Uncertainty in prediction and simulation of flow in sewer systems

Breinholt, Anders; Mikkelsen, Peter Steen; Madsen, Henrik; Grum, Morten

Publication date:
2012

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

Technical University of Denmark

Uncertainty in prediction and simulation of flow in sewer systems

Anders Breinholt

DTU Environment
Department of Environmental Engineering

PhD Thesis
May 2012

Uncertainty in prediction and simulation of flow in sewer systems

Anders Breinholt

PhD Thesis
May 2012

DTU Environment
Department of Environmental Engineering
Technical University of Denmark

Anders Breinholt

Uncertainty in prediction and simulation of flow in sewer systems

PhD Thesis, May 2012

The thesis will be available as a pdf-file for downloading from the homepage of the department: www.env.dtu.dk

Address:	DTU Environment Department of Environmental Engineering Technical University of Denmark Miljoevej, building 113 DK-2800 Kgs. Lyngby Denmark
Phone reception:	+45 4525 1600
Phone library:	+45 4525 1610
Fax:	+45 4593 2850
Homepage:	http://www.env.dtu.dk
E-mail:	reception@env.dtu.dk

Printed by:	Vester Kopi Virum, May 2012
Cover:	Torben Dolin
ISBN:	978-87-92654-60-1

Preface

This PhD thesis was prepared at the Technical University of Denmark (DTU) from 2007 to 2011 under the supervision of Associate Professor Peter Steen Mikkelsen (Department of Environmental Engineering, DTU Environment), Professor Henrik Madsen (Department of Informatics and Mathematical Modelling, DTU Informatics) and Dr. Morten Grum (Krüger A/S, a subsidiary of Veolia Water Solutions & Technologies). The PhD project was co-funded by DTU Environment, Krüger A/S and the Ministry of Science, Technology and Innovation through the graduate school for Urban Water Technology (UWT), with supplementary financial support from the Council for Strategic Research through the Storm and Wastewater Informatics (SWI) project.

The thesis consists of a summary report and a collection of six papers, which have been published or submitted to international peer reviewed journals or conferences. In the thesis the papers are referred to by their roman number, e.g. as **Paper I**.

Papers included in the thesis:

Paper I. Breinholt, A., Santacoloma, P.A., Mikkelsen, P.S., Madsen, H., Grum, M., Nielsen, M.K. *Evaluation framework for control of integrated urban drainage systems*, In: 11ICUD, Proceedings of 11th International Conference on Urban Drainage, Edinburgh, Scotland, 31st August-5th September 2008.

Paper II. Breinholt, A., Grum, M., Madsen, H. , Thordarson, F.Ø., Mikkelsen, P.S. *Informal uncertainty analysis (GLUE) of continuous flow simulation in a hybrid sewer system with infiltration inflow - consistency of containment ratios in calibration and validation?*, submitted.

Paper III. Breinholt, A., Thordarson, F.Ø., Møller, J.K., Grum, M., Mikkelsen, P.S., Madsen, H. *Grey-box modelling of flow in sewer systems with state dependent diffusion*, Environmetrics, Vol. 22 (8), pp. 946-961, 2011 (DOI: 10.1002/env.1135).

Paper IV. Thordarson, F.Ø., **Breinholt, A.**, Møller, J.K., Mikkelsen, P.S. Grum, M., Madsen, H. ,*Evaluation of probabilistic flow predictions in sewer systems using grey box models and a skill score criterion*. Stochastic Environmental Research and Risk Assessment. (in press)(DOI: 10.1007/s00477-012-0563-3).

Paper V. Breinholt, A., Møller, J.K., Madsen, H., Mikkelsen, P.S. *A formal statistical approach to representing uncertainty in rainfall-runoff modelling with focus on residual analysis and probabilistic output evaluation - distinguishing simulation and prediction*, submitted.

Paper VI. Breinholt, A., Thordarson, F.Ø., Møller, J.K., Grum, M., Mikkelsen, P.S., Madsen, H. *Identifying the appropriate physical complexity of stochastic gray-box models used for urban drainage flow prediction by evaluating their point and probabilistic forecast skill*, submitted.

Other publications:

The following papers and reports were also prepared during the project period. The scientific content is covered, or partly covered by the included papers. Therefore, these publications are not included in the monograph.

Breinholt, A., Grum, M., Madsen, H. , Mikkelsen, P.S. *Uncertainty Analysis of Storm- and Wastewater models*, 8UDM & 2RWHM. The 8th International Conference on Urban Drainage Modelling. The 2nd International Conference on Rainwater Harvesting and Management, 7-12 September, 2009, Tokyo, Japan, Proceedings.

Hansen, L.S, Borup, M., **Breinholt, A.**, Mikkelsen, P.S. *Performance of MOUSE UPDATE for level and flow forecasting in urban drainage systems*, MIKE by DHI International Conference Copenhagen 2010," Modelling in a World of Change" 6-8 September, 2010, DHI, Hørsholm, published in proceedings/book.

Breinholt, A., Sharma, A.K. *Case Area Baseline Report - Århus Public Water Utility*[1], Technical Report, DTU Environment. Department of Environmental Engineering. March 2010.

Breinholt, A., Sharma, A.K. *Case Area Baseline Report - Copenhagen Energy and Lynette Fællesskabet*[2], Technical Report, DTU Environment. Department of Environmental Engineering. March 2010.

Breinholt, A., Sharma, A.K. *Case Area Baseline Report - Avedoere Wastewater Services*[3], Technical Report, DTU Environment. Department of Environmental Engineering. July 2009.

[1] http://www.swi.env.dtu.dk/upload/swi/caserapport%20%C3%A5rhus%20270310.pdf
[2] http://www.swi.env.dtu.dk/upload/swi/casereport%20ke_lf%20270310.pdf
[3] http://www.swi.env.dtu.dk/upload/swi/case%20area%20baseline%20report%20aved%C3%B8re%20100327.pdf

The papers are not included in this www-version, but can be obtained from the Library at DTU Environment: Department of Environmental Engineering Technical University of Denmark Miljoevej, Building 113 2800 Kongens Lyngby, Denmark (library@env.dtu.dk)

Acknowledgements

First of all I would like to thank my supervisor Associate Professor Peter Steen Mikkelsen (DTU Environment) for valuable and highly appreciated support and contribution to this thesis.

Also many thanks go to my co-supervisor Professor Henrik Madsen (DTU Informatics) for many significant contributions to the statistical and modelling challenges and also for lending me a chair at DTU Informatics.

I would also like to thank my co-supervisor Dr. Morten Grum (Krüger A/S) for valuable comments and contributions to many of the papers included in this thesis.

I am very grateful for the cooperation with my fellow PhD student Fannar Örn Thordarson (DTU Informatics) who introduced me to CTSM and helped me out with various Latex and Matlab problems and for his willingness to discuss the statistical subjects. Also I am thankfull to Assistant Professor Jan Kloppenborg Møller (DTU Informatics) for his willingness to help with CTSM and to answer questions and discuss possible new approaches to solve the statistical and modelling problems that I came across during my PhD. I would also like to thank Associate Professor Niels Kjølstad Poulsen (DTU Informatics) for beeing willing to answer questions regarding control theory and for setting up a model platform for control of urban drainage system although it was not finalised within the project.

I am also very grateful for the data and support I received from Spildevandscenter Avedøre (SCA) and especially I would like to thank Jacob Nørremark for his support.

Summary

Models are commonly applied for design of urban drainage systems. Typically they are of deterministic nature although it is well accepted that they only reflect reality approximately. When measurements are available they can be used for calibration of models. However, deviations between model outputs and observations will often remain and should hence be quantified, especially when used for model predictive control.

The objective with this thesis has been to quantify and qualify the modelled output uncertainty. For this purpose a catchment in Ballerup (1,320 hectares) was selected and data included flow from downstream the catchment, rain measured at two rain gauges and monthly evaporation. The data period covered subperiods of 2007-2010. The catchment area consists of both combined and separated drainage systems and significant infiltration inflow enters the system through permeable surface areas. The simple serial linear reservoir flow routing principle was applied for modelling both the fast rainfall runoff from paved areas and the slow infiltration inflow from permeable areas. The wastewater flow variation was modelled by a harmonic function. Models of different complexity in terms of describing features such as flow constraints, basins and pumps were tested for their ability to describe the output with a time resolution of 15 minutes.

Two approaches to uncertainty quantification were distinguished and adopted, the stochastic and the epistemic method. Stochastic uncertainty refers to the randomness observed in nature, which is normally irreducible due to the inherent variation of physical systems. Epistemic uncertainty on the contrary arises from incomplete knowledge about a physical system. For quantifying stochastic uncertainties a frequentist approach was applied whereas the generalised likelihood uncertainty estimation method (GLUE) was adopted for the epistemic approach. Two different uncertainty estimates were furthermore distinguished: prediction and simulation uncertainty. To quantify the prediction uncertainty the model should accommodate an updating step thereby benefitting from observations that arrive in continuation of the predictions made. The simulation uncertainty on the other hand is calculated from data of a limited measuring campaign and the model does not accommodate a model correction step. The stochastic approach was applied for uncertainty quantification in both prediction and simulation whereas the epistemic uncertainty was assessed only in simulation. A maximum likelihood method was applied for parameter estimation in the stochastic approach, i.e. one optimal parameter set was

derived that minimises the errors between model outputs and observations. Conversely in GLUE, parameters are viewed as stochastic variables and many acceptable parameter sets were therefore identified.

The predictive stochastic models were built on stochastic differential equations that include a drift term containing the physical description of the model and a diffusion term describing the uncertainty in the state variables. Additionally the observation noise is accounted for by a separate observation noise term. This approach is also referred to as stochastic grey-box modelling. A state dependent diffusion term was developed using a Lamperti transformation of the states, and implemented to compensate for heteroscedastic state uncertainty and to avoid predicting negative states. A flow proportional observation noise term introduced by a log transform was furthermore used to avoid predicting negative flows.

In the simplest stochastic prediction models all parameters were estimated easily; however increasing the deterministic model complexity involved that some of the parameters had to be fixed. The statistical assumptions that require the residuals to correspond to a white noise process were fulfilled for the one-step prediction but beyond the one-step prediction auto-correlated residuals were obtained. The Akaike's (AIC) and the Bayesian (BIC) information criteria were used to identify preferred models for the one-step prediction whereas a skill scoring criterion addressing both the reliability and the sharpness of the confidence bounds was used when assessing the forecasting performance beyond the one-step. The reliability was satisfied for the one-step prediction but were increasingly biased as the prediction horizon was expanded, particularly in rainy periods.

GLUE was applied for estimating uncertainty in such a way that the selection of behavioral parameter sets continued until a required coverage of observations was obtained (targeting 90%). A likelihood measure were used for ranking the parameter sets and two different ways of drawing parameter sets were tested, a Latin Hypercube Monte Carlo method and a modified Monte Carlo Markov Chain method. When using the stochastic models for simulation, it was found that the simulation uncertainty was best described when estimating parameters by the output error minimisation method. In order to remove the heteroschedastic residuals structure it were necessary to apply a transformation of the observations, however autocorrelation remained in the simulation case. A skill scoring comparison of a simulation and a prediction model showed that a major improvement is gained by updating the model states continuously, i.e. updating of model states results in much lower

forecasting uncertainty at shorter prediction steps.

In the GLUE methodology there are no requirements to the residuals. Nevertheless the aim is the same as for the stochastic simulation models, namely to cover a proportion of observations consistent with the considered quantile with maximum sharpness, i.e. to minimise the skill score. In one calibration case, even though very broad prior parameter ranges were specified, it was difficult to acquire a 90% coverage of observations and the reliability in rainy periods was much lower than in dry weather. However the GLUE method proved quite consistent in the sense that similar coverage rates were obtained in both calibration and validation periods with the same set of retained parameter sets.

A comparison of the stochastic and epistemic approaches to uncertainty evaluation was conducted by comparing the sharpness, the reliability and the skill score on the same set of data. Very similar performance was obtained with the stochastic method as the preferred. The thesis has demonstrated that the statistical requirements to the formal stochastic approach are very hard to fulfill in practice when prediction steps beyond the one-step is considered. Thus the underlying assumption of the GLUE methodology, that uncertainty in modeling and simulation is not only of stochastic nature, seems fairly consistent with the results of this thesis.

A major drawback of the GLUE methodology as applied here is the lumping of total uncertainty into the parameters, which entails a loss of physicality of the model parameters. Conversely the parameter estimates of the stochastic approach are physically meaningful. This thesis has contributed to developing simplified rainfall-runoff models that are suitable for model predictive control of urban drainage systems that takes uncertainty into account.

Dansk sammenfatning

Modeller anvendes hyppigt til design af regn- og spildevandssystemer i urbane oplande. Sådanne modeller er normalt deterministiske omend det er velkendt, at modellerne sjældent afspejler virkeligheden perfekt. I de tilfælde hvor målinger er tilgængelige kan modellerne kalibreres. Men ofte vil der være afvigelser som ignoreres, og denne usikkerhed bør reelt kvantificeres hvilket særligt gælder i forbindelse med styringer der omfatter modelprædiktioner.

Formålet med afhandlingen har således været at finde metoder til at kvantificere og kvalificere modellernes output usikkerheder. Til dette formål blev et caseopland beliggende i Ballerup (1.320 ha) udvalgt, og flowdata målt nedstrøms fra oplandet blev anvendt til modellering og usikkerhedskvantificering. Input til modellerne omfattede regn fra to regnmålere samt månedlige fordampningsdata. Den samlede dataperiode udgjorde 2007-2010. Caseoplandet rummer både fælles og separat kloakerede oplande og er påvirket af uvedkommende infiltrationsvand fra permeable områder. Det lineære reservoir princip blev benyttet til modellering af regnafstrømning og infiltration og en harmonisk funktion beskrev spildevandsflowet. Derudover antog modellerne forskellig kompleksistet med hensyn til beskrivelsen af flowbegrænsninger, overløb samt pumper. Modellernes output blev sammenholdt med observationerne med en tidsopløsning på 15 minutter.

Der er skelnet mellem to overordnede tilgange til usikkerhedskvantificering, den stokastiske og den epistemiske . I den stokastiske verden opfattes usikkerheder som tilfældige, irreducerbare og som kvantificerbare størrelser, mens usikkerheder i den epistemiske tilgang beror på utilstrækkelig viden om et givent fysisk system. Til kvantificering af usikkerheder benyttes henholdsvis en frekventistisk statistisk metode til at repræsentere den stokastiske tilgang, og GLUE metoden (Generalised Likelihood Uncertainty Estimation) som repræsentant for den epistemiske tilgang. Derudover blev skelnet mellem usikkerhedsbestemmelse i simulation og prædiktion. Hvilken type usikkerhedsbestemmelse der kan estimeres vil afhænge af om der er on-line målinger tilgængelige for løbende opdatering/korrektion af modellen eller ej. Prædiktionsusikkerheden bestemmes ved løbende opdatering mens simulationsusikkerheden bestemmes ud fra en afgrænset måleperiode. Den stokastiske metode blev anvendt til bestemmelse af usikkerheden i både prædiktion og simulation, mens den epistemiske kun blev anvendt til bestemmelse af usikkerheden under simulation. I den stokastiske metode benyttes maksimum likelihood estimation til at finde det optimale parametersæt dvs. til at minimere fejlen mellem

modellens output og observationerne. I GLUE derimod opfattes parametersæt som stokastiske variable og mange egnede parametersæt findes.

I den stokastiske metode formuleres modellen ved hjælp af stokastiske differentialligninger som indeholder to vigtige led, et driftsled og et diffusionsled. Driftsledet rummer den fysiske beskrivelse af modellen mens diffusionsledet beskriver usikkerheden i tilstandsvariablene. Derudover findes et separat usikkerhedsled der beskriver observationsusikkerheden. Den benyttede stokastiske metode er i denne afhandling også benævnt stokastisk grey-box modellering. For at tage højde for tilstandsafhængig usikkerhed blev et tilstandsafhængigt diffusionsled implementeret ved hjælp af Lamperti transformation hvilket sikrede positive tilstandsvariable. Tilsvarende sikrede et flowafhængigt observationsled positive flows.

I de simple modeller kunne alle parametre estimeres, men med voksende modelkompleksitet blev det vanskeligere at estimere alle parametrene hvorfor nogle parametre måtte fikseres. De statistiske forudsætninger for de stokastiske modeller indebærer at residualerne skal svare til en hvid støjproces. Dette var en rimelig antagelse for et-trins prædiktionen men antagelsen var ikke opfyldt for flertrinsprædiktionen samt ved simulation pga. autokorrelerede residualled. Informationskriterierne Akaike's (AIC) og det Bayesianske (BIC) blev benyttet til udvælgelse af den mest velegnede model til et-trins prædiktionen, mens et interval skill score kriterium blev anvendt til udvælgelse af den mest velegnede model til flertrins prædiktioner samt simulation. Interval skill scoren sammenholder den faktiske procent af observationerne der falder indenfor et givent konfidensbånd med den forventede (reliabiliteten), og tager højde for konfidensbåndenes bredde. Reliabiliteten for konfidensbåndene viste god overenstemmelse for et-trinsprædiktionen men reliabiliteten faldt med voksende prædiktionshorisont og var generelt dårligere i regnvejr end i tørvejr.

I GLUE metoden bestemmes usikkerheden ved udvælgelse af parametersæt indtil en passende dækning af observationer opnås (f.eks. 90%). Rangordningen af og valget af egnede parametersæt blev foretaget ved hjælp af et uformelt likelihood mål, og udtrækning af parametersæt skete på to måder dels via en latin hypercube Monte Carlo metode, og dels med en Markov Chain Monte Carlo metode.

Det viste sig at den mest optimale simulationsmodel var en model hvor den samlede outputfejl blev minimeret, og hvor det var nødvendigt at transformere måledata for at fjerne de heteroskedastiske residualled. Autokorrelationen kunne dog

ikke fjernes, og kan næppe fjernes i en simulationsmodel, hvorfor det statistiske grundlag ikke var helt opfyldt. Parametrene var generelt signifikante i simulationsmodellerne men enkelte parametre måtte fikseres. En sammenligning af skill scoren for en simulationsmodel og en prædiktionsmodel viste, at en væsentlig forbedring i prædiktionsevne opnås ved løbende opdatering af modellen.

I GLUE metoden er der ingen krav til residualerne men målet er det samme, nemlig at opnå simulationsintervaller med høj reliabilitet og smalle usikkerhedsbånd. I ét tilfælde viste det sig vanskeligt at opnå 90% dækning af observationerne, på trods af brede prior parameterintervaller. Generelt var reliabiliteten meget dårligere i regnvejr end i tørvejr. Men metoden var målt på reliabilitet forholdsvist konsistent mellem kalibrerings- og valideringsperioderne.

De to metoder blev endvidere aftestet på det samme datagrundlag med hensyn til reliabilitet, bredde af konfidensbånd samt skill score, og gav nogenlunde samme resultat, omend den stokastiske metode alligevel var bedre målt på skill scoren. En fordel ved den stokastiske metode er at den kan bruges til at udlede parameterværdier.

De statistiske forudsætninger for den stokastiske metode kunne kun overholdes for et-trins prædiktionen. Under fler-trins prædiktionen og under simulation er der signifikant autokorrelation. På dette grundlag må det konkluderes at usikkerheden ikke kun er stokastisk men at den også beror på utilstrækkelig viden sådan som tilhængerne af den epistemiske anskuelse hævder. Denne afhandling har bidraget til udvikling af simple modeller som er velegnet til brug for modelprædiktiv styring af afløbssystemer når usikkerheden indregnes.

Contents

Preface . . . . . **i**

Acknowledgements . . . . . **v**

Summary . . . . . **vii**

Dansk sammenfatning . . . . . **xi**

1 Introduction **1**
1.1 Modelling within urban drainage engineering 1
1.2 Sources and approaches to uncertainty quantification in rainfall-runoff modelling 1
1.3 Distinguishing simulation and prediction models with focus on control 2
1.4 Key research aims 4
1.5 Outline 6

2 Literature review **9**
2.1 Approaches to uncertainty evaluation of model outputs and parameters 9
2.2 Data assimilation and model complexity 11
2.3 Benchmarking models and their uncertainty performance 13

3 Case study **15**
3.1 Catchment 15
3.2 Data 15

4 Simplistic deterministic sewer flow modelling **19**
4.1 Lumped conceptual modelling of sewer flow 19
4.2 Overview of simplistic models applied 20

5 The stochastic approach to uncertainty evaluation **27**
5.1 Stochastic grey box models 27
5.2 Maximum likelihood parameter estimation 28
5.3 Seeking an appropriate diffusion term 29
5.4 Seeking an appropriate observation noise term 32
5.5 The transformed grey box model with transformed observations . . . 35
5.6 Generating predictions and uncertainty limits using stochastic grey box models 35

6 The epistemic approach to uncertainty evaluation 37
6.1 Introducing stochastic parameters 37
6.2 The pseudo-Bayesian approach 37
6.3 Choosing a likelihood measure and the behavioral parameter sets . . 38
6.4 Searching the parameter space 39
6.5 Generating uncertainty limits 40

7 Benchmarking models and uncertainty approaches 43
7.1 Residual analysis 43
7.2 Probabilistic prediction and simulation measures 43
7.2.1 Reliability bias 44
7.2.2 Sharpness 44
7.3 Model performance comparison 45
7.3.1 Evaluation using information criteria 45
7.3.2 Evaluation using an interval skill score criterion 45

8 Results and discussion 47
8.1 Results of the stochastic approach 47
8.1.1 Checking the statistical assumptions 47
8.1.2 Selecting among model candidates by information criteria . . 49
8.1.3 Evaluating prediction and simulation bounds 50
8.2 Results of the epistemic approach 56
8.2.1 Extraction of behavioural parameter sets 56
8.2.2 Parameter uncertainty 57
8.2.3 Consistency of the GLUE generated bounds 58
8.3 Comparison of the stochastic and epistemic approaches 59

9 Conclusions 63

10 Outlook 67

11 References 69

12 Papers 77

1 Introduction

1.1 Modelling within urban drainage engineering

Models of urban drainage systems serve many important purposes, some of which are listed below:

- Design of drainage systems.
- Evaluation of an existing drainage systems performance by checking that the statutory requirements are met.
- Investigating upgrading or redesign proposals.
- Check where flooding of basements and terrain will occur.
- Investigating the consequences of climate change.
- Modelling of pollution discharges.
- For real-time control of pumps, gates, orifices, weirs and waste water treatment plants to optimise the mututal performance.

The models applied for the listed purposes are normally deterministic models assuming to model reality perfectly. However due to the presence of various uncertainties these models should not be expected to reflect reality perfectly. In this thesis different methods to quantify output uncertainty will be addressed in the case when data are available for model output comparison, and rainfall-runoff (RR) modelling will serve as illustrative examples.

1.2 Sources and approaches to uncertainty quantification in rainfall-runoff modelling

Although many uncertainty typologies exist each serving a specific purpose, in this thesis the focus is on uncertainty in model-based decision support, and uncertainty is recognised as any departure from the ideal of complete determinism, a definition also adopted by Walker et al. (2005).

It is generally accepted that errors and biases (or uncertainty) in RR modelling results from the following sources (Refsgaard et al., 2007; Renard et al., 2010; Peel and Blöschl, 2011):

1. Input uncertainty due to sampling and measurement errors and inadequate spatio-temporal rainfall variability coverage.
2. Output observation uncertainty, i.e. inaccurate flow/level measurements.

3. Model structural uncertainty due to incomplete understanding and simplified descriptions of modelled processes as compared to reality.

4. Parameter uncertainty, i.e. uncertainty related to parameter values.

A thorough account of references to each of these sources is given in Peel and Blöschl (2011). Within RR modelling many different approaches to uncertainty evaluation exists (Matott et al., 2009) but uncertainty can be broadly classified as *stochastic* or *epistemic*. Stochastic uncertainty refers to the randomness observed in nature, which is normally irreducible due to the inherent variation of physical systems. Epistemic uncertainty on the contrary arises from incomplete knowledge about a physical system (Refsgaard et al., 2007; Fu et al., 2011; Beven et al., 2011). In this thesis both a stochastic and epistemic approach to uncertainty evaluation will be adopted and compared.

1.3 Distinguishing simulation and prediction models with focus on control

In general we should distinguish models suitable for prediction from models suitable for simulation. A model tailored for long-term simulations should describe the important long-term phenomena of the system whereas a model tailored for prediction or forecasting (referred to also as a real-time- or on-line model) accommodates an updating step thereby benefitting from observations that normally arrive in immediate continuation of the predictions made. Due to this continuous updating/correction of the model a simple model structure will often suffice for predicting the short term (Carstensen et al., 1998; Dorado et al., 2003).

Simulation models are used for all the purposes listed in Section 1.1 while prediction models are relevant only in connection with real-time warning or control systems for urban drainage systems. Simulation models are often used to test and compare different control strategies (in a deterministic way) but outputs generated from simulation models are also normally used for model predictive control. In both cases the uncertainty of the output plays a significant role. An untapped model predictive control potential is identified in many sewer systems and generally considered to be a cheap alternative to traditional storage solutions (**Paper I**), but as the model predictions are uncertain, there is a risk that a wrong decision may be taken on the basis of the model predictions. In general predictive uncertainty can be defined as the uncertainty that a decision maker has on the future evolution

of a predictand that he uses to trigger a specific decision (Beven, 2009a). The requirement is then to provide predictions at the lead time of interest with minimum uncertainty, and to quantify this uncertainty.

Suppose we want to switch on some storm water control strategy at the WWTP in order to increase the hydraulic capacity if the predicted flow exceeds the current capacity of the WWTP. Assuming it takes 2-3 hours before the plant reaches its maximum hydraulic capacity, then this is our desired lead time. In sewer systems with long response time we may find sufficient lead time inside the sewer system by monitoring the level of a storage basin, pumping station or the flow in a large intercepting pipe, and use this information to trigger the storm water control. However, it may also well be that these measurements are insufficient for predicting the hydraulic load to the WWTP or that such measurements are unavailable in real time, and a model then will be needed.

If the sewer system does not facilitate sufficient lead time because of a fast system response time in the catchment (small catchment and/or steep slopes), it becomes necessary to extend the lead time by the use of model predictions from rain gauge input, or if this also provides too short lead time, from forecasting the rain input e.g. using radars. But as indicated in Figure 1.1 increasing the lead time by forecasting the rain input normally also entails that the model predictions become much more uncertain, and hence the risk that we may make a wrong decision will increase. Achleitner et al. (2009) used radar forecasts to extend the lead time and found uncertainty on rain volume increased up to some hundred percent for a lead time of 3 hours. Depending on the costs of making a wrong decision we would be more or less willing to increase the lead time and the risk of making a wrong decision. Considering again the WWTP control example two wrong decisions could be taken: (1) switching to wet weather control without the need occurring or (2) not switching to wet weather control but with the need occurring. Such a decision should essentially be subjected to risk analysis. In the first case the costs would be increased outlet concentrations of nutrients and organic substances (and probably also increased energy costs) for a prolonged period with associated extra tax expenses. In the second case the WWTP would be unprepared and wastewater would have to be bypassed without treatment with large impacts for the recipient. There may also be a model predictive control potential in optimising the utilisation of internal up-stream storage in the sewer system, and lead time will then be required for predicting the inlet to the storage tanks or a pumping station (see Figure 1.1)

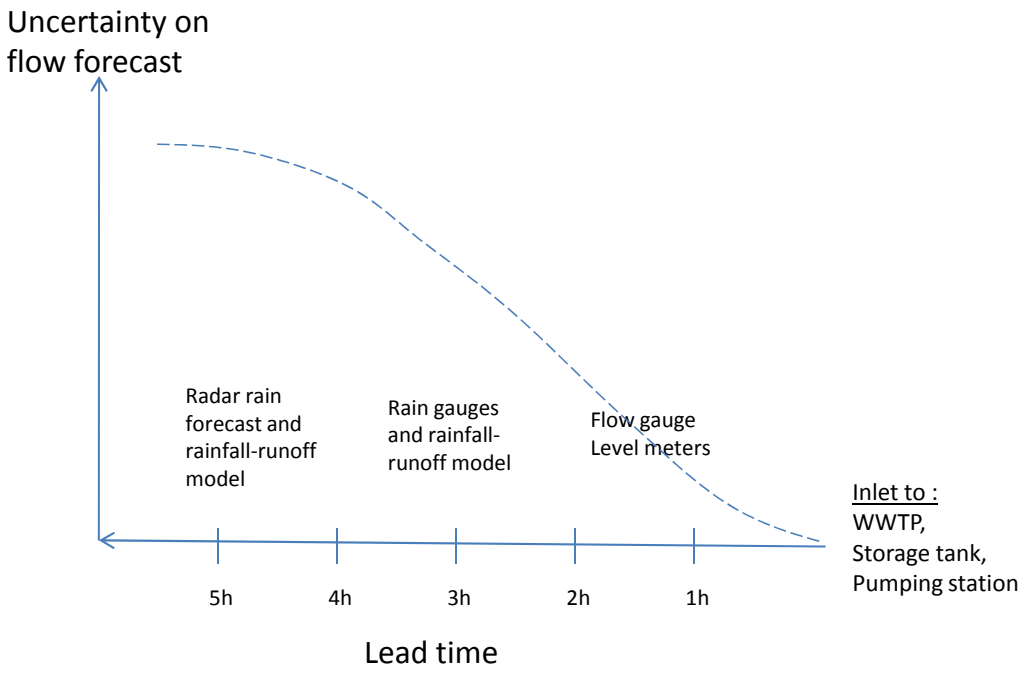

Figure 1.1: Uncertainty generally increases with lead time. Redrawn from (**Paper I**). See text for details.

which could be obtained from forecasted rain inputs as already discussed.

1.4 Key research aims

Because of the significant role uncertainty plays in model predictive control of sewer systems this thesis is dedicated to qualifying and quantifying uncertainty in RR modelling notably in simulating and predicting flows in sewer systems. The thesis covers the specific aims listed below.

Model development:

- Development of simple RR models for simulation and prediction of flow rates in sewer systems that are capable of describing the output uncertainty satisfyingly.

The stochastic grey-box approach for prediction:
The chosen stochastic approach is a frequentist approach that is based on stochastic differential equations formulated in state-space (Kristensen et al., 2004a,b) that

uses the maximum likelihood estimation method for parameter estimation. The system states that contains both a drift term and a diffusion noise term is continuously updated according to the measurements and the estimation method is therefore optimised to describe the one-step prediction error well. The model development involved the following steps:

- Testing of parameter significance.
- Developing suitable diffusion terms and observation noise terms.
- Checking that the actual residuals conform to the model assumed to define the likelihood function and suggesting model improvements on the basis of deviations from the assumptions.
- Using statistical criteria for evaluating model performance and model comparison.
- Testing of the model's suitability to describe the uncertainty when the prediction horizon is expanded beyond the one-step ahead prediction with focus on confidence bounds, i.e. probabilistic predictions rather than point predictions.
- Testing how the derived confidence bounds perform overall and in dry weather and wet weather periods, respectively.
- Applying a skill scoring criterion to select the best model among several model candidates when the prediction horizon is expanded beyond the first step.

The stochastic approach for simulation:
The simulation model is not continuously updated, because it is intended for long term investigations or because data are unavailable for real time updating. The simulation model development involved the following steps:

- Testing of parameter significance.
- Developing suitable observation and diffusion noise terms.
- Checking that the actual residuals conform to the model assumed to define the likelihood function and suggest model improvements on the basis of deviations from the assumptions.
- Testing the performance of the confidence bounds.

- Applying a skill scoring criterion to select the best model among several simulation model candidates.

The epistemic approach to simulation:
The chosen epistimic approach is based on the Generalised Likelihood Uncertainty Estimation (GLUE) approach (Beven and Binley, 1992). When applying this methodology the following items had particular focus:

- Defining a likelihood measure and a criterion for pinpointing the behavioural parameter sets, aiming to cover a large fraction of the observations.
- Choosing a method for sampling the model space to identify the behavioral parameter sets.
- Examining the assumption that behavioural parameter sets deduced from a calibration period enables the derivation of reasonable uncertainty limits in a validation period by checking of observation coverage.

Comparison of the stochastic and epistemic approach for simulation:
The two approaches to uncertainty evaluation are finally compared and the differences discussed by:

- Using a skill score criterion for comparison of the two uncertainty approaches when applied to the same data.
- Considering the underlying assumptions of each method.

1.5 Outline

This summary report includes nine chapters. Following this introduction, a litterature review is provided in Chapter 2 that presents an overview of the uncertainty methods commonly applied within RR modelling, some updating techniques are described and the level of model complexity discussed. Finally the litterature review overviews some of the typical benchmarking tools applied for model comparison. Then the catchment and data that underlie the research is presented in Chapter 3. Chapter 4 gives an introduction to the simplistic modelling concept that was pursued in all the papers and also reviews the different models applied for either simulation or prediction in a deterministic setting. Chapter 5 outlines the stochastic approach to uncertainty evaluation and Chapter 6 the epistemic approach. An

overview of the benchmark-indicators used for prediction and uncertainty assessment are provided in Chapter 7 and Chapter 8 presents the results and discusses the two approaches to uncertainty assessment. Subsequently, the conclusions of this thesis are drawn in Chapter 9 and some future research perspectives are outlined in Chapter 10.

2 Literature review

Assuming we are interested in realising a model predictive control potential and therefore want to quantify the uncertainty associated with some model prediction how should we approach that? A natural starting point is obviously to consult the literature on this aspect.

2.1 Approaches to uncertainty evaluation of model outputs and parameters

As mentioned in Chapter 1 uncertainty can be broadly classified as *stochastic* or *epistemic*. These distinct interpretations of uncertainty are reflected in two major methodologies used for uncertainty evaluation, on one hand we have the *formal statistical methods*, and on the other hand we have the *Generalized Likelihood Uncertainty Estimation (GLUE) methodology* (Beven and Binley, 1992; Beven, 2006; Beven et al., 2011). Within the formal statistical inference to uncertainty evaluation two fundamentally different approaches to the estimation problem are distinguished, the *frequentist* (or classical) approach that normally searches for a single optimal parameter set, and the *Bayesian* approach that allow probabilities to be associated with the unknown parameters (Gallagher and Doherty, 2007; Dotto et al., 2009). Another important difference is that the Bayesian approach requires a prior distribution of the parameters whereas the frequentist approach does not.

The Bayesian approach typically involves a Markov Chain Monte Carlo (MCMC) method (Engeland et al., 2005; Yang et al., 2007; Dotto et al., 2011; Schoups and Vrugt, 2010) with the DiffeRential Evolution Adaptive Metropolis (DREAM) scheme (Vrugt et al., 2009b,c; Vrugt, 2011) as the current state of the art for estimating the posterior parameter distribution. Both approaches apply a likelihood function that is based on formal statistical assumptions about model residuals and normally that they correspond to a white noise process (Dotto et al., 2011; Yang et al., 2008). A comparison between a frequentist approach and a Bayesian approach based on the Metropolis-Hastings algorithm showed that very similar parameter uncertainty and confidence bands are found when the residuals are stationary and ergodic (Engeland et al., 2005).

The GLUE method rejects the concept of an optimum model and parameter set, and instead acknowledges the existence of multiple likely models and parameter sets (in GLUE termed equifinality). In contrast to the stochastic approach GLUE also reject the use of statistical likelihood functions because they overestimate the

information content in a set of calibration data and increase the possibility of statistical type 1 and type 2 errors (Beven et al., 2011). Instead GLUE allows for the use of informal likelihoods (or fuzzy measures or likelihood measures) and treats residual errors implicitly in making predictions. The GLUE methodology has been criticized for being statistically incorrect and for generating prediction limits without statistical coherence (Mantovan and Todini, 2006; Mantovan et al., 2007; Stedinger et al., 2008; Vrugt et al., 2009a; Clark et al., 2011). This is due to the subjectivity in adopting a subjective likelihood measure, and in the choice of using a subjective threshold value to distinguish "behavioral" from "non-behavioral" parameter sets. In response hereto advocators of GLUE (Andréassian et al., 2007; Beven et al., 2007, 2008; Beven, 2009b; Beven et al., 2011) claim that the assumptions required for formal statistical analysis hardly ever are satisfied within hydrological modelling due to epistemic errors that leads to non-stationary model-residuals that are unsuitable for statistical likelihood functions.

In most GLUE applications all the uncertainty sources outlined in Chapter 1 are lumped into the parameters and GLUE will generally give the largest parameter uncertainty compared with the stochastic approaches. Comparisons of the Bayesian and the GLUE methodology for uncertainty evaluation of outputs and parameters (Jin et al., 2010; Li et al., 2010) suggest that the two methods can, given certain requirements to the cut-off threshold value for choosing the behavioural parameter sets, give more or less the same simulation output uncertainty. It is important to recognise that a GLUE analysis can equally well be carried out using MCMC algorithms to speed up the search for behavioural parameter sets (Blasone et al., 2008; Lindblom et al., 2011; Vezzaro and Mikkelsen, 2012), however in that case by replacing a formal likelihood with an informal likelihood measure.

Currently research in trying to unravel the individual sources to uncertainty is ongoing and the BATEA (Bayesian total error analyis) method (Kavetski et al., 2006; Thyer et al., 2007; Renard et al., 2010) is one such stochastic tool. According to Renard et al. (2010) this can be useful for (1) diagnosing the main causes of uncertainty, suggesting avenues for improving the predictive precision of RR models; (2) identifying RR model deficiencies indicating opportunities for model improvement; and (3) comparing RR models without obscuring the comparison by input/output data errors. However, even though the method can be shown to make statistically reliable inference and meaningfully disaggregate multiple sources of uncertainty, in practice, if no independent estimates of data uncertainties is avail-

able, the discrepancy between observed and simulated responses only provides information about total errors. Attempts to separate uncertainty sources from one another has also been investigated within GLUE (Liu et al., 2009; Westerberg et al., 2010; Krueger et al., 2010).

2.2 Data assimilation and model complexity

When using a prediction model for uncertainty quantification there is the need for correcting (updating) the model in real-time when a new observation is recieved. According to Refsgaard (1997) there are four variables that can be used either separately or in combination for updating. These are updating in inputs or parameters, output error prediction/correction and state updating. The most common of these are state updating and output error updating, where the error series is modelled in stochastic terms and this is used to improve the forecasts (Romanowicz et al., 2008).

A warning concerning the aim with the prediction model is issued in Beven (2009a), that decisions normally will involve the N-step ahead lead time and not just the one-step ahead prediction, the residuals of which is commonly minimised during parameter estimation. In cases when a particular decision relies on the N-step ahead forecast, it is therefore of upmost importance to check that the derived uncertainty limits are reliable at those prediction steps and not just the one-step ahead forecast.

Within RR two main modeling philosophies can be distinguished, namely the physically-based models (or white-box models), and the data-driven models (or black-box models) (Todini, 2007). The physically-based models has been criticized for resulting in models that are overly complex, leading to problems of overparameterisation and equifinality, which may manifest itself in large prediction uncertainty. On the other hand the data-driven approach has been criticised for beeing too reliant on the training sets.

In the middle of this modelling spectrum a data-driven approach has been advocated, where complexity is added to the model only when it improves the description of the data, without using an a priori defined model structure. The idea is to arrive at models that are complex enough to explain the data, but not more complex than necessary, a strategy often referred to as Occam's razor or the principle of parsimony (Schoups et al., 2009; Todini, 2011). Such models are also referred to as grey-box models (Kristensen et al., 2004a,b) or databased-mechanistic

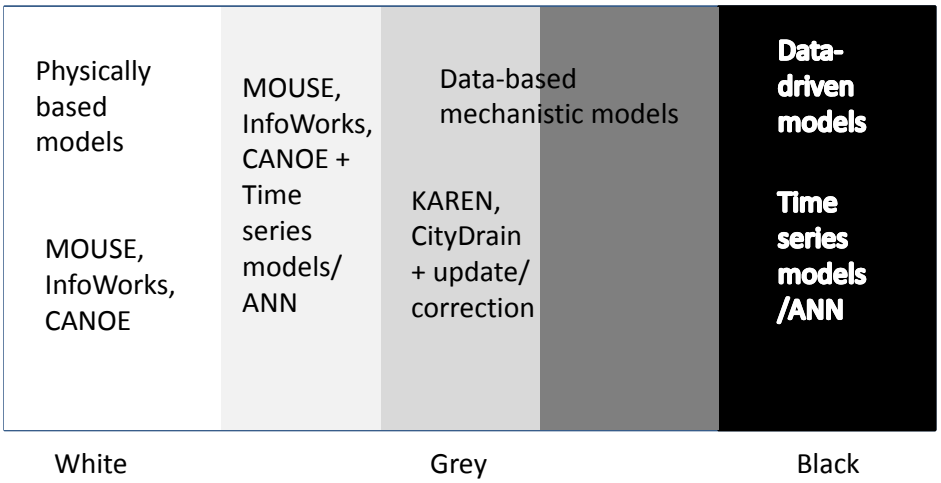

Figure 2.1: The modelling spectrum. White-box models are physically-based whereas black-box models utilise statistical methods and tools for estimating the model parameters and assessing the uncertainties. The two model approaches can be combined into a grey-box model with more or less white and black colour.

models (McIntyre et al., 2011), see the diagram in Figure 2.1 that picturise the modelling spectrum. Spatially distributed modelling is a typical example of the physically-based model, to construct a model that explicitly accounts for as much of the physics and the natural heterogeneity as computationally possible. Within urban drainage modelling we recognise this white-box model type as e.g. MOUSE, InfoWorks or CANOE that all build on the Saint-Venant equations for hydraulic calculation, see references to the models in Dotto et al. (2011). Such models are rarely applied in connection with model predictive control in real-time although currently the possibility of updating internal states are being investigated (Hansen et al., 2011). The possibility to fuse a white-box model with a black-box model using output error correction update has been pursued. Vojinovic et al. (2003) used a combination of a white-box model (MOUSE model) and a black-box model (a radial basis function neural network (RBFNN) model) as a stochastic error-correction model and obtained significant improvements in model predictions. Bruen and Yang (2006) used a full hydrodynamic model (HYDROWORKS, now called InfoWorks) together with different black box models (Artificial Neural Networks (ANNs) and linear time series models of Box et al. (2008)) to simulate and predict flow volume and attained significant improvement in overall efficiency. These output correction methods could equally well be applied using models of

less complexity with simplified conceptual hydrologic flow routing methods such as simple linear reservoirs as used in KAREN or Muskingum flow routing as used in CityDrain (Dotto et al., 2011). Another possibility would be to update the states of one of these more simple models using state-space filtering methods based on either the Kalman filter (KF), the extended Kalman filter (EKF) or the ensemble Kalman filter (EnKF) that updates through states and output errors. The Kalman filter methods demands assumptions to be made about the nature of the residuals and typically that they are Gaussian distributed, and various transformations are therefore normally necessary to facilitate this, i.e. the Normal Quantile Transform (NQT) or Box-Cox transformations (Beven, 2009a; Coccia et al., 2011).

Todini (2011) mentions the DBM (Data Based Mechanistic Modelling) approach of (Young and Garnier, 2006; Taylor et al., 2007; Young, 2011) as a tentative attempt to go beyond the black-box concept by selecting among the resulting model structures those that are considered physically meaningful. Alternatives without requirements to the residual distribution are the sequential Monte Carlo methods, also known as particle filters, that apply a Bayesian learning technique, and GLUE that can also be applied for assessing the uncertainties in real-time forecasting by recursively updating of the behavioural parameter sets. According to (Beven, 2009a) particle filters do have some limitations, they are computationally very expensive and exhibit problems with estimating the posterior distribution.

2.3 Benchmarking models and their uncertainty performance

To benchmark models and methods within RR modelling many more or less informative performance measures have been used. Deterministic model performance measures include the Nash & Sutcliffe efficiency coefficient (denoted E or NS) and the coefficient of determination R^2 (Dotto et al., 2011), the mean square error(MSE) (Achleitner et al., 2009) or root mean square error (RMSE). According to Franz and Hogue (2011) much of the hydrological modeling community is still performing model evaluation using standard deterministic measures such as those listed above, and these are deficient for fully analysing model performance and should be substituted with probabilistic assessment of model performance. Sometimes model performance is supported solely by hydrograph plots of a few events in which typically 95% or 90% uncertainty limits are plotted together with model observations (Yang et al., 2007; McMillan and Clark, 2009). When uncertainty limits or probabilistic forecasts are assessed the width (the sharpness) and the cov-

erage (reliability) of observations are calculated either for one or more uncertainty limits (Renard et al., 2010; Hostache et al., 2011), and the deterministic performance measures sometimes used to assess the performance of the median (Engeland et al., 2010). The NS has also been suggested for comparing upper and lower limits among models (Laloy et al., 2010). In a GLUE study Xiong et al. (2009) introduce seven indices for characterizing the prediction bounds from different perspectives and suggest that they be employed for assessing and comparing the uncertainty bounds in a more comprehensive and objective way. These indices takes the coverage, the width of the bounds and the symmetry of the bounds into account.

For probabilistic performance evaluation the discrete ranked probability score, that evaluates the squared difference between the cumulative distribution function of the forecast and the cumulative distribution function of a perfect forecast at some predefined thresholds, was applied by Morawietz et al. (2011). The Brier score was demonstrated by Engeland et al. (2010) as an uncertainty evaluation tool for model comparison. The Brier score is an attractive measure for quantifying performance of a probabilistic forecast since it combines the reliability, the resolution and the marginal distribution of a probabilistic forecast (Gneiting et al., 2007), however many other probabilistic evaluation methods exists (Gneiting and Raftery, 2007).

Quite commonly prediction limits or forecasting uncertainty are used in a context where they actually refer to simulation uncertainty. As noted by Andréassian et al. (2007) past misunderstandings on the uncertainty estimation issue would have been avoided if authors had clearly defined what type of model application they were discussing. That simulation and prediction models serve two different purposes was the subject of debate (Beven, 2009b; Vrugt et al., 2009d) following a paper by (Vrugt et al., 2009a) in which the prediction uncertainty of a formal statistical (Bayesian) approach was compared to the simulation uncertainty of an informal (GLUE) approach and used to conclude that the Bayesian approach gave smaller spread and higher coverage than the GLUE generated bounds. When making model- and/or uncertainty comparisons like these we should therefore compare like with like, that is, simulation models with simulation models and prediction models with prediction models (Beven, 2009b). To avoid misunderstandings the term "simulation uncertainty" will therefore be used when referring to uncertainty limits or confidence bounds generated by a simulation model and "prediction uncertainty" when generated from a prediction model.

3 Case study

3.1 Catchment

The catchment is located in the western part of greater Copenhagen in Ballerup Municipality, see Figure 3.1. The drainage system consists of both combined and separated sewer pipes, however most of the catchment is separated. Some catchment characteristics are given in Table 3.1. The drainage system facilitates a few detention basins and some pumping stations as indicated on Figure 3.1. When the system capacity is exceeded combined sewage overflows from the sewer system. A significant amount of infiltration inflow is occasionally entering the sewer system, and a flow meter has been installed downstream the catchment as indicated on Figure 3.1 to trace this infiltration contribution.

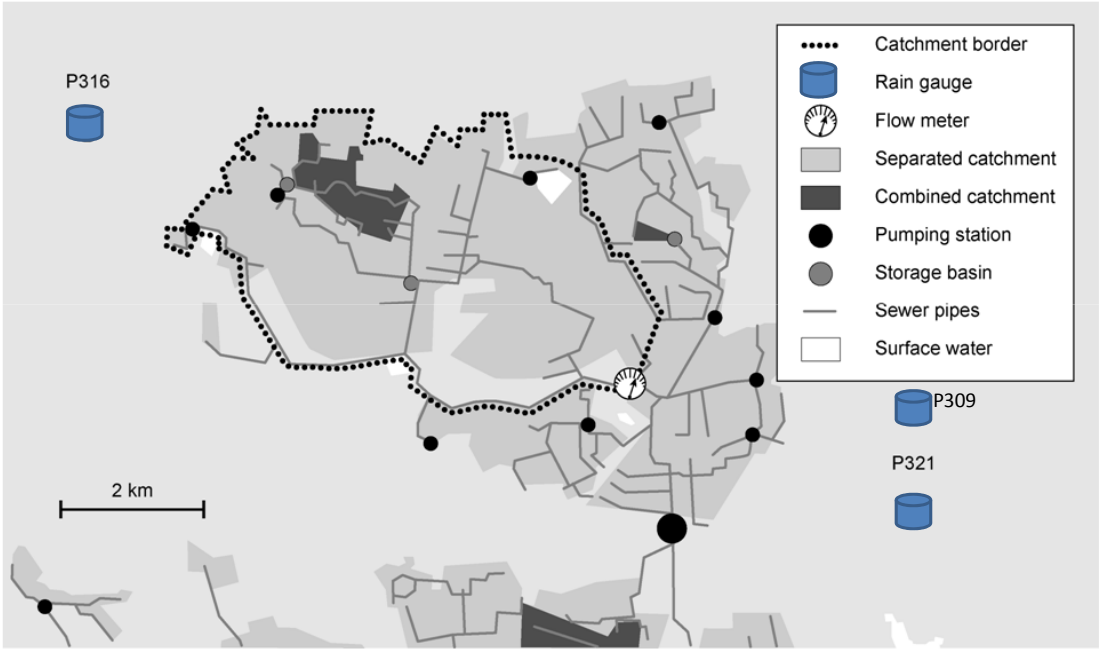

Figure 3.1: The Ballerup catchment area.

The three nearest rain gauges from the national Danish tipping bucket network (Jørgensen et al., 1998) are P_{321}, P_{316} and P_{309}, see Figure 3.1. All rain gauges are located outside the studied catchment.

3.2 Data

Flow data with a time resolution of 5 minutes was retrieved from two different flow meters. The flow data used in **Papers II**, **III** & **IV** were retrieved from a semi-mobile ultrasonic Doppler type which was considered to be less reliable for

Table 3.1: Catchment characteristics (Tomicic, 2006).

Ballerup Municipality	Total area		Imp.area	
	$[km^2]$	$[\%]$	$[km^2]$	$[\%]$
Combined	0.92	7	0.33	77
Separated	12.27	93	0.10	23
Total	13.20	100	0.43	100

measuring high flow rates. It was therefore replaced in 2010 with an Isco 2150 area-velocity type flow module that is considered more reliable especially during heavy rainfalls. Flow data from this flow meter was applied in **Papers V** & **VI**. All the flow data was retrieved from Avedøre Spildevandscenter.

In the **Papers II**, **III** & **IV** rain data from P_{321} and P_{316} were used whereas rain data from P_{309} replaced those from P_{321} in **Papers V** & **VI**, due to a longer outage period of P_{321}. The tipping bucket rain gauges have a resolution of 0.2 mm and rain data was retrieved from Danish Meteorological Institute (DMI).

Monthly evaporation data from 2007 was received from DHI, and applied in **Papers V** & **VI**.

The calibration period of **Papers II**, **III** & **IV** covered the period April-October 2007 whereas the validation period of **Paper II** encompassed April-October 2008 & 2009. In papers **Papers III** & **IV** the calibration period was also the validation period. In **Papers V** & **VI** the calibration and validation period was also identical covering June-September, 2010.

For application of the data in the models which run in 15-minute time steps the evaporation data was linearly interpolated between the months, rain data was summed in 15 minute windows, and every third flow measurement was used.

For discussion of the model results in **Papers II** & **IV** it was considered beneficial to distinguish dry from wet weather periods using an observed flow threshold of 540 m^3/h. In **Paper V** a flow threshold of 450 m^3/h separated dry from wet weather periods.

In all the models rain data from the two rain gauges are used for input and large spatial rainfall heterogeneity is clearly present. Figure 3.2 illustrates this by show-

ing the recorded rain event volume on a shifted log-scale from the two rain gauges P_{316} and P_{321}. The rain events were distinguished from one another by one hour dry periods.

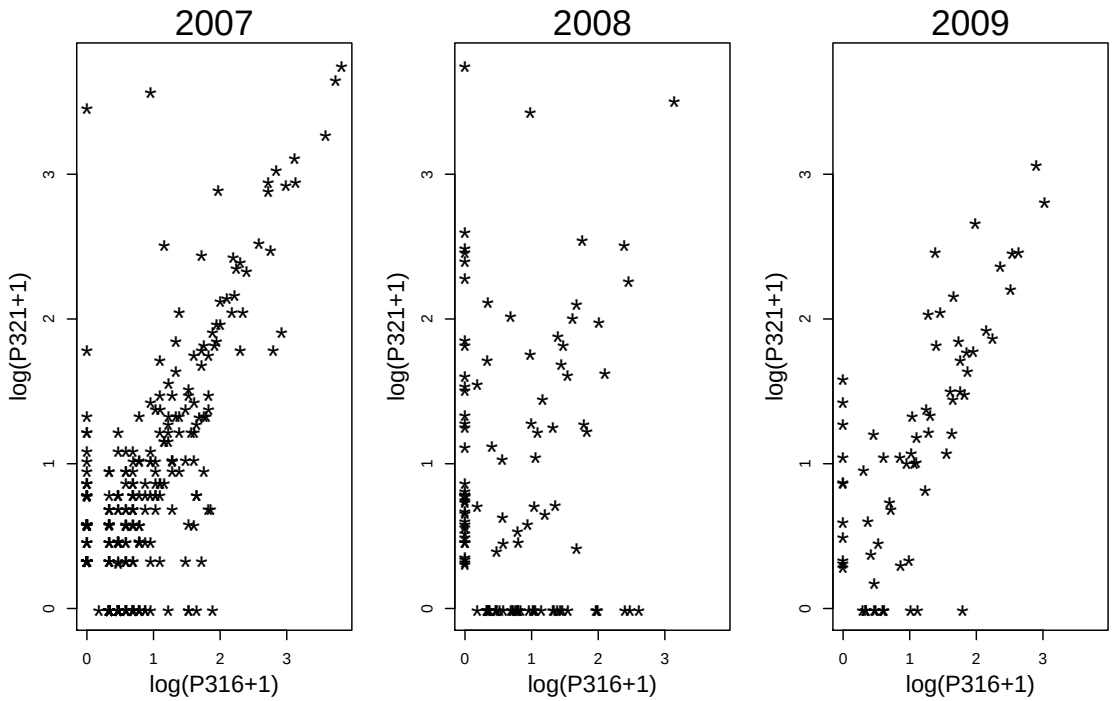

Figure 3.2: Rain events recorded at rain gauge P_{316} and P_{321} on a shifted log scale $[log(1 + acc.mm)]$, April-October, 2007-2009.

An example of the classification of flows into wet and dry weather periods are provided in Figure 3.3 and the slow infiltration inflow to the sewer system is recognised in the enhanced recession part of the hydrograph (08/24-09/04) shown in Figure 3.4.

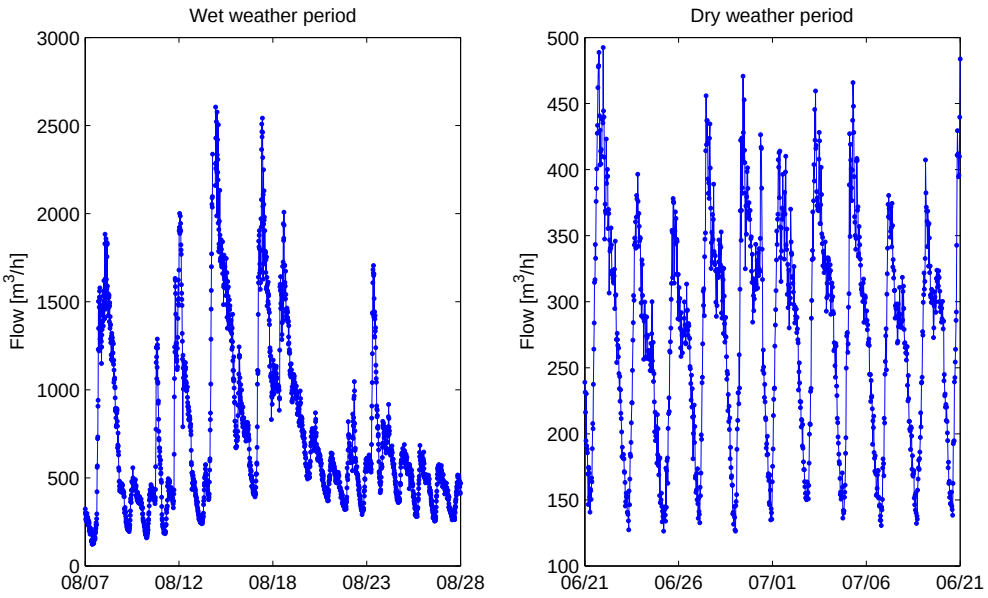

Figure 3.3: Illustration of the two flow subclasses dry and wet weather periods. The diurnal wastewater flow variation is clearly recognised. Flow data from 2010.

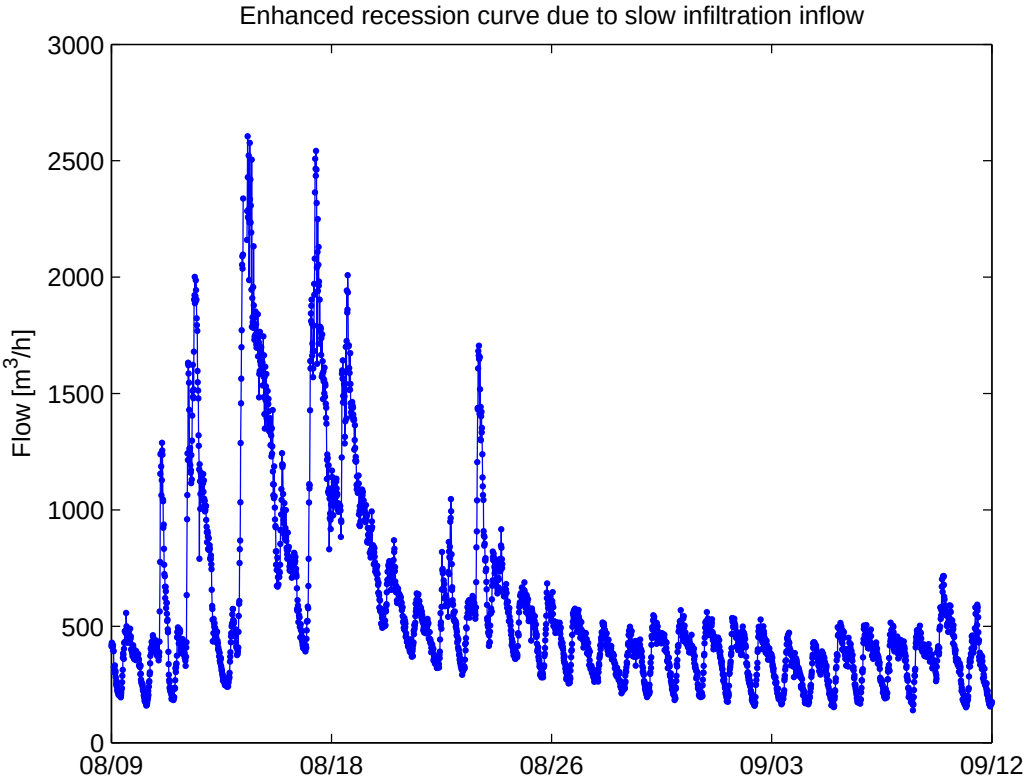

Figure 3.4: Infiltration inflow is recognised in the enhanced recession curve of the hydrograph. Flow data from 2010.

4 Simplistic deterministic sewer flow modelling

Since the aim with the modelling approach is to derive simple parsimonious models with identifiable parameters and ability to describe the output uncertainty well, a simple modelling approach was chosen and hence the stochastic grey-box modelling principle is adopted. However in this Chapter the models are described deterministically, and in the following two chapters 5 and 6 the two distinct approaches to uncertainty evaluation will be outlined.

4.1 Lumped conceptual modelling of sewer flow

For modelling the rainfall-runoff response the simple principle of flow routing of the Nash cascade is chosen which models flows in subcatchments by conceptually routing water through a series of reservoirs thereby achieving attenuation of the water wave in a simplistic manner compared to the dynamic wave of the Full Saint Venant equations or the kinematic/diffusive wave approximations (Schütze et al., 2002). This allows for rapid simulation but has the drawback that backwater effects cannot be directly simulated. Figure 4.1 shows the concept of the Nash cascade and illustrate how the outlet hydrograph change with the number of reservoirs in the cascade. From Figure 4.1 it appears that at least two reservoirs are necessary to mimic a typical rainfall-runoff hydrograph such as those presented in Figure 3.3 (left) and Figure 3.4. As the cascade is expanded with more reservoirs the time to peak flow is prolonged. The storage $S(t)$ $[m^3]$ in each reservoir can be described by the storage equation

$$S(t) = kQ(t), \tag{4.1}$$

where $Q(t)$ $[m^3/h]$ is the outflow and k $[h]$ is a storage constant. The continuity equation of the storage is given by

$$\frac{dS(t)}{dt} = I(t) - Q(t), \tag{4.2}$$

where $I(t)$ $[m^3/h]$ is the inflow and t is time. Assuming all reservoirs to have identical capacity, the time to flow peak is given by $t_p = (N-1)k$ and the time to half of the input volume has passed the reservoirs is given by $K = Nk$.

The system of linear reservoirs can be written on a state space form, where the nth state, S_n for $n = 2, \ldots, N$, is given by

$$\frac{dS_n(t)}{dt} = \frac{1}{k}S_{n-1}(t) - \frac{1}{k}S_n(t), \tag{4.3}$$

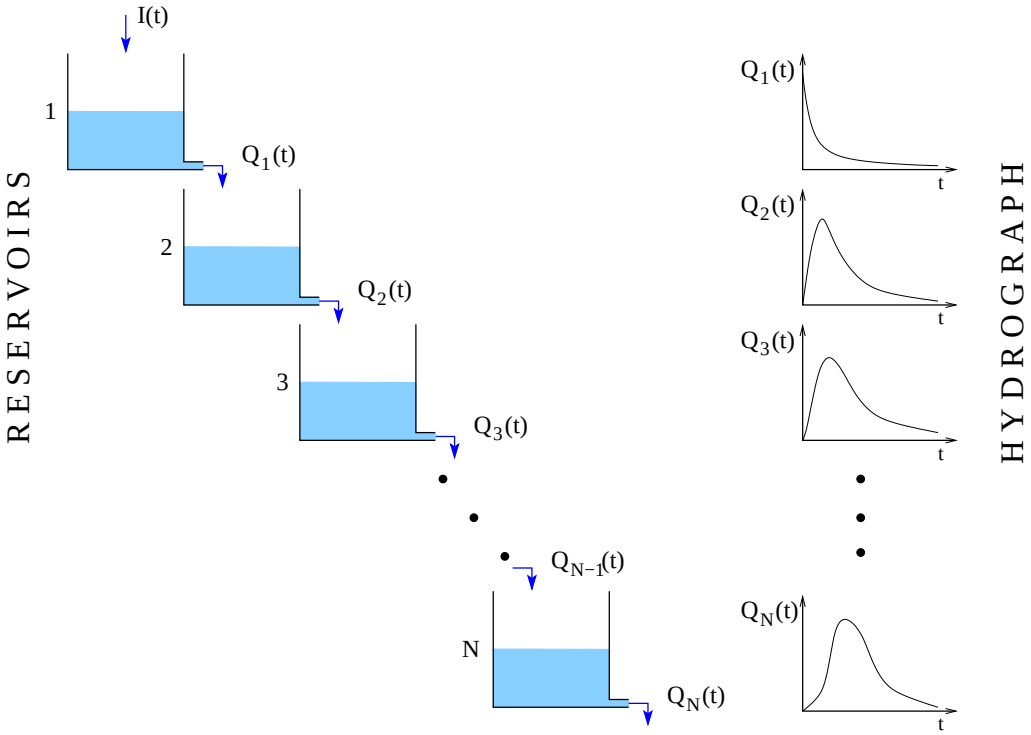

Figure 4.1: A cascade of N linear reservoirs and the corresponding hydrographs shown to the right. Figure from Thordarson (2012).

and 4.3 can then be generalised to describe the storage in N linear reservoirs

$$\frac{d}{dt}\begin{bmatrix} S_1(t) \\ S_2(t) \\ \vdots \\ S_N(t) \end{bmatrix} = \begin{bmatrix} -\frac{N}{K} & 0 & \cdots & 0 \\ \frac{N}{K} & -\frac{N}{K} & & \\ \vdots & \ddots & \ddots & \\ 0 & & \frac{N}{K} & -\frac{N}{K} \end{bmatrix} \begin{bmatrix} S_1(t) \\ S_2(t) \\ \vdots \\ S_N(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} I(t). \tag{4.4}$$

The wastewater flow variation D_t was described by a harmonic function reflecting the diurnal pattern of water discharge from households

$$D_t = \sum_{i=1}^{2} \left(s_i \sin \frac{i2\pi t}{L} + c_i \cos \frac{i2\pi t}{L} \right), \tag{4.5}$$

where L is the period of 24 hours, the parameters s_1, c_1, s_2 and c_2 are non-physical parameters and t is time. To fully account for the wastewater flow a constant average flow a_0 $[m^3/h]$ is added to D_t.

4.2 Overview of simplistic models applied

In all models the wastewater flow was described by (4.5). Statistical tests showed that two sinus and cosinus constants were optimal.

Conceptual model applied in Paper II

In Paper **II** a model suitable for *simulation* was developed and hence both fast

and slow runoff components were considered. The conceptual model (Figure 4.2) consists of two linear reservoirs for modelling the fast runoff (S_{f1}, S_{f2}) from paved areas and three linear reservoirs (S_{s1}, S_{s2}, S_{s3}) for modelling the slow infiltration inflow from permeable areas. Comparing the observed hydrographs of Figure 3.4 and

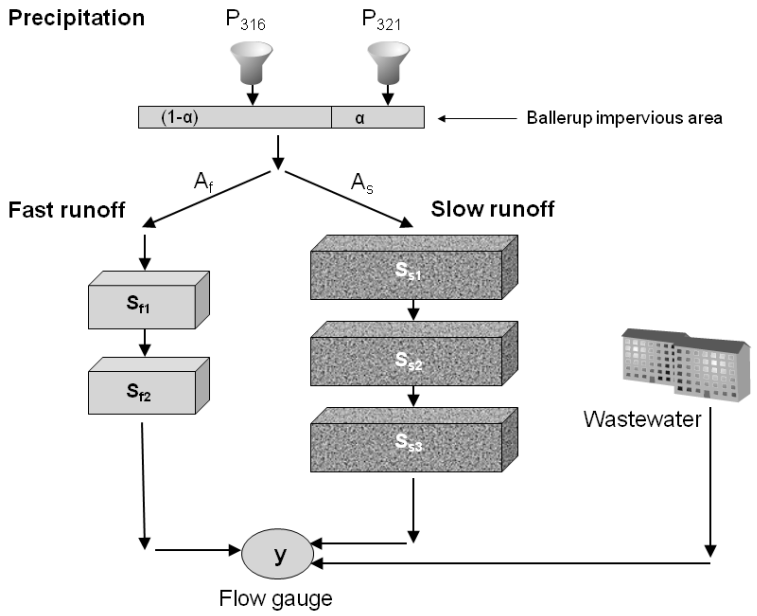

Figure 4.2: Graphical representation of the conceptual model used in paper **Paper II**.

Figure 3.3 (left) with the hydrographs in Figure 4.1 it appears that two reservoirs are sufficient for modelling the fast response because the observed hydrographs has a rather steep rising limb. The slow infiltration inflow is expected to have a more slow rising limb and a system of three reservoirs were therefore assumed adequate. Hence the model can be written on a state space form accordingly with (4.4):

Rainfall-runoff from paved areas

$$d\begin{bmatrix} S_{f1,t} \\ S_{f2,t} \end{bmatrix} = \begin{bmatrix} \alpha A_f P_{316,t} + (1-\alpha) A_f P_{321,t} + a_0 - \frac{2}{K_f} S_{f1,t} \\ \frac{2}{K_f} S_{f1,t} - \frac{2}{K_f} S_{f2,t} \end{bmatrix} dt, \tag{4.6}$$

where A_f $[ha]$ is the impervious fast runoff area, K_f $[h]$ is the retention time of the fast runoff, α $[-]$ is a rain gauge weighting coefficient, and P_{316} & P_{321} are the rain gauge inputs $[m/h]$.

Slow infiltration inflow from permeable areas

$$d\begin{bmatrix} S_{s1,t} \\ S_{s2,t} \\ S_{s3,t} \end{bmatrix} = \begin{bmatrix} \alpha A_s P_{316,t} + (1-\alpha) A_s P_{321,t} - \frac{2}{K_s} S_{s1,t} \\ \frac{2}{K_s} S_{s1,t} - \frac{2}{K_s} S_{s2,t} \\ \frac{2}{K_s} S_{s2,t} - \frac{2}{K_s} S_{s3,t} \end{bmatrix} dt, \tag{4.7}$$

where A_s $[ha]$ is the permeable area associated with the infiltration inflow and K_s $[h]$ is the slow infiltration retention time constant.

Observation equation

$$Y_t = \frac{2}{K_f} S_{f2,t} + \frac{3}{K_s} S_{s3,t} + D_t, \tag{4.8}$$

where Y_t $[m^3/h]$ is the modelled flow that is a sum of the outflow from the fast and slow runoff components and the wastewater flow.

Conceptual model applied in papers III & IV

In papers **III** & **IV** simplified *prediction models* were developed with a deterministic state space part equivalent to (4.6) (see Figure 4.3**a**), and thus a reduced observation equation was needed

Observation equation

$$Y_t = \frac{2}{K_f} S_{f2,t} + D_t, \tag{4.9}$$

Conceptual model applied in paper V

In paper **V** a *simulation model* was developed (Figure 4.3**e**). The model has both a fast and slow runoff component, however the slow infiltration description differ from the description used in **Paper II** by facilitating a storage S_E $[m^3]$ that mimics the degree of saturation of the infiltration surface. The basin is emptied by evaporation E $[m/h]$ and the inflow to the sewer system is controlled by the degree of saturation (the filling degree of S_E) using an exponential relationship. The model is also equipped with a detention basin S_d $[m^3]$ and a pumping station for emptying the detention basin. In case of heavy rainfall overflow will take place from S_{f1} to S_d if the maximum capacity of S_{f1} is exceeded, i.e. if $S_{f1} > S_{f1,max}$. Likewise if the maximum capacity of S_d is exceeded, i.e. if $S_d > S_{d,max}$ sewage overflow will take place. The detention basin is emptied by a pump with constant pumping rate Q_p, however to mimic the slow startup and slowdown of the pumping rate the pumped water is directed through three linear reservoirs (S_{p1}, S_{p2}, S_{p3}). The state space formulation then becomes:

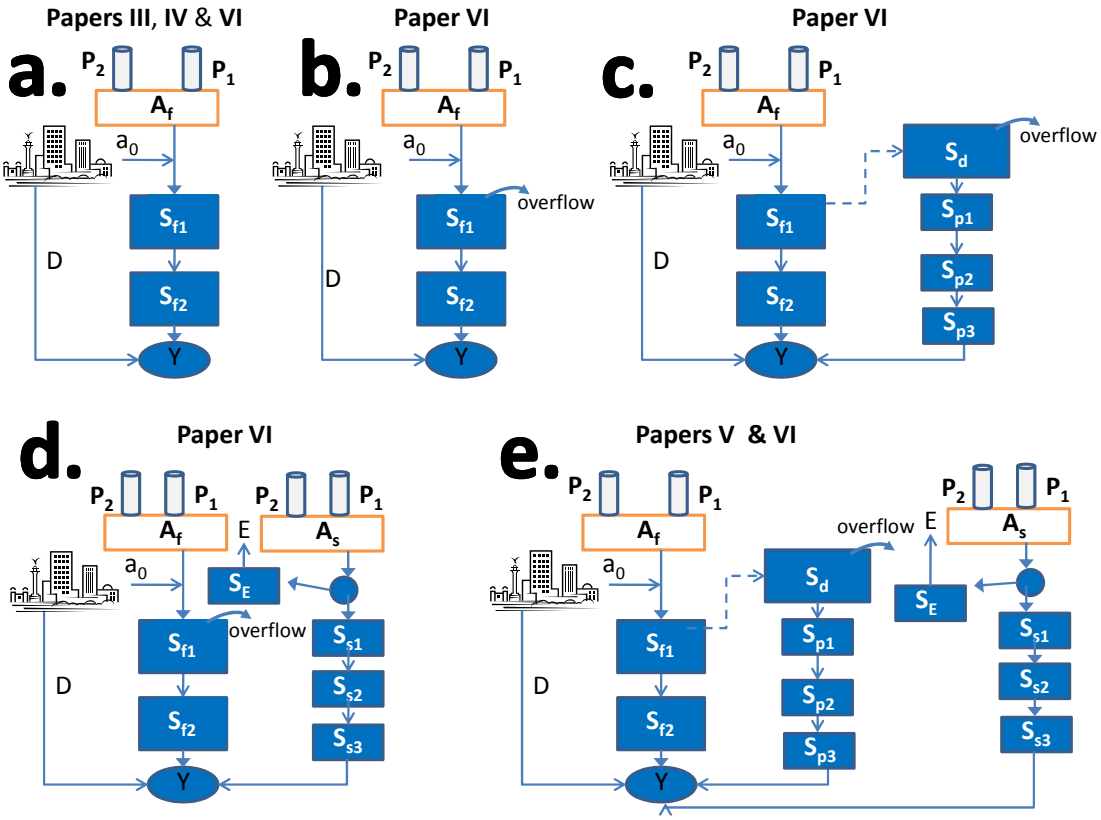

Figure 4.3: Conceptual representation of the models used in paper **Papers III –VI**. P_1 and P_2 are the rain gauge inputs from either P_{321} & P_{316} or P_{309} & P_{316}.

Rainfall-runoff from paved areas

$$d\begin{bmatrix} S_{f1,t} \\ S_{f2,t} \end{bmatrix} = \begin{bmatrix} C_{1,t}\left(\alpha A_f P_{316,t} + (1-\alpha)A_f P_{321,t} + a_0\right) - \frac{2}{K_f}S_{f1,t} \\ \frac{2}{K_f}S_{f1,t} - \frac{2}{K_f}S_{f2,t} \end{bmatrix} dt, \tag{4.10}$$

Where C_1 is a Kronecker delta function that governs the inlet to S_{f1} such that if the condition $S_{f1} < S_{f1,max}$ holds, $C_1 = 1$ and sewage is directed to S_{f1}, and if the opposite condition holds $S_{f1} > S_{f1,max}$ then $C_1 = 0$ and sewage is directed to S_d instead.

Slow infiltration inflow from permeable areas

$$d\begin{bmatrix} S_{E,t} \\ S_{s1,t} \\ S_{s2,t} \\ S_{s3,t} \end{bmatrix} = \begin{bmatrix} \exp\left(\frac{-S_{E,t}}{A_s\varsigma}\right)\left(\alpha A_s P_{316,t} + (1-\alpha)A_s P_{321,t}\right) - C_{2,t}A_s E_t \\ \left(1-\exp\left(\frac{-S_{E,t}}{A_s\varsigma}\right)\right)\left(\alpha A_s P_{316,t} + (1-\alpha)A_s P_{321,t}\right) - \frac{3}{K_s}S_{s1,t} \\ \frac{2}{K_s}S_{s1,t} - \frac{2}{K_s}S_{s2,t} \\ \frac{2}{K_s}S_{s2,t} - \frac{2}{K_s}S_{s3,t} \end{bmatrix} dt, \tag{4.11}$$

where ς $[m]$ is a saturation constant and C_2 is another Kronecker delta function

that secures a positive volume in S_E by assuming the value 1 when the condition $S_E > S_{Emin}$ is satisfied and zero when the opposite holds.

Detention basin and pumping station

$$d\begin{bmatrix} S_{d,t} \\ S_{p1,t} \\ S_{p2,t} \\ S_{p3,t} \end{bmatrix} = \begin{bmatrix} \big(1-C_{1,t}\big)\big(A_f P_{316,t} + (1-\alpha)A_f P_{321,t} + a_0\big) - C_{3,t}C_{4,t}Q_{p,t} - Q_{of,t} \\ C_{3,t}C_{4,t}Q_{p,t} - \frac{3}{K_p}S_{p1,t} \\ \frac{3}{K_p}S_{p1,t} - \frac{3}{K_p}S_{p2,t} \\ \frac{3}{K_p}S_{p2,t} - \frac{3}{K_p}S_{p3,t} \end{bmatrix} dt, \tag{4.12}$$

where $Q_{of,t} = C_{5,t}\frac{(S_{d,t}-S_{d,max})}{dt}$ is the overflow from the detention basin when the maximum volume of the detention basin $S_{d,max}$ is exceeded, and C_5 is zero when the condition $S_d < S_{d,max}$ is true, otherwise one. The initiation of the pump require C_3 and C_4 to take the value one. C_3 is one if $S_d > S_d, min$ otherwise zero, where S_d, min is the minimum volume of the detention basin. C_4 is one when $S_{f1} < \gamma S_{f1,max}$ otherwise zero which secures that the pump is activated after the rain has passed.

Observation equation

$$Y_t = \frac{2}{K_f}S_{f2,t} + \frac{3}{K_s}S_{s3,t} + \frac{3}{K_p}S_{p3,t} + D_t, \tag{4.13}$$

The observation equation is extended with the outlet from the pumping station.

Conceptual models applied in paper VI

In **paper VI** all the models of Figure 4.3 **a**-**e**) with various physical content incorporated (from Figure 4.3 **a**-**e**) were used for *prediction*.

The state space formulation of the model in Figure 4.3**a** is equivalent to (4.6) and the observation equation to (4.9). The only extension of the model shown in Figure 4.3**b** compared to Figure 4.3**a** is the overflow possibility in S_{f1} and this implementation changes the state space formulation to (4.10).

In Figure 4.3**c** the model is further extended with a detention basin and a pumping station and the overflow is now directed to the detention basin and the observation equation becomes:

Observation equation

$$Y_t = \frac{2}{K_f}S_{f2,t} + \frac{3}{K_p}S_{p3,t} + D_t. \tag{4.14}$$

In Figure 4.3**d** the model is equipped with the infiltration inflow description described for **paper V** and formulated in (4.11) and once again the fast runoff is given by (4.10) and the observation equation by (4.8). The last model shown in Figure 4.3**e** is already outlined (used in **paper V**).

5 The stochastic approach to uncertainty evaluation

The stochastic approach that was chosen for this thesis is a frequentist approach that is based on an Extended Kalman Filter (EKF) and a maximum likelihood function for parameter estimation. The models are formulated in state-space and uncertainty is accounted for in the states and in the output. How this works is explained in the following.

5.1 Stochastic grey box models

The deterministic models introduced in Chapter 4 can be formulated using a general notation

$$d\boldsymbol{X}_t = \boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{U}_t, t, \boldsymbol{\theta})dt \tag{5.1}$$

$$\boldsymbol{Y}_k = \boldsymbol{g}(\boldsymbol{X}_k, \boldsymbol{U}_k, t_k, \boldsymbol{\theta}), \tag{5.2}$$

where (5.1) is the *system equation*, describing the evolution of the states in continuous time and the function $\boldsymbol{f}(\cdot) \in \mathbb{R}^n$ corresponds to the deterministic state space formulations of the simple models introduced in Chapter 4, and (5.2) is the *observation equation* that relates the observations to the states in discrete time by the function $\boldsymbol{g}(\cdot) \in \mathbb{R}^l$ consistent with the observation equations introduced in Chapter 4. The time $t \in \mathbb{R}_O$ indicates the continuous time and k ($k = 1, \ldots, K$) are the discretely observed sampling instants for K number of measurements. $\boldsymbol{Y} \in \mathbb{R}^l$ is a vector of output variables, $\boldsymbol{X} \in \mathbb{R}^n$ a vector of state variables, $\boldsymbol{\theta} \in \mathbb{R}^p$ contains the unknown parameters of the system and $\boldsymbol{U}_t \in \mathbb{R}^m$ is a vector of input variables.

To address uncertainties in the system the model consisting of (5.1) and (5.2) is extended with a diffusion term and an observation noise term as follows

$$d\boldsymbol{X}_t = \underbrace{\boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{U}_t, t, \boldsymbol{\theta})}_{\text{drift term}} dt + \underbrace{\boldsymbol{\sigma}(\boldsymbol{X}_t, \boldsymbol{U}_t, t, \boldsymbol{\theta})}_{\text{diffusion term}} d\boldsymbol{\omega}_t \tag{5.3}$$

$$\boldsymbol{Y}_k = \boldsymbol{g}(\boldsymbol{X}_k, \boldsymbol{U}_k, t_k, \boldsymbol{\theta}) + \underbrace{\boldsymbol{e}_k}_{\text{obs. noise term}}, \tag{5.4}$$

where (5.3) again is the *system equation* and (5.4) the *observation equation*. $\boldsymbol{f}(\cdot)$ is denoted the drift term and $\boldsymbol{\sigma}(\cdot) \in \mathbb{R}^{n \times n}$ is denoted the diffusion noise term or the process noise function which represents the uncertainty of the states in the system, and $\boldsymbol{\omega}_t$ is an n-dimensional standard Wiener process, which simply means that the errors between the predicted states and the indirectly observed states are

assumed to be Gaussian distributed. The measurement error $\boldsymbol{e}_k$ is assumed to be an l-dimensional white noise process with $\boldsymbol{e}_k \in N(\boldsymbol{0}, \boldsymbol{S}(\boldsymbol{U}_k, t_k, \boldsymbol{\theta}))$. The system formulation of (5.3) and (5.4) is sometimes referred to as a *stochastic grey box model* due to the coupling of system knowledge (from white box models) and information from data (black box models) that are jointly utilised for estimation of the parameters.

5.2 Maximum likelihood parameter estimation

To estimate the parameters given N number of measurements $[\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \ldots, \boldsymbol{y}_N]$, and by introducing the notation $\boldsymbol{\mathcal{Y}}_k = [\boldsymbol{y}_k, \boldsymbol{y}_{k-1}, \ldots, \boldsymbol{y}_1, \boldsymbol{y}_0]$, the likelihood function is expressed as a product of conditional densities

$$L(\boldsymbol{\theta}; \boldsymbol{\mathcal{Y}}_N) = P(\boldsymbol{\mathcal{Y}}_N|\boldsymbol{\theta}) = \left(\prod_{k=1}^{N} P(\boldsymbol{y}_k|\boldsymbol{\mathcal{Y}}_{k-1}, \boldsymbol{\theta}) \right) P(\boldsymbol{y}_0|\boldsymbol{\theta}), \tag{5.5}$$

where Bayes theorem $P(A \cap B) = P(A|B)P(B)$ is repeatedly used at each time step to formulate the likelihood function as a product of the one step ahead conditional densities and where $P(\boldsymbol{y}_0|\boldsymbol{\theta})$ is a parameterisation of the starting conditions. It is assumed that the system equations are driven by a Wiener process which have Gaussian increments and thus the conditional probabilities in (5.5) can be approximated by Gaussian densities.

The Gaussian density is completely characterised by its mean and covariance of the one step prediction, which are denoted by $\hat{\boldsymbol{y}}_{k|k-1} = E\{\boldsymbol{y}_k|\boldsymbol{\mathcal{Y}}_{k-1}, \boldsymbol{\theta}\}$ and $\boldsymbol{R}_{k|k-1} = V\{\boldsymbol{y}_k|\boldsymbol{\mathcal{Y}}_{k-1}, \boldsymbol{\theta}\}$, respectively, and, by introducing an expression for the innovation formula, $\boldsymbol{\varepsilon}_k = \boldsymbol{y}_k - \hat{\boldsymbol{y}}_{k|k-1}$ the likelihood function can be rewritten as (Madsen, 2008)

$$L(\boldsymbol{\theta}; \boldsymbol{\mathcal{Y}}_N) = \left(\prod_{k=1}^{N} \frac{\exp\left(-\frac{1}{2}\boldsymbol{\varepsilon}_k^\top \boldsymbol{R}_{k|k-1}^{-1} \boldsymbol{\varepsilon}_k\right)}{\sqrt{\det(\boldsymbol{R}_{k|k-1})}\left(\sqrt{2\pi}\right)^l} \right) P(\boldsymbol{y}_0|\boldsymbol{\theta}), \tag{5.6}$$

where the conditional mean and covariance are calculated using an EKF. The conditional likelihood function (5.6) minimises the one step prediction uncertainty and hence the estimation is optimised for making predictions. In the simulation case there is no new information for conditioning as this information is not available during simulation. For a given set of calibration data this changes the likelihood function to minimise the output error for the whole considered period

$$L(\boldsymbol{\theta}; \boldsymbol{\mathcal{Y}}_N) = \frac{1}{(\sqrt{2\pi \boldsymbol{S}})^N} \exp\left[\sum_{k=1}^{N} \frac{1}{2\boldsymbol{S}} (\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k)^2 \right]. \tag{5.7}$$

The parameter estimates can be obtained by conditioning on the initial values and solving the optimisation problem

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \{\log\left(L(\boldsymbol{\theta}; \boldsymbol{\mathcal{Y}}_N|\boldsymbol{y}_0)\right)\}. \tag{5.8}$$

Numerical methods are needed to optimise the likelihood function (Kristensen and Madsen, 2003).

The maximum likelihood method also provides an assessment of the uncertainty for the parameter estimates in (5.8) since the maximum likelihood estimator is asymptotically normal distributed with mean $\boldsymbol{\theta}$ and covariance matrix

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}} = \boldsymbol{H}^{-1}.$$

The matrix $\boldsymbol{H}$ is the Fisher Information Matrix (Madsen and Thyregod, 2011) given by

$$h_{ij} = -E\left\{ \frac{\partial^2}{\partial\theta_i \partial\theta_j} \log(L(\boldsymbol{\theta}|\boldsymbol{\mathcal{Y}}_{k-1})) \right\} \qquad i,j = 1,\ldots,p. \tag{5.9}$$

In practice an approximation for $\boldsymbol{H}$ is obtained by the observed Hessian h_{ij} evaluated for $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Due to the asymptotic Gaussianity of the estimator in (5.8), a t-test can be performed to check if the estimated parameters are statistically significant (Madsen and Thyregod, 2011).

To solve the estimation problem the open source software CTSM[1] is used (Kristensen and Madsen, 2003; Kristensen et al., 2004a,b). The program was developed at the department of Informatics and Mathematical Modelling (IMM) at the Technical University of Denmark (DTU). CTSM is capable of estimating parameters in both linear and non-linear models using respectively an ordinary Kalman filter for linear models and an extended kalman filter for non-linear models.

5.3 Seeking an appropriate diffusion term

As can be seen from (5.3) the diffusion term can be a function of the states, the inputs, the time, and the parameters. One may readily expect the uncertainty of the states to be somehow related to the rain input and more or less constant in dry weather periods suggesting the following diffusion term to apply

$$\boldsymbol{\sigma}(\cdot) = \left[\sigma_{dry} + \sigma_{wet}(P_{1,t} + P_{2,t})\right], \tag{5.10}$$

where σ_{dry} represents the dry weather state uncertainty and σ_{wet} represents the added rain proportional uncertainty in rainy periods. The rain dependent diffusion

[1] Continuous-Time Stochastic Modelling - www.imm.dtu.dk/ctsm

term was tested in combination with the drift term of (4.6) but was however found unsuitable for two reasons. Firstly because of the time delay between a rainfall is recorded to a state impact is observed (this time delay would be different for every reservoir), and secondly because rapidly changing precipitation is frequently observed during rain, especially during convective rainfall events, whereas the states would exhibit much more damped changes due to the equalisation of the runoff from a large catchment. To overcome these problems different time lags and mean precipitation values were tested in CTSM (the mean precipitation of the last half hour, three quarters, 1 hour) and also various low pass filters of the rain data using the following threshold function

$$P_{lp} = \frac{P_{max}}{1 + exp(b_0 - b_1(P_1 + P_2))}, \tag{5.11}$$

where P_{lp} is the low pass filtered rain input, P_{max} the maximum rain input and b_0 and b_1 are threshold parameters. However, none of these proposals turned out entirely satisfying for describing the one step state prediction uncertainty, either it was impossible to estimate the parameters or the residuals of the one-step prediction did not assume a white noise process.

Instead in **Paper III** different state-dependent diffusion terms were tested in combination with (4.6) using a state exponentiated diffusion term

$$\boldsymbol{\sigma}(\cdot) = \left[\sigma_i X_{i,t}^{\gamma_i}\right], \tag{5.12}$$

where $\sigma_i^{\gamma_i}$ represents the standard deviation of the one step state uncertainty that is scaled with the reservoir volume X_i. A state proportional uncertainty (γ_i=1) was shown to best and adequately describe the one step prediction uncertainty. The advantage of using a state-dependent diffusion term is not just related to a better connection between rain input and state uncertainty, that is implicitly accounted for by a state dependent diffusion term, but furthermore due to the natural physical restrictions of the system. The state variables in the models correspond to the stored volume of water in the reservoirs and obviously these have a lower limit of zero since the water volume cannot turn negative. This lower restriction implies that the diffusion of the states must approach zero as the water volume in the drift term approaches this lower boundary of zero and the state proportional diffusion term ensures this. The state proportional diffusion term was also applied in **Paper V** in combination with the drift term of the states S_{f2} in (4.10), S_{s3} in (4.11), and S_{p3} in (4.12). In **Paper VI** different models was tested and the state proportional diffusion term was added to various drift terms (see **Paper VI** for details).

Similarly to the lower restriction of the reservoirs an upper restriction can also be argued in heavy rain storms where the sewer system reaches its maximum conveyance capacity or in connection with a storage basin that reaches its maximum volume. One way to deal with such upper constraints of the system is to use a logistic state dependency, which was applied in **Papers V & VI**. This state logistic diffusion term can be written

$$\boldsymbol{\sigma}(\cdot) = \left[\sigma_i X_{i,t}(X_{i,max}\kappa - X_i)\right]. \tag{5.13}$$

where $X_{i,max}$ is the upper restriction and κ is a shaping factor to control the state uncertainty around the upper restriction. Figure 5.1 shows how the two different state dependent diffusion terms impacts the total uncertainty of the state as the water volume in the reservoir increases for different σ_i values (here shown for a fast rainfall-runoff reservoir S_{f1}).

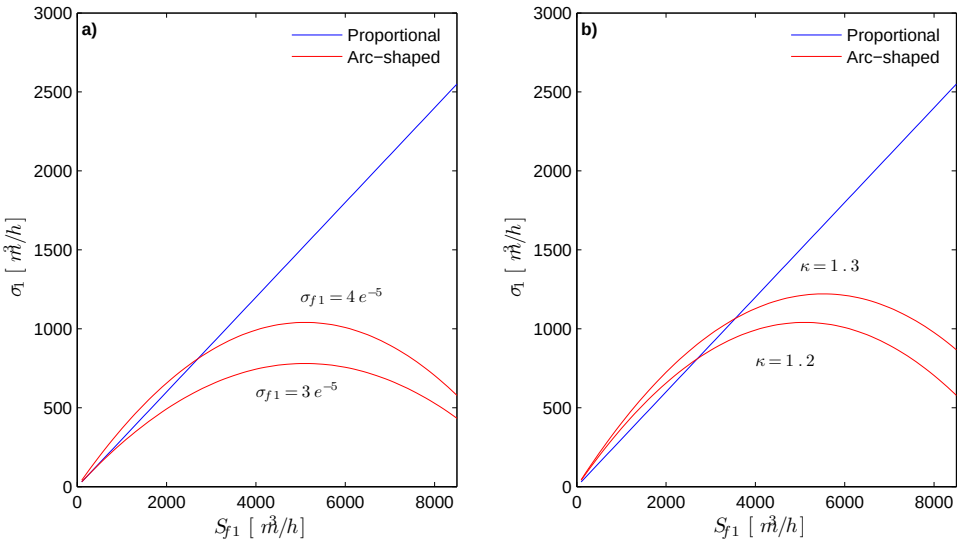

Figure 5.1: Proportional and logistic state dependency (Arc-shaped). In (a) $\kappa = 1.2$ and in (b) $\sigma_{f1} = 4e^{-5}$. Figure from **Paper V**.

It is clearly demonstrated that a logistic diffusion term leads to a decreasing state uncertainty as the water filled volume approaches the maximum capacity of the reservoir (in Figure 5.1**a** at 10,000 m^3). It is also shown that the shape of the logistic state dependency depends on σ_{f1}. Conversely the state proportional uncertainty is unrestricted and keeps rising with the filling of S_{f1} which is more appropriate in cases where there is no upper constraint such as for the slow infiltration reservoirs in **Paper V**. In Figure 5.1**b** it is shown how the shaping also depends on the value of κ. The logistic diffusion term (5.13) was applied to the fast runoff state S_{f1} and

added to the drift term of (4.10) in **Paper V** and **Paper VI** to constrain the flow during heavy rainfall where the maximum conveyance capacity is reached.

In general it is not possible to implement state dependent diffusion terms in CTSM directly because they require higher order filtering techniques that have been shown to become numerically unstable (Vestergaard, 1998). But if a transformation is available for the SDE that has a state dependent diffusion term, such that the diffusion term of the transformed process becomes independent of the state, the filtering techniques in CTSM can be applied to obtain efficient and numerically stable parameter estimates (Baadsgaard et al., 1997). A detailed account of how to apply the Lamperti transformation to obtain a state independent diffusion term is given in paper **Paper III** and will not be repeated here.

5.4 Seeking an appropriate observation noise term

As mentioned in Chapter 2, within RR modelling a transformation of the observations $\Psi(\mathcal{Y})$ is often required to stabilise the variance and obtain a Gaussian distribution of the residuals. In CTSM the observation noise is additive but can be a function of the input, time or parameters, however not a function of the output $\boldsymbol{g}(\cdot)$ in (5.4) which is often desirable. A transformation of the observations to $\boldsymbol{Z}_k = [\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, \ldots, \boldsymbol{z}_1, \boldsymbol{z}_0]$ will change the observation equation to

$$\boldsymbol{Z}_k = \Psi\left(\hat{\boldsymbol{y}}_k\right) + \boldsymbol{\epsilon}_k, \tag{5.14}$$

where $\epsilon_k \in N(0, S_z)$ is then assumed to be a Gaussian white noise process with variance S_z. An output proportional observation noise term was applied in the prediction models of **Papers III-VI** requiring a log-transformation of the observations. This changed the observation equation accordingly

$$z_k = \log\left(\hat{y}_k\right) + \epsilon_k. \tag{5.15}$$

The output proportional observation noise term is argued based on the assumption that flow meters typically become less accurate with increasing flow size. The effect of a log-transformation is visualised in Figure 5.2 with 1000 random numbers from the flow interval 0-3000 m^3/h. The desired flow proportional observation noise is demonstrated in Figure 5.2**a+b** using an output proportional observation noise with $\sqrt{S} = 0.1$ as an example, whereas Figure 5.2**c+d** shows the log-transformed $\mathcal{Z}$ domain in which $\sqrt{S_z} = 0.1$ is constant, which shows that the proportionality can be estimated in CTSM. Hence when $\sqrt{S_z}$ has been estimated in CTSM, taking the inverse $\Psi^{-1}(\mathcal{Z}) = \exp(\mathcal{Z})$ to return to the $\mathcal{Y}$ domain, it is seen that the standard deviance increases proportionally with the output.

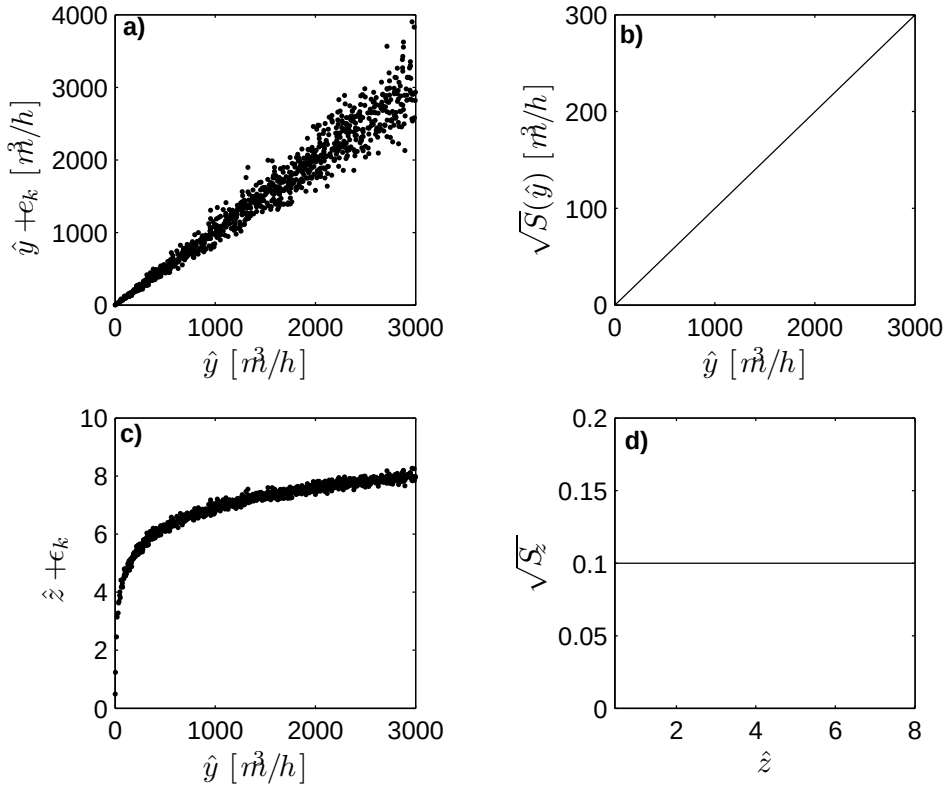

Figure 5.2: $\mathcal{Y}$ and $\mathcal{Z}$ domain using a log-transformation of the observations. Illustrated with 1000 randomly generated flow values of $\hat{y}$ in the interval 0-3000 m^3/h. (a) A flow proportional noise term of $\sqrt{S} = 0.1$ is added to $\hat{y}$ and the result is an increasing uncertainty with the flow magnitude. (b) Illustrates the proportionally increasing standard deviance of the observation noise with the flow magnitude. (c) The $\mathcal{Z}$ domain after log-transformation of the values in (a). (d) A constant standard deviance $\sqrt{S_z} = 0.1$ is observed in the $\mathcal{Z}$ domain.

In **Paper V** a different observation transformation was applied that was derived from a residual analysis of a simulation model with additive constant observation noise and from considering the constraints of the system. Again the aim with the transformation was to stabilise the variance and obtain homoscedasticity. A logistic observation transformation function of the following form was considered

$$z_k = \log\left(\frac{\hat{y}_k - y_{min}}{y_{max} - \hat{y}_k}\right) + \epsilon_k, \tag{5.16}$$

where y_{min} and y_{max} represented respectively a lower and upper flow boundary that was chosen by inspection of the flow data. The idea is to reduce the noise as the flow approaches these lower and upper boundaries. The effect of this observation noise term is illustrated in Figure 5.3 with 1000 randomly generated numbers from the flow range 30-3000 m^3/h. In Figure 5.3**a+b** we observe that the uncertainty is

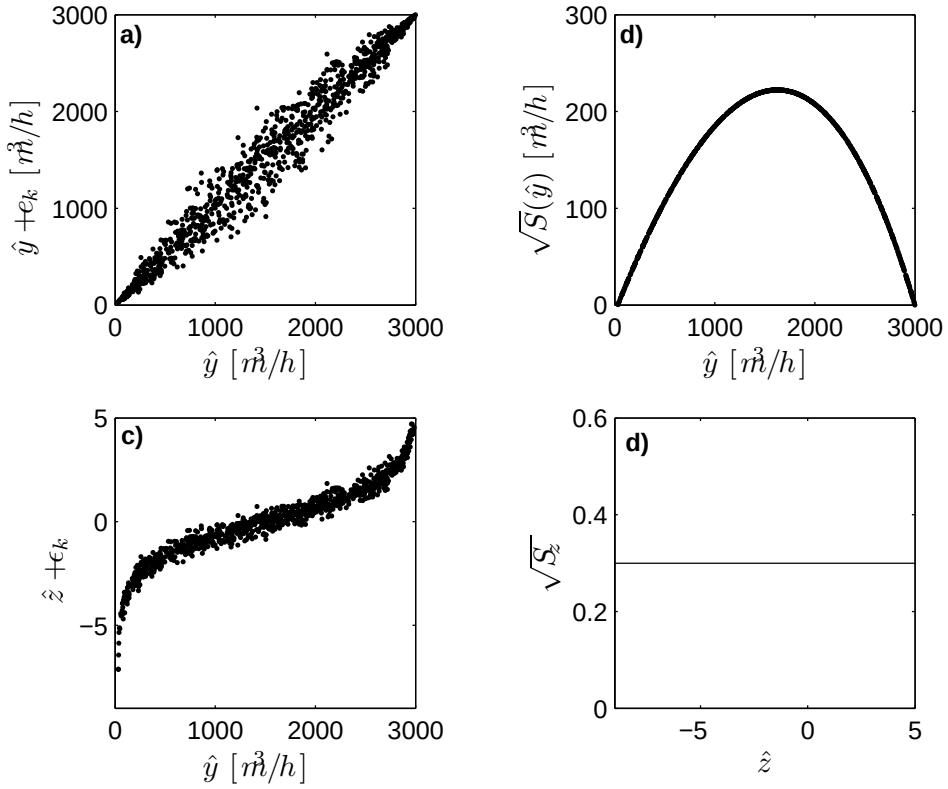

Figure 5.3: Illustration of how an arc-shaped observation noise term can be estimated using a logistic observation transformation that leads to a constant observation noise term in the $\mathcal{Z}$ domain. Illustrated with 1000 randomly generated flow values in the interval 30-3000 m^3/h. (a) Shows how the uncertainty is small around lower (30 m^3/h) and upper (3000 m^3/h) boundaries and is largest in between. (b) Shows the arc-shaped development of the observation noise term as a function of flow magnitude. (c) In the transformed $\mathcal{Z}$ domain the uncertainty is constant. (d) The observation noise term in the $\mathcal{Z}$ domain is constant (with $\sqrt{S_z} = 0.3$). Modified figure from **Paper V**.

small at the upper and lower boundaries and the noise increases with the flow rate until the middle of the flow range (around 1500 m^3/h) and then starts to decrease again. The resulting standard deviance is arc-shaped. When applying the transformation in (5.16) a constant standard deviance of $\sqrt{S_z} = 0.3$ (Figure 5.3**c+d**) is obtained meaning that the arc-shaped noise term can now be determined using CTSM. Having estimated $\sqrt{S_z}$ the inverse can be taken

$$y_k = \left(\frac{e^{z_k} y_{max} + y_{min}}{1 + e^{z_k}} \right), \tag{5.17}$$

and the observation noise in the $\mathcal{Y}$ domain calculated.

5.5 The transformed grey box model with transformed observations

The transformed grey box model with a transformed observation equation can now be written

$$d\tilde{\boldsymbol{X}}_t = \tilde{\boldsymbol{f}}(\tilde{\boldsymbol{X}}_t, \boldsymbol{u}_t, t, \boldsymbol{\theta})dt + \tilde{\boldsymbol{\sigma}}(\boldsymbol{u}_t, t, \boldsymbol{\theta})d\boldsymbol{\omega}_t \tag{5.18}$$

$$\boldsymbol{Z}_k = \Psi\big(\tilde{\boldsymbol{g}}(\tilde{\boldsymbol{X}}_k, \boldsymbol{u}_k, t_k, \boldsymbol{\theta})\big) + \boldsymbol{\epsilon}_k, \tag{5.19}$$

where the functions $\boldsymbol{f}(\cdot)$, $\boldsymbol{\sigma}(\cdot)$ and $\boldsymbol{g}(\cdot)$ in Eq's. (5.3) and (5.4) have been reformulated, respectively to $\tilde{\boldsymbol{f}}(\cdot)$, $\tilde{\boldsymbol{\sigma}}(\cdot)$ and $\tilde{\boldsymbol{g}}(\cdot)$ in relation to the Lamperti transformation of the state space and it is noticed that $\boldsymbol{X}$ has been removed from $\tilde{\boldsymbol{\sigma}}(\cdot)$. Furthermore $\boldsymbol{X}$ have been transformed to $\tilde{\boldsymbol{X}}$ and the observations to $\mathcal{Z}$. It should be emphasised that the parameters of the drift term and the diffusion term are unaffected by the Lamperti transformation, and hence the parameter estimates can be directly inserted into the original system equation (5.3). In the **Papers III-VI** a system consisting of (5.18) and (5.19) was applied.

5.6 Generating predictions and uncertainty limits using stochastic grey box models

With the given sequence of input up to time $k+1$ as $\mathcal{U}_{k+1} = [\boldsymbol{U}_{k+1}, \cdots, \boldsymbol{U}_0]^\top$ and the output sequence up to time k as $\boldsymbol{\mathcal{Y}}_k = [\boldsymbol{Y}_k, \ldots, \boldsymbol{Y}_0]^\top$, the optimal prediction in a least square sense is equal to the conditional mean (see proof by Madsen, 2008). If we disregard transformations for now the one step ahead prediction $\hat{\boldsymbol{Y}}_{k+1|k}$ can be written

$$\begin{aligned}\hat{\boldsymbol{Y}}_{k+1|k} &= E\big\{\boldsymbol{Y}_{k+1}|\boldsymbol{\mathcal{Y}}_k, \boldsymbol{\mathcal{U}}_{k+1}\big\} \\ &= \boldsymbol{g}\left(\boldsymbol{\mathcal{Y}}_{k+1}, \boldsymbol{\mathcal{U}}_{k+1}, t_{k+1}, \boldsymbol{\theta}\right).\end{aligned}$$

The one step predictions are generated from the conditional expectation of the future states

$$\hat{\boldsymbol{X}}_{k+1|k} = E\big[\boldsymbol{X}_{k+1}|\hat{\boldsymbol{X}}_{k|k}, \boldsymbol{U}_{k+1}\big], \tag{5.20}$$

where $\hat{\boldsymbol{X}}_{k|k}$ is the filtered reconstructed state of the EKF (Kristensen and Madsen, 2003) at time k , given all measurements up to k providing a mean and a variance for the normally distributed state (Madsen, 2008).

It is important to recognise that the normal assumption for the model output is valid only for the one step ahead prediction and if more than the one step ($h > 1$) ahead predictions are considered, the $k+h$ probabilistic forecasts can be generated using

a numerical approach (the Euler scheme) for simulating the SDEs of the system equation (5.3) (Kloeden and Platen, 1999)

$$\begin{aligned}\hat{\boldsymbol{X}}_{k+h|k} =& \hat{\boldsymbol{X}}_{k|k} + \left(\sum_{i=1}^{h/\Delta} \boldsymbol{f}\left(\hat{\boldsymbol{X}}_{k+(i-1)\Delta|k}, \boldsymbol{U}_{k+i\Delta}, \boldsymbol{\theta}\right) \right)\Delta \\ &+ \sum_{i=1}^{h/\Delta} \boldsymbol{\sigma}\left(\hat{\boldsymbol{X}}_{k+(i-1)\Delta|k}, \boldsymbol{U}_{k+i\Delta}, \boldsymbol{\theta}\right) \Delta W_{k+(i-1)\Delta}.\end{aligned} \tag{5.21}$$

where Δ is the time step for the Euler approximation (assumed small), and ΔW_k is a randomly generated increment of the Wiener process $\{W\}_k$, i.e. $\Delta W_k = W_{k+\Delta} - W_k$. In **Papers IV**, **V** & **VI** the Euler scheme was applied with $\Delta = 1\ minute$ which means that 15 euler steps were taken per prediction step in the models.

With an increasing prediction horizon ($h > 1$) the variance of the stochastic term $\hat{\boldsymbol{X}}_{k+h|k}$ increases and the accuracy of a single point prediction, generated from (5.21), is reduced. To obtain a predictive distribution of the h step ahead prediction a high number of state simulations are needed for each Δ step and 1000 was considered an adequate number in the papers mentioned above. Thus at all the Δ's the required 1000 simulations are run with (5.21). At the desired prediction step $k + h$ the 1000 predicted states are inserted in the observation equation (5.19) to obtain 1000 predicted output values $\hat{\boldsymbol{Z}}_{k+h|k}$.

A predictive empirical distribution can be profiled by sorting the 1000 predicted outputs from the highest (rank 1) to the lowest (rank 1000) value. For example to obtain the 90% bounds the lower bound is found at rank 950 and the upper at rank 50 and the median is found at rank 500. Following this example all the desired quantiles can be estimated.

For a simulation model without diffusion terms the uncertainty bounds are simply derived from the assumed normally distributed observation noise term (5.4).

6 The epistemic approach to uncertainty evaluation

The generalised likelihood estimation (GLUE) methodology is considered an epistemic method to uncertainty evaluation (Beven et al., 2011) and is described in this Chapter.

6.1 Introducing stochastic parameters

In the GLUE methodology the deterministic model formulation is once again considered

$$d\boldsymbol{X}_t = \boldsymbol{f}(\boldsymbol{X}_t, \boldsymbol{U}_t, t, \underbrace{\boldsymbol{\theta}}_{\text{stoch.var.}})dt \tag{6.1}$$

$$\boldsymbol{Y}_k = \boldsymbol{g}(\boldsymbol{X}_k, \boldsymbol{U}_k, t_k, \underbrace{\boldsymbol{\theta}}_{\text{stoch.var.}}), \tag{6.2}$$

and stochasticity is introduced by interpreting the parameters as stochastic variables. By doing so the need for stochastic model terms such as the diffusion and observation noise terms in the grey box model is unnecessary and model errors are instead implicitly accounted for by uncertainty in the parameters. This is in contrast to the interpretation of the stochastic approach outlined in Chapter 5 in which the parameters have constant but uncertain values. GLUE acknowledges the existence of many (almost) equally well performing parameter sets, a phenomenon referred to as equifinality by GLUE users (Beven, 2006).

6.2 The pseudo-Bayesian approach

In GLUE, a prior likelihood of each parameter set $\boldsymbol{\Theta}_i$ is specified $L_0\left(\boldsymbol{\Theta}_i\right)$, typically a non-informative uniform prior is chosen, and a likelihood measure $L\left(\boldsymbol{\Theta}_i|\boldsymbol{\mathcal{Y}}\right)$ is selected such that a posterior likelihood $L_p\left(\boldsymbol{\Theta}_i|\boldsymbol{\mathcal{Y}}\right)$ can be found from Bayes equation

$$L_p\left(\boldsymbol{\Theta}_i|\boldsymbol{\mathcal{Y}}\right) = \frac{L_0\left(\boldsymbol{\Theta}_i\right) L\left(\boldsymbol{\Theta}_i|\boldsymbol{\mathcal{Y}}\right)}{C}, \tag{6.3}$$

where C is a scaling constant to ensure that the cumulative of all the parameter sets is unity (Beven and Binley, 1992; Beven and Freer, 2001; Beven, 2012).

Most often an informal likelihood measure is applied, and many different variants have been suggested (Freer et al., 1996; Beven and Freer, 2001; Thorndahl et al., 2008; Freni et al., 2009a). Having chosen a likelihood measure for parameter set evaluation a choice regarding which parameters to retain for predictive analysis

needs to be taken before the posterior parameter distribution can be summarised. Because of these subjective choices the term pseudo-Bayesian was used by Mantovan and Todini (2006); Freni et al. (2009b).

6.3 Choosing a likelihood measure and the behavioral parameter sets

In **Paper II** two different likelihood measures was chosen; one for dry weather periods L_{dw} and one for wet weather periods L_{ww}, and the periods were differentiated by a flow threshold of 0.15 m^3/s

$$L_{dw} = 1 - \frac{\sigma_\epsilon^2}{\sigma_o^2}, \qquad \sigma_o^2 > \sigma_\epsilon^2 \quad \& \quad \boldsymbol{\mathcal{Y}_1} \in \left(\boldsymbol{\mathcal{Y}} < 0.15 m^3/s\right) \tag{6.4}$$

$$L_{ww} = e^{-H\left(\frac{\sigma_\epsilon^2}{\sigma_o^2}\right)}, \qquad \boldsymbol{\mathcal{Y}_2} \in \left(\boldsymbol{\mathcal{Y}} \geq 0.15 m^3/s\right), \tag{6.5}$$

where σ_ϵ^2 is the residual error variance, σ_o^2 is the observation variance, $\boldsymbol{\mathcal{Y}_1}$ denotes the dry weather observations, $\boldsymbol{\mathcal{Y}_2}$ denotes the wet weather observations and H is a shaping factor that was fixed to 1. A combined likelihood measure inspired by Choi and Beven (2007) was then calculated by multiplication of the dry and wet weather likelihoods

$$L\left(\boldsymbol{\theta}|\boldsymbol{\mathcal{Y}}\right) = \varpi_1 L_{dw} \varpi_2 L_{ww}, \tag{6.6}$$

where ϖ_1 and ϖ_2 are weighting coefficients both set to 1, hence the model performance in dry and wet weather periods is equally weighted even though the dry weather period amounts 80% of the whole period considered. The more positive the likelihood values are, the better. Parameter sets returning negative likelihood values are not included because in that case the observed mean would be a better predictor than the model.

In a comparative study for describing the flow simulation uncertainty in Section 8.3, the GLUE method is compared with the grey box approach. In this study the Nash-Sutcliffe model efficiency coefficient (NS) is used as likelihood measure which is defined in (6.4), however in this case without subdividing flow data into dry and wet weather periods. This was chosen because the scope was to compare a traditional GLUE uncertainty analysis normally based on NSE with the stochastic approach.

The choice of which parameter sets to accept and which to reject is completely subjective but commonly GLUE users resort to a "statistical" approach in this choice

and demands that the GLUE generated simulation limits encompass a consistent proportion of the observations (Lindblom et al., 2011; Vezzaro and Mikkelsen, 2012). This approach was adopted in both GLUE studies (**Paper II** and the comparative study to be reported in Section 8.3). In **Paper II** we aimed at covering 90% of the observations by the 90% uncertainty bounds and in the comparative study this was extended to include ten "quantiles" , see Chapter 7.

6.4 Searching the parameter space

Before the parameter space can be searched an appropriate distribution and parameter range for each parameter needs to be specified. In **Paper II** a uniform distribution was used for all parameters but a reasonable prior range for each parameter was not easily decided and a broad range was therefore initially selected. In general two methods for searching the parameter space are applied, either a plain Monte Carlo method and its more efficient counterpart Latin Hypercube Sampling (LHS) that both search randomly (Thorndahl et al., 2008; Freni et al., 2009b; Jin et al., 2010; Li et al., 2010), or a Markov Chain Monte Carlo method (MCMC) (Blasone et al., 2008; McMillan and Clark, 2009; Lindblom et al., 2011; Vezzaro and Mikkelsen, 2012) that is designed to locate the high performing areas of the parameter space in an efficient and intelligent way.

The LHS technique was used in **Paper II**, and initially 100,000 random parameter sets were generated and so-called dotty plots (Beven, 2009a) were then studied to locate the higher likelihood areas. The parameter ranges were then adjusted accordingly and a new LHS started. Finally likelihood values of 200,000 parameter sets were produced of which 18,720 returned positive values and therefore were retained for further analysis.

In the comparative study it was decided to take advantage of an MCMC sampling method as other GLUE users have advocated for, the advantage being less computation time and avoidance of iterative prior range specification. Two MCMC schemes were investigated, initially the DRAM schme developed by Haario et al. (2006) and secondly the DREAM (DiffeRential Evolution Adaptive Metropolis) scheme developed by Vrugt et al. (2009c), however DREAM was found to be much more effective in finding the high performance parameter sets and this scheme was therefore chosen. The algorithm will not be detailed here and the reader is referred to Vrugt et al. (2009a,c) and Vrugt (2011) for more information. With its default settings and a formal likelihood function the DREAM scheme proved to be ex-

tremely effective in locating the best performing parameter sets and adjustments were therefore necessary to ensure more frequent visits of lower likelihood areas. McMillan and Clark (2009) describe a way to implement the NS likelihood measure in the SCEM-UA algorithm which is the predecessor of DREAM. They argue that in order for NS to be used in the algorithm it must be nonnegative and monotonically increasing with improved performance. To meet the former condition, the NS is set to zero when negative values are returned. The NS is only used via the posterior density ratio R of two samples, which can be expressed in the following form

$$R = \frac{1 - \frac{\sigma_{\epsilon_1}^2}{\sigma_o^2}}{1 - \frac{\sigma_{\epsilon_2}^2}{\sigma_o^2}} = \frac{\sigma_o^2 - \sigma_{\epsilon_1}^2}{\sigma_o^2 - \sigma_{\epsilon_2}^2} = \frac{K - SSE_1}{K - SSE_2}, \tag{6.7}$$

where K is a constant. However they found that the chain only slowly migrated towards high-performance regions due to two issues: (1) Poor representation of relative model performance and (2) Lack of ability to order poor model fits. For example the probability of moving from 0.9 to 0.8 is quite high as the posterior density ratio is 0.8/0.9 = 0.89. To deal with this issue K can be reduced which causes higher weight to be placed on small improvements in NSE. To address issue 2, the exact sums of squared error scores (SSE) were retained such that all model fits could be correctly ordered, even though this information was not used to calculate the ratio R. Lindblom et al. (2011) suggested a slightly different posterior density ratio

$$R = \frac{\exp(-SSE_1/T)}{\exp(-SSE_2/T)}, \tag{6.8}$$

where T is a scaling factor that is manually adjusted until a certain proportion of the observations are covered by the simulation limits. None of the suggestions resulted in a completely satisfactory exploration of the parameter space but inspired to the following posterior density ratio

$$R = \left(\frac{1 - SSE_1/T}{1 - SSE_2/T}\right)^{-J}, \tag{6.9}$$

where T was chosen lower than σ_o^2 and J is an even number manually adjusted until a satisfactory distribution of low and high likelihood values are found. The original SSE values were retained as recommended.

6.5 Generating uncertainty limits

To obtain the uncertainty limits the posterior likelihood values of all the K behavioural parameter sets $\boldsymbol{\Theta}_{\boldsymbol{B},\boldsymbol{i}}$ were calculated using (6.3). To each time step k the

ith simulated flow $Y^k_{sim,i}$ produced by the behavioural parameter set $\boldsymbol{\Theta_{B,i}}$ and its associated posterior likelihood $L_p\left(\boldsymbol{\Theta_{B,i}}\right)$ value is sorted in descending order with respect to flow magnitude. For example the 95% upper simulation limit can be found by summation of posterior likelihood values from the highest flow value and downwards until

$$\sum_{i=1}^{K} L_p\left(\Theta_{B,i}\right) = 0.05,$$

and likewise the lower 95% simulation limit can be obtained by

$$\sum_{i=1}^{K} L_p\left(\Theta_{B,i}\right) = 0.95.$$

The lower and upper bounds form the 95% simulation limits. Similarly other simulation limits can be derived.

7 Benchmarking models and uncertainty approaches

7.1 Residual analysis

For the stochastic approach the statistical assumptions underlying the model should be tested by analysing the model residuals for structure, dependence, normality and homoscedasticity. The following tests are all described in Madsen (2008).

- The *structure* of the residuals should be inspected by plotting the standardised residuals (std.residuals) as a function of the observations. If the statistical assumptions are met the std.residuals appear randomly scattered about zero, there should not be many std.residuals with magnitudes greater than 3, and most of the std.residuals should have values less than 2.
- The std.residuals should also be checked for *dependence* and ideally they are independent of each other. The autocorrelation function (ACF) and the partial autocorrelation function (PACF) can be used to check that the std.residuals are not serially autocorrelated.
- The std.residuals should also be *normally distributed*, and this can be examined by a quantile-quantile (QQ) plot or by the cumulative periodogram. In both tests a straight line indicates a normal distribution.
- Finally the std.residuals should be tested for *variance homogeneity* (homoscedasticity) which means that the variance of the std.residuals is constant across observations. If the variance changes with the flow magnitude (heteroscedasticity) some transformation of the observations (log-transform, square root, or some Box-Cox transformation) may be required to stabilise the variance.

Residual analyses was performed in **Papers III** and **V**.

7.2 Probabilistic prediction and simulation measures

The scope with model development is to use the models for making predictions and/or for simulation and assess the associated uncertainty (Gneiting et al., 2007; Gneiting and Raftery, 2007). Therefore it is meaningful to evaluate the probabilistic prediction and simulation performance, and hence not only consider the residuals. To do so we should first consider the reliability of the predictions or the simulation.

7.2.1 Reliability bias

A prediction or simulation interval $\hat{I}^{(\beta)}$ of a given quantile level β is defined as

$$\hat{I}^{(\beta)} = [\hat{Y}^{(l)}; \hat{Y}^{(u)}] \tag{7.1}$$

where $\hat{Y}^{(l)}$ and $\hat{Y}^{(u)}$ are respectively the lower and upper limits of the model output at levels $\beta/2$ and $1 - \beta/2$. The interval is *reliable* if the observed coverage $n^{(\beta)}$ matches the nominal coverage $1 - \beta$ $\beta \in [0; 1]$. For example if we consider the 90% prediction interval $\hat{I}^{(0.1)}$ we should expect to cover 90% of the observations. The reliability bias can be defined as the discrepancy between the nominal and observed coverage

$$b^{(\beta)} = 1 - \beta - n^{(\beta)} \tag{7.2}$$

and hence the ideal reliability bias is $b^{(\beta)} = 0$. It is seen from (7.2) that if the prediction interval covers more observations than expected then $b^{(\beta)} < 0$, and a negative bias is observed, and conversely if the prediction interval covers less observations than expected, i.e. $b^{(\beta)} > 0$, this would entail a positive bias. The reliability bias was calculated in **Papers IV**-**VI** wheras in **Paper II** the term *coverage ratio* (CR) was applied which is simply the fraction of observations covered by the prediction limits.

7.2.2 Sharpness

The sharpness (Gneiting et al., 2007) is a measure of the average size of the prediction/simulation interval which is relevant because in general we are interested in obtaining as narrow bounds as possible, while still reliable. Thus the sharper the intervals the more narrow. Sharpness is defined as

$$\overline{\delta}^{(\beta)} = \frac{1}{K} \sum_{k=1}^{K} \left(\hat{y}_k^{(u)} - \hat{y}_k^{(l)} \right), \tag{7.3}$$

where $\hat{y}_k^{(u)}$ and $\hat{y}_k^{(l)}$ represent, respectively, the upper and lower limits of the output at any given time step k and quantile level β of the entire period K. The term sharpness were used in **Papers IV**-**Paper VI** and in **Paper II** another designation *the average band width* (ABW) was used synonymously and the related term *average relative interval length* (ARIL) that measures the relative uncertainty of the bounds

$$ARIL = \frac{1}{N} \sum_{t=1}^{N} \left(\frac{\hat{y}_k^{(u)} - \hat{y}_k^{(l)}}{y_k} \right), \tag{7.4}$$

where y_k is the observed output. Hence ARIL can be used to compare uncertainties obtained from different study areas or study periods (Jin et al., 2010).

As both the reliability and the sharpness are average values calculated for the whole evaluation period, it is important to consider also the *resolution* which means that the performance measures should be calculated also in different subperiods corresponding to different output domains, which is relevant because the performance of the models may differ considerably during the two very different conditions dry and wet weather periods.

7.3 Model performance comparison

7.3.1 Evaluation using information criteria

Aikaike's information criterion (AIC) and the Bayesian information criterion (BIC) (Madsen, 2008) can be used two score different model candidates suitability for generating one step predictions or simulations and are thus tools for model selection. The information criteria are defined as

$$AIC = 2n - 2\ln(L) \tag{7.5}$$

$$BIC = -2\ln(L) + n\ln(K), \tag{7.6}$$

where n is the number of free parameters, L is the likelihood value, and K is number of observations. The preferred model is the model with the minimum AIC and BIC values. AIC not only rewards goodness of fit (as measured by the likelihood), but also includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages over-parameterisation. In calculating BIC over-parameterisation is penalised even more than in AIC because the penalisation depends on the number of observations. Both criteria have been used for model performance comparison in **Papers III** and **VI**.

7.3.2 Evaluation using an interval skill score criterion

The interval skill scoring criterion is used to assess the ability of the model to generate probabilistic predictions and simulations. The sharpness and reliability of given quantile is combined in the interval skill score criterion $Sc^{(\beta)}$ (Gneiting and Raftery, 2007) that rewards narrow and reliable confidence bounds. For any given quantile level for evaluation and K measurements the interval skill score is defined as

$$\begin{aligned} Sc^{(\beta)} &= \frac{1}{K}\sum_{k=1}^{K}(y_k^{(u)} - y_k^{(l)}) + \\ &\quad \frac{2}{\beta}\left[(y_k^{(l)} - y_k)\mathbf{1}\{y_k < y_k^{(l)}\} + (y_k - y_k^{(u)})\mathbf{1}\{y_k > y_k^{(u)}\}\right], \end{aligned} \tag{7.7}$$

where the indicator $\mathbf{1}$ is one if the condition is true. In case the observation misses the quantile interval a penalty is incurred which depends on the considered quantile level, meaning that an observation miss at the 95% confidence interval is much more expensive than a miss at e.g. the 5% confidence interval. Thus, the smaller the interval skill score the better the probabilistic prediction performance of the model, and so $Sc^{(\beta)}$ can be used to evaluate the different models' probabilistic performance and eventually to select a preferred model. It is noted that the skill score is suitable for comparison at different prediction horizons, which the AIC and BIC criteria are not. The skill score was used for model comparison in **Papers IV**, **V** and **VI** and also used to compare the simulation performance of the stochastic approach and the GLUE approach in Section 8.3.

8 Results and discussion

8.1 Results of the stochastic approach

8.1.1 Checking the statistical assumptions

Before the model can be used for simulation or prediction with confidence it is of upmost importance to check that the assumptions that underly the likelihood function holds in practice. The statistical assumptions of the grey-box models demand the model residuals to correspond to a white noise process, which means that the residuals should be serially uncorrelated, have constant variance and resemble a Gaussian distribution. The assumptions were tested in **Papers III** and **VI**. For

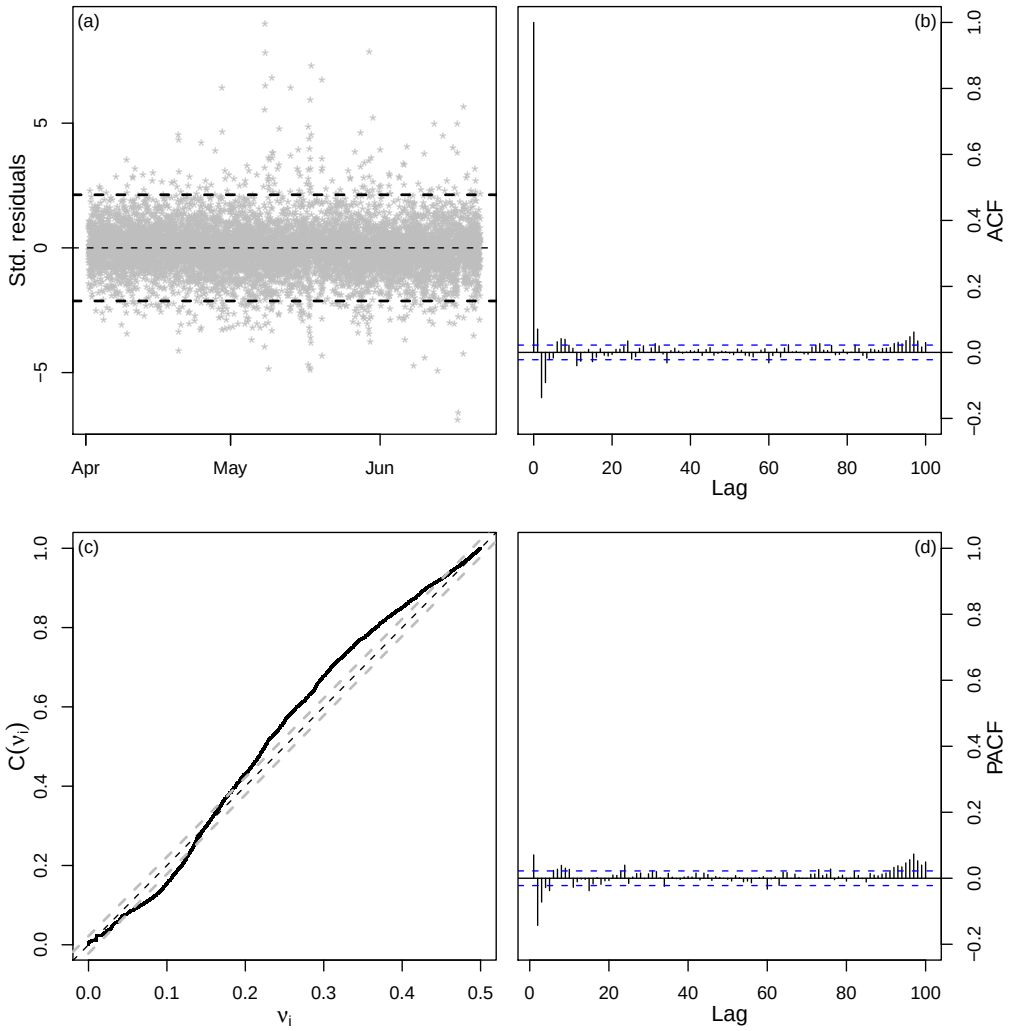

Figure 8.1: Validation of model a in Figure 4.3. (a) standardised residual plot; (b) autocorrelation function (ACF) (c) cumulative periodogram; and (d) partial autocorrelation function (PACF). Plot from **Paper III**.

the chosen model in **Paper III** the assumption of Gaussianity and mutual inde-

pendence were tested (for model a in Figure 4.3) by investigating the standardised residual plots, the autocorrelation function (ACF), the cumulative periodogram, and the partial autocorrelation function (PACF) as seen in Figure 8.1.

The actual residuals conform fairly well with the statistical assumptions of the residuals for the *one-step ahead prediction* although we did not check for heteroscedasticity in **Paper III**. There is a small departure from the whiteness assumption as seen from the cumulative periodogram, and we also observe some small spikes outside the 95% confidence interval at several lags in both ACF and PACF plots, which indicates that the periodicity of the wastewater variation is not perfectly described by the harmonic function.

In **Paper V** a different calibration data set and a different model (See model **e** in Figure 4.3) was employed and hence the residual tests provided other results. From an investigation of the std.residuals of the *one-step ahead prediction* shown

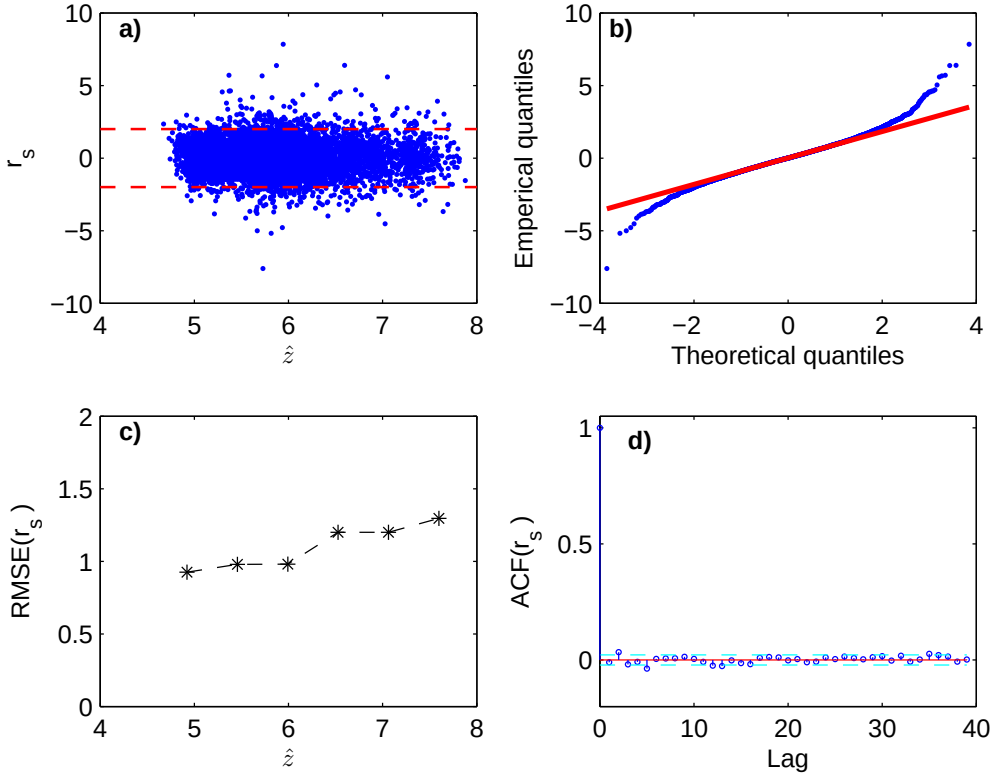

Figure 8.2: Investigation of the standardised one-step prediction residuals of Model M3 in **Paper V**.

in Figure 8.2, from the root mean square error of the std.residuals (Figure 8.2**c**), it is seen that homoscedasticity is nearly achieved, and from the Q-Q plot shown in

Figure 8.2**b** that the residuals does not resemble a Gaussian distribution very well at the tails, but up to the 95% confidence bounds the assumption is reasonable. From the ACF plot it is concluded that autocorrelation has been satisfactorily removed.

Figure 8.3 displays a residual investigation of the preferred *simulation model* in **Paper V**. Again the Gaussian distribution is not resembled too well at confidence levels higher than 95%, but below the residuals looks more in line with the assumed Gaussianity (see Figure 8.3**b**). There is clearly some heteroscedasticity left in Figure 8.3**c** which indicates that other observation transformations could possibly be tested. Autocorrelation is highly significant (Figure 8.3**d**) and it should be recognised that autocorrelation is unavoidable in simulation models with such short time resolution (15 minutes).

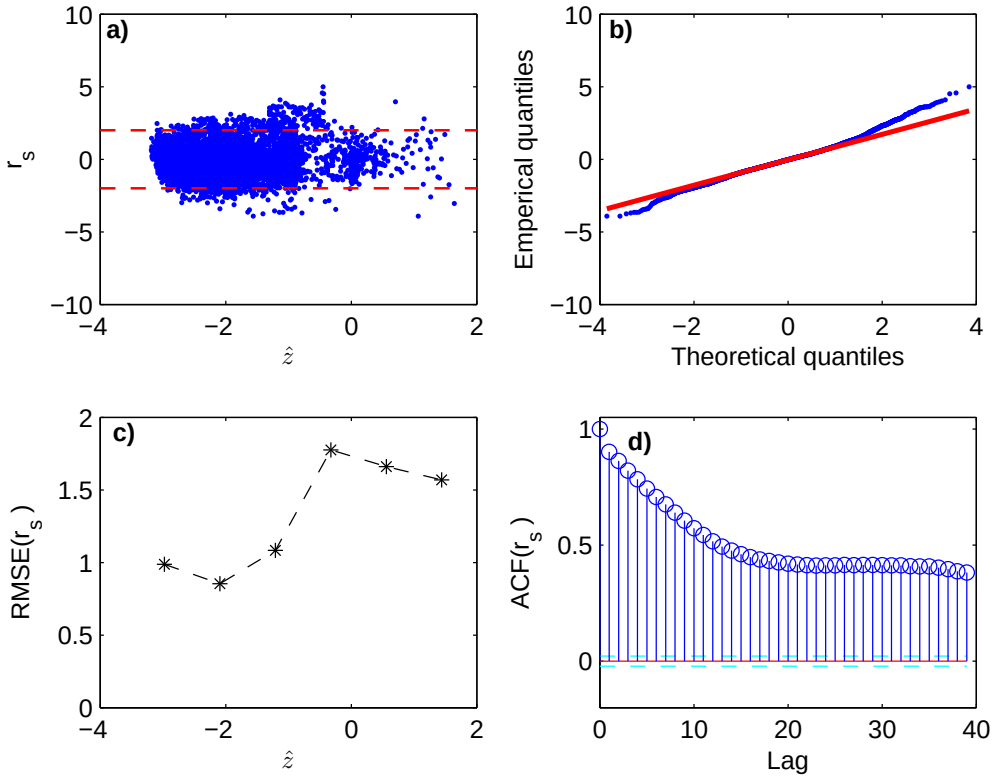

Figure 8.3: Investigation of the standardised residuals of simulation model M2 in **Paper V.**

8.1.2 Selecting among model candidates by information criteria

In paper **Paper III** the Aikaike information criterion (AIC) and the Bayesian Information Criterion (BIC) were used to choose the preferred model among the model candidates that differed with respect to the diffusion term only. It is recalled that

Table 8.1: Model comparison by information criteria in **Paper III**.

	$\log(L)$	DF	AIC	BIC	Diffusion terms
Model 1	11379.81	13	-22733.62	-22643.12	State independent
Model 2	12555.67	13	-25085.34	-24994.84	State proportional
Model 3	12461.81	15	-24893.62	-24789.19	State exponentiated

the model with the minimum AIC and BIC values is the preferred and both information criteria pointed at Model 2 which was the model with state proportional diffusion terms (see Table 8.1).

In **Paper VI** these criteria were also applied to select a preferred model among the model candidates at the *one-step ahead prediction* and in this paper the models differed from each other primarily with respect to the drift term, i.e. the amount of physical knowledge incorporated (see overview of models tested in **Paper VI**). Table 8.2 gives the obtained ranking order in terms of information criteria for the model candidates. As the model selection criteria are based on a likelihood func-

Table 8.2: Model selection from information criteria in **Paper VI**.

	$\log(L)$	DF	AIC	BIC	Priority, AIC	Priority, BIC
Model 1a	9,864	11	-19,706	-19,629	5	5
Model 1b	9,874	12	-19,724	-19,639	4	4
Model 1c	9,876	11	-19,730	-19,653	3	3
Model 2	9,843	7	-19,672	-19,623	6	6
Model 3	9,997	14	-19,966	-19,867	2	2
Model 4	10,027	18	**-20,018**	**-19,891**	1	1

tion that is a product of conditional densities for the one step prediction the criteria cannot be used to evaluate model performance at longer prediction horizons than the one-step (nor for simulation). Therefore we need another performance criterion for model selection if the model is to be applied for the N-step ahead prediction.

8.1.3 Evaluating prediction and simulation bounds

As the ultimate scope with stochastic model building is to enable decision-making on a more informative basis, we should also check the models probabilistic prediction and simulation performance, and hence we are especially concerned with the

correspondence between observed and nominal coverage of our confidence bounds (the reliability bias). In **Paper IV** the reliability biases were examined for the 90% confidence bounds up to 16 prediction steps (4 hours) ahead, see Figure 8.4. The models differed with respect to the diffusion term only, indicated by γ_1 and γ_2 (see **Paper IV** for details). All of the model candidates 90% prediction bounds at the

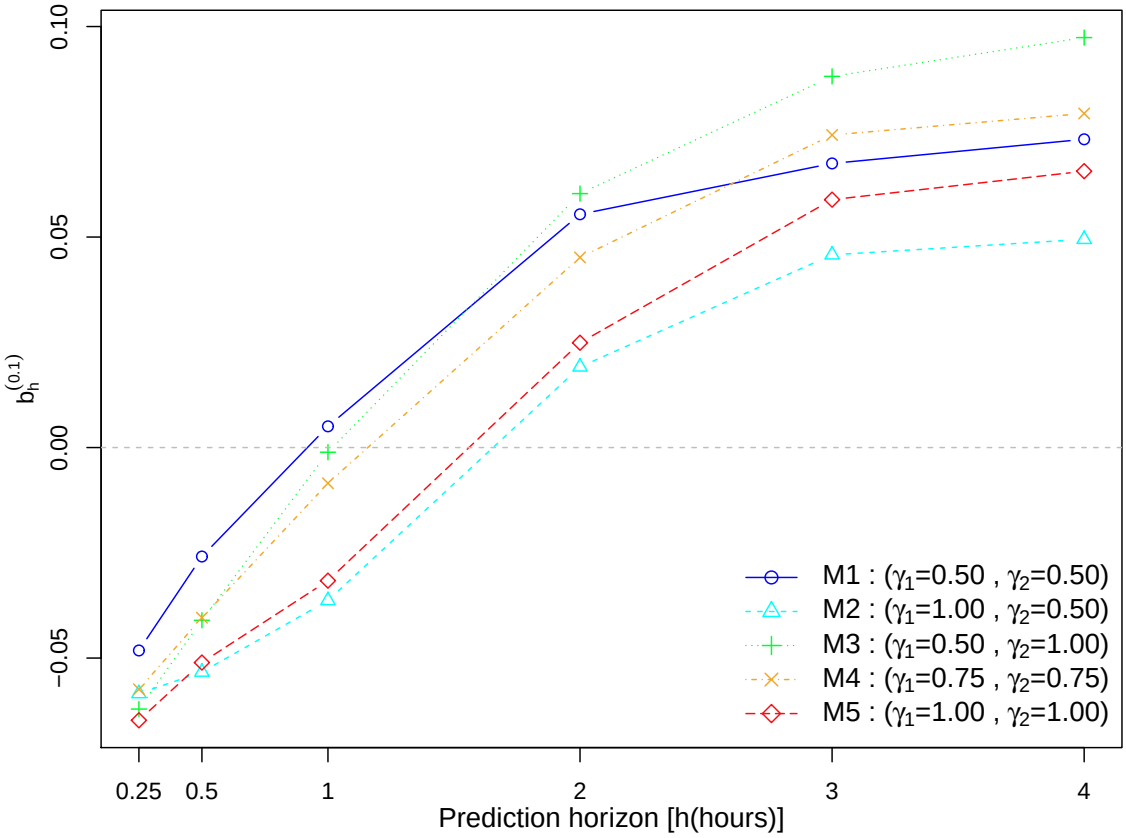

Figure 8.4: Examining the reliability bias of the model candidates 90% prediction bounds as a function of prediction horison. Figure from **Paper IV**.

one step ahead prediction are too wide covering approximately 5% more observations than prescribed. As the prediction horizon increases the agreement between the expected and observed coverage actually improves up to about 1-1.5 hour (the lines approaches zero), then an increasing coverage deficit materialises as the prediction horizon is further expanded up to 4 hours where the deficit is 5-10 %. Generally these results are encouraging, and it seems at first glance rather tempting to conclude that although some reliability biases are observed, this should not prevent us from putting a reasonable degree of confidence in the models prediction ability also at longer horizons.

Before doing so the important term "resolution" is considered and it is acknowl-

edged that two distinct flow classes should be differentiated: dry and wet weather periods. In Figure 8.5 this classification into dry and wet weather periods for separate examination was realised by an observed flow threshold of 540 m^3/h. It shows

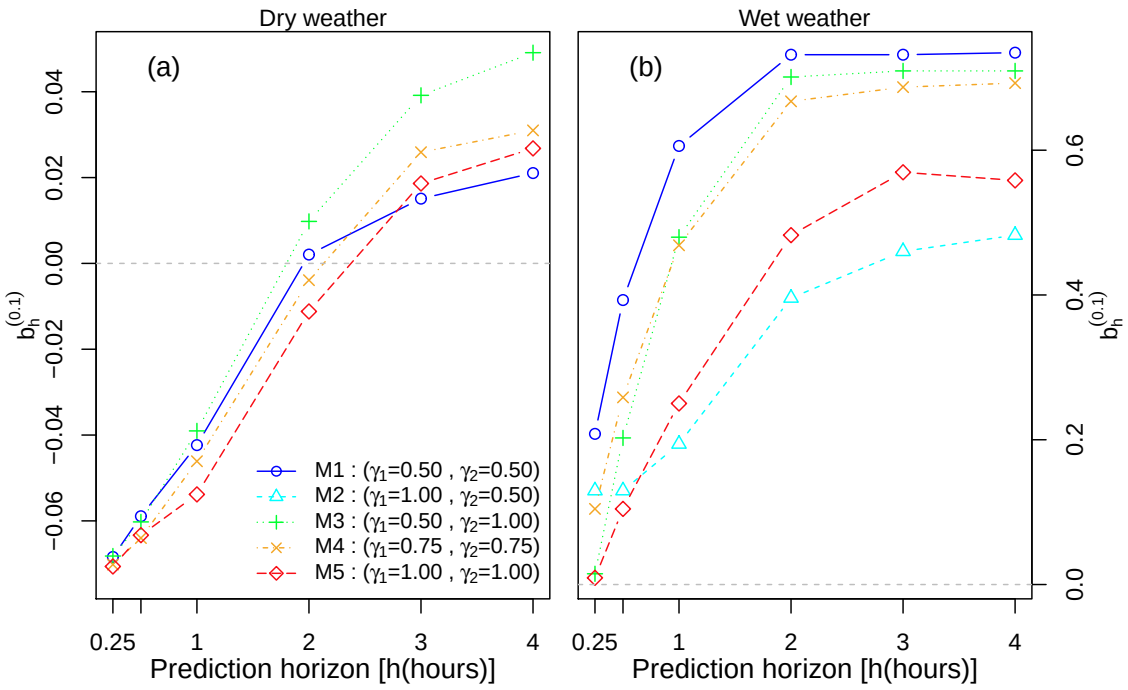

Figure 8.5: Reliability of the 90% prediction limits examined for dry and wet periods separately. Plots from **Paper IV.**

very clearly that the models prediction bounds perform fairly well in dry weather periods (biases in the range 0-6%) as opposed to wet weather periods where only the one-step prediction bounds are reliable for two of the models (M3 & M5). Already at prediction step 2 coverage deficits of 10-40% is observed, and at 4 hour predictions this worsens to 50-70%, suggesting that we should not put too much confidence in any of the proposed models when making predictions in wet weather beyond the one step.

To overcome these problems we may add more physical knowledge to the models as was done in **Paper VI** but another way would be to increase the uncertainty contained in the diffusion term. This was however not pursued due to instability related problems when a state exponentiated diffusion term of more than one ($\gamma_i > 1$) was used. Most of the uncertainty origins from a non-representative rain input and e.g. by placing rain gauges inside the catchment it is likely that more reliable prediction bounds would be found.

If the prediction bounds were to be applied for activating a WWTP's wet weather

operation as discussed in Chapter 1, the model would be tailored for this, i.e. if this decision requires a one-hour total volume prediction, then a one-hour total volume prediction is what we would aim for in the likelihood function. Hence the reliability bias for the one-hour prediction might not be as significant as observed here for a model that is optimised for a 0.25-hour flow prediction. Another possibility to improve the reliability of the predictions would be to set up a likelihood function for minimising the error of several prediction steps.

In paper **Paper V** the reliabilities for both a prediction model (M3 one-step) and three *simulation models* (M1-M3) were examined, see Figure 8.6 (left column) that shows the reliability bias overall and classified into wet and dry weather, as a function of nominal coverage. Considering the overall coverage, M3 appears to

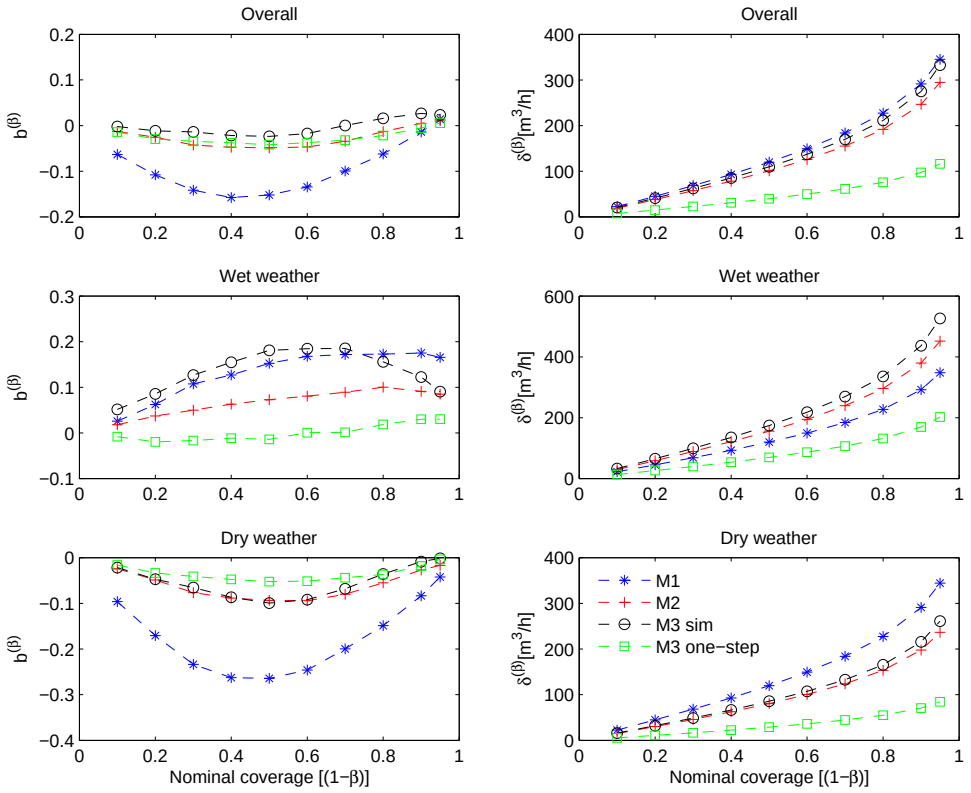

Figure 8.6: Reliability bias (left) and sharpness (right) for models M1-M3 calculated for the whole period (overall) and in dry and wet weather periods separately. Plot from **Paper V.**

be the favoured model, but again we observe reliability variations when focusing on dry and wet weather periods separately. The uncertainty bounds are generally too wide in dry weather and too narrow in wet weather. Of the simulation models M2 has the lowest reliability bias in wet weather, whereas in dry weather M2 and

M3 show more or less equivalent reliability bias. If comparing the reliability bias for M2 with the equivalent in Figure 8.5, we observe that a model tailored for simulation (with biases of less than 10%) perform much better in the long-run compared to a prediction model that exhibits large biases (in the range of 10-70% in wet weather) at prediction horizons of more than two or three prediction steps in Figure 8.5. In drawing this conclusion we should recall though that the two figures are not directly comparable, because M2 contain more physical knowledge and is estimated on a different calibration set than the models in Figure 8.5.

Notice also the improving effect of updating by comparing the reliability bias of the simulation models with the prediction model (M3 one-step) in Figure 8.5.

As mentioned in Chapter 1 models are just approximations of reality and we should perhaps add that so are stochastic models and their confidence bounds. Confronted with this fact and challenge we are still interested in finding one preferred model among the model candidates. In this choice we should not just consider reliability but also the width of the prediction bounds (sharpness) because a trade off exists between reliability and sharpness. Consider now Figure 8.7 that shows the dry and wet weather sharpness for the 90% prediction bounds as a function of the prediction horizon. Note the substantial difference in sharpness between the two weather classes and between the models and note also how the sharpness evolves as the prediction horizon increases. To determine which model to prefer given some prediction horizon and some reliability bias we adopt the interval skill score criterion as defined in Chapter 7, and recall that the model with the minimum skill score value is the preferred. Table 8.3 shows the skill score for several prediction steps conditioned on wet weather periods. It is noteworthy how the model choice de-

Table 8.3: Skill score calculated for the 90% prediction interval conditioned on wet weather periods. Table from **Paper IV**.

			Prediction Horizon						
	γ_1	γ_2	0.25h	0.5h	1h	2h	3h	4h	Average
M1	0.50	0.50	397.3	709.9	1243.0	1859.4	1996.8	2015.7	1370.4
M2	1.00	0.50	572.9	820.9	1044.5	**1328.0**	**1439.9**	**1456.1**	1110.4
M3	0.50	1.00	**289.1**	**489.3**	913.1	1607.8	1825.9	1874.5	1166.6
M4	0.75	0.75	372.3	631.8	1051.3	1619.2	1785.3	1815.3	1212.5
M5	1.00	1.00	365.5	583.0	**903.5**	1374.9	1547.9	1583.5	**1059.7**

pends on the prediction step, but if one single model is to be selected for the whole

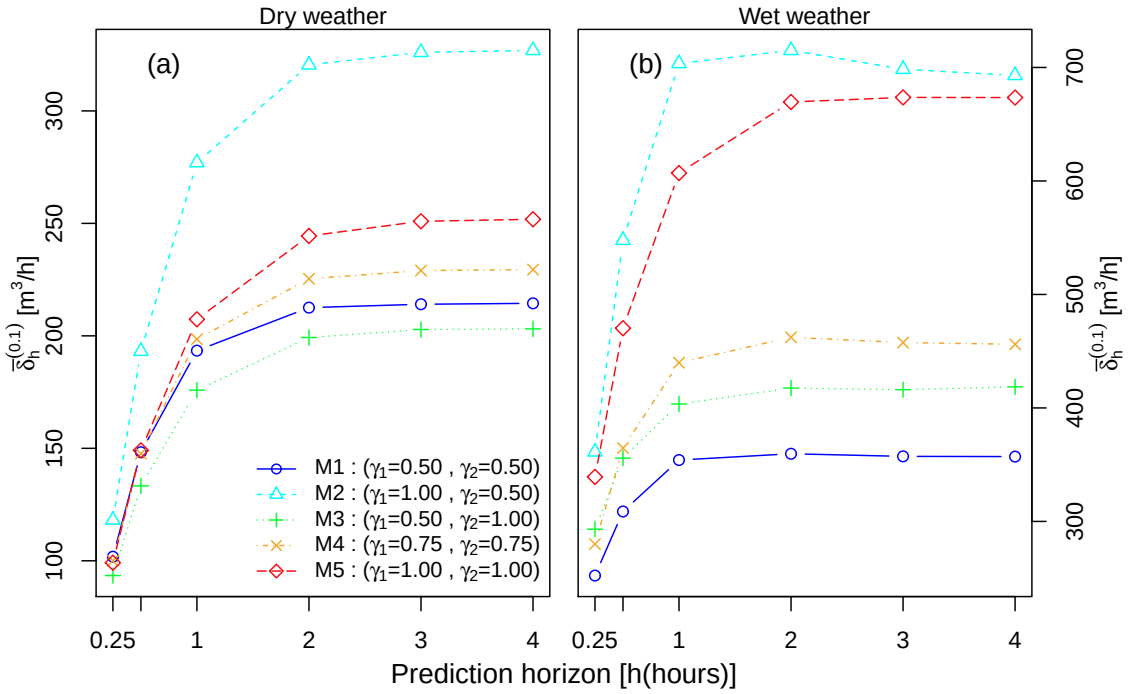

Figure 8.7: Sharpness of the 90% coverage, as a function of the prediction horizon and conditioned on the flow: (a) for dry weather periods ; (b) for wet weather periods. An observed flow threshold of 540 m^3/h separated the two weather periods. Figure from **Paper IV**.

prediction horizon M5 is the choice due to the best overall performance. In **Paper V** the skill score was used to compare the performance of three different simulation models and a one-step prediction model. Figure 8.8 shows this comparison and notice again here the effect updating has on the skill score. In all three cases (overall, dry weather, wet weather) the one-step prediction model outperforms the simulation models. However if the prediction model is launched in simulation mode i.e. if used without updating, then the this model is inferior to a model tailored for simulation (compare M2 and "M3 sim" in Figure 8.8). The skill score difference between M2 and "M3 sim" becomes larger with increasing nominal coverage. The obvious question for investigation is then at what prediction step the simulation model M2 starts to outperform the prediction model. This was not investigated in **Paper V** but stresses the importance of distinguishing between the two model types.

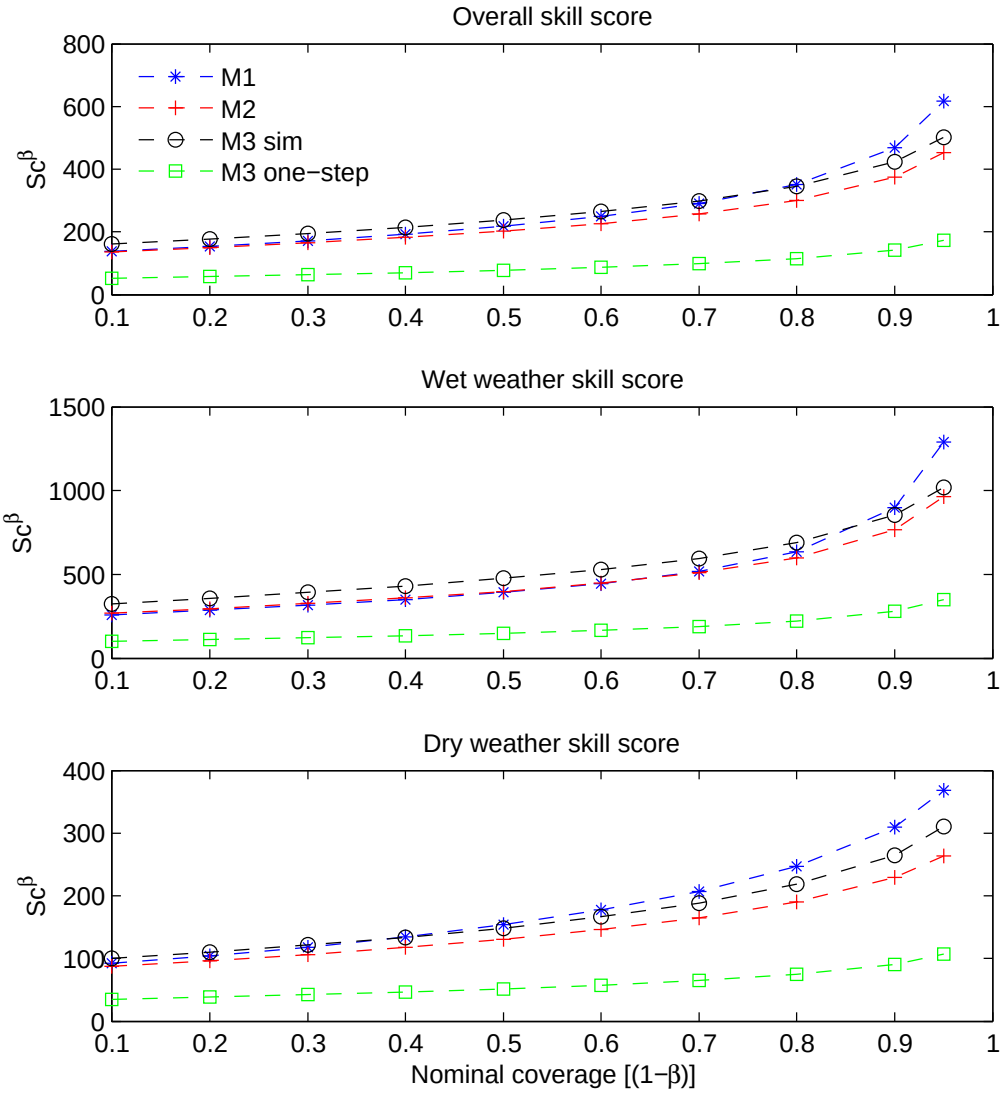

Figure 8.8: Interval skill score as a function of nominal coverage for all models, whole period and separately for dry and weather weather periods. Plot from **Paper V.**

8.2 Results of the epistemic approach

8.2.1 Extraction of behavioural parameter sets

In **Paper II** the aim was to obtain a 90% coverage of the GLUE generated 90% simulation limits, overall, in dry weather and in wet weather periods in the calibration period (2007). Again an observed flow threshold of 540 m^3/h were applied to separate the two flow classes. Two separate likelihood measures L_{dw} (for dry weather) and L_{ww} (for wet weather) were multiplied to obtain the overall likelihood measure L for extraction of the behavioural parameter sets, and the 10,000

best performing parameter sets (according to L) out of 200,000 Latin Hypercube Monte Carlo sampled sets were retained for further analysis. Figure 8.9 shows how L_{dw} and L_{ww} varies as L decreases. It was found impossible to obtain the desired

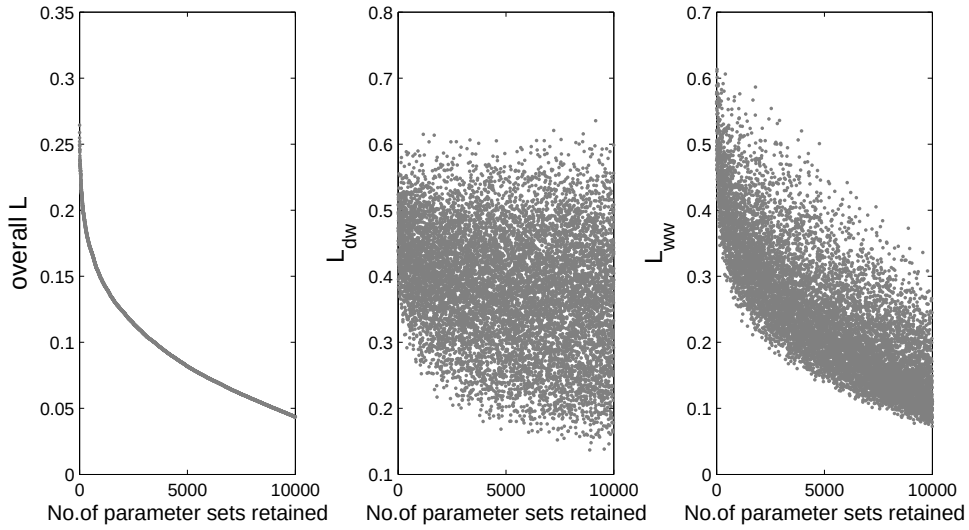

Figure 8.9: Likelihood vs. number of retained parameter sets. Shown for overall likelihood (L), dry weather likelihood (L_{dw}) and wet weather likelihood (L_{ww}). Plot from **Paper II.**

90 % coverage using 10,000 parameter sets, and it was pointless to retain more, as this had only minor influence on the coverage rate of the 90% simulation limits (see Figure 8.11 top left, data of 2007).

In the second GLUE study a different model (model **e** in Figure 4.3 and data (from 2010) were applied. This time the traditional Nash & Sutcliffe efficiency likelihood measure was used to rank the parameter sets and the aim was to obtain coverage rates consistent with the nominal coverages selected for the study in **Paper V**. The search for behavioural parameter sets was enhanced by the DREAM scheme and 20,000 parameter sets with L in the range 0.56-0.9 were retained and proved to fit the purpose.

8.2.2 Parameter uncertainty

As all sources to uncertainty are transferred to the parameter sets a large parameter uncertainty is generally expected. Figure 8.10 shows a dotty plot of the retained parameter sets and it is seen how increasing the number of behavioural parameter sets from 500 to 10,000 entails that the posterior parameter range approaches the prior parameter range for many of the wet weather parameters, which is a clear sign of equifinality as the prior parameter range initially was chosen to be quite

wide. Generally the dry weather parameters are more well determined as evidenced by the location of the dots that are concentrated in a smaller region of the prior parameter space.

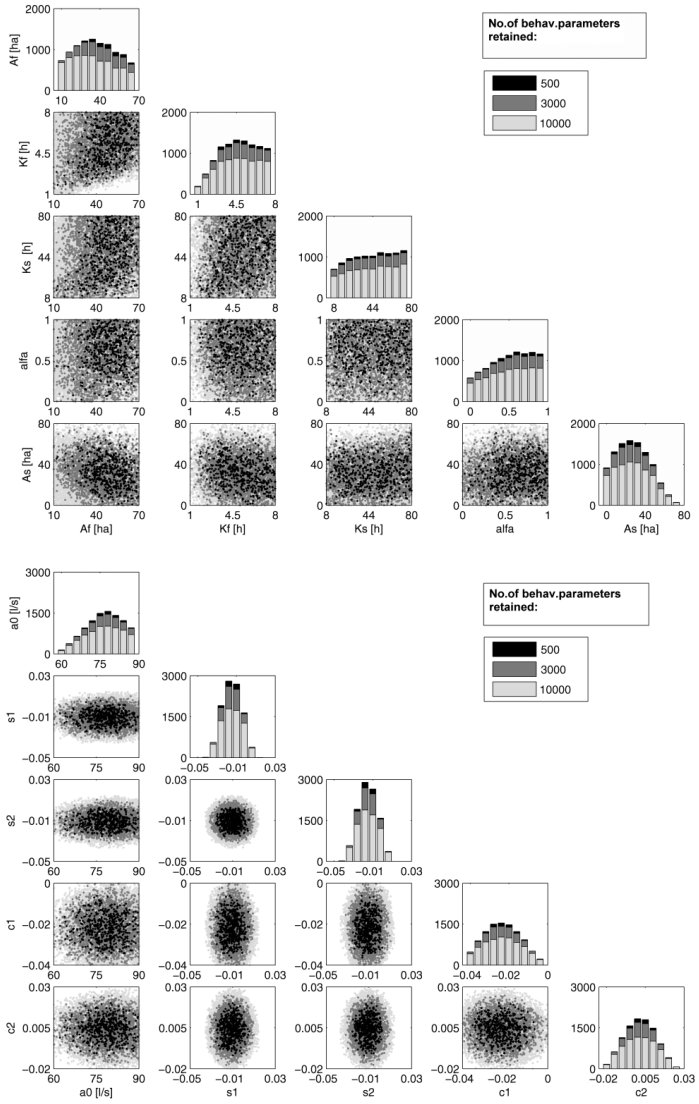

Figure 8.10: Dotty plots of wet weather parameters (top) and dry weather parameters (bottom). Plot from **Paper II.**

8.2.3 Consistency of the GLUE generated bounds

In **Paper II** the consistency of the GLUE generated simulation bounds was investigated, i.e. it was tested if the coverage obtained during calibration corresponded to the coverage obtained in validation. Two validation periods (2008 and 2009) were included for this consistency analysis as seen in Figure 8.11. The coverage (CR) increases and the curve flattens off as the number of retained parameter sets is in-

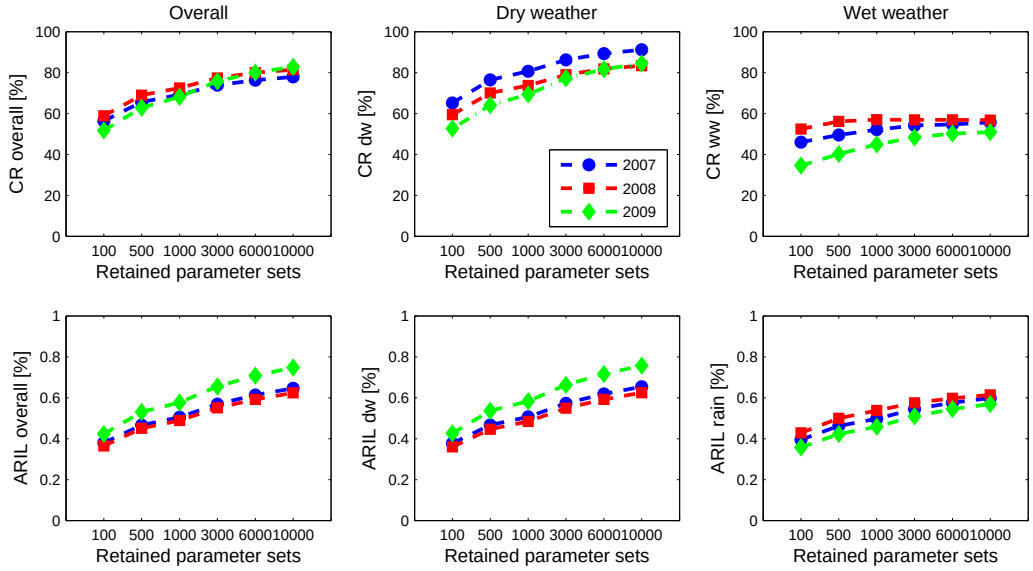

Figure 8.11: CR (upper panels) and ARIL (lower panels) vs. the number of retained parameter sets in the calibration year (2007) and the two validation years (2008 and 2009) for the total 6 months period (left panels), the dry weather periods (middle panels) and the wet weather periods (right panels).

creased. Good consistency between the calibration year (2007) and the validation years for the overall coverage is observed, however when looking separately at dry and wet weather periods, less good consistency is clearly observed. Figure 8.11 also shows the average interval length (ARIL) of the 90% simulation bounds as a function of the number of parameter sets included. It is seen that the width of the simulation limits rises with the the number of behavioural parameter sets and good consistency between ARIL values obtained in 2007 and 2008 is observed whereas ARIL values of 2009 differ from 2007 and 2008 values.

8.3 Comparison of the stochastic and epistemic approaches

The best simulation model of **Paper V** (M2) was chosen to represent the stochastic approach and a GLUE uncertainty analysis was conducted with the same model as M2 (Figure 4.3) although without statistical noise terms representing the epistemic approach. The uncertainty performance of the two approaches was then compared with respect to reliability, sharpness and interval skill score calculated overall and in dry and wet weather periods separately.

Figure 8.12 shows the comparison of reliability and sharpness as a function of

nominal coverage rates. First of all it is seen that deriving the behavioural param-

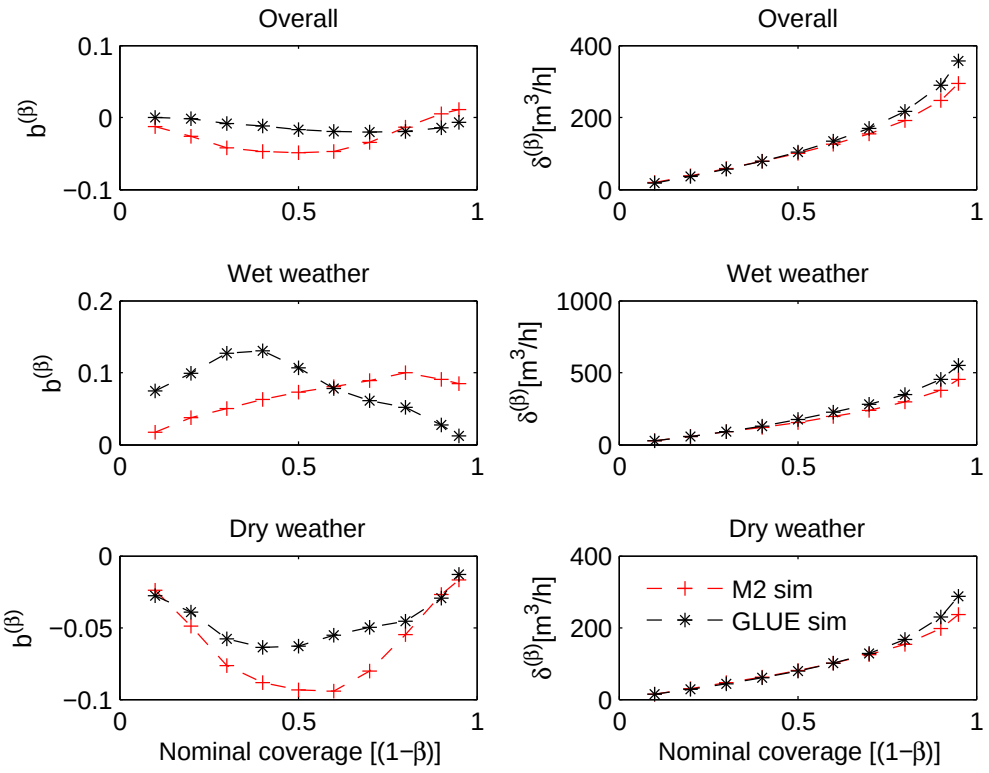

Figure 8.12: Comparison of reliability bias (left) and sharpness (right) for the stochastic (M2) and the epistemic (GLUE) approach calculated for the whole period altogether and in dry and wet weather periods separately.

eter sets using the DREAM scheme so that an overall coverage consistent with the nominal coverages succeeded. There is nearly no visible overall reliability bias. In both dry and wet weather some bias is observed for both approaches but the stochastic approach produces the sharper uncertainty limits in all three cases. The skill score comparison in Figure 8.13 suggests that the stochastic approach performs a bit better than the epistemic approach in this particular case study. In concluding this it should be recalled that the GLUE derived uncertainty limits are quite influenced by the chosen likelihood measure and therefore the conclusion found here cannot be generalised but shows that GLUE is able to generate somewhat similar skill scores as the stochastic approach. It should also be mentioned that the statistical assumptions behind M2 were not entirely satisfied and hence it may be possible to find a transformation of the observations that would improve the skill score for M2.

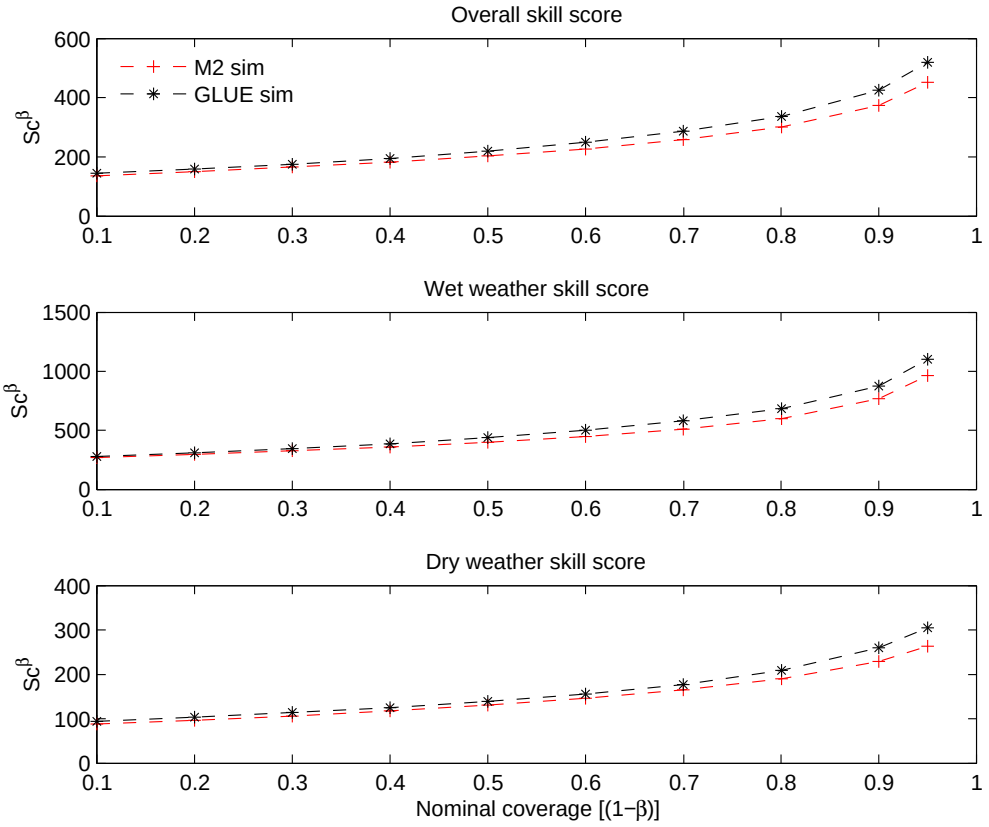

Figure 8.13: Skill score comparison between the stochastic (M2) and epistemic approach (GLUE) calculated for the whole period altogether and in dry and wet weather periods separately.

Table 8.4 compares the interpretation of parameter uncertainty of the two approaches. As all the uncertainty is lumped into the parameters in the GLUE approach much wider parameter intervals are found in GLUE compared to the stochastic approach. In the stochastic approach the assumption is that a single optimal parameter set exists and this parameter set can be found within the specified limits with 95% confidence.

Table 8.4: Parameter uncertainty of the stochastic and epistemic approaches. $[\hat{\theta}_{min}; \hat{\theta}_{max}]$ indicates the 95% confidence interval for the stochastic approach and for the epistemic approach the minimum and maximum parameter values used in the posterior parameter distribution.

Parameter	M2		GLUE	
	$\hat{\theta}_{min}$	$\hat{\theta}_{max}$	$\hat{\theta}_{min}$	$\hat{\theta}_{max}$
Wastewater:				
s_1	-46.8	-42.8	-227.3	135.0
c_1	-87.7	-92.1	-248.0	94.0
s_2	-46.1	-40.7	-179.5	98.8
c_2	20.0	25.2	-100.0	148.4
a_0	275.6	278.8	150.7	458.4
Fast rainfall-runoff:				
A_f	65.2	66.8	12.4	150.0
K_f	4.83	5.03	1.36	15.0
α	0.34	0.38	0.0	1.0
$S_{f1,max}$	8445	9209	2032	20,000
γ	0.55	0.67	0.0	0.9
$S_{d,max}$	18,165	25,813	116	40,000
Q_p	884.6	895.4	10	2,500
K_p	2.1	2.9	0.2	14.0
Slow infiltration:				
K_s	82.7	87.5	24.1	10,000
ς	1,634	1,710	103.9	8,000

9 Conclusions

This main objective of this thesis has been to qualify and quantify output uncertainty in simulation and prediction using simple rainfall-runoff models for urban drainage systems because such models are attractive for control purposes and because uncertainty plays a significant role in the choice of an optimal control strategy. To assess the prediction and simulation uncertainty in a realistic setting a case catchment with real flow data was needed. The catchment of Ballerup was chosen and inputs included data from point rain gauges and monthly evaporation data. Flow data sampled downstream from the catchment were used for model conditioning and uncertainty evaluation. The data period covered 2007-2010. The 1,320 ha catchment consists of both separated and combined sewage pipes, is significantly affected by infiltration inflow and facilitates a few detention basins and some pumping stations.

To quantify the model output uncertainty two uncertainty approaches were distinguished, the stochastic and the epistemic approach. A frequentist approach was taken for the stochastic approach and the Generalised Likelihood Uncertainty Estimation (GLUE) methodology was adopted for the epistemic approach. Both methods were applied for assessing the simulation uncertainty but only the stochastic approach was applied for making predictions from continuous updating of model states.

Model development

The simple serial linear reservoir flow routing principle was applied for modelling both the fast rainfall runoff from paved areas and the slow infiltration inflow from permeable areas. The wastewater flow variation was modelled by a harmonic function. Models of different complexity in terms of describing features such as flow constraints, basins and pumps were tested for their ability to describe the output with a time resolution of 15 minutes.

The stochastic grey-box approach for prediction

The stochastic grey-box modelling approach that incorporates only the most important physical knowledge in the models and uses information from data to describe the uncertainty and for estimating the model parameters was chosen. Uncertainty is accounted for explicitly by two noise terms, the observation noise term and the diffusion noise term. The parameters are estimated using the prediction error method by maximising a likelihood function. All the estimated model parameters were tested for significance by a t-test and although most of the parameters could be de-

termined, some had to be fixed. Increasing the model complexity generally implied that more parameters had to be fixed. A flow proportional observation noise term was preferred to an additive observation noise term because the flow metre accuracy generally decreased with the flow magnitude and because the risk of receiving negative flows is avoided. A state dependent diffusion term was also preferred to an additive noise term because the risk of obtaining negative states could thereby be avoided. The state dependent diffusion term that was implemented by a Lamperti transformation of the states is also justified by the fact that the flow uncertainty generally increases with the volume of the rain. The residual tests contributed to developing the observation and diffusion noise terms and for the eventual approval of the models. The Akaikes information criterion (AIC) and the Bayesian information criterion (BIC) were applied to choose a preferred model among the model candidates when considering the one-step prediction performance corresponding to 15 minutes lead time, and the two criteria agreed well. In general the models were unsuitable for making probabilistic forecasts beyond the one-step especially during wet weather periods. For the case when a longer prediction horizon is required a skill scoring criterion was applied to select a preferred model out of several model candidates. Even though a preferred model can be pinpointed from this criterion it should be recognised that the model do not necessarily provide reliable forecasts. Therefore the sharpness and the reliability bias that evaluates the difference between the nominal and observed coverage of a particular quantile were also used to compare the models.

The stochastic approach for simulation

In the case that real time data are unavailable for continuous updating of the model states a simulation model can be applied instead. Assuming flow data from a measuring campaign are available for calibration it is possible to estimate the parameters by maximum likelihood estimation and estimate the total output uncertainty. The parameter estimation can be conducted in two ways, either using the prediction error method (as discussed above) or by minimising the total output error for the whole calibration period. In the latter case the diffusion term is removed and only the observation noise term is kept. Three different models were tested. Two of the models contained only an observation noise term, one which was additive and one which was logistic, whereas the last one contained both an observation noise term and a diffusion noise term. Parameters were generally significant although a few parameters had to be fixed. In order to remove the heteroschedastic residuals structure it was necessary to apply a logistic transformation of the observations,

however auto-correlation remained. The simulation model with a diffusion term embedded did not remove the auto-correlation either and hence it seems unrealistic to assume residual independency in the simulation case. The logistic observation noise term improved the reliability of the confidence bounds significantly compared to an additive observation noise term. This model also proved to be the most reliable simulation model and the skill scoring criterion pointed to this model as the preferred model of the three simulation models. The reliability of the preferred models confidence bounds were similar in dry and wet weather periods. A skill scoring comparison of a simulation and a prediction model showed that a major improvement is gained by updating the model states continuously, i.e. updating of model states results in much lower uncertainty.

The epistemic approach for simulation

The GLUE methodology is very different from the chosen stochastic approach because of the absent requirements to the residuals and because parameters are seen as stochastic variables. Nevertheless the aim is the same, namely to cover a proportion of observations consistent with the considered quantile with maximum sharpness. Two different likelihood measures were applied for qualifying the model parameters, the traditional Nash & Suttcliffe model efficiency coefficient and an exponential variant. For sampling the model space a latin hypercube monte carlo method (LHS) and a markov chain monte carlo method (DREAM) were both applied. The algorithm in DREAM is very efficient in locating the high performing areas of the model space and the posterior density ratio had to be adjusted to secure visits of lower performing areas. Generally the use of DREAM saves computational time. To select the behavioral parameter sets a coverage criterion of 90% by the generated 90% uncertainty limits was proposed. It was however impossible to reach such a high coverage in the calibration period and the coverage were much lower during rain than in dry weather. The consistency of the GLUE generated uncertainty limits were then evaluated by comparing obtained coverage rates between the calibration period and two validation periods and a reasonable consistency was generally found.

Comparison of the stochastic and epistemic approach for simulation

To facilitate a comparison of the two approaches to uncertainty evaluation the sharpness, reliability and skill score were calculated for both using the same data set for calibration and evaluation. Results showed that very similar performance was obtained, with the stochastic method as the preferred. It should be stressed that

in order to test this more properly a data set other than the calibration data should be used for validation. The work of this thesis has demonstrated that the statistical requirements to the formal stochastic approach, i.e. that residuals should conform to a white noise process, are very hard to fulfill in practice, especially in simulation and for predictions beyond one step where significant auto-correlation remained. The requirements were easier to fulfill for the one-step prediction which was due to the continuous updating of the model states. As a consequence hereof the reliability of prediction steps beyond the one-step prediction was quite biased, especially during rain. Although the GLUE derived uncertainty limits did not prove completely consistent nor superior to the stochastic derived uncertainty description, the underlying assumption of the GLUE methodology that uncertainty in modeling and simulation is not only of stochastic nature seems fairly consistent with the results of this thesis. A major drawback of the GLUE methodology is the lumping of total uncertainty into the parameters which entails a loss of physicality of the model parameters. Conversely the parameter estimates of the stochastic approach are physically meaningful. This thesis has contributed to developing simplified rainfall-runoff models that are suitable for stochastic model predictive control of urban drainage systems.

10 Outlook

The benchmarking tools and the grey box models developed in connection with this thesis may well be applied to other catchments for assessing the uncertainty on model outputs in connection with model predictive control investigations. The modelled output are in that case not limited to flow forecasting but could just as well be applied for forecasting of water level, volume or concentrations.

It should also be mentioned that measurements from levels of storage tanks or pumping stations could be directly used in the observation equation of the grey box model which would be desirable in many cases for improving the prediction performance if such measurements are available.

The predictions made in connection with this thesis all assumed future rain inputs to be known. However because of the retention time of the system, i.e. between a rainfall is recorded at the rain gauges until a corresponding runoff is observed at the flow meter in the sewer system, it is often possible to obtain a flow forecast without the need for rainfall input forecasts. This potential should be investigated further.

As it was shown that a simulation model was more reliable than a prediction model (estimated by minimsing the one step ahead errors) when considering several prediction steps into the future it would be desirable to know up to which prediction step we should apply the prediction model and hence at what prediction step the simulation model becomes the preferred model. The skill score benchmaring tool outlined in this thesis can be used for this investigation.

It was also shown that if the reliability were calculated separately for dry and wet weather periods, the reliability in dry weather periods (with reliability biases in the order of 5% at prediction steps up to 4 hours) were generally higher than in wet weather periods in which a rapidly deteriorating reliability as a function of prediction horizon was found (with reliability biases in the order of 50-80% at 4 hour prediction steps). These findings suggest that the models predictive performances for prediction steps larger than one needs to be improved, and one way could perhaps be to replace the minimisation of the one step prediction error in the extended Kalman filter with a k step error minimisation in accordance with whatever lead time is required for the desired model predictive control implementation.

It is in connection with model predictive control of a waste water treatment plant perhaps more important to predict the volume load to the plant over the next half

hour and in that case a grey box model optimised for this might perform better than a model optimised for predicting the inflow with 15 minutes time steps.

The use of rain radars to extend the lead time and obtain a more adequate spatio-temporal rain input description also holds a significant flow prediction potential and the use of the performance indicators outlined in this thesis could be used to conclude on this potential.

Finally the next step should be to show how the probabilistic forecasts can be utilised in model predictive control.

11 References

Achleitner, S., Fach, S., Einfalt, T., Rauch, W., 2009. Nowcasting of rainfall and of combined sewage flow in urban drainage systems. Water Science & Technology - WST 59 (6), 1145–1151.

Andréassian, V., Lerat, J., Loumagne, C., Mathevet, T., Michel, C., Oudin, L., Perrin, C., 2007. What is really undermining hydrologic science today? Hydrological Processes 21 (20), 2819–2822.

Baadsgaard, M., Nielsen, J. N., Spliid, H., Madsen, H., Preisel, M., 1997. Estimation in stochastic differential equations with a state dependent diffusion term. SYSID '97 - 11th IFAC symposium of system identification, IFAC.

Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology 320 (1-2), 18–36.

Beven, K., 2009a. Environmental modelling: An uncertain future / An introduction to techniques for uncertainty estimation in environmental prediction. Routledge Taylor & Francis Group.

Beven, K. J., 2009b. Comment on "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" by Jasper A. Vrugt, Cajo J. F. ter Braak, Hoshin V. Gupta and Bruce A. Robinson. Stochastic Environmental Research and Risk Assessment 23 (7), 1059–1060.

Beven, K. J., 2012. Rainfall - runoff modelling / The primer. John Wiley & Sons.

Beven, K., Binley, A., 1992. The future of distributed models - model calibration and uncertainty prediction. Hydrological Processes 6 (3), 279–298.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology 249 (1-4), 11–29.

Beven, K., Smith, P., Freer, J., 2007. Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. Journal of Hydrology 338 (3-4), 315–318.

Beven, K. J., Smith, P. J., Freer, J. E., 2008. So just why would a modeller choose to be incoherent? Journal of Hydrology 354 (1-4), 15–32.

Beven, K., Smith, P. J., Wood, A., 2011. On the colour and spin of epistemic error (and what we might do about it). Hydrology and Earth System Sciences 15 (10), 3123.

Blasone, R.-S., Vrugt, J. A., Madsen, H., Rosbjerg, D., Robinson, B. A., Zyvoloski, G. A., 2008. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Advances in Water Resources 31 (4), 630–648.

Box, G. E., Jenkins, G. M., Reinsel, G. C., 2008. Time series analysis / Forecasting and control. Wiley, wiley series in probability and statistics.

Bruen, M., Yang, J., 2006. Combined hydraulic and black-box models for flood forecasting in urban drainage systems. J. Hydrologic Eng. 11 (6), 589–596.

Carstensen, J., Nielsen, M. K., Strandbæk, H., 1998. Prediction of hydraulic load for urban strom control of a municipal wwt plant. Water Science & Technology 37 (12), 363–370.

Choi, H. T., Beven, K., 2007. Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in an application of topmodel within the glue framework. Journal of Hydrology 332 (3-4), 316–336.

Clark, M. P., Kavetski, D., Fenicia, F., 2011. Pursuing the method of multiple working hypotheses for hydrological modeling. WATER RESOURCES RESEARCH 47, –.

Coccia, G., Todini, E., Coccia, G., 2011. Recent developments in predictive uncertainty assessment based on the model conditional processor approach. Hydrology and Earth System Sciences 15 (10), 3253–3274.

Dorado, J., Rabuñal, J. R., Pazos, A., Rivero, D., Puertas, A. S. J., 2003. Prediction and modeling of the rainfall-runoff transformation of a typical urban basin using ann and gp. Applied Artificial Intelligence 17, 329–343, DOI: 10.1080/08839510390198673.

Dotto, C., Deletic, A., Fletcher, T., 2009. Analysis of parameter uncertainty of a flow and quality stormwater model. Water Science and Technology 60 (3), 717–725.

Dotto, C., Kleidorfer, M., Deletic, A., Rauch, W., McCarthy, D., Fletcher, T., 2011. Performance and sensitivity analysis of stormwater models using a bayesian approach and long-term high resolution data. Environmental Modelling and Software 26 (10), 1225–1239.

Engeland, K., Renard, B., Steinsland, I., Kolberg, S., 2010. Evaluation of statistical models for forecast errors from the hbv model. Journal of Hydrology - Amsterdam 384 (2-2), 142.

Engeland, K., Xu, C.-Y., Gottschalk, L., 2005. Assessing uncertainties in a conceptual water balance model using bayesian methodology. Hydrological Sciences Journal - Journal des Sciences Hydrologiques 50 (1), 45.

Franz, K., Hogue, T., 2011. Evaluating uncertainty estimates in hydrologic models: Borrowing measures from the forecast verification community. Hydrology and Earth System Sciences 15 (11), 3367–3382.

Freer, J., Beven, K., Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. Water Resources Research 32 (7), 2161–2173.

Freni, G., Mannina, G., Viviani, G., 2009a. Uncertainty in urban stormwater quality modelling: The influence of likelihood measure formulation in the GLUE methodology. Science of the Total Environment 408 (1), 138–145.

Freni, G., Mannina, G., Viviani, G., 2009b. Urban runoff modelling uncertainty: Comparison among bayesian and pseudo-bayesian methods. Environmental Modelling and Software 24 (9), 1100–1111.

Fu, G., Butler, D., Khu, S.-T., Sun, S., Khu, S.-T., 2011. Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory. Water Resources Research 47 (2), W02534.

Gallagher, M., Doherty, J., 2007. Parameter estimation and uncertainty analysis for a watershed model. Environmental Modelling and Software 22 (7), 1000–1020.

Gneiting, T., Balabdaoui, F., Raftery, A. E., 2007. Probabilistic forecasts, calibration and sharpness. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 69 (2), 243–268.

Gneiting, T., Raftery, A. E., 2007. Strictly proper scoring rules, prediction and estimation. Journal of the American Statistical Association 102 (477), 359–378.

Haario, H., Laine, M., Mira, A., Saksman, E., 2006. Dram: Efficient adaptive mcmc. Statistics and Computing 16 (4), 339–354.

Hansen, L. S., Borup, M., Møller, A., Mikkelsen, P. S., 2011. Flow forecasting in urban drainage systems using deterministic updating of water levels in distributed hydraulic models. 12th International Conference on Urban Drainage.

Hostache, R., Matgen, P., Montanari, A., Montanari, M., Hoffmann, L., Pfister, L., 2011. Propagation of uncertainties in coupled hydro-meteorological forecasting systems: A stochastic approach for the assessment of the total predictive uncertainty. Atmospheric Research 100 (2-3), 263–274.

Jin, X., Xu, C.-Y., Zhang, Q., Singh, V., 2010. Parameter and modeling uncertainty simulated by glue and a formal bayesian method for a conceptual hydrological model. Journal of Hydrology - Amsterdam 383 (4-4), 147.

Jørgensen, H. K., Rosenørn, S., Madsen, H., Mikkelsen, P. S., 1998. Quality control of rain data used for urban runoff systems. Water Science and Technology 37 (11), 113–120.

Kavetski, D., Kuczera, G., Franks, S., 2006. Bayesian analysis of input uncertainty in hydrological modeling: 2. application. WATER RESOURCES RESEARCH 42 (3), –.

Kloeden, P. E., Platen, E., 1999. Numerical Solution of Stochastic Differential Equations. Springer-Verlag.

Kristensen, N. R., Madsen, H., 2003. Continuous Time Stochastic Modeling - CTSM 2.3 - Mathematics Guide. Technical Unversity of Denmark.

Kristensen, N. R., Madsen, H., Jørgensen, S. B., 2004a. A method for systematic improvement of stochastic grey-box models. Computers and Chemical Engineering 28 (8), 1431–1449.

Kristensen, N. R., Madsen, H., Jørgensen, S. B., 2004b. Parameter estimation in stochastic grey-box models. Automatica 40 (2), 225–237.

Krueger, T., Freer, J., Quinton, J. N., Macleod, C. J. A., Bilotta, G. S., Brazier, R. E., Butler, P., Haygarth, P. M., 2010. Ensemble evaluation of hydrological model hypotheses. WATER RESOURCES RESEARCH 46, –.

Laloy, E., Fasbender, D., Bielders, C., 2010. Parameter optimization and uncertainty analysis for plot-scale continuous modeling of runoff using a formal Bayesian approach. Journal of Hydrology 380 (1-2), 82–93.

Li, L., Xia, J., Xu, C.-Y., Singh, V., 2010. Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. Journal of Hydrology - Amsterdam 390 (4-4), 210.

Lindblom, E., Ahlman, S., Mikkelsen, P., 2011. Uncertainty-based calibration and prediction with a stormwater surface accumulation-washoff model based on coverage of sampled zn, cu, pb and cd field data. Water Research 45 (13), 3823–3835.

Liu, Y., Freer, J., Beven, K., Matgen, P., 2009. Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error. Journal of Hydrology - Amsterdam 367 (2-2), 93.

Madsen, H., 2008. Time Series Analysis. Chapman & Hall/CRC.

Madsen, H., Thyregod, P., 2011. Introduction to general and generalized linear models. CRC Press, texts in statistical science.

Mantovan, P., Todini, E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology 330 (1-2), 368–381.

Mantovan, P., Todini, E., Martina, M. L. V., 2007. Reply to comment by Keith Beven, Paul Smith and Jim Freer on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology". Journal of Hydrology - Amsterdam 338 (3-4), 319.

Matott, L. S., Babendreier, J. E., Purucker, S. T., Matott, L. S., 2009. Evaluating uncertainty in integrated environmental models: A review of concepts and tools. Water Resources Research 45 (6), W06421.

McIntyre, N., Young, P., Orellana, B., Marshall, M., Reynolds, B., Wheater, H., 2011. Identification of nonlinearity in rainfall-flow response using data-based mechanistic modeling. WATER RESOURCES RESEARCH 47, –.

McMillan, H., Clark, M., 2009. Rainfall-runoff model calibration using informal likelihood measures within a markov chain monte carlo sampling scheme. WATER RESOURCES RESEARCH 45, –.

Morawietz, M., Xu, C.-Y., Gottschalk, L., Tallaksen, L. M., 2011. Systematic evaluation of autoregressive error models as post-processors for a probabilistic streamflow forecast system. Journal of Hydrology 407 (1-4), 58–72.

Peel, M. C., Blöschl, G., 2011. Hydrological modelling in a changing world. Progress in Physical Geography 35 (2), 249–261.

Refsgaard, J., 1997. Validation and intercomparison of different updating procedures for real-time forecasting. Nordic Hydrology 28, 65–84.

Refsgaard, J. C., Højberg, A. L., van der Sluijs, J. P., Vanrolleghem, P. A., Vanrolleghem, P. A., 2007. Uncertainty in the environmental modelling process - a framework and guidance. Environmental Modelling and Software 22 (11), 1543–1556.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks, S. W., 2010. W05521 Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors (doi 10.1029/2009WR008328). Water Resources Research 46 (5).

Romanowicz, R. J., Young, P. C., Beven, K. J., Pappenberger, F., 2008. A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. Advances in Water Resources 31 (8), 1048–1056.

Schoups, G., van de Giesen, N. C., Savenije, H. H. G., 2009. Model complexity control for hydrologic prediction (doi 10.1029/2008wr006836). Water Resources Research 45 (12).

Schoups, G., Vrugt, J. A., 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-gaussian errors. Water Resources Research 46 (10), W10531.

Schütze, M., Butler, D., Beck, M., 2002. Modelling, simulation and control of urban wastewater systems. Springer-Verlag.

Stedinger, J. R., Vogel, R. M., Lee, S. U., Batchelder, R., 2008. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resources Rsearch 44, 1–17.

Taylor, C. J., Pedregal, D. J., Young, P. C., Tych, W., 2007. Environmental time series analysis and forecasting with the captain toolbox. Environmental Modelling and Software 22 (6), 797–814.

Thordarson, F. Ö., 2012. Grey box modelling of hydrological systems with focus on uncertainties. Ph.D. thesis, DTU Informatics, Technical University of Denmark.

Thorndahl, S., Beven, K., Jensen, J., Schaarup-Jensen, K., 2008. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. Journal of Hydrology 357 (3-4), 421–437.

Thyer, M. A., Renard, B., Kavetski, D., Kuczera, G., Srikanthan, S., 2007. Bayesian total error analysis for hydrological models: Preliminary evaluation using multi-site catchment rainfall data. MODSIM 2007: INTERNATIONAL CONGRESS ON MODELLING AND SIMULATION, 2459–2465.

Todini, E., 2007. Hydrological catchment modelling: past, present and future. Hydrology and Earth System Sciences 11 (1), 468.

Todini, E., 2011. History and perspectives of hydrological catchment modelling. Hydrology Research 42 (2-3), 73–85.

Tomicic, B., 2006. Kapacitetsplanlægning. Opdatering af MOUSE/MIKE URBAN Model. Modeltilpasningsrapport for spildevandscenter Avedøre. Tech. rep., Danish Hydraulic Institute (DHI), Hørsholm, Denmark.

Vestergaard, M., 1998. Nonlinear filtering in stochastic volatility models, Master Thesis. Technical University of Denmark. Department of Mathematical Modelling, Lyngby, Denmark.

Vezzaro, L., Mikkelsen, P. S., 2012. Application of global sensitivity analysis and uncertainty quantification in dynamic modelling of micropollutants in stormwater runoff. Environmental Modelling and Software 27-28, 40–51.

Vojinovic, Z., Kecman, V., Babovic, V., 2003. Hybrid approach for modeling wet weather response in wastewater systems. Journal of Water Resources Planning and Management 129 (6), 511–521.

Vrugt, J. A., 2011. DREAM_(D): an adaptive markov chain monte carlo simulation algorithm to solve discrete, noncontinuous, posterior parameter estimation problems. Hydrology and Earth System Sciences Discussions 8 (2), 4025.

Vrugt, J. A., Braak, C. J. F. t., Gupta, H. V., Robinson, B. A., 2009a. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment 23 (7), 1011–1026.

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., Robinson, B. A., 2009b. W00B09 - Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation (DOI 10.1029/2007WR006720). Water Resources Research 45 (12).

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., Higdon, D., 2009c. Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. International Journal of Nonlinear Sciences and Numerical Simulations 10 (3), 273.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., Robinson, B. A., 2009d. Response to comment by Keith Beven on "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?". Stochastic Environmental Research and Risk Assessment 23 (7), 1061–1062.

Walker, W., Harremoes, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P., von Krauss, M. K., 2005. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. Integrated Assessment 4 (1).

Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., Xu, C.-Y., 2010. Calibration of hydrological models using flow-duration curves. Hydrology and Earth System Sciences Discussions 7 (6), 9467.

Xiong, L., Wan, M., Wei, X., O'Connor, K. M., 2009. Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation. Hydrological Sciences Journal 54 (5), 852–871.

Yang, J., Reichert, P., Abbaspour, K., Xia, J., Yang, H., 2008. Comparing uncertainty analysis techniques for a swat application to the chaohe basin in china. Journal of Hydrology 358 (1-2), 1–23.

Yang, J., Reichert, P., Abbaspour, K. C., Yang, H., 2007. Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference. Journal of Hydrology 340 (3-4), 167–182.

Young, P., Garnier, H., 2006. Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems. Environmental Modelling and Software 21 (8), 1055–1072.

Young, P. C., 2011. Gauss, Kalman and advances in recursive parameter estimation. Journal of Forecasting 30 (1), 104–146.

12 Papers

Paper I. Breinholt, A., Santacoloma, P.A., Mikkelsen, P.S., Madsen, H., Grum, M., Nielsen, M.K. *Evaluation framework for control of integrated urban drainage systems*, In: 11ICUD, Proceedings of 11th International Conference on Urban Drainage, Edinburgh, Scotland, 31st August-5th September 2008.

Paper II. Breinholt, A., Grum, M., Madsen, H. , Thordarson, F.Ø., Mikkelsen, P.S. *Informal uncertainty analysis (GLUE) of continuous flow simulation in a hybrid sewer system with infiltration inflow - consistency of containment ratios in calibration and validation?*, submitted.

Paper III. Breinholt, A., Thordarson, F.Ø., Møller, J.K., Grum, M., Mikkelsen, P.S., Madsen, H. *Grey-box modelling of flow in sewer systems with state dependent diffusion*, Environmetrics, Vol. 22 (8), pp. 946-961, 2011 (DOI: 10.1002/env.1135).

Paper IV. Thordarson, F.Ø., **Breinholt, A.**, Møller, J.K., Mikkelsen, P.S. Grum, M., Madsen, H. *Evaluation of probabilistic flow predictions in sewer systems using grey box models and a skill score criterion*. Stochastic Environmental Research and Risk Assessment.(in press)(DOI: 10.1007/s00477-012-0563-3).

Paper V. Breinholt, A., Møller, J.K., Madsen, H., Mikkelsen, P.S. *A formal statistical approach to representing uncertainty in rainfall-runoff modelling with focus on residual analysis and probabilistic output evaluation - distinguishing simulation and prediction*, submitted.

Paper VI. Breinholt, A., Thordarson, F.Ø., Møller, J.K., Grum, M., Mikkelsen, P.S., Madsen, H. *Identifying the appropriate physical complexity of stochastic gray-box models used for urban drainage flow prediction by evaluating their point and probabilistic forecast skill*, submitted.

The papers are not included in this www-version, but can be obtained from the Library at DTU Environment: Department of Environmental Engineering Technical University of Denmark Miljoevej, Building 113 2800 Kongens Lyngby, Denmark (library@env.dtu.dk)

The Department of Environmental Engineering (DTU Environment) conducts science-based engineering research within four themes:
Water Resource Engineering, Urban Water Engineering,
Residual Resource Engineering and Environmental Chemistry & Microbiology.
Each theme hosts two to four research groups.

The department dates back to 1865, when Ludvig August Colding, the founder of the department, gave the first lecture on sanitary engineering as response to the cholera epidemics in Copenhagen in the late 1800s.

DTU Environment
Department of Environmental Engineering
Technical University of Denmark

Miljoevej, building 113
DK-2800 Kgs. Lyngby
Denmark

Phone: +45 4525 1600
Fax: +45 4593 2850
e-mail: reception@env.dtu.dk
www.env.dtu.dk

ISBN 978-87-92654-60-1