management decisions on knowledge of both the expected return periods and their corresponding confidence limits. It is further argued that the traditional pooling of inherent variation and uncertainty leads to meaningless return period curves with no engineering value.

The presented methodology is described as Embedded Error Propagation and its current implementation as Embedded Monte Carlo Simulations. This new approach is demonstrated in an integrated setting involving models for rainfall characteristics, combined sewer overflow loads and impacts on the surface water dissolved oxygen (DO). CSO loads are modelled using event lumped non-linear regression models with rainfall as input and with overflow volume, duration and relevant event mean concentrations as output. Oxygen depletion in the surface water is described using a dynamic model including oxidation of dissolved chemical oxygen demand (COD) and nitrification. Conversion models had to be developed to integrate the output variables of the CSO model with the input variables of the surface water model. The parameters of all the models were estimated from observed data on rainfall, CSO load and surface water impacts. The data are obtained from a combined sewer system discharging to a pond in Loenen, the Netherlands. In this chapter focus is on chemical effects of CSO on surface water. The argued distinction between event to event variation and uncertainty and the associated methodology are equally valid to the return period analysis of flooding.

This chapter is based on Grum, M. and Aalderink, R. H. (1999). Uncertainty in Return Period Analysis of Combined Sewer Overflow Effects Using Embedded Monte Carlo Simulations. *Wat. Sci. Tech.* 39 (4), pp. 233-240. Published after oral presentation at the AQUATECH conference in Amsterdam, September 1998.

INTRODUCTION

The return periods of combined sewer overflow (CSO) effects are often used as design criteria in urban storm water management. There are two main reasons why models are used to calculate return periods of CSO effects. Firstly, because calculation of for example events with a mean time between occurrence of one year by direct measurement would require observation periods of several time longer than just one year. Attempts are then made to use models to predict the return periods of CSO impacts as a function of a relatively well observed explanatory variables namely one or more point measurements of rainfall. The poorer the explanatory value of the rainfall the larger will be the uncertainty introduced as a result of short monitoring periods of the combined sewer itself, its overflows and the surface water effects. The second main reason for using models is that their mechanistic structure permits the analysis of proposed engineering scenarios. The use of models in return period analysis is treated in more detail in the following chapter.

Taking a reductionist approach to uncertainty in return period analysis it could be split up into a multitude of sources. However, essentially uncertainty can be considered to result from two main aspects; a limited number of observations and, on extrapolation, uncertainty in model structure. The present study takes a parametric approach to evaluate uncertainty caused by a limited number of observations of a process. Models used in the return period analysis have their parameters estimated from observed data. Parameter uncertainty, including that of parameters describing inputs and the variances of random errors, are estimated and considered in the analysis.

Uncertainties in input characterisation and system description provide information on the certainty of a given expected return period value. They do not change it. It is therefore argued that variability and uncertainty should be handled separately using an embedded error propagation method as the one outlined in the following section. Similar proposals and methodologies distinguishing between uncertainty and variability have been made in risk and frequency analysis in other fields (e.g. Brattin et al. 1996, Hession et al. 1996).

METHODOLOGY

Error Propagation and Return Period

In a many-to-one input-output system, the cumulative probability distribution function of the output variable can be obtained by the integration of the joint probability density function over the region of the input space for which the output is greater than or equal to the evaluated value of the output. This integration can sometimes be done analytically but is in most practical applications solved numerically. In this study the integration has been solved using Monte Carlo simulations which involves the creation of a large number of input data sets by pseudo-random sampling from the input variables' parametric probability distributions. For a set of input variables an output variable is calculated using the model(s).

Return period, $T(y)$, of the exceedance of a variable y, is found from the exceedance probability using (9.1).

$$T(y) = \frac{1}{n \cdot p_y} = \frac{1}{n \cdot \left(1 - \left(r_y - 1\right)/\mathrm{N}\right)} \tag{9.1}$$

where p_y is its probability of exceedance, n is the mean number of events per year, r_y is the rank of y after sorting the output variable from the N samplings and following Monte Carlo simulations. In this study the number of runs used to generate a single return period curve was 34200 which corresponds to a coefficient of variation (due solely to imprecision in the numerical integration) of 0.1 for events with a return period of 5 years. Portielje et al. (2000) give a detailed treatment in the context of water quality management of how to reduce the numerical error or the required number of model simulations to attain a desired level of accuracy. In the two evaluated methods a higher sampling density in the region of interest is attained by combining the First Order Reliability Method with Latin Hypercube Sampling and Directional simulation respectively. At the cost of more model simulations, the Monte Carlo procedure was chosen for the current study due alone to its simplicity in description and implementation.

Uncertainty and Inherent Variation

All quantities, which are described by a probability distribution, are placed in either of the two layers:

- an inner layer consisting of quantities varying from event to event,
- an outer layer consisting of uncertain, but constant quantities.

For each set of random realisations of the quantities in the outer uncertainty layer a full set of Monte Carlo simulations for the inner inherent variations layer is performed. Each set of realisations in the outer uncertainty layer thus results in a single return period curve. This could for example be a curve of minimum oxygen concentrations against return period (i.e. mean time between occurrence). Multiple samplings in the outer layer thus results in a band of return period curves representing the uncertainty distribution of the return periods. Having generated a sufficiently large number of return period curves, their spread may be summarised into curves giving the confidence limits of the return periods. A schematic illustration of the sampling procedure is given in Figure 9.2 case C.

On making the distinction between uncertainty and inherent variation in practice the following descriptions may be used. If the variation of a given quantity ...

- results from changes that actually take place then we speak of inherent variation and this belongs in the inner sampling layer,
- results from our lack of knowledge then we speak of uncertainty and this belongs in the outer sampling layer.

Intensifying a measurement campaign improves our knowledge (and thereby reduces uncertainty) but it does not alter the variations that actually take place (the inherent variations). In a given practical situation it may be difficult to make this hard-cut distinction and in such cases it may even be necessary to evaluate the sensitivity of the final results to the decision in question.

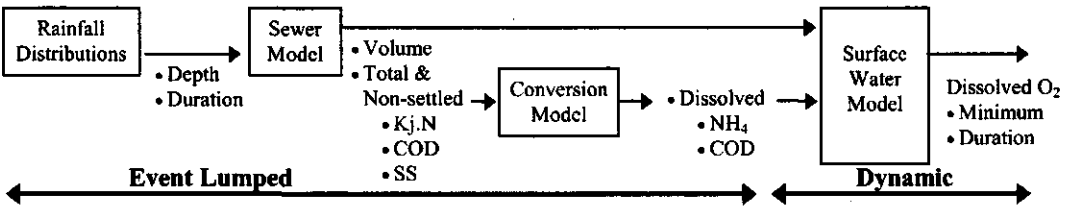


Figure 9.1 Schematic overview of the integrated model components.

In this study 34200 realisations and model simulations were performed in the inner inherent variations layer for each set of realisations in the outer layer. With 72 realisation of the quantities in the outer layer, a total of over 2.5 million model simulations were performed.

Estimating the Uncertainty

The uncertainties in model and distribution parameters are often not readily estimated. Here uncertainty in the form of covariance matrices were estimated using the non-parametric Jack-knife method. The essence of this method, which is well described in several statistical text books such as Stuart and Ord (1994), is that by repeated exclusion of each observation and re-estimating the parameters of the distribution an estimate of their variance and covariance can be found. How this is implemented for a given model depends very much on the underlying assumptions and is therefore presented with the individual models in the following section.

THE INTEGRATED SYSTEM

The studied integrated system comprises a combined sewer and a pond. The components of the integrated model include rainfall distributions, a combined sewer overflow model, a conversion model for water quality parameters and a surface water model. This section briefly outlines each of the components/sub-models which are schematically presented in Figure 9.1.

Rainfall, sewer system and conversion models are all event lumped models (i.e. as total rainfall depth, event mean concentration, etc.) whereas the surface water model is a dynamic model. The time scales of interest in the first 3 models are typically in the order of hours, whereas receiving water impacts extend over periods of days to weeks and errors in linking event lumped with dynamic are therefore likely to be small.

Rainfall Distributions

Duration and mean intensity of a rainfall event have been described by a joint lognormal distribution. The two means and the three elements of the covariance matrix were estimated from a point rainfall series of 12 years excluding rainfall events of less than 3 mm. The time between rainfall events was described by an exponential distribution. These rainfall distributions describe variations from event to event and belong therefore in the inner variability layer of the embedded Monte Carlo.

A six by six covariance matrix describing the uncertainty in the parameters of the rainfall distribution was estimated using the Jack-knife technique. The parameters of the above distributions were assumed to be

normally distributed. The joint distribution of these six parameters describes parameter uncertainty and not variation and therefore belongs in the outer uncertainty layer of the Embedded Monte Carlo procedure.

Sewer Models

Quantity and quality of the Combined Sewer Overflow CSO have been described using event lumped models with rainfall characteristics as input. These models are briefly outlined below but a fuller description of the data and models can be found in Chapter 4 and Grum and Aalderink (1998).

Overflow volume, V_{OVERFLOW}, was modelled using (9.2) where A is the catchment area, H_{RAIN} and D_{RAIN} are

$$V_{\text{EFF}} = A \cdot \alpha \cdot H_{\text{RAIN}} - Q_{\text{PUMP}} \cdot D_{\text{RAIN}} - V_{\text{STORE}} + \varepsilon_V \qquad \text{and} \qquad V_{\text{OVERFLOW}} = \begin{cases} V_{\text{EFF}} & \text{for } V_{\text{EFF}} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{9.2}$$

rainfall depth and duration respectively, Q_{PUMP} is pump capacity, V_{STORE} is static storage volume and ε_V is a random error term with zero mean. The runoff coefficient, α, was found to vary from event to event based on a wetness index which in turn was dependent on a varying drying rate. Parameters in the model including those of the wetness index and those of the drying rate were estimated on the basis of rainfall events of over 3mm during a period of just over four years including 56 events leading to overflow.

The combined sewer event mean concentrations were modelled as shown in (9.3) as a function of rainfall intensity with i = 1,2,3 for COD, Kjeldahl Nitrogen and suspended solids respectively.

$$C_i = \exp\left(a_i + b_i \cdot I_{\text{RAIN}} + \varepsilon_i\right) \tag{9.3}$$

where I_{RAIN} is the mean rainfall intensity, a_i and b_i are constant coefficients and ε_i are random error terms with zero mean. Event mean slow-settling fractions had been calculated on the basis of concentrations before and after 1 hours settling of all sample. These event mean slow-settling fractions, f_i, were modelled as shown in (9.4) and (9.5) with i = 4,5,6 for COD, Kjeldahl Nitrogen and suspended solids respectively.

$$\ln\frac{1}{f_i} = a_i + b_i \cdot \ln\left(C_{i-3}\right) + \varepsilon_i \tag{9.4}$$

$$f_i = 1 \Big/ \exp\left(a_i + b_i \cdot \ln\left(C_{i-3}\right) + \varepsilon_i\right) \tag{9.5}$$

The parameters of all six equations, i = 1, ..., 6, were estimated simultaneously using a maximum likelihood criterion in which the joint probability distribution of the six error terms was also estimated. The non-settling fraction of Kjeldahl nitrogen was found not to be related to its total event mean concentration and the corresponding coefficient b_3 was set to zero.

In the above models the error terms vary from event to event and consequently describe inherent variation. The model parameters were assumed to have a constant value during all events and do therefore not contribute to inherent variation. They and the variances of the error terms are however uncertain quantities whose joint covariance matrix was estimated using the earlier outlined Jack-knife technique. This covariance matrix describes uncertainty and belongs in the outer layer of the Embedded Monte Carlo procedure.

Conversion Models

The CSO models predict Kjeldahl nitrogen, COD concentrations and suspended solids before and after one hours settling. Conversion of these to ammonium and dissolved COD, required for the surface water model, has been based on three presumed relationships.

Firstly, it has been assumed that Kjeldahl nitrogen, C_{KJ-N}, comprised only particulate organic nitrogen, X_{ORG-N}, and the completely soluble ammonia, S_{NH_4-N}.

$$C_{KJ-N} = X_{ORG-N} + S_{NH_4-N} \tag{9.6}$$

Secondly, the suspended solids concentration was assumed to be proportional to that of organic nitrogen. Thus the ratio of the concentration after settling, *slow*, to that before settling, *total*, is the same for the unobserved organic nitrogen as for the observed suspended solids. Note that subscript *total* is here used to symbolise the sum of *fast* and *slow* settling particulate concentrations.

$$X_{ORG-N,\,slow} / X_{ORG-N,\,total} = X_{SS,\,slow} / X_{SS,\,total} \tag{9.7}$$

Thirdly, the following linear relationship was assumed to exist between the slow settling fraction of Kjeldahl nitrogen and that of suspended solids.

$$C_{KJ-N,\,slow} / C_{KJ-N,\,total} = a \cdot X_{SS,\,slow} / X_{SS,\,total} + b + \varepsilon \tag{9.8}$$

Combining the three relationships, (9.6), (9.7) and (9.8), results in the following expression for the conversion of the available CSO concentrations to ammonia. An equivalent expression was used for dissolved COD.

$$S_{NH_4-N} = \frac{\left(X_{ORG-N,\,slow} + C_{KJ-N,\,total} - C_{KJ-N,\,slow}\right) \cdot \left(a \cdot X_{SS,\,slow} + (b+\varepsilon) \cdot X_{SS,\,total}\right) - X_{ORG-N,\,slow} \cdot X_{SS,\,total}}{\left(1-(b+\varepsilon)\right) \cdot X_{SS,\,total} - a \cdot X_{SS,\,slow}} \tag{9.9}$$

where the slow settling organic nitrogen is given by

$$X_{ORG-N,\,slow} = \frac{\left(C_{KJ-N,\,total} - C_{KJ-N,\,slow}\right) \cdot X_{SS,\,slow}}{X_{SS,\,total} - X_{SS,\,slow}} \tag{9.10}$$

The values of the two constants, a and b, and the error variance, σ_ε^2, were estimated by standard linear regression on (9.8) using the sample data (i.e. not on event mean concentrations). During an event flow proportional sampling took place at the overflow structure. Their respective variances were estimated using the earlier outlined Jack-knife technique and the results are presented in Table 9.1. To avoid the amplification of measurement errors which occurs when taking ratios of very small values, samples with suspended solids concentrations below 130 mg/l were excluded from the regression. This was found to be the cut-off value after which the regression coefficients and coefficient of determination (ca. 0.6) stopped changing.

Table 9.2. Ordinary differential equations defining the processes included in the surface water model.

Dissolved Oxygen:

	Reaeration	CSO-COD degradation	Background-COD degradation	Nitrification ammonia	Sediment oxygen demand	
$\frac{dS_{DO}}{dt} =$	$k_L \cdot (S_{DO,SAT} - S_{DO})$	$- r_{OX} \cdot S_{COD}$	$- r_{OX,B} \cdot S_{COD,B}$	$- r_{NIT} \cdot S_{NH_4\text{-}N}$	$- \frac{r_{SOD}}{H} \cdot \sqrt{S_{DO}}$	(9.11)

Dissolved COD of combined sewer origin:

$$\frac{dS_{COD}}{dt} = - r_{OX} \cdot S_{COD} \qquad (9.12)$$

Dissolved COD from background load:

$$\frac{dS_{COD,B}}{dt} = - r_{OX,B} \cdot S_{COD,B} + B_{V,COD,B} \qquad (9.13)$$

Ammonia from background load:

$$\frac{dS_{NH_4\text{-}N}}{dt} = - r_{NIT} \cdot S_{NH_4\text{-}N} + B_{V,NH_4\text{-}N,B} \qquad (9.14)$$

Table 9.1. Estimates of the constants and error variance in the straight line regressions defined by (9.8).

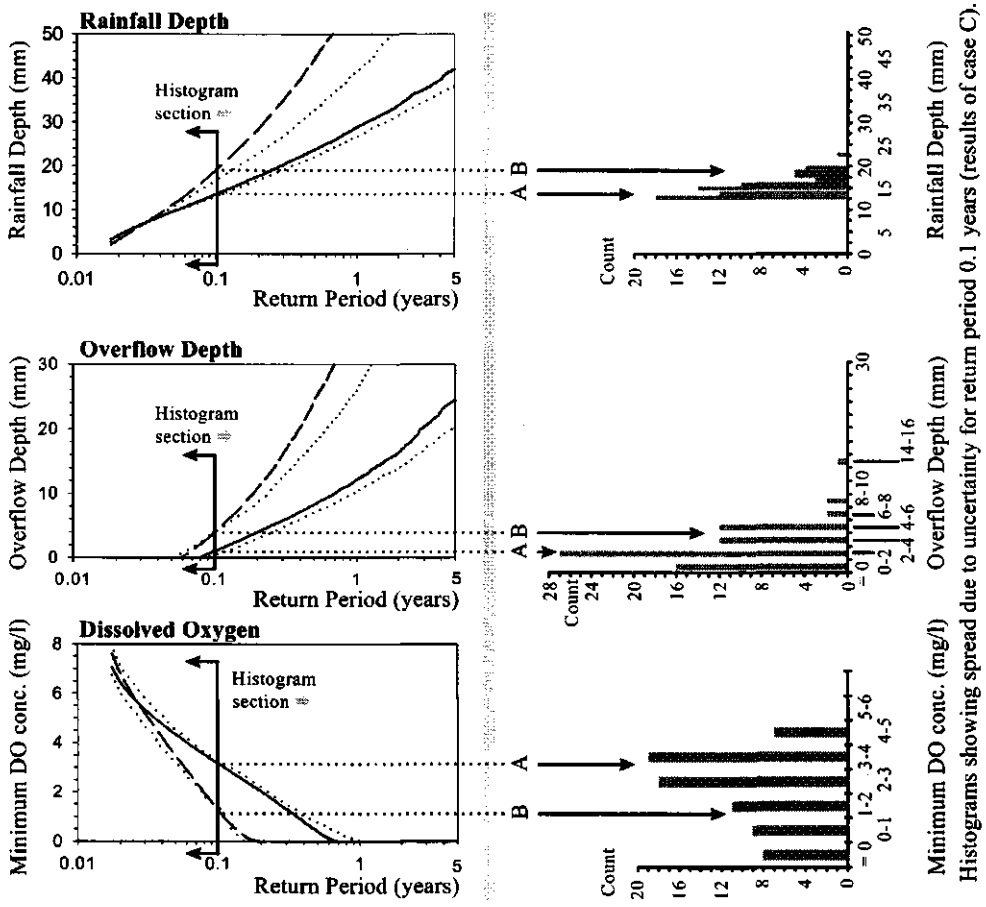
	Conversion to Ammonium		Conversion to Dissolved COD	
Parameters	Mean	Variance	Mean	Variance
Constant a	0.772	7.97e-3	0.752	2.26e-3
Constant b	0.279	1.67e-3	0.235	5.42e-4
Variance σ_ε^2	1.15e-2	9.13e-6	1.88e-2	4.39e-6

Most of the uncertainty introduce through these conversions presumably lies in the constants and error term of the relationship (9.8) which have therefore been included in the uncertainty layer of the Embedded Monte Carlo simulations. It should however be noted that the uncertainty introduced by the assumptions behind relationships (9.6) and (9.7) have had to be ignored because no reasonable estimates of their magnitude could be found given the available data. It is however expected that these uncertainties are much smaller than those of (9.8) and than the rest of the uncertainty in this return period analysis.

Surface Water Model

The surface water oxygen dynamics was defined by the four differential equations (9.11), (9.12), (9.13) and (9.14). Dissolved oxygen is lost on COD oxidation, ammonia nitrification and to a sediment oxygen demand and is added to the pond by re-aeration. Note that CSO-COD and background COD are treated separately with oxidation rates r_{OX} and $r_{OX,B}$ respectively. The background COD represents the organic matter brought to the pond by the infiltration flow into the pond.

S indicates dissolved concentrations, B_V, volumetric loading rates, r reaction rates and k_L is the reaeration coefficient. The additional subscript B indicates a background component.



Case	A	B	C
Line(s)	———	– – – – –	.....
Description	Ignoring uncertainty	Pooling of uncertainty and event to event variation	50% confidence limits by distinguishing between uncertainty and variability
Sampling Scheme of the Monte Carlo Integration: Outer layer §	None	None	
Inner layer	Fixed to expected values		
	... to model simulation	... to model simulation	... to model simulation
Distributions	* Distributions describing inherent variations from event to event, ** Distributions describing uncertainty of constant parameters of the inherent variation distributions, *** Distributions describing uncertainty of constant model parameters.		

Figure 9.2. Results of the analysis. § For each set of random realisations of variables in the outer layer a complete Monte Carlo integration (of 34200 simulations) was done for the variables in the inner layer.

Though dynamically deterministic, the model was calibrated as a random coefficient model whose parameters are characterised by joint probability distributions describing inherent variation from event to event. The parameters were estimated using observed data of six overflow events, an expanded model and the downhill simplex optimisation procedure on a maximum likelihood criteria. Aalderink et. al. (1998) present a more complete outline of the data and the surface water model.

The parameter set for a given overflow event was assumed to be the realisation of a random variable characterised by a selected joint probability distribution. Of the tested distributions, normal, lognormal, exponential, weibull and gamma, the weibull distribution had the highest likelihood and was consequently used. The parameters of the distributions (including covariance) were calculated on the basis of the six sets of model parameters obtained from each of the six overflow events. The uncertainty in these inherent variability distributions was in turn characterised by a joint probability distribution whose parameters were estimated using the earlier outlined Jack-knife technique.

The observed surface water data originated from a pond with unusually high background loading of COD and ammonia, due to an atypically large infiltration of oxygen-free ground water. As a result the CSOs had no visible impact on the extreme statistics of DO in the surface water. A hypothetical surface water was defined differing from the existing pond by scaling the background loading and by excluding the oxygen-free infiltration into the pond. Mixing of the CSO in the surface water pond was assumed to lie at random between the two extreme cases of plug-flow and completely mixed reactor. Test runs using each of the two extreme mixing cases suggest that the thus non-described uncertainty has little effect on the results.

RESULTS AND DISCUSSION

Results for rainfall depth, overflow depth and minimum dissolved oxygen are presented in Figure 9.2. Each of the three graphs on the left contains curves corresponding to the following three cases:

A. *Ignoring Uncertainty (solid line).* All uncertain quantities were fixed to their expected value and random sampling was done only from the distributions of varying quantities.

B. *Pooling Uncertainty with Variability (dashed line).* Distributions of uncertain quantities were sampled along with those of varying quantities resulting also in a single return period curve.

C. *50% Confidence Limits (two dotted lines).* For each of the 72 sets of randomly sampled parameters in the outer uncertainty layer a complete return period curve was generated by 34200 Monte Carlo simulations. The spread of these 72 lines has been summaries to the presented 50% confidence limits. Reasonable plots of the usual 90% or 95% confidence limits would have required either roughly another 5 million model simulations or further assumptions concerning the parametric distribution types.

On the right in the figure are histograms showing the spread due to uncertainty for a return period of 0.1 years. The spread is here represented through the 72 return period curves resulting from case C.

From the 50% confidence limits plotted in Figure 9.2 it is clear that the return period analysis of combined sewer overflow effects are associated with a substantial amount of uncertainty. From the figure it can be seen that with only 50% confidence (equal chance of being right and wrong) we are able to predict that the DO concentration will fall below 1.0 mg/l between 2.5 and 10 times a year. Viewed in terms of a given return

period the dissolved oxygen concentration is in the histogram (lower right Figure 9.2) seen 10 times a year to fall below a concentration of anything between 1 and 3.5 mg/l (50% confidence limits).

In agreement with the skewness of the return period curve distributions, which is apparent from the histograms in Figure 9.2, the curves resulting from case A lies just above the lower 50% confidence limit for rainfall and overflow depths, and just below the upper 50% confidence limit for dissolved oxygen. According to case A a DO concentration of 1 mg/l is expected to occur just over 3 times a year whereas pooling uncertainty with variability (case B) suggests that this would happen about 9 times a year. For DO the results of case B are seen to lie close to the lower 50% confidence interval. For rainfall pooling (case B) is seen to result in the 50 mm rainfall event occurring twice a year.

Pooling uncertainty with inherent variation (case B) leads to the awkward situation that reduced uncertainty through for example increased monitoring *systematically* leads to all extremes becoming less frequent. Increased uncertainty through for example data "loss" would similarly *systematically* increase the frequency of extreme events.

Ranking of sources contributing to uncertainty has not been treated. However, also here it is important that uncertainty and variability are kept apart. Thus, the ranking of sources could be based on estimates of the linear regression coefficients of the return period of a given dependent variables level (e.g. of 2 mg/l DO) with appropriately scaled values of the uncertain quantities as independent variables. Rainfall depth would thus not be an independent variable but its two distribution parameters would.

Distributions of variables describing the initial state of the surface water were estimated on the basis of only 6 events resulting in very uncertain distribution parameters. Uncertainty in these distribution parameters could be reduced through monitoring of the surface water's response to a greater number of overflow events. Reduced uncertainty could also be achieved by shifting the influence of various uncertainty sources. For example, creating a better CSO event mean concentration model would shift influence away from the poorly characterised random error of the sewer system model to the relatively more certain rainfall characterisation.

As mentioned, parameters of distributions describing surface water variability were estimated on the basis of six observations. The resulting return period curve (irrespective of the incorporated parameter uncertainty) proved to be very sensitive to the chosen distribution type. This would suggest that the results of both return period analysis and uncertainty analysis rest rather heavily on these presumed parametric probability distribution types. In other parts of the system much more data was available for choosing the type of probability distribution and, though not tested, is there expected to be less of a problem.

In order to limit the complexity of the integrated system description the seasonality has not been taken into account. By thus not including correlation between rainfall and certain surface water variables it is likely that the frequencies of extremes has been slightly overestimated. This comes from the fact that neglecting co-variation between variables or parameters is equivalent to including the some variation more than once.

CONCLUSION

In this chapter it has been argued that a distinction ought to be made between uncertainty and inherent variability in return period analysis. It has been shown how this distinction can be made using Embedded Monte Carlo simulations. This was done in an integrated approach to evaluating the effects of combine sewer overflow on a surface water pond. The substantial uncertainty is illustrated by results of the dissolved oxygen concentration: the mean 10 times per year lowest concentration was determined to lie between 1 and 3.5 mg/l with only 50% confidence. Expressed as an uncertainty on the return period for a given dissolved oxygen concentration: the return period of a concentration of 2.0 mg/l lies between 0.08 years (i.e. 12.5 times a year) and 0.2 years (i.e. 5 times a year). Note again that these are only the 50% confidence limits.

Whilst distinguishing between uncertainty and variability the contributions of individual sources should be studied in order to optimise efforts to reduce uncertainty of the end result. Both a shift in the influence of the various uncertainty sources and a reduction in the sources themselves could result from, for example, dynamic descriptions of both rainfall and the sewer system.

REFERENCES

Aalderink, R. H. et al. (1998). Modelling impacts of combined sewer overflows on receiving water dissolved oxygen. In preparation.

Brattin, W.J., Barry, T.M. and Chui, N. (1996). Monte Carlo modeling with uncertain probability density functions. *Human and Ecological Risk Assessment*, 2 (4), 820-840.

Grum, M. and Aalderink, R.H. (1998). Event lumped modelling of combined sewer overflow pollution. In preparation.

Hession, W.C., Storm, D.E., Haan, C.T., Burks, S.L. and Matlock, M.D. (1996). A watershed-level ecological risk assessment methodology. *Wat. Res. Bulletin*, 32 (5), 1039-1054.

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*, Vol.34(1), pp. 153-170.

Stuart, A. and Ord, J. K. (1994). Kendall's Advanced Theory of Statistics. Vol. 1. Distribution Theory. Edward Arnold.

CHAPTER 10

DISCUSSION ON RETURN PERIOD ANALYSIS AND ITS UNCERTAINTY

ABSTRACT

Combined sewers overflow during extreme rainfall causes pollution and flooding in the surrounding surface waters. Over the past decades the measure of severity for the discharge of oxygen depleting organic matter has shifted from average annual loads and number of discharges, through return periods for discharged loads, to return periods of surface water effects.

The objective of this chapter is to put together a framework encompassing the many methods, approaches and types of models that can be used to calculate return periods and to discuss examples from literature in the context of the presented framework. The subsequent discussion on uncertainty in return period analysis has the additional objective of placing the methodology and results of the previous chapter in a broader perspective.

INTRODUCTION

Discharge from combined sewer overflows during rainfall constitutes a limiting factor to the amelioration of surrounding surface waters. The high concentrations of organic matter can cause oxygen depletion in the surface water with consequences on the aquatic ecology such as fish death. Depending on local conditions sewer overflows may also constitute a major source of nutrients leading to eutrification problems in the receiving waters and/or to the surface waters further downstream of the overflow structure.

Runoff pollutants should be considered in relation to the time scale of the surface water's response to the pollutant effects (Aalderink and Lijklema, 1985, Harremoës, 1989). The division most often made is between acute and accumulative pollutants. A more thorough classification of runoff pollutants is presented in Lijklema et al. (1993a). Accumulative pollutants such as the principle nutrients phosphorous and nitrogen should be evaluated in terms of their average annual loads whereas acute pollutants such as degradable organic matter should be evaluated on the basis of the return periods of their effects. This chapter deals primarily with methods and approaches to return period analysis of the effects of acute pollutants such as those resulting in oxygen depletion or acute ammonia poisoning.

Extreme statistics may be performed directly on the depth or intensity of events of a rainfall series and then events representing selected return periods can be used to calculate corresponding effects in the surface waters. This approach is illustrated in Figure 10.1 (left). The extremeness of a rainfall event is generally based on characteristics such as total volume, mean or maximum intensity. A first draw back of this approach is that the extremeness of a rainfall event, based on total volume, mean or maximum intensity, does

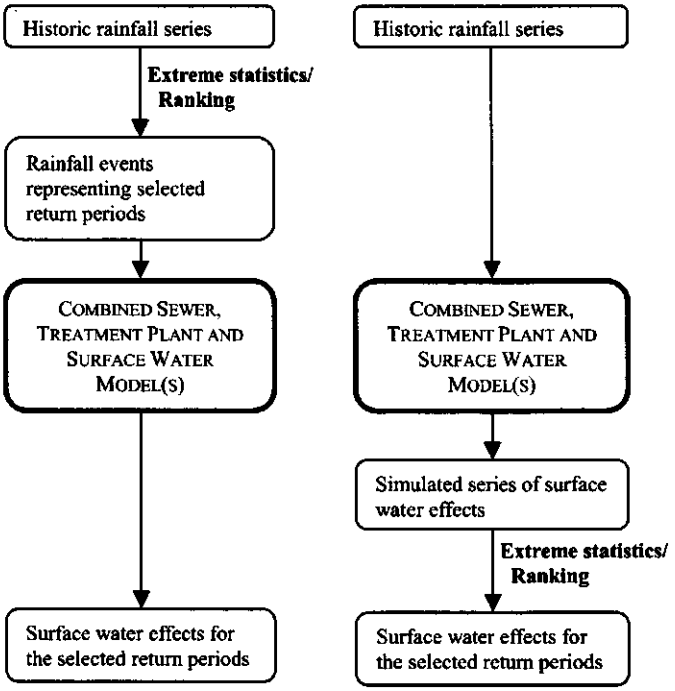


Figure 10.1. Extreme statistic may be performed either before or after simulation of the integrated urban water system.

generally not correspond to the extremeness of the resulting effects. In other words the return period of the minimum oxygen concentration resulting from a given rainfall event does generally not correspond to the return period of the rainfall event. A second draw back of the approach is that the definition of an event has to be related to the effect and not to the rainfall. That is, the decision "now the event has ended" ought to be based on the studied effect supplemented by other information such as the degree of filling of the sewer storage capacities. In the integrated system the time constants of the treatment plant secondary clarifier and the surface water are often such that a definition of an event based on a maximum amount of rainfall in a given period would be inadequate and may often result in an underestimation of effects. The effects would often be underestimated because the coupling effect of adjacent events would be ignored in such an approach.

The alternative approach is to perform the extreme statistics on series of the surface water effects. This approach is illustrated in Figure 10.1 (right). These series could have been generated from simulations with historic rainfall series, rainfall models or in theory even by direct measurement. This chapter deals exclusively with methods and approaches where the extreme statistics are performed on the series of surface water effects.

The first section of the chapter presents a framework of methods and approaches to calculation of the return periods of given surface water effects. The approach applied in Chapter 9 and selected approaches from literature are discussed within the presented framework. The chapter's second section deals with uncertainty in return period analysis. This includes a more general discussion of the distinction between inherent variation and uncertainty (as applied earlier in Chapter 9), a discussion on the major sources of uncertainty and how uncertainty should be handled on scenario comparison.

Terminology

A distinction will be made here between the terms "method" and "approach". The term "method" will here refer to aspects of the return period analysis detached from the environmental engineering problems to which it may be applied. In contrast the term "approach" will refer to the way in which the different methods may be applied to problems relating to the effects of combined sewer overflow.

The term "model" will here be used to describe a collection of mathematical expressions aimed at imitating physical, biological or chemical systems. Most often reference will be to models that transform one or more inputs to one or more outputs as illustrated in Figure 10.2. The system could for example be an urban drainage system or a surface water system. Examples of models would thus be input-output black box

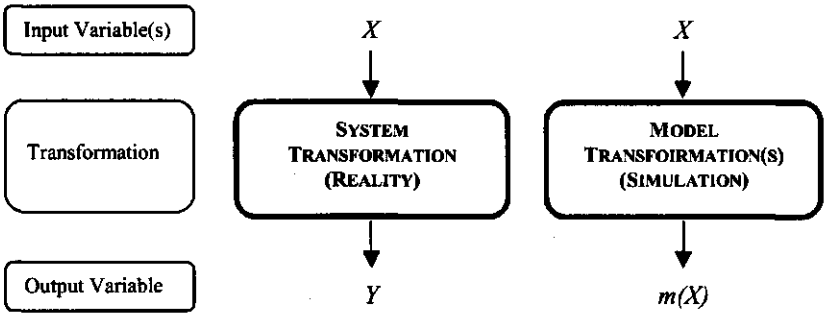


Figure 10.2 Distinction between transformation by reality and transformation by model simulation(s).

models, regression models, a model defined by the mixing equation or a complex hydraulic pipe flow model based on the Saint Venant equation. An input variable is not a prerequisite for a model. The probability distributions of rainfall characteristics and stochastic dynamic rainfall models are examples of models that have no tangible input.

Models can be dynamic, accounting for a system's changes with respect to time, in which case they will be described as *time based* models. Alternatively, models may consider a complete rainfall or pollution event as a whole, in which case the term *event lumped* models will be used. An example of an event lumped model is one that simulates the total overflow volume as a function of variables such as depth and duration of the rainfall event. Chapter 4 of this thesis looked in detail at event lumped modelling of a combined sewer system.

RETURN PERIOD ANALYSIS

The return period of a given extreme event can be found as the reciprocal of the product of the probability of exceeding the given extreme value and the mean number of occurrences per year. This has been outlined in Chapter 9. However, there are many different methods and approaches that can be used to obtain the required probability.

Methods

Three different methods of obtaining the probability distribution of system or model outputs can be identified. These have been depicted in Figure 10.3. The last two methods make use of an explanatory variable, X, and a mathematical description of the system, namely a model. The model that is used in any one of these two methods could be an input-output black box model, a regression model or a deterministic surface and pipe flow model. The need for this sharp distinction between the return period analysis method and the model that accounts for the physical system has been emphasised earlier by Medina (1986).

Return period and uncertainty analysis in the context of rainfall-runoff pollution modelling both deal with the transformation of distribution functions or error propagation.

The sections that follow is an outline of different approaches to the practical application of the methods in the context of urban runoff pollution and a discussion on the use of the methods and approaches as they have been encountered in the literature.

Method I: Direct fitting.

Given a number of observations of the output variable, Y, a histogram can be plotted in order to view the empirical distribution of the data. The data can be fitted to statistical distributions and the precision of the parameter estimates thus found will depend on the amount of data available and on the number of parameters defining the theoretical distribution.

Method II: Moments transformation.

Many parametric distributions can be characterised by their moments. The number of moments needed to characterise a probability distribution corresponds to the number of parameters in the parametric expressions of the cumulative distribution function or the probability density function. When dealing with linear systems for which the outputs can be described by linear transformations of the input series, the moments of the output series can be calculated as functions of the moments of the input variables.

As a simple illustrative example consider the interdependent stochastic variables X_1 and X_2 with means and variances given by (μ_{X_1} , $\sigma^2_{X_1}$) and (μ_{X_2} , $\sigma^2_{X_2}$) respectively and with crosscorrelation given by ρ_{X_1,X_2}.

Let Y be the sum of the X_1 and X_2 then Y's mean and variance can be calculated as

$$\mu_Y = \mu_{X_1} + \mu_{X_2} \tag{10.1}$$

and

$$\sigma^2_Y = \sigma^2_{X_1} + \sigma^2_{X_2} + 2\rho_{X_1,X_2}\sigma_{X_1}\sigma_{X_2} \tag{10.2}$$

respectively. The moments method is applicable only to linear problems. There are however many cases in which preliminary studies can be based on calculations using approximate linear formulations of non-linear relationships.

The moments method (II) does not require that assumptions are made as to the type of distribution of the input and output series. The assumption made is that the distributions can be characterised by the same number of parameters as the number of moments transformed. In the above example the first two moments (i.e. mean and variance) are transformed. The underlying assumption is thus that the inputs and outputs can be appropriately described by two parameter distributions. If the input random variables are defined by 3, 4 or 5 parameter distributions, similar expressions can be found for the third, fourth and fifth moments. It may also, in some cases, be necessary to include more than simply the crosscorrelation to describe the interdependency between input variables to a model.

Method III : Analytical or numerical integration.

Consider again the realisation x_X of the random variable X and the realisation $m(x_X)$ of the random variable $m(X)$. Suppose that the function is one-to-one mapping of range of X to the range of $m(x_X)$. If the probability density function of the variable X is given by $f_{X(x)}$ then the cumulative probability distribution function of the variable $m(X)$ can be expressed as

$$F_{m(X)}(m(x_X)) = \text{Prob}[m(X) \le m(x_X)] = \int_{-\infty}^{m^{-1}(m(x_X))} f_X(u)du \tag{10.3}$$

In the more general case where the function m is a many-to-one mapping the cumulative probability distribution function of the variable $m(X_1, X_2, X_3, \ldots)$ can be expressed as

$$\begin{aligned} F_{m(X_1, X_2, X_3, \ldots)}\left(m\left(x_{X_1}, x_{X_2}, x_{X_3}, \ldots\right)\right) \\ = \text{Prob}\left[m(X_1, X_2, X_3, \ldots) \leq m\left(x_{X_1}, x_{X_2}, x_{X_3}, \ldots\right)\right] \\ = \int_{\Re_{m\left(x_{X_1}, x_{X_2}, x_{X_3}, \ldots\right)}} f_{X_1, X_2, X_3, \ldots}(u_1, u_2, u_3, \ldots) du_1 du_2 du_3 \ldots \end{aligned} \tag{10.4}$$

where $\Re m(x...)$ is the region of the X_1 - X_2 - X_3 - ... vector space for which $m(X_1, X_2, X_3, \ldots)$ is less than or equal to realisation $m(x_{X1}, x_{X2}, x_{X3}, \ldots)$ and where $f_{X_1, X_2, X_3, \ldots}\left(x_{X_1}, x_{X_2}, x_{X_3}, \ldots\right)$ is the joint probability density function of all the random variables that are inputs to the function $m(...)$.

Depending on the nature of the functions involved, (10.4) can be evaluated either analytically or numerically.

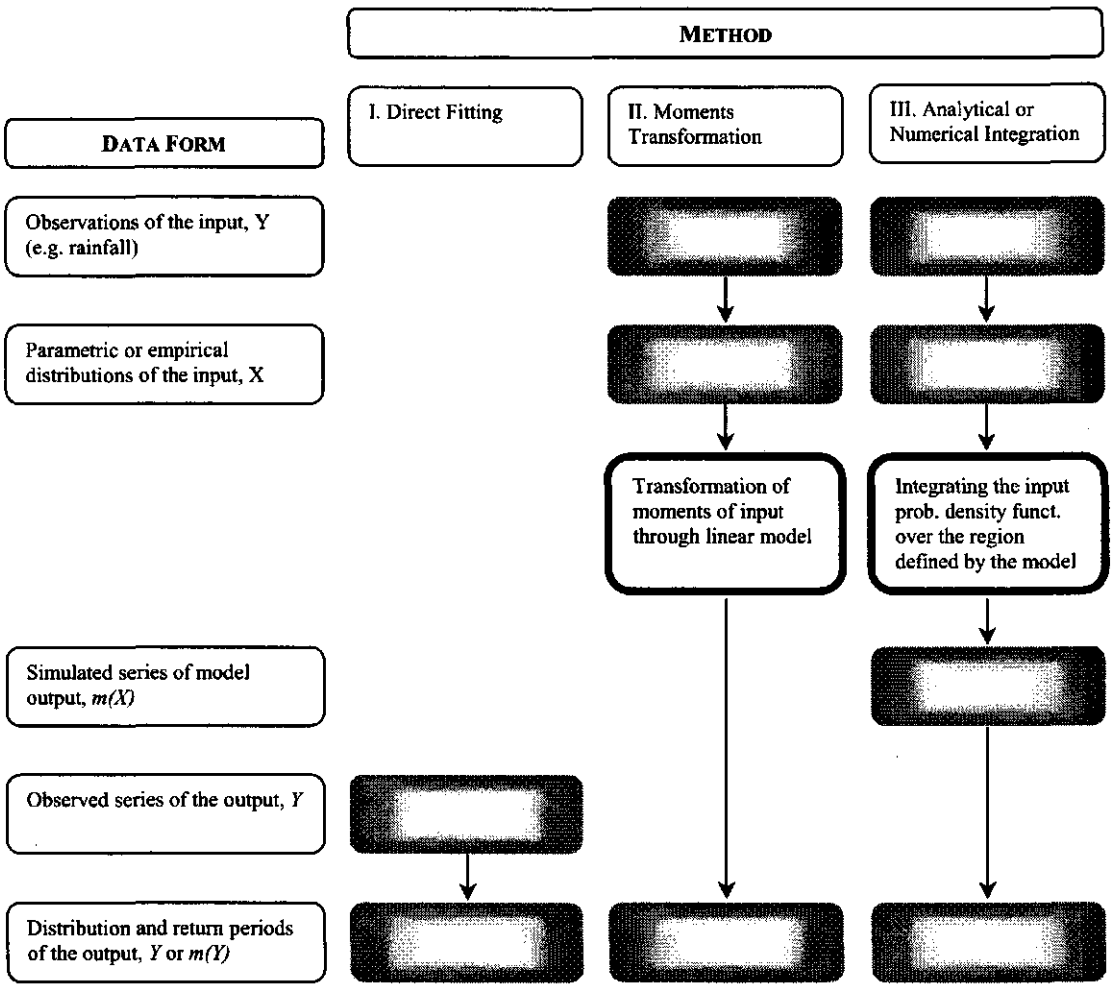


Figure 10.3. Diagram summarising the three methods of obtaining return periods of effect.

If the inputs to the function *m(...)* are independent of one another or if they can be expressed as functions of each other then the integration in (10.4) can often be simplified to a multiple of univariate numerical integrations avoiding the multivariate integration problem. In some cases the joint probability density function could take a form that would allow an analytical solution to the integration. The better known numerical integration techniques used in water quality engineering are the univariate quadrature method and the multivariate Monte Carlo methods.

The use of historic series is a special case of this method. Complete simulations are made using for example a historic rainfall series and the distribution of the output is thus found. The historic rainfall series, which may be either event lumped or as a time series, is in the context of this method an empirical distribution of the input variable.

Approaches

This section aims to formulate a framework within which the above methods could operate. Over recent years it has become increasingly clear that there is a need for an integrated systems approach in runoff pollution (and flooding) analysis and management (Lijklema, 1993b, Lijklema et al., 1993a, Harremoës et al., 1993, Tyson et al., 1993, Harremoës, 1994).

Although the framework is valid for the integrated urban waste water system as a whole, only two parts of the system have actually been included in the present version of the framework. These are the combined sewer system and the surface waters. Including the treatment plant would give more possible combinations but the approach would essentially remain the same.

Inputs to models may be in the form of time series or in the form of event lumped variables. This is illustrated schematically in Figure 10.4. Dynamic combined sewer models are examples of models that take a rainfall time series, a pluviograph, as input and produce a flow time series as output and thus could be symbolised by T→T. Another example would be a dynamic surface water model with time series of combined sewer overflow flow rates and concentrations as input and minimum oxygen concentrations as output. However, due to the difference in time constants between the sewer system (in the order of minutes) and the surface water (in the order of hours), event lumped overflow characteristics such as total overflow volume and event mean concentration may be used. Thus the same mathematical model would be used with event lumped input and time series output (i.e. E→T). This is often done when the dynamics of the input series are much faster than those of the system being modelled.

Rainfall is the driving force and will always be the principal input to the kind of systems and occurrences studied. A historic rainfall series can be applied in the following four ways:

1. used directly as simulation input,
2. modelled dynamically in time and thereafter used to simulate synthetic rainfall used as simulation input,
3. converted to an event lumped series (i.e. depth, duration, maximum intensity and so on),
4. parametric distributions can be fitted to the event lumped series from 3.

The possible uses of these four forms of rainfall data is illustrated in Figure 10.5 (left). Depending on the model used output from the combined sewer system will again be in the form of time series of in the form of event lumped data and will in turn consequently be appropriate for a surface water model taking the available form as input. The output of the surface water model will often be in the form of a time series such as the concentration of dissolved oxygen, which is then converted into an event lumped variable such as the

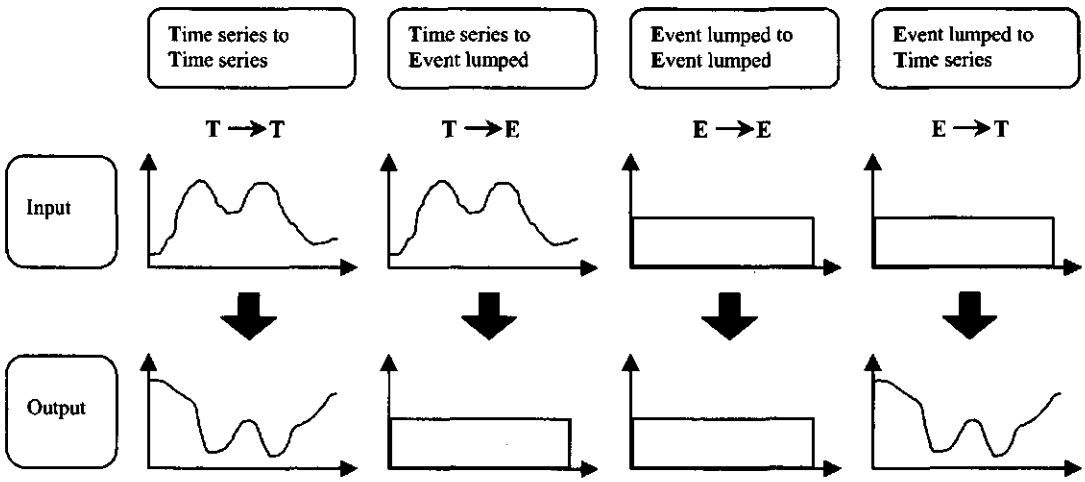


Figure 10.4. The inputs and outputs to models used for return period analysis in integrated urban water management may be in the form of time series or event lumped values.

minimum dissolved oxygen concentration or a one hour minimum. The event lumped data would then be used to calculate empirical or parametric probability distributions from which the return periods can be found. In cases where there is a sufficiently linear relationship between the event lumped rainfall series and the surface water effects, the moments method could theoretically be used. However, if such a linear relationship were present one would generally choose to carry out the statistics on the rainfall series and work with selected design rains. In general this is not the case.

The approach applied in Chapter 9 is highlighted on the right in Figure 10.5. A historic rainfall series was first converted to an event lumped series of rainfall depth and duration, which were fitted to a joint lognormal distribution. Using sampled event lumped rainfall characteristics as input a model, similar to that presented in Chapter 4, was used to generate synthetic series of event lumped overflow characteristics: volume, total and settled nitrogen, chemical oxygen demand (COD) and suspended solids. These event lumped variables were used as input to a dynamic surface water model calculating time series of the dissolved oxygen concentration. From this series minimum oxygen concentrations were calculated and converted to empirical dissolved oxygen concentration.

In spite of the fact that output distributions are generally found using Method III it is clear that there exists a vast number of alternative approaches that can be followed when doing a return period analysis.

Discussion

The method of direct fitting has limited use in the context of return period analysis in urban runoff pollution. Pollution effects are most often very dependent on the quantity of polluting fluid either in terms of flow rate or in terms of total volume. The major input variable to runoff systems is rainfall, which is a meteorological variable. The long term characterisation of meteorological variables would require an observation period of about 30 years. Depending on the relevant return period and on the accuracy required, direct monitoring would have to take place for roughly the same length of time. Another clear disadvantage is that the method is inapplicable to ungauged sites and has no value for the evaluation of proposed system alterations of gauged sites.

Considering water quality variables such as event mean concentration the method of direct fitting may well prove more accurate than any of the other methods outlined. Event mean concentration has sometimes little relation to the event lumped rainfall characteristics (Chapter 4, Harremoës, 1994, Driver and Troutman, 1989) and rainfall may barely be considered an explanatory variable. Thus even directly measured data from a relatively short period may give more accurate estimates than deriving a distribution from the probability characteristics of process inputs via fairly poor model simulation.

The fitting of overflow event mean data to theoretical distributions has been done by Driscoll (1986), Brizio et al. (1989), Hall et al. (1990) and Bomboi et al. (1990). The mentioned articles do not present direct fitting as an alternative method in return period analysis but aim at revealing the shape of the probability distribution of overflow event mean concentrations. This is done with the intention of finding appropriate theoretical distributions because assumptions as to the distribution type must be made in order to apply methods II and III. Brizio et al. (1989) and Hall et al. (1990) show particular interest in the fact that pollutant's event mean concentration often fit better to bimodal or mixture probability distributions (i.e. with two peaks in the probability density function) and can therefore not be described by two parameter distributions such as the often used lognormal distribution (see also Results and Discussion in Chapter 3).

Fitting parametric distributions to data is often part of a return period analysis both with methods I and III (and sometimes with method II). The actual curve fitting is most often done by methods of moments, maximum likelihood estimation or least square estimation. Brizio et al. (1989) found that maximum likelihood estimates gave the best results when fitting distributions with more than two parameters to event mean concentration data. The most important step of distribution fitting is the use of adequate statistical goodness-of-fit tests to compare fits made using different distributions. The most important of these are the chi-square test and the Kolmogorov-Smirnov tests. The methods of parameter estimation are outlined in Patel et al. (1976) and the goodness-of-fit tests are described in most textbooks on statistical distributions such as Hastings and Peacock (1975) and Hodge and Seed (1972).

In the context of water resources analysis, Vogel (1986) presents the probability plot correlation coefficient test as a suitable goodness-of-fit test when dealing with theoretical distributions having only two parameters such as the normal, lognormal, extreme value and gumbel distributions. In water quality management the upper tail of the probability distribution can often have great influence on the decisions made. In this context Ochoa et al. (1980) have studied the tail behaviour of distributions of hydrological phenomena such as flooding. As in the cases of Brizio et al. (1989) and Hall et al. (1990), particular interest is paid to the relatively high probability density of the right (extreme) tail of the probability density functions.

Special attention is needed if the model used (with the moments method) has a discontinuity in the form of threshold values such as the event based sewer overflow model in Chapter 4 or those used by Van der Heijden et al. (1986) and again by Benoist and Lijklema (1989). These references do not use a moments method and are therefore discussed later in this section.

Up until the early 1980s efforts were geared towards finding the ideal design storm. In recognition of the inherent random characteristics of rainfalls a design storm was found on the basis of a chosen return period (Arnell et al., 1984). The authors note, that the result thus obtained, was identical to that obtained from long term simulations when dealing with linear model outputs such as the peak-flow found using the time-area runoff model. The design storm approach became less used in the combined recognition that during full pipe flow in part of a drainage system the drainage of rainfall was dominated by non-linear transformations and that it is the effects of an event that should set design criteria. The increased numerical calculation capacity also played a role in providing realistic alternatives.

Interestingly, the ideas of the design storm lies in a sense close to those of the transformation of the statistical properties of the rainfall, the moments. It is somewhat unfortunate that in water quality engineering the moments method of obtaining output statistics were presented first in conjunction with a non-linear system models that needed a number of controversial assumptions in order to be linearised (e.g. Warn and Brew, 1980, Di Toro, 1984). The moments method is simple and well suited for obtaining fast results using event based linear or log-linear models.

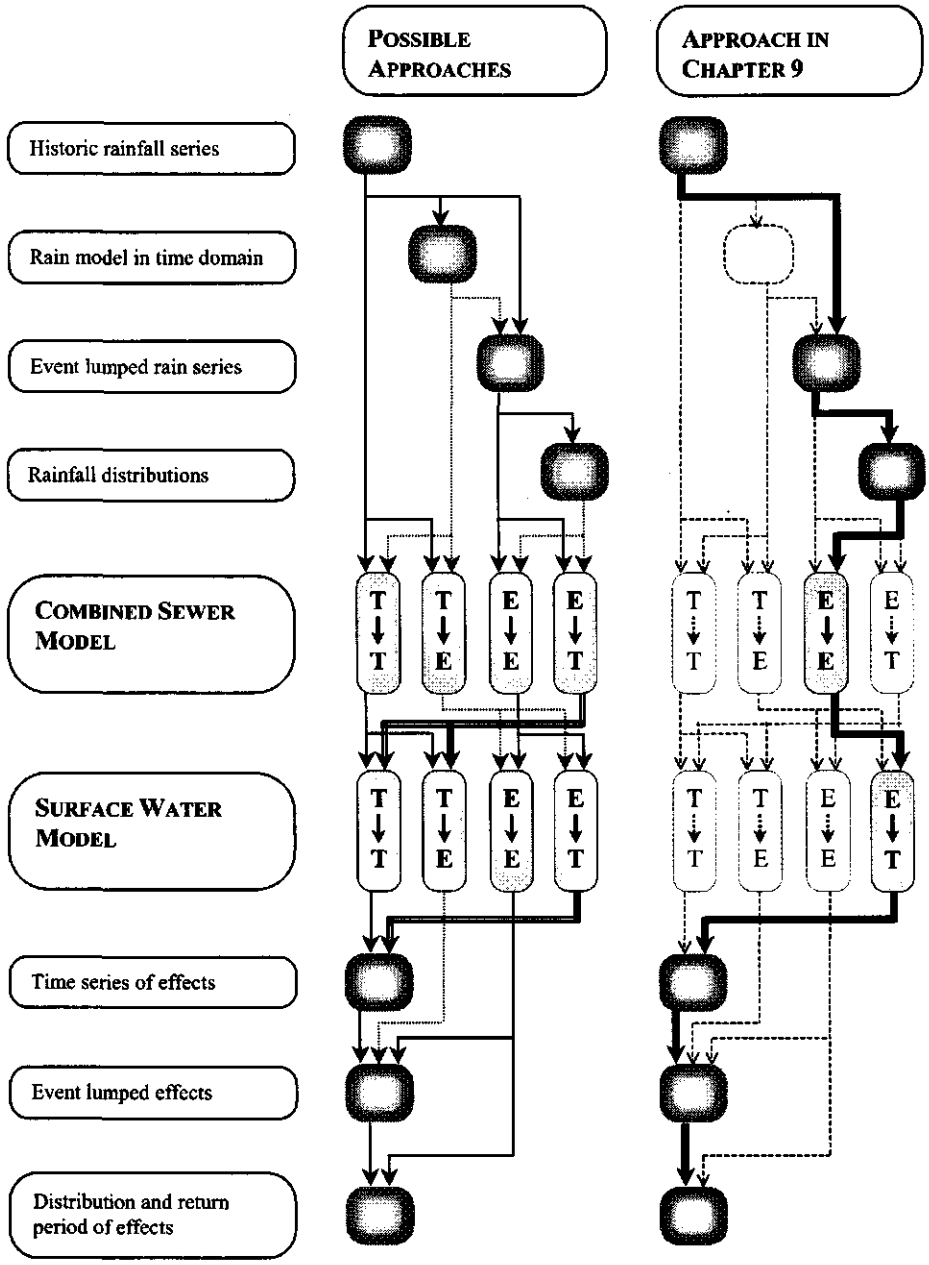


Figure 10.5. Return period analysis approaches (only those having the extreme statistics carried out on the discharge *effects*). The flow scheme to the right illustrates the approach applied in Chapter 9.

Thus if using linear or log-linear regression models, such as those presented in Driver and Troutman (1989) and in Arnbjerg-Nielsen and Harremoës (1994a), the idea of a design storm, which was to some extent dropped in the early 1980s, would give similar solutions as both the long term simulation method and the integration method. Not forgetting the purpose of our anxieties, it is the surface water(s) that sets or ought to set pollution design criteria. When using linear or log-linear event based models, the moments of the output distribution can be calculated directly from those of the inputs without any simulation or integration needed. Note that for single input-single output models, to assume linearity would also be to assume identical distribution types for the input and output variables of the model.

Two methodologies for studying the probability properties of stream quality due to runoff are presented in Di Toro (1984). The first method is an approximate moments transformation method and thus falls under method II. The second is a simplification of a multivariate integration problem to obtain multiple univariate integration and thus falls under method III. Both methods use a dilution equation to describe the mixing of the stream flow and the sewer overflow. This mass balance equation is of fundamental importance to the evaluation of water pollution problems.

For the sake of simplicity in illustration, consider the mixing equation (10.5) as an example of a system model. Let C_R and Q_R be the event mean urban runoff concentration and flow rate, respectively. Let C_S and Q_S be the event mean stream concentrations and flow rate, respectively. The downstream concentration is then given by the mixing equation (10.5) given by and illustrated graphically in Figure 10.6.

$$C_T = \frac{Q_S}{Q_S + Q_R} C_S + \frac{Q_R}{Q_S + Q_R} C_R \tag{10.5}$$

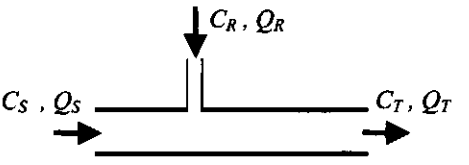


Figure 10.6. Diagram showing variables of the mixing equation (10.5).

The event mean downstream concentration, C_T, is thus given as a function of four random variables. The objective is to obtain the probability density function of the event mean downstream concentration in order to evaluate the return period for exceeding given pollutant concentrations. A complete analytical solution to this problem does not exist and a thorough solution would thus be to resort to numerical multivariate integration techniques.

Di Toro's first method is an approximate moments method (i.e. method II). Assumptions have to be made in order to apply the moments method to an non-linear transformation. The first assumption made is that both event mean concentrations and flow rates are lognormaly distributed. Whether this is a reasonable assumption or not, relates to the earlier referenced findings of Brizio et al. (1989) and Hall et al. (1990). The second, and perhaps more precarious, assumption is that both the concentrations, C_S and C_R, are uncorrelated to their respective flows, Q_S and Q_R, and, therefore, to the dilution ratio. The only crosscorrelation considered is that between the two flow rates. A third assumption is that the runoff flow rate, Q_R, is small compared to upstream flow rates, Q_S. In practice combined sewer overflows discharge rate is often

comparable to, or even larger than, the recipient flow rates. The second and third assumptions are needed to linearise the dilution problem so that a transformation of moments method can be applied.

As mentioned earlier Di Toro's second method is a simplification of a multivariate integration problem to multiple univariate integrations which are solved numerically using quadrature methods. Confronted with a multivariate integration problem, it is a good idea to try to simplify the problem first to see if other solutions can be found or if the problem can be broken down into a series of univariate integrations (Press et al., 1992). Di Toro's second method is just such a simplification. The assumptions made in this second method are much fewer than those made in the first approximated moments transformation method. The input probability distributions and the conditional probability distributions of certain inputs with respect to others are however still assumed to be lognormal distributions. The simplifications made in this second quadrature numerical integration solution would however most likely be possible for a number of other parametric distributions. It could, however, be argued that for fast event based models the computation time gained by using the quadrature numerical integration rather than crude Monte Carlo integration is insignificant compared to the loss of generality in assuming variable independence and in fixing the distributions types.

In a generally positive discussion paper Novotny (1985) emphasises certain limitations concerning the applicability of Di Toro's approach. The first two points made, relate directly to general limitations of the dilution equation itself. Local conditions must be such that the substance in question can be considered conservative and, dispersion and mixing must be such that complete mixing is a reasonable approximation. The third point also relates to the dilution equation but is of greater interest to the return period analyst. Novotny (1985) illustrates that in slow mixing waters the attenuation of high frequency events is greater. In practice this would affect the probability distribution of the downstream water and thus the quality of the approximation would depend on the actual speed of mixing.

Roesner and Dendrou (1985) accuse Di Toro's methodology of totally ignoring cause-effect relationships in the modelling approach. In a discussion paper Roesner and Dendrou (1985) assume that the probabilistic characteristics of the four input variables, flow rate and concentration of the runoff and upstream respectively have to be obtained by measurement. Roesner and Dendrou (1985) seek parameters relating the runoff flow rate and pollutant concentrations to the characteristics of urban runoff area such as in the model STORM. Di Toro presents two simplified solutions to a surface water quality problem. As outlined more clearly later on in this chapter the probability characteristics of the flow rate and concentration of the input urban runoff could well have been computed using dynamic models such as STORM, MOUSE and HydroWorks. This could be done either by simulation using historic rainfall series or by numerical integration using the statistical properties of the rainfall.

As mentioned earlier, Di Toro (1984) presents two *approximate methods* of calculating the probability characteristics of the dilution equation output based on those of the input. Criticism was unfortunately directed more to the use and validity of the dilution equation than to the presented *methods* and *approximations*.

Garboury et al. (1987) and Strecker et al. (1990) have adapted Di Toro's approximate moments method to a general and practical use in highway storm water runoff studies in the USA. The latter authors have incorporated the method as a first level analysis in a highway water quality software package. Phillips (1989) uses the method in conjunction with a model that characterises the surface water's flow further downstream.

Since the early 1980s the method of long term simulation using historic rainfall series has become one of the most used methods in urban runoff return period analysis. The method is most often used in connection with computationally fast time-based runoff models but is also used with event based models. Arnell et al. (1984)

present long term simulation as an alternative in comparison to the then used methods of design storms. The method had however been presented in the context of urban runoff pollution as early as Geiger (1975).

Johansen et al. (1983) used long term simulation in conjunction with a time based modified version of the time-area method to compute overflow series from which the probability characteristics of the overflow can be calculated. The authors then point out that these statistics can be used to generate the input for a surface water model. The authors have further developed graphs and tables meant for manual calculation of both annual and extreme discharges on the basis of interceptor capacity and concentration time of the runoff system.

In a four-stage calculation-detail approach to acute runoff pollution studies, Harremoës et al. (1983) suggested long term simulation with a time based unit-hydrograph model determined using the time-area method as the third level of detail. The fourth level of detail uses iterative hydraulic solutions (i.e. the full Saint Venant equations), whose computation time compels the engineer to retreat to the principle of return period analysis on the inputs (rainfall) rather than on the outputs (the effects). The fourth stage is a compromise between precision and computation time.

Fischer and Buczek (1988) used long term simulation as a reference point when studying the output's dependency on the definition of "a rainfall-runoff event" when using an event based model.

The information contained in event based historic rainfall or overflow series is equivalent to that contained in the series' empirical probability distributions. If statistical tests suggest that the distributions could be adequately described by two or three parameters then the use of the empirical distribution or the equivalent event based historical series could well be considered as a case of over-parameterisation. The use of event based historical series could thus be considered as using thousands of parameters to describe phenomena that could be described using only two or three parameters. Analogous reasoning could lead to similar suggestions concerning the use of time based historic rainfall and overflow series. No literature has been found dealing with the relevance of either of these problems. Long term simulation with both event based and time based historic data series is a well established technique and will probably continue to form the basis of urban runoff return period analysis for some time into the future.

The simplified numerical integration method presented in Di Toro (1984) has been discussed earlier in this chapter. A few other studies applying the integration method (III) will be briefly outlined below.

Working with event based urban runoff models Van der Heijden et al. (1986) and Benoist and Lijklema (1989) have used simplified integration formulations comparable to Di Toro's second method. There is however one very important difference between Di Toro's second method and the method presented by Benoist and Lijklema (1989). The latter use the empirical distributions of rainfall duration and intensity as inputs. A method is then presented in an attempt to compensate for the fact that the duration and intensity are not independent. As mentioned earlier the information contained in the empirical cumulative distribution functions is identical to that which is contained in the historic series on which they are based. The result obtained, the computation time and data storage needed would be similar to that of a long term simulation. One of the essential advantages of the analytical or numerical integration method (III) is that the probability characteristics of the output can be generated (in a relatively small number of model runs) from a limited number of parameters that describe the probability characteristics of the inputs.

Akan (1988) developed a non-linear event based accumulation-washoff model for urban runoff surfaces. The inputs to the model are the time passed since last rainfall and the rainfall depth (volume). Using the model the author presents an integral expression equivalent to that in (10.4) but with the relevant integration region specified. The author does not attempt to simplify the integration problem but envisages that this will have to

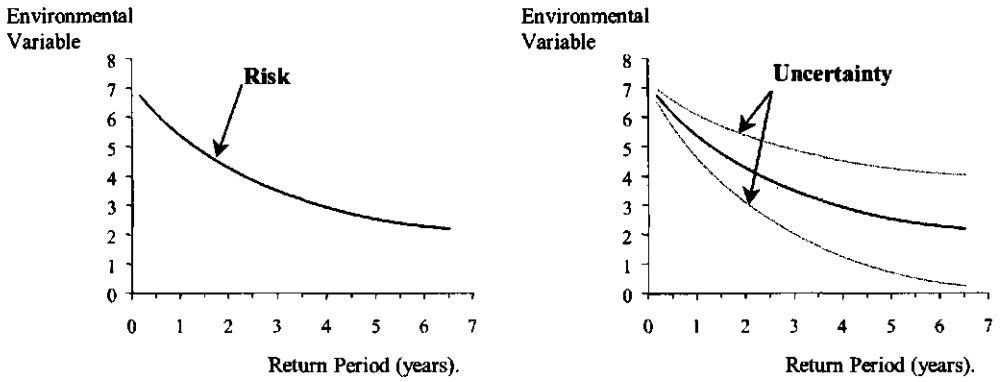


Figure 10.7. The information sought during return period analysis (left) and the analysis of its uncertainty (right). The curves give (fictitious) return periods for minimum values of an environmental variable of interest.

be solved numerically. The joint probability density function of the time passed since last event and the rainfall depth, which should be estimated from historic rainfall series, is foreseen to "have a rather complex form". In the example, Akan (1988) has, for illustrative reasons, used a very simple but unrealistic joint probability density function.

The integration method (III) is used by Cadavid et al. (1991) in a flood-frequency analysis with a kinematic wave model. The authors assume that the joint probability density function of the rainfall intensity and duration follow a bivariate exponential distribution. After accounting for infiltration, a fairly complex joint probability density function for an effective duration and effective intensity results. The integration problem is solved numerically.

In the earlier mentioned article by Warn and Brew (1980) the results of Monte Carlo integrations are used as a reference point in evaluating the presented approximated moments method. The method was used to solve problems of the dilution type.

The response time of the surface water stream, to inputs of rainfall, is often several days or weeks if the stream is dominated by the entries from subsurface flow (Duysings et al., 1983). This would mean that even after corrections for seasonality of the stream characteristics, using the event definition used for the runoff model would result in very high autocorrelations in the event series of the stream characteristics. The moments method (II) or the integration method (III) used with an event lumped surface water quality model would require some method of handling this "overlapping of events". One way would be to estimate the parameters of a distribution for the time between consecutive sewer system events and to incorporate this in the moments transformation or numerical integration. An alternative could be the use of what in this chapter is called an event-to-time model for the surface water response. This could either be done with historic series using the long term simulation method or by characterising the probability distribution of the time between events and then incorporating this into the integration method (i.e. both method III).

An event-to-time model is in this chapter defined as one that has a dynamic description of the physical system with at least one of its inputs being arranged as an event series. Examples could be the great variety of runoff models that used rectangular, triangular, bell-shaped, trapez rainfall profiles to define the shape of a design storm input (Arnell et al., 1984). Event-to-time runoff models have not been used to any significant degree since the beginning of the 1980s but are still common in the evaluation of surface water response to event based pollution runoff input. This is because the time constants of rainfall variations are comparable to

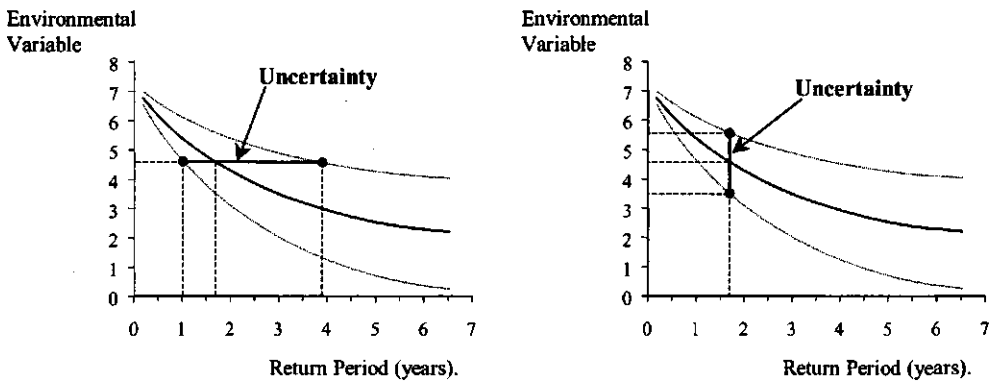


Figure 10.8. Return period uncertainty can be considered in terms of the uncertainty in the return period for a given value of the environmental variable (left) or as a uncertainty in the environmental variable for a given return period (right).

those of most urban drainage systems whereas the characteristic time constants of combined sewer overflow is often shorter than those of the processes controlling the relevant receiving water variables (e.g. dissolved oxygen). An example of an event-to-time model would be the surface water model used by Portielje et al. (1996) where the runoff characteristics of total event overflow volume, event overflow duration and event mean concentration are used as input to a dynamic river water quality model.

UNCERTAINTY IN RETURN PERIOD ANALYSIS

In the previous chapter it was emphasised how a distinction should be made between inherent variation and uncertainty and a methodology for doing this was presented. The framework and discussion above dealt only with handling of inherent event to event variation resulting in return period curves as shown in Figure 10.5 (left). This section contains a general discussion on uncertainty in return period analysis aimed at placing the uncertainty analysis presented in the previous chapter in a broader perspective. What is understood by uncertainty in return period analysis is illustrated graphically in Figure 10.7 (right).

Return period uncertainty can be expressed in terms of an uncertain return period for a given level of the environmental variable or the same uncertainty can be expressed in terms of an uncertain level of the environmental variable for a given return period. These two ways of expressing the same information are illustrated in Figure 10.8. Which representation is most appropriate will depend on the decision to be made. In some cases it may be a good idea to evaluate both.

Relevance of Uncertainty

The presence of uncertainty in return period analysis is clear. The extent, relevance and consequences of the uncertainty may be less clear. Management decisions concerning effects of acute pollutants or sewer flooding are often supported by return period curves to clarify the extent of the problem and the amelioration that can be expected after proposed modifications. If the uncertainty associated with such return period curves is high then the uncertainty should be presented together with the return period curve in the form of a standard deviation or confidence band.

Consider the example depicted in Figure 10.9. The uncertainty of the minimum concentration relating to a given return period has been found and the upper and lower 95% confidence limits are indicated as dashed lines. Assuming that a major design criteria is related to the return period of a minimum concentration of 4 mg/l, the graph in Figure 10.9a appropriately indicates that the expected return period this value is 2 years and that the lower and upper confidence limits are 1.5 and 3.6 years respectively. Figure 10.9b illustrates how a criterion related to the return period of a minimum concentration of 3 mg/l could have an expected return period of 5 years, the lower confidence limit 2.6 years and *no upper confidence limit*. The investor will then have to be content with the somewhat awkward situation of not knowing whether the intended improvement will lead to a detrimental oxygen concentration every 2½ years or every 250 years. Note that for larger return periods, the occurrence of an upper confidence limit at infinity may sometimes say more about low rate of change of the effect for increasing return periods than about the absolute magnitude of the uncertainty of the effect for a given return period.

Uncertainty in Rainfall Characterisation

The main dynamic inputs to urban runoff models are the rainfall variables. It is evident that the accuracy of rainfall description highly depends on the length of the monitoring period. Assuming that trends due to for example climatic change or urbanisation are small, the certainty with which one can predict a rainfall variable's value (depth or intensity) associated with a given return period depends largely on the ratio between the length of monitoring period and the return period in question. A detailed non-parametric study of the uncertainties in the characterisation of rainfall has been made by Arnbjerg-Nielsen (1993) and further reported in Arnbjerg-Nielsen et al. (1994b) and Harremoës (1994). The data used originated from 58 rain gauges in Denmark that had been monitored for durations of 2 to 14 years. Using a resampling method, Arnbjerg-Nielsen (1993) determines the magnitude of the uncertainties associated with a given return period. For a rain gauge with an observation period of about 13 years a design depth for a return period of two years is found to have an expected value of 30 mm with 95% confidence limits from 23mm and 40mm. Owing to the relatively short observation period of only 13 years and owing to the small rate of change of the rainfall depth with respect to return period, it is (at the referenced gauge) only realistic to consider the upper confidence limit of the return period for rainfall depths that have expected return periods of less than one year. For longer periods the uncertainty simply becomes too large.

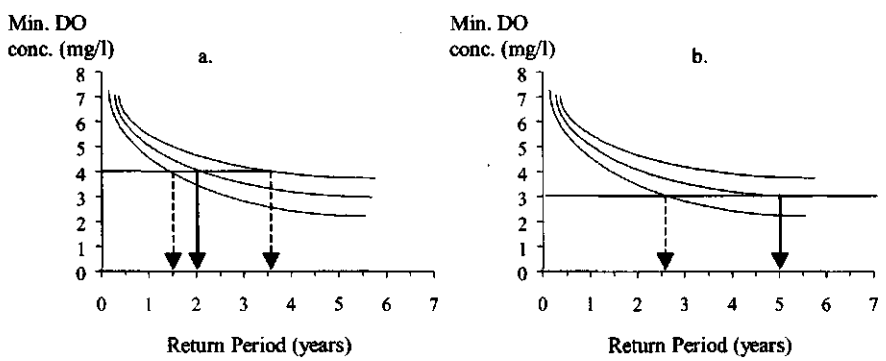


Figure 10.9. Fictitious return period curves for a surface water's minimum dissolved oxygen concentration (see text for details on a. and b.).

Similar results were obtained during the uncertainty analysis of the previous chapter for which the rainfall results have been reprinted in Figure 10.10. Here the parametric joint lognormal distribution was fitted to rainfall depth and duration and variances of the parameters estimated from the data using a jack-knife estimation procedure as described in Chapter 9. For the studied rainfall series the return period of an event of 34 mm had an expected value of 2 years and to lie between 0.55 years (i.e. almost twice a year) and 3.0 years (estimated 50% confidence limits). Equivalently the two years event was found to have an expected depth of about 34 mm with estimated 50% confidence limits at 31 mm and 52 mm.

Model Uncertainty

In urban runoff pollution problems, uncertainties in transformation are most often associated with the model parameter uncertainties. These could arise in the determination of the initial rainfall loss, the runoff coefficient, the conduit roughness, quantity of the pollutants on the surface, dry weather flow volumes and concentrations and many others.

Traditional approaches to the evaluation of the effects of uncertainties has been to quantify the uncertainty of the models output, $m(X)$, with respect to the uncertainties of the inputs and of the model parameters. Structural uncertainties of the model such as those associated with conduit geometry, spatial distribution of rainfall, rainfall movement, various sediment accumulation expressions and other unknown uncertainties are ignored. The confidence range thus found is the range within which the 'true' value of the model's predicted output can be expected and not the confidence range within which the 'true' value of reality can be expected. The uncertainty that ought to be of interest is uncertainty of the model's ability to predicted the actual system's output. The distinction here between system and model transformation has been defined earlier in Figure 10.2.

This is one of the most essential reasons why stochastic modelling presents better options for the handling of uncertainties in risk analysis of urban runoff pollution problems. The stochastic models may well include conceptual and physical relationships rather than be purely empirical. The potentials of stochastic modelling in obtaining improved handling of uncertainties in urban runoff pollution have been emphasised by Harremoës et al. (1993), House et al. (1993) and Harremoës (1994). Further reasons for use stochastic models are given below.

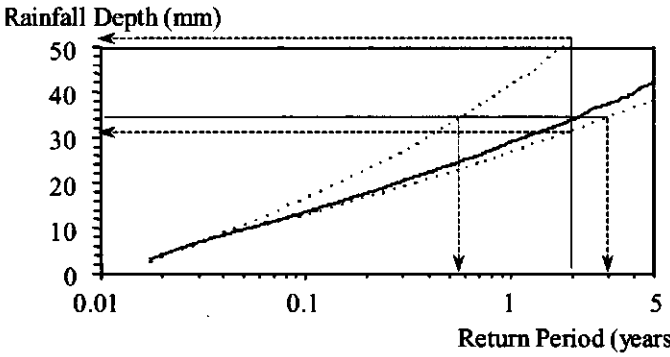


Figure 10.10. These results from Chapter 9 showing 50% confidence limits for the return period of rainfall depth demonstrate the rainfalls substantial contribution to the overall uncertainty.

Handling Uncertainty

An uncertainty analysis can be carried out analogous to the return period analysis method III (and in some very simple case as method II). This is most often done using Monte Carlo integration techniques such as has been done by Pedersen (1993) (extensively discussed by Harremoës, 1994). For a two year return period, the variability (coefficient of variance) of fluid quantities was found to be in the order of 30% and the variability of the pollutant concentrations was found to be in the order of a factor two (Harremoës, 1994). The studied uncertainties included those of rainfall characterisation.

Portielje et al. (2000) present two alternatives to using crude Monte Carlo integration in environmental risk and uncertainty analysis. The first method combines a first order reliability method with directional simulation using importance sampling and the second method combines it with Latin hypercube simulation. For extreme events with exceedance probabilities smaller than 0.1, the authors found that both methods were more efficient in terms of the number of model evaluations and more accurate than the crude Monte Carlo integration. The methods have been applied in the context of oxygen depletion in a sewer overflow surface water in Portielje et al. (2000).

Return period analysis can be carried out using either deterministic or stochastic models. As discussed in detail earlier in the thesis the underlying assumption during deterministic estimation is that only observation error is present. During simulation as part of the return period and uncertainty analysis this observation error is generally not simulated. However, during stochastic modelling the system is itself assumed to behave partly in a random manner. This random variation is simulated during the return period and uncertainty analysis and is treated as an inherent random variation of the system. In general one would therefore expect the deterministic models to have a tendency to underestimate the frequency of extreme occurrences. This is another reason why efforts should be made to use stochastic models for return period analysis.

Uncertainty in Scenario Comparison

Let (μ_A, σ_A) and (μ_B, σ_B) be the expected values and the standard deviations of the minimum dissolved oxygen concentration with a return period of one year for the modelled scenarios A and B respectively. One may often plot the mean and confidence range as has been done in Figure 10.11. Assuming normality the 95% confidence range ($\pm 2\sigma$) has been indicated for each scenario.

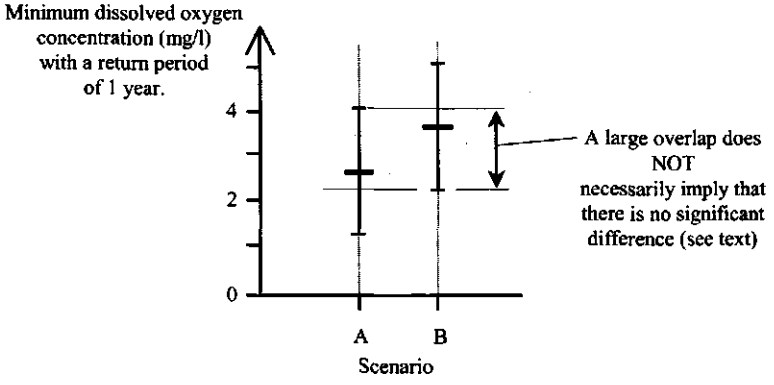


Figure 10.11. Diagram showing the expected value and range of the minimum dissolved oxygen concentration with a return period of one year for the modelled scenarios A and B respectively (fictitious example).

In decision making one would be interested in knowing whether scenario B is significantly better than A and if so (assuming B to be more costly) with what probability is B at least a given absolute amount or percentage better than A.

Looking at the range overlaps in Figure 10.11 one may be tempted conclude that a difference between the scenarios A and B cannot be confirmed with the available tools and data. However, the expected values and their respective standard deviations are most often not independent. The two probability distributions of the minimum concentrations for a given return period have most often been calculated using the same models with a large number of common parameters and inputs. Consequently much of the uncertainty is common to both estimates.

Consider two random variables Y_A and Y_B with mean and standard deviation, (μ_A, σ_A) and (μ_B, σ_B) respectively. The variable of interest is however their difference $Y_{B-A} = Y_B - Y_A$ whose mean and standard deviation can be expressed as

$$\mu_{B-A}=\mu_B-\mu_A \tag{10.6}$$

$$\sigma^2_{B-A}=\sigma^2_B+\sigma^2_A-2\rho_{A,B}\sigma_A\sigma_B \tag{10.7}$$

where $\rho_{A,B}$ is the correlation coefficient between Y_A and Y_B. Thus it is clear that there is simply not enough information available if the interdependence is not available.

In many cases the output probability distributions for the two scenarios have been generated by doing Monte Carlo simulations with the same model having some different inputs or different input distributions. A great number of the uncertainties are often the same for both scenarios. The exact same random sampled realisations from the input uncertainty distributions should thus be used for both scenarios, which would result in paired outputs from A and B. For each pair the difference should be calculated. The distribution of these differences should then be used to answer question concerning the comparison of A and B. A similar procedure should be used for cases where the interest is in the ratio rather than the difference between the scenarios.

Note that the fact that a given source of uncertainty is the same for both scenario A and B does not mean that it can be omitted in the comparative uncertainty analysis. The transformation of the uncertainty by the model (and in reality) will most often depend on the specific conditions of the scenario.

CONCLUSION

Return period analysis of the effects of urban runoff pollution can be performed in several different ways and it is important to be conscious of this in selecting approaches for a given purpose. A framework that encompasses most methods and approaches has been presented and discussed. The approach to return period analysis implemented in Chapter 9 was discussed in the perspective of the presented framework. Practice and water quality engineering tools ought to master a selection of these methods and approaches aimed at different levels of effort corresponding to different stages of an urban water management project.

The uncertainty associated with the return period curve of an environmental variable such as dissolved oxygen is large and its quantification would improve the grounds for decision making. In quantifying return period uncertainty a distinction should be made between inherent variation resulting from randomness of the system described and uncertainty resulting from our lack of knowledge. Though the uncertainty in return period analysis is large the uncertainty on scenario comparison is much lower and ought to be evaluated separately.

In return period and uncertainty analysis there are two main reasons why stochastic models incorporating physical, chemical and/or biological features should be preferred compared to deterministic models. These are:

- to avoid the bias on parameter estimates that results from deterministic calibration (see Chapter 5 & 8),
- to avoid underestimation of the frequencies of extreme occurrences due to the exclusion of random behaviour of the system (see page 137).

REFERENCES

Aalderink, R.H. and Lijklema, L. (1985). Water Quality Effects in Surface Waters Receiving Storm Water Discharge. *Hydrological Res.*, Proc., No.33.

Aalderink, R.H. (1996). Uncertainty analysis of recipient at Loenen. In preparation.

Akan, A.O. (1988). Derived Frequency Distribution for Storm Runoff Pollution. *J. Envir. Engrg.*, ASCE 114(6), 1344-1351.

Arnbjerg-Nielsen, K. (1993). Non-Parametric Statistics on Extreme Rainfall. Masters Thesis at IMSOR & LTH, Technical University of Denmark, IMSOR MSc. Thesis no. 5/93.

Arnbjerg-Nielsen, K. and Harremoës, P. (1994a). Prediction of Hydrological Reduction Factor and Initial Loss in Urban Surface Runoff from Small Ungauged Catchments. International Workshop Closing the Gap Between Theory and Practice in Urban Rainfall Applications, St. Moritz, Switzerland, November 1994.

Arnbjerg-Nielsen, K., Harremoës, P. and Spliid, H. (1994b). Non-Parametric Statistics on Extreme Rainfall, *Nordic Hydrology*, Vol. 25, 267-278

Arnbjerg-Nielsen, K., Harremoës, P. and Spliid, H. (1994c). Interpretation of regional variation of extreme values of point precipitation in Denmark. International Workshop Closing the Gap Between Theory and Practice in Urban Rainfall Applications, St. Moritz, Switzerland, November 1994.

Arnell, V., Harremoës, P., Jensen, M., Johansen, N.B. and Niemczynowicz, J. (1984). Review of Rainfall Data Application for Design and Analysis. *Wat. Sci. Tech.*, Vol. 16, 1-45.

Benoist, A.P. and Lijklema, L. (1989). A Methodology for the Assessment of Frequency Distributions of Combined Sewer Overflow Volumes. *Wat. Res. Vol.*, 23(4), 487-493.

Bomboi, M.T., Hernandez, A., Marino, F. and Hontoria, E. (1990). Distribution of Hydrocarbon Concentrations from Urban Runoff. *The Science of the Total Environment*, Vol. 93 (1990) 465-480.

Brizio, M.G, Hall, M.J. and Ellis, J.B. (1989). The Statistical Distribution of Event Mean Concentrations of Pollutants in Stormwater Runoff. Second National Hydrology Symposium, University of Sheffield, 4-6th Sept. 1989. British Hydrological Society.

Cadavid, L., Obeysekera, J.T.B. and Shen, H.W. (1991). Flood-Frequency Derivation from Kinematic Wave. *J. Hydr. Engrg.*, ASCE, 117(4), 489-510.

Di Toro, D.M. (1984). Probability Model of Stream Quality Due to Runoff. *J. Envir. Engrg.*, ASCE 110(3), 607-628.

Driscoll, E.D. (1986). Lognormality of Point and Non-point Source Pollutant Concentrations, in Urbonas, B. and Roesner, L.A. (ed.), Urban Runoff Quality, Impact and Quality Enhancement Technology, Am. Soc. Civ. Engrs., New York, 438-458.

Driver, N.E. and Troutman, B.M. (1989). Regression Models for Estimating Urban Storm-Runoff Quality and Quantity in the United States. *J. Hydrol.*, 109(3/4), 221-236.

Duysings, J.J.H.M., Verstraten, J.M. and Bruynzeel, L. (1983). The Identification of Runoff Sources of a Forested Lowland Catchment: A Chemical and Statistical Approach. *J. Hydrol.*, Vol. 64(1983), 357-375.

Fischer, F. and Buczek, H. (1988). Pollution Load Calculations based on Statistical Model Rainfall as Compared to Longterm Simulation. Korrespondenz Abwasser : Informationsblatt fuer das Abwasserwesen, 1/88, 35. Jahrgang. 41-46.

Garboury, D.R., Driscoll, E.D. and Sartor, J.D. (1987). A Probability Methodology for Estimating Water Quality Effects from Highway Stormwater Runoff. *The Science of the Total Environment*, Vol. 59(JAN.) 447-456.

Geiger, W.F. (1975). Urban Runoff Pollution Derived from Long-Time Simulation. Nat. Symp. on Urban Hydrology and Sediment Control, University of Kentucky, Lexington, KY., July 28-31, 1975.

Grum, M. (1993). Stochastic Modelling of Urban Drainage Systems. Masters Thesis at LTH & IMSOR, Technical University of Denmark.

Hall, M.J., Ellis, J.B., and Brizio, M.G. (1990). On the Statistical Distribution of Event Mean Concentrations of Pollutants in Stormwater Runoff. Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Harremoës, P., Jensen, M. and Johansen, N.B. (1983). A Staged Approach to the Application of Rainfall Data to Urban Runoff Calculations. In Harremoës, P. (ed.). Rainfall as the Basis for Urban Run-off Design and Analysis. IAWPRC/IAHR Specialised Seminar Preprints.

Harremoës, P., Hansen, O.B. and Sund, C. (1989). Rain Run-off from Sewer Systems & Treatment Plants. In Torno, H.C. (ed.). Urban Stormwater Quality Enhancement - Source Control, Retrofitting and Combined Sewer Technology, Engineering Foundation Conference, Davos, Switzerland, 1989, 420-440 ASCE, New York.

Harremoës, P., Capodaglio, A.G., Hellstrom, B.G., Henze, M., Jensen, K.N., Lynggaard-Jensen, A., Otterpohl, R. and Soeborg, H. (1993). Wastewater Treatment Plants under Transient Loading - Performance, Modelling and Control. *Wat. Sci. Tech.*, Vol. 27(12), 71-115.

Harremoës, P. (1994). Integrated and Stochastic Features of Urban Drainage Systems. *Wat. Sci. Tech.*, Vol. 30(1), 1-12.

Hastings, N.A.J. and Peacock, J.B. (1975). Statistical Distributions: A handbook for students and practitioners. Butterworth & Co. (Publishers) Ltd., 88 Kingsway, WC2B 6AB, London.

Hodge, S.E. and Seed, M.L. (1972). Statistics and Probability. Blackie & Son Limited, Bishopbriggs, Glasgow G64 2NZ, 14-18 High Holborn, London WC1V 6BX.

House, M.A., Ellis, J.B., Herricks, E.E., Hvitved-Jacobsen, T., Seager, J., Lijklema, L., Aalderink, H. and Clifforde, I.T. (1993). Urban Drainage - Impact on Receiving Water Quality. *Wat. Sci. Tech.*, Vol. 27(12), 117-158.

Johansen, N.B., Harremoës, P. and Jensen, M. (1983) Methods for Calculation of Annual and Extreme Overflow Events for Combined Sewer Systems. In Harremoës, P. (ed.). Rainfall as the Basis for Urban Run-off Design and Analysis. IAWPRC/IAHR Specialised Seminar Preprints.

Lijklema, L., Tyson, J.M. and Lesouef, A. (1993a). Interaction between sewers, treatment plants and receiving waters in urban areas: a summary of the Interurba '92 workshop conclusions. *Wat. Sci. Tech.*, Vol. 27(12), 1-29.

Lijklema, L. (1993b). Integrated Management of Urban Waters - A Preface. *Wat. Sci. Tech.*, Vol. 27(12), vii-x.

Medina, M.A. (1986). State-of-the-Art, Physical-Based and Statistically-Based Water Quality Modeling. In Torno, H.C., Marsalek, J. and Desbordes, M. (eds.), Urban Runoff Pollution, NATO ASI series G: *Ecological Science*, Vol. 10. Springer, Berlin (1986).

Novotny, V. (1985). Discussion: Probability Model of Stream Quality Due to Runoff. *J. Envir. Engrg.*, ASCE 111, 737.

Ochoa, I.D., Bryson, M.C. and Shen, H.W. (1980). On the Occurrence and Importance of Paretian-tailed Distributions in Hydrology. *J.Hydrol.*, Vol. 48(1980), 53-62.

Patel, J.K., Kapadia, C.H. and Owen, D.B. (1976). Handbook of Statistical Distributions. Marcel Dekker, Inc., 270 Madison Avenue, New York 10016.

Pedersen, K. (1993). Statistics on Overflow and Flooding. MSc. Thesis, Dept. Eng., Technical University of Denmark.

Phillips, J.D. (1989). Nonpoint Source Pollution Risk Assessment in Watershed Context. *Envir. Management*, Vol. 13(4), 493-502.

Portielje, R., Schaarup-Jensen, K. and Hvitved-Jacobsen, T. (1996). Extreme event statistics of dissolved oxygen depletion in streams receiving multiple COD loadings from sewer systems using stochastic reliability methods. Draft (in preparation). To be submitted to *J. Env. Eng.*, Aalborg University, Dept. of Civil Engineering, Sohngaardsholmsvej 57, DK 9000 Aalborg, Denmark.

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*, Vol.34(1), pp. 153-170.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992). Numerical Recipes in C: The Art of Scientific Computation. Second edition. Press Syndicate of the University of Cambridge, The Pitt Building, Trumpington Street, Cambridge CB2 1RP.

Roesner and Dendrou (1985) Discussion: Probability Model of Stream Quality Due to Runoff. J. Envir. Engrg., ASCE 111, 738-740.

Strecker, E.W., Driscoll, E.D., Shelley, P.E., Garboury, D.R. and Sartor, J.D. (1990). The U.S. Federal Highway Administrations Receiving Water Impact Methodology. *The Science of the Total Environment*, Vol. 93(1990) 489-498.

Tyson, J.M., Guarino, C.F., Best, H.J. and Tanaka, K. (1993). Management and institutional aspects. *Wat. Sci. Tech.*, Vol. 27(12), 159-172.

Van der Heijden, R.T.J.M., Lijklema, L. and Aalderink, H. (1986). A Statistical Methodology for the Assessment of Water Quality Effects of Storm Water Discharges. In Torno, H.C., Marsalek, J. and Desbordes, M. (eds.), Urban Runoff Pollution, NATO ASI series G: Ecological Science, Vol. 10. Springer, Berlin (1986).

Vogel, R.M. (1986). The Probability Plot Correlation Coefficient Test for the Normal, Lognormal, and Gumbel Distributional Hypotheses.

Warn, A.E. and Brew, J.S. (1980). Mass Balance. *Water Research*, Vol. 14, 1427-1434.

SUMMARY

SUMMARY

Since the construction of wastewater treatment plants combined sewer overflows have become an increasingly important limitation to the quality of the surrounding surface waters. Over the years urban water resources have often been so modified by anthropogenic activity that water quality management requires an integrated approach both at an evaluation and an investment level. Effects of acute pollutants, such as oxygen depletion caused by excessive organic material, should be evaluated on the basis of their return periods. Models of the relevant water systems are then used both to calculate the present return periods and to predict those of proposed amelioration projects. In traditional deterministic modelling of combined sewer systems there has been a tendency to continuously add new processes to the model structure in an attempt to improve the quality of the predictions made. This has often resulted in many model parameters with unknown values and the inclusion of processes much less significant than others that are not known or well understood. An alternative approach is to describe only the most essential processes and to include stochastic terms to describe the remaining variation. The present study has focused on comparing and contrasting deterministic and stochastic approaches to modelling of urban runoff pollution and water quality in general. Methodologies surrounding the application of models in return period analysis and its uncertainty have also been studied.

A multivariate analysis was made on event mean concentrations data sets from three Dutch and two Danish combined sewer catchments (Chapter 3). This was done to examine the underlying structure of variations in event mean concentrations. Results confirmed expectations that the most pronounced common variations relate to the groups of particulate pollutants and dissolved pollutants. The distribution of the principal factors clearly reconfirmed the bimodal or mixed distribution that have earlier been reported for event mean concentrations of particulate substances.

Non-linear event lumped models were developed to predict combined sewer discharged volume and event mean concentrations as a function of rainfall variables (Chapter 4). The aim was to combine basic understanding of the physical system with information held in the data. The discharged volume was well described with a wetness dependent runoff coefficient. Seasonality revealed initially by the data and then described using an empirical "cut-off" sinusoidal expression exhibited a remarkable agreement with average monthly open water evaporation data. Using open water evaporation as an input variable to the model improved the prediction whilst at the same time reducing the number of model parameters. The event lumped rainfall variables were only able to explain very little of the variations in the event mean concentrations of the combined sewer overflow and subsequently some of the water quality variables were characterised by their probability distributions alone.

An analysis of the underlying assumptions made during mathematical modelling of water systems in time has resulted in a new portrayal of the essential differences between deterministic and stochastic modelling (Chapter 5). The implicit assumption made during deterministic modelling is that our model gives a perfect description of reality and that all deviation between modelled and observed values is a result of observation error. During stochastic modelling the implicit assumption is that the model only gives a partial description of reality and that deviation between modelled and observed values results from unexplained random behaviour of the system being modelled as well. Having isolated the core differences between deterministic and stochastic modelling allows for more interchange of methods and approaches, thus enhancing the quality of water and water quality modelling. Knowledge of the dominating physical, chemical and biological processes of our system can be built into the traditionally empirical stochastic models. Parameter statistics, experimental design, empirical elements and concepts of identifiability can be applied to deterministic models. Quantitative *a priori* knowledge of given model parameter values can be incorporated into the estimation procedure. In the long term it is the aim that the selected approach will depend more on the appropriateness of the assumptions made (viewed also in relation to the available resources and the possible

consequences of a poor model) than on background of the modeller, as is often the case today. Parameters of a combined sewer rainfall-runoff model have been estimated both in a deterministic and in a stochastic model to study and illustrate the main points of the chapter.

Using a stochastic differential equations approach water quantity and quality models for a combined sewer system were formulated and their parameters estimated (Chapter 6). The aim was to evaluate the potentials and limitations of this approach where the sewer system is defined by a set of differential equations that is solved stochastically in continuous time. Parameter estimation was possible for the water quantity model and a very small observation error confirmed the relevance of a stochastic modelling approach. Results from the water quality modelling suggest that more work is needed in order to fully appreciate potentials and limitations of the approach.

A non-linear random coefficient model to describe suspended chemical oxygen demand in a combined sewer system was identified and its parameters were estimated (Chapter 7). In random coefficient modelling certain selected parameters are assumed to vary from event to event and a value for these parameters is estimated for each event. In the present study a critical soft threshold flow at which resuspension begins is assumed to be a random coefficient. Although there is a lack of data in the period before overflow begins, the results suggest that there is a high potential for random coefficient modelling in urban runoff pollution both as an alternative to and in combination with stochastic modelling. The recipient water quality model used in the uncertainty analysis of Chapter 9 was also estimated using this approach.

Methods and approaches studied in the preceding chapters have been discussed in a broader perspective whilst drawing attention to some interesting developments within the field of water and water quality modelling (Chapter 8). Structuring our physical, chemical and biological theory in stochastic state space models we acknowledge that the deviation between "what we model" and "what we see" is the result of both unexplained random behaviour of the system being modelled and observation error. This acknowledgement will reduce bias in parameter estimates and therefore improve the models' abilities to predict and extrapolate in time and to new circumstances. Although stochastic state space modelling using the Kalman filter had its main entry into hydrology and water quality modelling in the late 1970s, this was mostly with empirical formulations based entirely on observed data and therefore of little use to the engineer wishing to examine and compare alternative scenarios. To avoid over-parameterised models with highly interchangeable parameters it is important that model structure is identifiable on the basis of data being used to estimate the model parameters. The *a posteriori* estimation criteria incorporating quantitative *a priori* knowledge present an interesting formalised method of introducing the engineer's intuition and experience into the parameter estimation procedure.

A new methodology for evaluating the uncertainty of a return period analysis is presented and exemplified in an integrated approach to urban runoff pollution involving models of both the combined sewer and the receiving water (Chapter 9). The underlying hypothesis of the presented methodology is that a distinction has to be made between inherent variation and uncertainty resulting from a lack of knowledge. This distinction is attained through embedded error propagation, which was here implemented as Embedded Monte Carlo Simulations. It is argued that pooling uncertainty with inherent variation systematically increases the frequency of extreme events resulting in return period curves with little or no engineering value. The study also demonstrates that efforts are needed to implement faster alternatives to the crude Monte Carlo simulations to reduce computation time, which would be necessary for use in practice.

A review of methodologies surrounding return period analysis in urban runoff pollution and its uncertainty was carried out with the aim of viewing the new methodology presented in Chapter 9 in its broader perspective (Chapter 10). Three principally different methods of calculating return periods of given effects have been described: direct fitting, moments transformation and analytical or numerical integration. Combining these methods with the different types of models (in terms of input and output being time series or event lumped variables) results in a framework encompassing most approaches to return period analysis. Uncertainty in engineering work becomes particularly

relevant when design criteria are based on return periods of very rare events. Because they are rare the precision with which they are described is poor and cannot be ignored when large investments and consequences are at stake. A distinction should be made between inherent variation and uncertainty due to a lack of knowledge. Furthermore, an effort should be made to use stochastic models in return period analysis to reduce bias resulting from inappropriate assumptions during parameter estimation and to avoid underestimation of the frequencies of extreme occurrences due to the exclusion of certain inherent random behaviour.

SAMENVATTING

Sinds de aanleg van rioolwaterzuiveringsinstallaties zijn overstortingen uit gemengde rioolstelsels in toenemende mate beperkend geworden voor de kwaliteit van de omliggende oppervlaktewateren. In de loop der jaren zijn de stedelijke wateren dermate gewijzigd door menselijke ingrepen dat het waterkwaliteitsbeheer nu een integrale benadering vereist met betrekking tot het functioneren van het systeem en de te investeren kosten. De gevolgen van op korte termijn werkende verontreinigingen, zoals zuurstofuitputting door overmatige belasting met organische stoffen, dienen te worden geevalueerd op basis van hun herhalingstijd. Modellen van het betreffende watersysteem worden dan gebruikt om zowel de huidige herhalingstijd te berekenen alswel die na uitvoering van voorgenomen verbeteringsprojecten. Bij de traditionele deterministische modellering van gemengde rioolstelsels is er een neiging geweest om steeds meer nieuwe processen aan de modelstruktuur toe te voegen in een poging om de gemaakte voorspellingen te verbeteren. Dit leidde dikwijls tot veel parameters met onbekende waarde en het opnemen van processen die veel minder van belang zijn dan andere, die onbekend zijn of slecht begrepen. Een alternatieve benadering is om alleen de meest wezenlijke processen te beschrijven en de resterende variatie in stochastische termen op te nemen.
De onderhavige studie concentreert zich op een vergelijking en tegenoverstelling van de deterministische en stochastische benaderingen in het modelleren van de verontreiniging door afvloeiing uit het stedelijke gebied en van de resulterende waterkwaliteit. Ook methodologieën rond modeltoepassingen om herhalingstijden te analyseren en de onzekerheid daarin te bepalen zijn bestudeerd.

Een multivariate analyse van de over de overstortings gebeurtenissen gemiddelde concentraties werd gemaakt van de gegevensbestanden van drie nederlandse en twee deense gemengde rioleringsgebieden (Hoofdstuk 3). Dit werd gedaan om de onderliggende struktuur van variaties in de per gebeurtenis gemiddelde concentraties te achterhalen. De resultaten bevestigden de verwachting dat de meest uitgesproken voorkomende variaties betrekking hebben op de groepen particulaire en opgeloste verontreinigingen. De distributie van de meest bepalende factoren herbevestigde vroegere bevindingen dat bimodale of gemengde distributies de gemiddelde concentraties van particulaire stoffen kenmerken.

Niet-lineaire modellen werden ontwikkeld om per overstort gebeurtenis het totale volume en de gemiddelde concentraties daarin te voorspellen als functie van variabelen die de neerslag kenmerken (Hoofdstuk 4). Het doel hievan was om het basis begrip van het fysieke systeem te combineren met de informatie die in de meetgegevens schuilt. Het overstort volume werd goed beschreven met een vochtigheidsindex. De seizoen variatie die in eerste instantie uit de gegevens naar voren kwam en vervolgens werd beschreven met een empirische sinusvormige uitdrukking toonde een opmerkelijke overeenstemming met gegevens van de maandgemiddelde open water verdamping. Gebruik van deze open water verdampings gegevens als een invoer grootheid van het model verbeterde de voorspelling terwijl tegelijkertijd het aantal model parameters verminderde. De per gebeurtenis samengevoegde neerslag gegevens konden slechts in zeer beperkte mate de variatie in de per gebeurtenis gemiddelde concentraties verklaren. Daardoor konden sommige waterkwaliteits variabelen alleen door hun waarschijnlijkheids distributie worden gekenmerkt.

Een analyse van de achterliggende vooronderstellingen die worden gemaakt bij de wiskundige modellering in de tijd van water systemen leidde tot een nieuw beeld van de essentiële verschillen tussen deterministisch en stochastisch modelleren (Hoofdstuk 5). De impliciete aanname bij deterministisch modelleren is dat het model een perfecte beschrijving geeft van de werkelijkheid en dat alle afwijkingen tussen gemodelleerde en waargenomen waarden het gevolg zijn van waarnemingsfouten. Bij stochastisch modelleren is de impliciete aanname dat het model slechts een gedeeltelijke beschrijving van de werkelijkheid geeft en dat het verschil tussen gemodelleerde en waargenomen waarden mede wordt veroorzaakt door onverklaard toevallig gedrag van het systeem. Door de kenmerkende

kernverschillen tussen deterministische en stochastische modellen zo te identificeren ontstaat de mogelijkheid tot meer uitwisseling van methoden en benaderingen hiertussen.
Daarmee verbetert het modelleren van water en van waterkwalitieit. Kennis van de dominante fysische, chemische en biologische processen in het betreffende water systeem kan worden verwerkt in de gewoonlijk empirische stochastische modellen. Parameter statistiek, concepten van "experimental design", empirische elementen en begrippen betreffende identificeerbaarheid kunnen worden toegepast in deterministische modellen. Kwantitatieve a priori kennis omtrent bepaalde model parameters kan worden opgenomen in de schattings procedure. In de verdere toekomst is het de bedoeling dat de te kiezen benadering meer afhangt van het passend zijn van de gemaakte vooronderstellingen (mede gezien de beschikbare hulpbronnen en de mogelijke gevolgen van het gebruik van een pover model) dan van de achtergrond van de modelleur, zoals nu vaak het geval is. De parameters van een neerslag-afvoer model voor een gemengd rioolstelsel werden zowel in een deterministische als een stochastische versie geschat, teneinde de belangrijkste punten uit dit hoofdstuk te bestuderen en te illustreren.

Gebruik makend van stochastische differentiaalvergelijkingen werden water-kwantiteits en kwaliteits modellen voor een gemengd rioolstelsel geformuleerd en hun parameters geschat (Hoofdstuk 6). Het doel was de mogelijkheden en beperkingen vast te stellen van deze benadering waarbij het rioolstelsel wordt gedefinieerd door een set differentiaal vergelijkingen die stochastisch in het tijdsdomein worden opgelost. Voor het kwantiteitsmodel bleek parameter schatting mogelijk en een heel kleine waarnemingsfout bevestigde de toepasselijkheid van de stochastische benadering. De resultaten van het modelleren van de water kwaliteit duidden er op dat meer werk nodig is om ten volle de mogelijkheden en beperkingen van deze benadering te waarderen.

Een niet-lineair "random" coefficient model voor de beschrijving van het gesuspendeerde chemisch zuurstof verbruik (CZV) in een gemengd rioolstelsel werd geïdentificeerd en de parameters daarin werden geschat (Hoofdstuk 7). Bij "random" coefficient modellering wordt aangenomen dat bepaalde, geselecteerde parameters van gebeurtenis tot gebeurtenis verschillende waarden kunnen aannemen. Voor elke gebeurtenis afzonderlijk wordt deze waarde geschat. In de onderhavige studie werd een critische stroming, waarboven resuspensie geleidelijk toeneemt, verondersteld een "random" coefficient te zijn. Hoewel er een tekort is aan gegevens voor de periode voorafgaande aan de overstorting, suggereren de resultaten toch dat er goede mogelijkheden zijn voor de toepassing van random coefficient modellering van stedelijke vuilwater systemen; zowel als een alternatief voor alsook in combinatie met stochastische modellering.

De bestudering van de methoden en benaderingen beschreven in de voorgaande hoofdstukken worden in een breder perspectief besproken waarbij tevens de aandacht wordt gericht op een aantal interessante ontwikkelingen op het terrein van het modelleren van water kwantiteit en kwaliteit (Hoofdstuk 8). Door onze fysische, chemische en biologische theorie te structureren in stochastische toestands modellen wordt erkend dat de afwijking tussen model en waarneming het gevolg is van zowel onverklaard willekeurig gedrag van het gemodelleerde systeem als van waarnemingsfouten. Deze vaststelling zal de systematische afwijking in de geschatte parameters verminderen en derhalve het vermogen van het model tot voorspelling verbeteren, ook onder nieuwe condities. Hoewel stochastische toestandsmodellen die Kalman filters gebruikten vooral in de hydrologie en de waterkwaliteits modellering hun intrede deden in de late 70-er jaren, betroffen deze voornamelijk empirische formuleringen die geheel op meetgegevens waren gebaseerd. Daardoor waren deze van weinig nut voor de ingenieur die alternatieve scenarios wilde beoordelen en vergelijken. Om te vermijden dat modellen te veel parameters bevatten die in hoge mate uitwisselbaar zijn, is het van belang dat hun modelstruktuur kan worden geïdentificeerd op basis van de gegevens die werden gebruikt om de parameters te schatten. De a posteriori schattings criteria die kwantitatieve a priori kennis insluiten, vormen een interessante geformaliseerde methode om de intuitie en ervaring van de ingenieur te introduceren in de parameter schattings procedure.

Een nieuwe methodologie om de mate van de onzekerheid in de analyse van herhalingstijden te bepalen wordt gepresenteerd en toegelicht in een geintegreerde benadering van de verontreiniging van stedelijk water met modellen van zowel het gemengde rioolstelsel als van het ontvangende oppervlakte water (Hoofdstuk 9). De achterliggende

vooronderstelling in de gepresenteerde methodologie is dat een onderscheid moet worden gemaakt tussen variaties die aan het systeem eigen zijn en onzekerheid die voortkomt uit gebrek aan kennis. Dit onderscheid werd verkregen door een ingebedde voortplanting van fouten, die hier geimplementeerd werd als een "Embedded Monte Carlo Simulation". Er wordt beargumenteerd dat het samenvoegen van onzekerheid en inherente variaties systematisch leidt tot een verhoging van de frequentie van extreme gebeurtenissen met frequentieverdelingen van herhalingstijden die weinig of geen waarde voor de ingenieur hebben. Het onderzoek toont ook dat aandacht nodig is om snellere alternatieven te implementeren dan de ruwe Monte Carlo simulaties om de rekentijd te bekorten. Dit zou voor praktische toepassing nodig zijn.

Een inspectie van methoden om herhalingstijden te analyseren voor de vervuiling van stedelijke wateren en de onzekerheid daarin werd uitgevoerd met de bedoeling om de in Hoofdstuk 9 gepresenteerde methodes in een breder perspectief te beschouwen (Hoofdstuk 10). Drie principieel onderscheiden methoden om herhalingstijden van bepaalde effecten te berekenen worden beschreven: direkt fitten; transformatie van momenten en analytische of numerieke integratie. De combinatie van deze methoden met verschillende model typen (in termen van input en output als tijdreeksen of als "lumped" grootheden), leidt tot een raamwerk waarin de meeste benaderingen tot de analyse van herhalingstijden zijn begrepen. In ingenieurs werk wordt onzekerheid bijzonder belangrijk wanneer ontwerp criteria worden gebaseerd op de frequentie van zeldzame, extreme gebeurtenissen. Juist omdat zij zeldzaam zijn is de precisie waarmee ze worden beschreven gering, hetgeen niet kan worden genegeerd wanneer het gaat om grote investeringen en belangrijke consequenties. Een onderscheid dient te worden gemaakt tussen inherente variatie en onzekerheid als gevolg van ontbrekende kennis. Daarnaast dient te worden geprobeerd om stochastische modellen te gebruiken in de analyse van herhalingstijden om systematische afwijkingen tengevolge van ongeschikte aannames bij de parameterschatting te beperken. Ook de onderschatting van de frequentie van extreme gebeurtenissen ten gevolge van het niet meerekenen van bepaald systeem-eigen toevallig gedrag dient te worden vermeden.

About the Author

CURRICULUM VITAE

Morten Grum was born on the 16th December 1965 in Copenhagen, Denmark. From the age of 2½ he has grown-up in Africa with his parents, brothers and sister in Kenya, The Gambia and Swaziland. In 1986 he obtained his International Baccalaureate diploma at Waterford Kamhlaba United World College of Southern Africa, Swaziland. In 1991, he completed his B.Sc. in civil engineering at the Engineering Academy of Denmark with research work on "Biogas in the Tropics" involving experimental work both in the laboratory and on full scale biogas plants in Tanzania. These studies included a seven months practical period as site engineer and assistant laboratory engineer in the road construction company Jean Lefebvre in Lyon, France. Morten Grum pursued his studies at the Technical University of Denmark with a focus on environmental water quality engineering and mathematical statistics. He completing his M.Sc. thesis on the topic of "Stochastic Modelling of Urban Drainage Systems" jointly at the Department of Environmental Science and Technology and at the Institute for Mathematical Statistics and Operations Research in 1993. In the mid 1994, after a years civilian national civil service working with statistics and accident modelling at the Danish Road Safety Research Council, Morten Grum began research on "Stochastic Properties of Urban Runoff Pollution" at the then Department of Water Quality Management and Aquatic Ecology, Agricultural University of Wageningen. The research work, which was carried out in the period from mid 1994 until the mid 1998 and which is presented in this thesis, formed part of the EU sponsored research MATECH network with partners at universities in several European countries. Subsequently, Morten Grum work as research assistant at the Department of Environmental Science and Technology, Technical University of Denmark. He was engaged with the incorporation of complex biofilm kinetics into stochastic models for use in real time control. Since late 1999, he has been a full time employ at the water-engineering consultants PH-Consult in Copenhagen where he among other is engaged in research where stochastic state space modelling is used in order to assimilate a wide variety of rainfall and flow times series data from various parts within the hydrological system.

LIST OF PUBLICATIONS

Scientific and Technical Journals

Grum, M. and Aalderink, R. H. (1998). Uncertainty in Return Period Analysis of Combined Sewer Overflow Effects using Embedded Monte Carlo Simulations. Presented at Aquatech '98, Amsterdam: Conference on application of models in water management. *Wat. Sci. Tech.* 39 (4), pp. 233-240.

Grum, M. (1997). Incorporating Concepts from Physical Theory into Stochastic Modelling of Urban Drainage Runoff Pollution. In the proceedings of 2. International Conference on "Sewer as a physical, chemical and biological reactor" in Aalborg, Denmark, May 25-28, 1997. *Wat. Sci. Tech.* 37 (1), pp. 179-185.

Grum, M. and Aalderink, R. H. (1997). A Statistical Approach to Urban Runoff Pollution Modelling. In the proceedings of 4. International Symposium on "Systems Analysis and Computing in Water Quality Management" WATERMATEX '97, Quebec City, Canada. *Wat. Sci. Tech.* 36 (5) pp. 117-124

Grum, M., Aalderink, R. H., Lijklema, L. and Spliid, H. (1996). The Underlying Structure of Systematic Variations in the Event Mean Concentrations of Pollutants in Urban Runoff. In the proceedings of 7. International Conference on Urban Storm Drainage, Hannover. *Wat. Sci. Tech.* 36 (8-9), pp. 135-140.

Moens, M., Grum, M. and Aalderink, H. (1999). Inzet van troebelheidsmeters ter bepaling van CZV en drogestof in gemengde rioolstelsel. *H 2 O* 32 (8), pp. 22-25.

Professional Magazines and Conferences

Grum, M. (1997). WATERMATEX '97: A Symposium Review. *Water Quality International*, Sept./Oct. 1997, pp 37-39.

Grum, M. and Aalderink, R. H. (1997). Trends and Traditions at the 7. International Conference on Urban Storm Drainage, Hannover 1996. *European Water Pollution Control*, Vol. 7, No. 1, pp 69-71.

Grum, M., Puznava, N. and Harremoës, P. (1999). Modelling of simultaneous nitrifiction-denitrification for design and predictive control. In: Conference on biofilm systems. New York, October 17-20, 1999. Program and proceedings, IAWQ.

Harremoës, P., Falkentoft, C.M. and Grum, M. (1999). Biofilm kinetics and reactor operation for nutrient removal. Invited lecture. In: 18th Federal Convention, Australian Water and Wastewater Association, Adelaide, April 1999. Proceedings. CD-ROM, Australian Water & Wastewater Association, Artarmon, Australia.

APPENDIX A

EQUATIONS OF THE LINEAR KALMAN FILTER

This appendix contains the equations of the linear Kalman filter as presented in Madsen, H and Holst, J. (1998). Modelling Non-linear and Non-stationary Time Series. Lecture notes used at Institute for Mathematical Modelling Technical University of Denmark, University of Copenhagen, Copenhagen Business School, and the University of Iceland. 284 pp., www.imm.dtu.dk. Detailed outline of the linear Kalman filter can also be found in Harvey. A. C. (1993). Forecasting, Structural Time Series Models and the Kalman Filter. 310 pp.

MODEL

System equation:

$$\underline{X}_t = \underline{\underline{A}}\,\underline{X}_{t-1} + \underline{\underline{B}}\,\underline{u}_{t-1} + \underline{e}_{1,t} \tag{A.1}$$

Observation equation:

$$\underline{Y}_t = \underline{\underline{C}}\,\underline{X}_t + \underline{e}_{2,t} \tag{A.2}$$

KALMAN FILTER

Up-date or reconstruction

$$\underline{\hat{X}}_{t|t} = \underline{\hat{X}}_{t|t-1} + \underline{\underline{K}}_t \left(\underline{Y}_t - \underline{\underline{C}}\,\underline{\hat{X}}_{t|t-1} \right) \tag{A.3}$$

$$\underline{\underline{\Sigma}}^{XX}_{t|t} = \underline{\underline{\Sigma}}^{XX}_{t|t-1} - \underline{\underline{K}}_t\,\underline{\underline{\Sigma}}^{YY}_{t|t-1}\,\underline{\underline{K}}^{\mathbf{T}}_t \tag{A.4}$$

where the Kalman gain is given by

$$\underline{\underline{K}}_t = \underline{\underline{\Sigma}}^{XX}_{t|t-1}\,\underline{\underline{C}}^{\mathbf{T}} \left(\underline{\underline{\Sigma}}^{YY}_{t|t-1} \right)^{-1} \tag{A.5}$$

Prediction

$$\underline{\hat{X}}_{t+1|t} = \underline{\underline{A}}\,\underline{\hat{X}}_{t|t} + \underline{\underline{B}}\,\underline{u}_t \tag{A.6}$$

$$\underline{\underline{\Sigma}}^{XX}_{t+1|t} = \underline{\underline{A}}\,\underline{\underline{\Sigma}}^{XX}_{t|t}\,\underline{\underline{A}}^{\mathbf{T}} + \underline{\underline{\Sigma}}_1 \tag{A.7}$$

$$\underline{\hat{Y}}_{t+1|t} = \underline{\underline{C}}\,\underline{\hat{X}}_{t+1|t} \tag{A.8}$$

$$\underline{\underline{\Sigma}}^{YY}_{t+1|t} = \underline{\underline{C}}\,\underline{\underline{\Sigma}}^{XX}_{t+1|t}\,\underline{\underline{C}}^{\mathbf{T}} + \underline{\underline{\Sigma}}_2 \tag{A.9}$$

Innovation

$$\underline{\tilde{Y}}_{t+1|t} = \underline{Y}_{t+1} - \underline{\hat{Y}}_{t+1|t} \tag{A.10}$$

$$\underline{\underline{R}}_{t+1} = \underline{\underline{\Sigma}}^{YY}_{t+1|t} \tag{A.11}$$

Maximum Likelihood Estimation Criteria

$$\log L(\theta; Y_N) = -\frac{1}{2}\sum_{i=1}^{N}\left[\log\left(\det \underline{\underline{R}}_i\right) + \underline{\tilde{Y}}_i^{\mathrm{T}}\underline{\underline{R}}_i^{-1}\underline{\tilde{Y}}_i\right] + \text{constant} \tag{A.12}$$

SYMBOLS

$\underline{X}_t$ is a vector containing the state variables at time t,

$\underline{u}_t$ is a vector containing the values of the input variables at time t,

$\underline{\underline{A}}$ is the system matrix,

$\underline{\underline{B}}$ is the input matrix,

$\underline{e}_{1,t}$ is the system noise terms at t which is assumed to be normally distributed with mean zero and variance $\underline{\underline{\Sigma}}_1$ (a diagonal matrix),

$\underline{Y}_t$ is the observation at time t,

$\underline{\underline{C}}$ is an observation matrix indicating the observed state variables,

$\underline{e}_{2,t}$ is the observation (measurement) error at time t which is assumed to be normally distributed with mean zero and variance $\underline{\underline{\Sigma}}_2$ (a diagonal matrix),

$\underline{\underline{K}}_t$ is the Kalman amplification at time t.

APPENDIX B

EQUATIONS OF THE EXTENDED KALMAN FILTER

This appendix contains the equations of the discrete time extended Kalman filter that has formed the basis for the modelling in Chapter 5. The filter equations are from Madsen, H and Holst, J. (1998). Modelling Non-linear and Non-stationary Time Series. Lecture notes used at Institute for Mathematical Modelling Technical University of Denmark, University of Copenhagen, Copenhagen Business School, and the University of Iceland. 284 pp., www.imm.dtu.dk. Details of the Kalman filtering can also be found in Harvey. A. C. (1993). Forecasting, Structural Time Series Models and the Kalman Filter. 310 pp.

MODEL

System equation:

$$\underline{X}_t = \underline{f}\left(\underline{X}_{t-1}, \underline{u}_{t-1}\right) + \underline{e}_{1,t} \tag{B.1}$$

Observation equation:

$$\underline{Y}_t = \underline{h}\left(\underline{X}_t\right) + \underline{e}_{2,t} \tag{B.2}$$

KALMAN FILTER

Up-date or reconstruction

$$\hat{\underline{X}}_{t|t} = \hat{\underline{X}}_{t|t-1} + \underline{\underline{K}}_t \left(\underline{Y}_t - \underline{h}\left(\hat{\underline{X}}_{t|t-1}\right)\right) \tag{B.3}$$

$$\underline{\underline{\Sigma}}^{XX}_{t|t} = \underline{\underline{\Sigma}}^{XX}_{t|t-1} - \underline{\underline{K}}_t \, \underline{\underline{\Sigma}}^{YY}_{t|t-1} \, \underline{\underline{K}}^{\mathbf{T}}_t \tag{B.4}$$

where the Kalman gain is given by

$$\underline{\underline{K}}_t = \underline{\underline{\Sigma}}^{XX}_{t|t-1} \, \underline{\underline{H}}^T\left(\underline{X}_{t|t-1}\right)\left(\underline{\underline{\Sigma}}^{YY}_{t|t-1}\right)^{-1} \tag{B.5}$$

where $\underline{\underline{H}}\left(\underline{X}_{t|t-1}\right)$ is the partial derivatives of $\underline{h}\left(\underline{X}\right)$ with respect to $\underline{X}$.

Prediction

$$\hat{\underline{X}}_{t+1|t} = \underline{f}\left(\hat{\underline{X}}_{t|t}, \underline{u}_t\right) \tag{B.6}$$

$$\underline{\underline{\Sigma}}^{XX}_{t+1|t} = \underline{\underline{F}}\left(\hat{\underline{X}}_{t|t}, \underline{u}_t\right) \underline{\underline{\Sigma}}^{XX}_{t|t} \, \underline{\underline{F}}^T\left(\hat{\underline{X}}_{t|t}, \underline{u}_t\right) + \underline{\underline{\Sigma}}_1 \tag{B.7}$$

where $\underline{\underline{F}}\left(\hat{\underline{X}}_{t|t}, \underline{u}_t\right)$ is the partial derivatives of $\underline{f}\left(\underline{X}, \underline{u}\right)$ with respect to $\underline{X}$.

$$\underline{Y}_{t+1|t} = \underline{h}\left(\hat{\underline{X}}_{t+1|t}\right) \tag{B.8}$$

$$\underline{\underline{\Sigma}}^{YY}_{t+1|t} = \underline{\underline{H}}\left(\hat{\underline{X}}_{t+1|t}\right) \underline{\underline{\Sigma}}^{XX}_{t+1|t} \, \underline{\underline{H}}^T\left(\hat{\underline{X}}_{t+1|t}\right) + \underline{\underline{\Sigma}}_2 \tag{B.9}$$

Innovation

$$\underline{\tilde{Y}}_{t+1|t} = \underline{Y}_{t+1} - \underline{\hat{Y}}_{t+1|t} \tag{B.10}$$

$$\underline{\underline{R}}_{t+1} = \underline{\underline{\Sigma}}^{YY}_{t+1|t} \tag{B.11}$$

Maximum Likelihood Estimation Criterion

$$\log L(\theta; Y_N) = -\frac{1}{2}\sum_{i=1}^{N}\left[\log\left(\det \underline{\underline{R}}_i\right) + \underline{\tilde{Y}}_i^T \underline{\underline{R}}_i^{-1} \underline{\tilde{Y}}_i\right] + \text{constant} \tag{B.12}$$

SYMBOLS

$\underline{X}_t$ is a vector containing the state variables at time t,

$\underline{u}_t$ is a vector containing the values of the input variables at time t,

$\underline{f}(\,)$ is the system function and $\underline{h}(\,)$ is the observation function (see model definition at top),

$\underline{e}_{1,t}$ is the system noise terms, normally distributed with mean zero and variance $\underline{\underline{\Sigma}}_1$,

$\underline{Y}_t$ is the observation at time t,

$\underline{e}_{2,t}$ is the observation (measurement) , normally distributed with mean zero and variance $\underline{\underline{\Sigma}}_2$,

$\underline{\underline{K}}_t$ is the Kalman amplification at time t.