Probabilistic Forecasting for On-line Operation of Urban Drainage Systems

Roland Löwe

Kongens Lyngby 2014
PHD-2014-324

Technical University of Denmark
Department of Applied Mathematics and Computer Science
Building 303B, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253031
compute@compute.dtu.dk
www.compute.dtu.dk
PHD-2014-324
ISSN: 0909-3192

Summary (English)

This thesis deals with the generation of probabilistic forecasts in urban hydrology. In particular, we focus on the case of runoff forecasting for real-time control (RTC) on horizons of up to two hours.

For the generation of probabilistic on-line runoff forecasts, we apply the stochastic grey-box model approach. Building on previous work concerning the development of conceptual stochastic rainfall-runoff model structures, we

- investigate approaches for the calibration of model parameters that tune the models for multistep predictions,
- develop an approach for generating probabilistic multistep predictions of runoff volume in an on-line setting,
- develop a new approach for dynamically modelling runoff forecast uncertainty.

We investigate how rainfall inputs can be optimally combined for runoff forecasting with stochastic grey-box models and what effect different types of radar rainfall measurements and forecasts have on on-line runoff forecast quality.

Finally, we implement the stochastic grey-box model approach in a real-world real-time control (RTC) setup and study how RTC can benefit from a dynamic quantification of runoff forecast uncertainty.

Summary (Danish)

Denne afhandling omhandler metoder for beregning af probabilistiske prognoser for afløbssystemer. Vi fokuserer især på flowprognoser for horisonter op til 2 timer med henblik på styring af afløbssystemer.

Vi bruger stokastiske grey-box modeller for at generere probabilistiske online flow prognoser. Tidligere forskning har beskæftiget sig med udviklingen af stokastiske, konceptuelle model strukturer for afløbssystemer. Med udgangspunkt heri fokuserer denne afhandling på at:

- udvikle nye metoder for parameter kalibrering som fokuserer på modellens evne til beregning af prognoser af flow på langtidshorisonter,
- udvikle nye metoder for at generere probabilistiske prognoser af strømningsvolumen på langtidshorisonter til online formål,
- udvikle nye metoder for en dynamisk beskrivelse af prognoseusikkerheden.

Vi undersøger hvorledes forskellige typer regn målinger bedst kan kombineres når formålet er at lave probabilistiske afløbsprognoser med grey-box modeller, samt hvilken type regn måling og -prognose er mest velegnet som grundlag for afløbsprognoser.

Endelig implementerer vi stokastiske grey-box modeller i en online styring og tester hvordan en styring af afløbssystemer kan profitere fra en dynamisk beskrivelse af prognose usikkerheden.

Preface

This thesis was prepared at the department of Applied Mathematics and Computer Science (DTU Compute) at the Technical University of Denmark in partial fulfilment of the requirements for acquiring a Ph.D. degree.

The thesis deals with the development of methods for probabilistic forecasting of runoffs from urban drainage systems and their application for real-time control.

The thesis consists of a summary report and 7 papers, documenting the work carried out during the period between January 2011 and January 2014. Three of these papers are published or accepted in international peer-reviewed journals, two papers will be submitted shortly after the hand-in of this thesis and two papers are under preparation.

København, 28-February-2014

Roland Löwe

Acknowledgements

This thesis was prepared during an intense period of 3 years and 2 months of Ph.D. study at the Department of Applied Mathematics and Computer Science at the Technical University of Denmark. It is the outcome of an inspiring research environment, support from my supervisors and a friendly and supportive atmosphere, both at the institute and amongst the project partners.

The thesis is a part of the Storm- and Wastewater Informatics (SWI) project and was financed by the Danish Council for Strategic Research, Programme Commission on Sustainable Energy and Environment, the Technical University of Denmark and the utility companies HOFOR (former Københavns Energi), Aarhus Vand, Lynettefællesskabet and Spildevandscenter Avedøre and involved significant contributions from Krüger AS and DHI. I further acknowledge the UrbanWaterTech and ITMAN graduate schools for the possibility of interacting with other researchers at scientific seminars.

I wish to thank my supervisors Henrik Madsen (DTU Compute) and Peter Steen Mikkelsen (DTU Environment) for their support during this period, for the always open doors, for the freedom to try new ideas and for the patience with a newcomer in this field. In retrospect, I was lucky to receive important guidance methodologically and in the choice of collaborators.

Amongst our collaborators I want to thank Søren Thorndahl and Michael Rasmussen (both Department of Civil Engineering, Aalborg University) for providing radar rainfall data and extensive support for the interpretation of results and for discussions with reviewers in paper B. I am very grateful to Luca Vezzaro (DTU Environment and Krüger AS) for his endurance, the many discussions and feedbacks during our collaboration on papers E and F over a period of almost 2 years. Finally, I would like to thank Morten Grum (Krüger AS) for always pushing this work towards the problems relevant in practice and Rune Juhl (DTU Compute) for the discussions on parameter estimation and his improvements in the *CTSM* software.

A different but no less important aspect is the very friendly atmosphere at the former Time Series (now Dynamical Systems) group at DTU Compute, which made it very much enjoyable for me to go to work. I wish to name Anne Katrine Duun-Christensen and Julija Tastu with whom I was lucky to share the office for several years and with whom I had many discussions and talks.

Finally, I wish to very much thank my girlfriend Rasa Narkevičiūtė for her patience with the increasingly frequent long nights over the last year, for moral support and for proof-reading this thesis.

List of Publications

Papers Included in the Thesis

A Roland Löwe, Peter Steen Mikkelsen, Michael Rasmussen, Henrik Madsen (2013). "‘State-space adjustment of radar rainfall and skill score evaluation of stochastic volume forecasts in urban drainage systems"’, *Water Science and Technology*, 68(3): 584–590, 2013.

B Roland Löwe, Søren Thorndahl, Peter Steen Mikkelsen, Michael Rasmussen, Henrik Madsen (2014). "‘Probabilistic online runoff forecasting for urban catchments using inputs from rain gauges as well as statically and dynamically adjusted weather radar"’, accepted by *Journal of Hydrology*, 2014.

C Roland Löwe, Peter Steen Mikkelsen, Henrik Madsen (2014). "‘Stochastic rainfall-runoff forecasting: parameter estimation, multi-step prediction, and evaluation of overflow risk"’, *Stochastic Environmental Research and Risk Assessment*, 28: 505–516, 2014.

D Dario Del Giudice, Roland Löwe, Henrik Madsen, Peter Steen Mikkelsen, Jörg Rieckermann (2014). "‘Comparing two stochastic approaches to predict urban rainfall-runoff with explicit consideration of model bias"’, in preparation for *Water Resources Research*, 2014.

E Roland Löwe, Luca Vezzaro, Peter Steen Mikkelsen, Morten Grum, Henrik Madsen (2014). "‘Investigating the use of probabilistic forecasts for RTC of urban drainage systems - A Layout for Probabilistic Online Forecasting of Sewer Flows"’, in preparation for *Environmental Modelling and Software*, 2014.

F Luca Vezzaro, Roland Löwe, Henrik Madsen, Morten Grum, Peter Steen Mikkelsen (2014). "'Investigating the use of probabilistic forecasts for RTC of urban drainage systems – Full scale testing in the city of Copenhagen, Denmark"', in preparation for *Environmental Modelling and Software*, 2014.

G Roland Löwe, Rune Juhl, Peter Steen Mikkelsen, Henrik Madsen (2014). "'Forecasting Operational Runoff Forecast Uncertainties - State, Rainfall and Error Dependencies"', in preparation for *Hydrology and Earth System Science Discussions.*

Other Publications not Included

The following publications have also been prepared during the course of the Ph.D. study. They are omitted in this thesis because they either resulted from work not directly related to the thesis or overlap with contents in the above publications.

H Katja Seggelke, Roland Löwe, Thomas Beeneken, Lothar Fuchs (2013). "'Implementation of an integrated real-time control system of sewer system and waste water treatment plant in the city of Wilhelmshaven"', *Urban Water Journal*, 10, 330–341, 2013.

I Roland Löwe, Peter Steen Mikkelsen, Henrik Madsen (2012). "'Forecast generation for real-time control of urban drainage systems using greybox modelling and radar rainfall"', in *Proceedings of the 10th International Conference on Hydroinformatics*, Hamburg, Germany, 2012.

J Roland Löwe, Peter Steen Mikkelsen, Michael Rasmussen, Henrik Madsen (2012). "'State-space adjustment of radar rainfall and stochastic flow forecasting for use in real-time control of urban drainage systems"', in *Proceedings of the 9th International Conference on Urban Drainage Modeling*, Belgrade, Serbia, 2012.

K Luca Vezzaro, Roland Löwe, Henrik Madsen, Morten Grum, Peter Steen Mikkelsen (2013). "'Investigating the use of stochastic forecast for RTC of urban drainage systems"', in *Proceedings of NOVATECH 2013*, Lyon, France, 2013.

L Roland Löwe, Henrik Madsen, Peter Steen Mikkelsen (2013). "'Stokastiske prognoser for afløb og real tids styring"', *EVA : Erfaringsudveksling i vandmiljøteknikken*, 26(2),11–18, ISSN 1901-3663, 2013.

Glossary

Abbreviation	Explanation
ACF	autocorrelation function
ARIL	average interval length
ANN	artificial neural network
CCF	cross correlation function
CRPS	continuous ranked probability score
CSO	combined sewer overflow
CTSM	Continuous Time Stochastic Modeling software
DORA	dynamic overflow risk assessment
EKF	extended Kalman filter
iid	independent and identically distributed
LAWR	local area weather radar
ML	maximum likelihood
MAP	maximum a posteriori
NSE	Nash-Sutcliffe efficiency index
ODE	ordinary differential equation
QPE	quantitative precipitation estimate
QPF	quantitative precipitation forecast
PI	Persistence Index
RTC	real-time control
SC	Interval Score
SDE	stochastic differential equation
SS	skill score (generic)
SWI	Storm and Wastewater Informatics project
WWTP	wastewater treatment plant

Contents

Part I

Summary Report

Chapter 1

Introduction

1.1 Models in Urban Hydrology and their Objectives

In general, models are created with different complexities and based on different principles. In hydrology, it is reasonable to classify models according to their complexity. For example, this approach can be found in [Ref96] for deterministic hydrologic models and, more generally, in [Lju10]. [Bre12] classifies urban hydrological models according to this approach (figure 1.1).

The exact classification varies among authors. Yet, the general line of thought is that white-box models are based on first order principles which describe the actual physical process, while black-box models are purely data-based and typically do not allow for a physical interpretation of model structure and parameters. Grey-box models cover the area between the two extremes.

Other than depicted in Figure 1.1, models applied in urban hydrology are generally not white. Even rather detailed model structures as MOUSE ([DHI03]) or WEST ([VMA+03]) build on conceptual assumptions for example in the runoff formation and transformation processes or in the description of treatment processes.

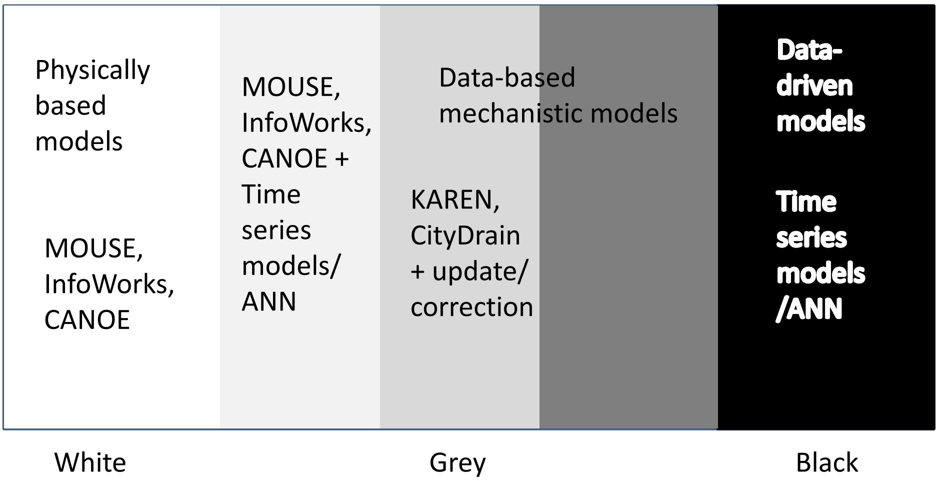


Figure 1.1: Classification of urban hydrological models according to their structural complexity (from [Bre12])

The task of the hydrologist is thus, to identify a model that makes optimal use of the available data and is suitable for a given purpose rather than aiming at a perfect physical process description.

Modelling tasks in urban hydrology can be separated into off-line and on-line applications. Off-line applications typically focus on design or analysis of the system and comprise decision horizons of several years. They very often include 'What if?' tasks, leading to an analysis of effects resulting from an anticipated change in the system. In this field, complex model structures have their main range of application as they permit for the use of a maximum of physical (static) information such as pipe diameters, the location of sealed area in a catchment or the dimensions of treatment facilities.

On-line applications, on the other hand, are related to supervision of the system and decision making over short horizons in the range from minutes to days. Computational speed of the models is often relevant in this area, limiting the use of complex model structures. In turn, some current measurement information about the system is typically available and should be used in the model. In this case, the use of statistical techniques such as auto-calibration or state updating is typically simplified by simple model structures. Examples of on-line modelling tasks in urban hydrology are

- forecasting runoff or treatment capacity for real-time control ([VG14])
- software sensors deriving information about the system from an indirect measurement ([DDR13])
- automated cleaning and correction of measurements ([ATCV13])

Most of these tasks profit from model structures which are simple, and thus numerically and statistically easy to handle, but have a physical interpretation. Such a structure allows for the automatic estimation of parameters from measurements as well as for algorithms to update model states and parameters. At the same time, physical information can be included using for example prior parameter estimates. We denote such models as grey-box models.

The thesis is centred around the application of a special class of grey-box models, so called stochastic grey-box models (see chapter 3), for runoff forecasting in urban hydrology. Generally, every model that is fitted to data may be considered as a stochastic model, but the different approaches distinguish themselves in whether and how they explicitly account for the time-dynamic and autocorrelated model error structures. One of the major advantages of the stochastic grey-box modelling approach is that it allows us to describe these error structures time-dynamically in a way which is suitable for on-line purposes.

1.2 Forecasting and Real-time control of Urban Drainage Systems

An important area of application for on-line models is real-time control of urban drainage systems. Following the definition of [SCC$^+$04], we denote as real-time control the operation in real time of actuators in the system based on the monitoring of process variables.

The most common objective for real-time control is to use existing drainage infrastructure efficiently. This permits the operators to reduce environmental (and economic) impacts (such as combined sewer overflows, sludge escape from the wastewater treatment plant, energy consumption or use of treatment chemicals) with moderate investments into the infrastructure.

While existing implementations have focused largely on the reduction of combined sewer overflows and thus the reduction of investments into storage basins (for example [FPM11, FB05, SLBF13, SGT$^+$13]), foreseen challenges such as the inclusion of water systems into smart cities ([RTSS13, MB10]) or the efficient utilisation of storage space in sustainable urban drainage are likely to contribute to a development where real-time operation of drainage systems becomes a standard rather than an exception.

Implementations of real-time control systems on the global or system ([SCC+04]) level can coarsely be distinguished into strategies that were developed off-line and strategies where an optimal control decision is determined on-line ([SBB02]). Off-line strategies can be seen as the traditional approach to real-time control and have, as a result of their robustness and usually easy interpretation, been implemented in a multitude of urban water systems. Typically such strategies are implemented in the form of 'IF-THEN' rules, possibly supplemented by fuzzy systems. Examples of such implementations can be found in [FB05] and [SLBF13]. With suitable decision rules, such systems can also make use of for example precipitation forecasts ([MMI05]).

An on-line optimisation strategy dynamically determines the actuator settings on-line based on (model) forecasts of future system states (such as basin fillings) and a model-based evaluation of the effects of different actuator set points. Forecasts are required in such a system to gain information about parameters such as expected future runoff, treatment capacity or energy consumption.

This type of optimisation approach is often considered problematic because the underlying forecast models need to be simple enough to run in an optimization routine but at the same time provide realistic forecasts of the system states ([BS05]). Yet, these systems are appealing due to the objectivity of the derived control decisions, the use of forecast information and thus a potentially higher efficiency as well as the rather simple adaptability of the system to new requirements. Recently developed schemes have successfully combined on-line optimization and forecast models for the operation of drainage systems ([PCL+05, PCR+09]) and partly combined these schemes with decision rules to determine boundary conditions ([GTC+11]).

New developments also account for the uncertainty of forecasts during decision making in the control setup ([VG14]). This is an important step due to high uncertainty related to rainfall measurements and forecasts, and the uncertainty inherent in the model structures. However, a missing link is the development of models that can provide probabilistic forecasts (which quantify forecast uncertainty) in an on-line setting over a multitude of horizons from 10 minutes to 12 hours. Stochastic grey-box models may provide such a means as suggested by [BMPN00, BTM+11, CNH96, TBM+12].

1.3 Objective of the PhD project

The aim of this PhD project is to develop probabilistic forecast models that can be applied on-line in urban hydrology. Requirements for such models are:

- Forecast models need to be fast in order to be applicable in on-line routines.
- Forecast models need to provide accurate forecasts that are reliable and sharp (see Section 2.2)
- Forecast models need to provide probabilistic forecasts over a multitude of forecast horizons and account for correlation between the different horizons.

This thesis focuses on the generation of probabilistic runoff forecasts for horizons up to 120 minutes using stochastic grey-box models. However, the developed methods can and will also be applied to other problems such as forecasting pollutant loads and the capacity of the wastewater treatment plant. The thesis builds on and extends previous work in particular by [BTM+11] and [TBM+12]. It particularly focuses on practical applicability and aims to answer the following questions:

1. What type of rainfall inputs should be used for short-term runoff forecasting and, in particular, do we benefit from using quantitative precipitation estimates (QPE) from weather radar?
2. Do quantitative precipitation forecasts (QPF) provide benefits for short term runoff forecasts?
3. Do short-term runoff forecasts benefit from a combined rainfall input, making use of both rain gauge and radar rainfall measurements?
4. How can forecast models and parameters be identified in the context of noisy data and providing forecasts over a multitude of horizons?
5. How can dynamically changing forecast uncertainties be correctly captured in a probabilistic model structure?
6. How can probabilistic forecasts be generated for decision making in real-time control?
7. What effect does the consideration of forecast uncertainty have on the efficiency of real-time control schemes?

1.4 Thesis Outline

This thesis is structured as follows. Part I is a report which provides the background and introduces and summarizes the papers. Within this part, Chapter 2 gives an overview of the general requirements for probabilistic forecast and introduces ways to evaluate probabilistic forecast quality. The methods introduced in Chapter 2 are used in the later chapters and the articles.

Chapter 3 introduces the grey-box modelling approach. This approach is the basis for the generation of probabilistic runoff forecasts throughout the thesis. The chapter explains how to apply grey-box models for rainfall runoff modelling, how to estimate model parameters, how to generate probabilistic forecasts etc. and discuss how the approach relates to other uncertainty techniques applied in hydrology.

Chapter 4 discusses how and what rainfall data should be used for on-line runoff forecasting over short horizons. We focus in particular on the questions whether rain gauge or radar data should be applied and how the two data sources can be combined.

In Chapter 5 we discuss the practical implementation of probabilistic forecasts from stochastic grey-box models in real-time control and what implications forecast uncertainty has on decision making.

Finally, in Chapters 6 and 7 we conclude the thesis with reference to the objectives defined in section 1.3 and provide an outlook for future work.

Part II is a collection of publications and it contains the following papers.

Paper A is a journal article published in *Water Science and Technology.* It deals with the combination of radar and rain gauge measurements using a state-space modelling approach and the application of different rainfall data for probabilistic runoff forecasting.

Paper B is a journal article accepted by *Journal of Hydrology.* It assesses the impact on on-line runoff forecasting skill from time-constant and time-varying radar adjustment and investigates whether an improved spatial model resolution is desirable for on-line runoff forecasting.

Paper C is a journal article published in *Stochastic Environmental Research and Risk Assessment.* It develops a method for estimation of probabilistic on-line runoff forecasting models based on the skill of multi-step-ahead forecasts and assesses the effect of forecast uncertainty on the expected risk of combined sewer overflow (CSO).

Paper D is a journal article in preparation for *Water Resources Research*. It compares the stochastic grey-box modeling approach with a Bayesian bias approach for uncertainty modelling in a common urban case study.

Paper E is a journal article in preparation for *Environmental Modelling and Software*. It describes the implementation of stochastic grey-box models for probabilistic runoff forecasting into an existing real-time control scheme and evaluates the forecast quality that can be obtained in six different subcatchments.

Paper F is a journal article in preparation for *Environmental Modelling and Software*. It evaluates the effect of runoff forecast uncertainty on decision making in real-time control.

Paper G is a manuscript in preparation for *Hydrology and Earth System Science Discussions*. It describes different model structures for describing forecast uncertainty in stochastic grey-box models and their effect on the calibration of multi-step probabilistic forecasts.

Chapter 2

Verification of Probabilistic Forecasts

As explained in Section 1.3 the thesis focuses on the generation of probabilistic forecasts for short horizons (<2 hours) in urban hydrology. The most common approach to forecasting in hydrology is still the generation of deterministic or point forecasts. In this case, the forecast quality can be measured in terms of the error of the forecasts as compared to observations. Different measures such as the root mean square error, the delay to peak or the volume error can be applied here.

For probabilistic forecasts we obtain not a single forecast value but probabilities that the forecasted quantity takes a given value. This can for example be expressed parametrically by generating a probability distribution of the forecasted quantity or non-parametrically by generating various realisations or ensembles of the forecasted quantity. In either case, we evaluate probabilistic forecast quality in terms of

- the location of the forecasted distribution with respect to the observations, quantified for example by the mean absolute error of the forecast median,
- the spread of the forecasted distribution, quantified for example by the width of prediction intervals,

- and, in the case of multivariate forecasts over multiple forecast horizons, the correlation between different horizons.

The purpose of this section is to introduce the measures applied for forecast verification in this work. For a good overview on measures that can be used for evaluating point forecast quality, we refer the reader to [BCG+13]. Scoring rules for probabilistic forecast evaluation are summarized by [GR07].

In this work we largely omit the verification of the correlation structure of forecasts generated for different forecast horizons as the forecasts are always generated as multi-step forecasts from a single model. A correlation between different horizons is thus implicit in the forecasting method.

2.1 Scoring Rules for Point Forecast Quality

Following a definition by [Wil11], a generic forecast skill SS can be defined as

$$SS = \frac{A - A_{ref}}{A_{perf} - A_{ref}}. \tag{2.1}$$

Here A is a measure of forecast accuracy, A_{ref} the accuracy obtained for a reference (or benchmark) forecast and A_{perf} the accuracy obtained for a perfect forecast. Positive values of SS indicate that the evaluated forecast performs better than the reference. The score values can range between negative infinity (if our forecast performs much worse than the reference) and 1 (if our forecast performs equivalently to the perfect forecast).

In hydrology, the most commonly applied forecast skill is the **Nash-Sutcliffe efficiency** NSE ([NS70]). For a point forecast $\hat{y}_i$ which is compared to an observation y_i we obtain

$$NSE = \frac{\frac{1}{n}\sum(y_i - \hat{y}_i)^2 - \frac{1}{n}\sum(y_i - \overline{y})^2}{-\frac{1}{n}\sum(y_i - \overline{y})^2}. \tag{2.2}$$

The measure of forecast accuracy A in this case is the mean squared error. The mean squared error obtained for the perfect forecast is $A_{perf} = 0$. As reference forecast, the average $\overline{y}$ over all observations y_i is applied. This, however, is a very weak reference which will lead to NSE values indicating positive forecast skill even for very badly performing models as is also argued by [SKZ+12].

We can define a more critical forecast skill by replacing the mean of the observations in equation 2.2 by the last observed value. This leads to the **persistence**

index PI ([BCG$^+$13])

$$PI = \frac{\frac{1}{n}\sum(y_i - \hat{y}_i)^2 - \frac{1}{n}\sum(y_i - y_{i-1})^2}{-\frac{1}{n}\sum(y_i - y_{i-1})^2}. \tag{2.3}$$

The application of the persistence index is difficult in situations where the observations are very noisy. This problem can be solved by replacing the last known observation y_{i-1} in equation 2.3 by the result of an exponential smoothing y_{i-1}^S of the observations up to the last known observation:

$$y_{i-1}^S = \lambda \cdot y_{i-2}^S + (1 - \lambda) \cdot y_{i-1} \tag{2.4}$$

The parameter λ must be tuned in a calibration period. In this tuning, the squared forecast error of the reference forecast from exponential smoothing is minimized. We follow this approach in paper E and denote it as **smoothed persistence index**.

When dealing with probabilistic forecasts, we often evaluate a point forecast derived from the median of the probabilistic forecasts. We choose the median to reduce the influence from long tails of the distribution on the forecast error.

2.2 Elements of Probabilistic Forecast Quality

Generally, the aim of probabilistic forecasting is to maximize the sharpness of predictive distributions subject to calibration ([GR07, Pin07]). "'Calibration refers to the statistical consistency between the distributional forecasts and the observations, and is a joint property of the forecasts and the events or values that materialize. Sharpness refers to the concentration of the predictive distributions and is a property of the forecasts only"' ([GR07]).

The above definition implies that probabilistic forecasts should be evaluated by checking if they are probabilistically calibrated (see for example [Pin13]). This requires assessing over the whole distribution, whether the observed probabilities match the predicted (or nominal) ones. Subsequently the sharpness of the distributions should be assessed.

In the following section, we introduce the measures we have applied for the evaluation of calibration and sharpness of probabilistic forecasts in this thesis. In particular, the evaluation of probabilistic calibration over a whole distribution can be hard to communicate to practitioners. We have therefore in many cases used a simplification. We extracted the median from the probabilistic forecast and evaluated the corresponding point forecast error. This will give an indication on how good the forecast model captures the physical behaviour of the considered system. To evaluate the calibration and sharpness of the distribution, we then focus on the reliability and width of for example 90% prediction intervals. In addition, a probabilistic score can be applied to evaluate the overall fit of the predicted distribution.

2.2.1 Calibration

A probabilistic forecast can be considered calibrated or reliable in a probabilistic sense, if the forecasted probability distribution matches the distribution of the observations of the considered variable. A simple way to evaluate reliability of a forecast is to analyse if the predicted (nominal) probabilities match the observed ones, i.e. whether, for example, for a given confidence level α between 0 and 1, a prediction interval with coverage $(1-\alpha)\cdot 100$ % indeed includes $(1-\alpha)\cdot 100$ % of the observations.

We express **reliability** Rel for a confidence level α as the portion of observations included in a $(1-\alpha)\cdot 100\%$ prediction interval. Initially, in paper C the inverse definition, describing reliability as the portion of observations not included in a $(1-\alpha)\cdot 100\%$ prediction interval was used. We switched to the former approach later as it is more intuitive.

Following [TBM+12], a **reliability bias** for confidence level α can be defined as

$$RB = Rel - (1-\alpha) \tag{2.5}$$

and becomes negative if the forecasted distribution is unreliable (a $(1-\alpha)\cdot 100$ % prediction interval includes less than $(1-\alpha)\cdot 100$ % of the observations) and positive if it is overreliable (a $(1-\alpha)\cdot 100$ % prediction interval includes more than $(1-\alpha)\cdot 100$ % of the observations).

Reliability Rel can be analysed for different confidence levels and plotted in a reliability diagram ([MW77]). A perfectly reliable forecast will result in a straight line in this diagram. This is exemplified in Figure 2.1. If the reliability bias RB is plotted instead, a perfectly reliable forecast will yield a constant line at 0. The forecast evaluated in Figure 2.1 is unreliable for predicted coverages between 10 and 80% and overreliable for predicted coverages above 80%.

The reliability bias has an upper and a lower bound depending on the considered level $(1-\alpha)$. A more general approach for the evaluation of predictive distributions is the **predictive QQ-plot** which is well described in [RKK+10]. From

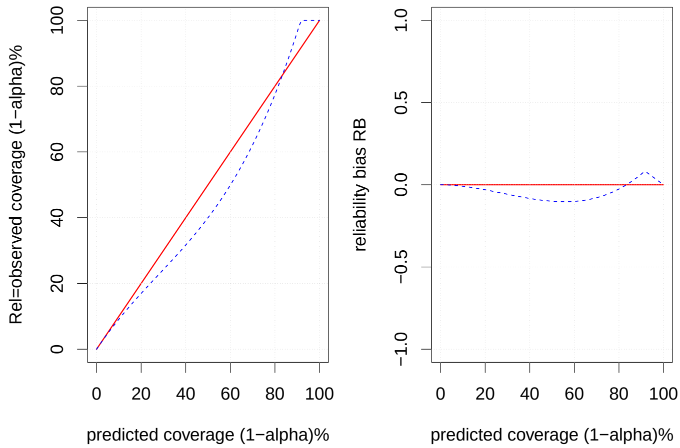


Figure 2.1: Observed reliability Rel (left) and corresponding reliability bias (RB) for an example forecast (blue) and a perfect forecast (red)

the probabilistic predictions we can for each observation y_t derive the value of the cumulative predicted distribution $F(y_t) = p(Y_t \leq y_t)$. If the observations y_t are consistent with the predicted distributions, these p values follow a standard uniform distribution, in the interval $[0, 1]$. This can be illustrated graphically in a quantile-quantile plot (see Figure 2.2). A calibrated probabilistic forecast will lead to a straight line in this plot.

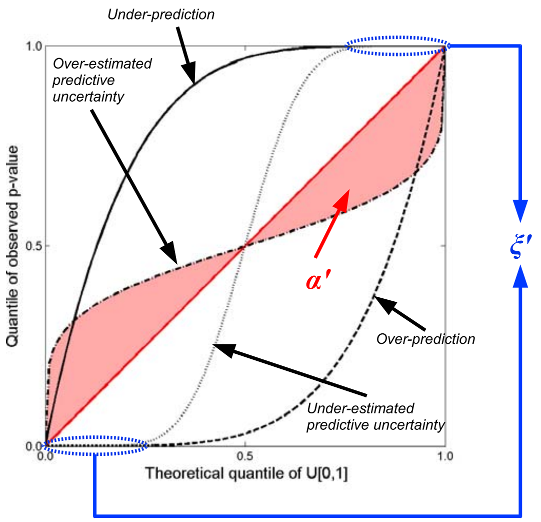


Figure 2.2: Predictive QQ-Plot comparing quantiles from 'p values' of observations and a standard uniform distribution U[0,1] (from [RKK+10])

2.2.2 Sharpness

Assuming that a probabilistic forecast is calibrated, a good forecast will require only a very narrow spread of the predictive distribution to capture the observations, while a bad forecast will yield a wide spread of the predictive distribution and thus only little information content to the decision maker.

In this work we assess the **sharpness** (Sh) of probabilistic forecasts by measuring the average width of a $(1-\alpha)\cdot 100$ % prediction interval. Defining $\hat{u}_t$, $\hat{l}_t$ as the upper and lower prediction bounds generated for such an interval at time step t and considering observations for N time steps, sharpness can be expressed as:

$$Sh = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{u}_t - \hat{l}_t\right). \tag{2.6}$$

Lower sharpness values indicate a higher information content in the probabilistic forecast. The sharpness measure Sh, however, depends on the absolute value of the forecasted quantity. [JXZS10] therefore introduced the **average interval length** $ARIL$ as a normalized measure by dividing the sharpness value of the forecast generated at time step t by the corresponding observation y_t:

$$ARIL = \frac{1}{N}\sum_{t=1}^{N}\frac{\hat{u}_t - \hat{l}_t}{y_t}. \tag{2.7}$$

2.2.3 Scoring Rules for Probabilistic Forecast Quality

In practice, probabilistic forecasts will rarely be perfectly calibrated. A comparison of different forecasts must thus take both, the calibration and the sharpness of the forecasts into account.

There are different approaches to combine these two qualities into a single value. We consider two of these approaches in this work: the interval score SC focuses on the evaluation of prediction intervals and the continuous ranked probability score $CRPS$ can be considered as a mean squared error measure of the whole predictive distribution.

For a single time step, the **interval score** of a $(1-\alpha)\cdot 100$ % prediction interval with upper bound u and lower bound l and the corresponding observation y is found as

$$SC^{\alpha} = u - l + \frac{2}{\alpha}(l-y)\cdot H(l-y) + \frac{2}{\alpha}(y-u)\cdot H(y-u) \tag{2.8}$$

where H denotes the Heaviside function and takes the value of one if its argument is greater than zero and zero otherwise. The score for a whole time series is the average of the score values derived for the single time steps. This score evaluates the width of the prediction interval and adds a penalty for observations not included in the prediction interval. Smaller score values correspond to better forecasts.

For a single time step the **continuous ranked probability score** for a probabilistic forecast s with cumulative distribution $F(s)$ and observation y is defined as

$$CRPS = -\int_{-\infty}^{\infty}(F(s) - H(s \geq y))^2 ds. \tag{2.9}$$

Again, H denotes the Heaviside function and the score for a whole time series is the average over the different time steps. This score can be depicted graphically as shown in Figure 2.3. The observation is represented as a stepwise cumulative distribution function (red). The $CRPS$ measures the area between the observation and the predicted cumulative distribution function (black). Better forecasts yield lower differences between the two distributions and thus lower score values.

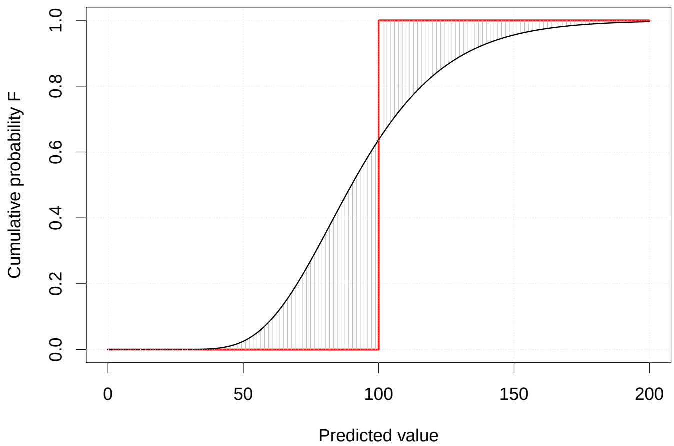


Figure 2.3: Graphical interpretation of the CRPS - the score corresponds to the marked area between the observation (red) and the predictive distribution (black), both expressed as cumulative probability functions F

Analytical expressions for the $CRPS$ are available if the distribution of the probabilistic forecast is known. However, in most cases we have used a scenario (or ensemble) approach for the generation of probabilistic forecasts. [Brö12] proposed an approach for the evaluation of the CRPS using ensembles which is based on approximating the cumulative distribution function of the forecast with a piecewise constant function.

Considering $k = 1, ..., K$ ensemble members with values e_k, we obtain

$$F(s) = \sum_{k=1}^{K} \omega_k H(s - e_k). \tag{2.10}$$

H denotes the Heaviside function. The weights ω_k are assumed > 0 for all k and $\sum_k \omega_k = 1$ and correspond to the probability that the forecasted value equals any ensemble member e_k. Thus $\omega_k = \frac{1}{K}$. This expression of $F(s)$ can be inserted in equation 2.9 to evaluate the $CRPS$ for an observation y.

Probabilistic scores can be normalized according to equation 2.1. The measure of accuracy A is then the CRPS, for example. In this case the score value A_{perf} for the perfect forecast is 0. A reference can be derived from the so-called "‘climatological"’ forecast. We use all the observations in the considered dataset to derive an empirical distribution. This same "‘climatological"’ forecast distribution is then applied at every time step and the $CRPS$ is derived for each observation. Averaging over all time steps, we obtain the reference score A_{ref}.

Chapter 3

Grey-box Modelling of Sewer Flows

3.1 Introduction to Stochastic Grey-box Models

The thesis focuses on the generation of probabilistic runoff forecasts using stochastic grey-box models (c.f. Section 1.3). As described in Section 1.1 we consider a grey-box model to be a (strongly) simplified representation of reality which can, nevertheless, be interpreted physically. In this work a stochastic grey-box model is defined as the implementation of such models in a state-space framework using stochastic differential equations (SDE). This terminology was largely developed at the Department for Applied Mathematics and Computer Science at the Technical University of Denmark and can be found in numerous previous works (for example [BNMP99, BTM$^+$11, KMJ04, Møl10, NM06]).

Generally, such a model structure consists of a system of time-continuous state (equation 3.1) and time-discrete observation equations (equation 3.2). The former describe the evolution of the considered system and often may be considered the "‘actual model structure"’. The latter relate the system states (for example filling of virtual basins in a reservoir cascade) to actual observations (for example flow measurements).

$$dX_t = \underbrace{f(X_t, u_t, t, \theta)dt}_{Drift\ term} + \underbrace{\sigma(X_t, u_t, t, \theta)d\omega_t}_{Diffusion\ term} \tag{3.1}$$

$$Y_k = h(X_k, u_k, t_k, \theta) + e_k \tag{3.2}$$

In equation 3.1 X corresponds to a vector of (unobserved) system states, u to a vector of external forcings (model inputs), t to the considered time point and θ to a vector of model parameters. Other than an ordinary differential equation (ODE), the system equations are divided into a (physical) drift term (expressed as function f) and a (stochastic) diffusion term (expressed as function σ). The latter can be used to model uncertainties resulting from insufficient model structures and uncertain inputs. The diffusion term is driven by a Wiener process with increments $d\omega_t$. These are normally distributed with mean 0 and variance dt (c.f. [Mad08] and paper D). It is important to note that drift and diffusion term represent a coupled system and that generally the expected value for the states X_t is not equal to the solution of the ODE represented by the drift term.

In equation 3.2 the system states for discrete time steps k are related to a vector of observations Y_k by the function h. The observations are assumed to be subject to a normally distributed error e_k with mean 0.

An extended Kalman filter (EKF) is applied to adjust the system states based on current observations Y_k of the system ([KMJ04]). The models applied in this thesis are implemented in the open source software *CTSM* ([JKB+13]) and the routines implemented in this software are applied mainly for filtering and (with exceptions) for parameter estimation.

3.2 Rainfall Runoff Modeling with Stochastic Grey-box Models

3.2.1 The Linear Reservoir Cascade

In this work, we exclusively apply a linear reservoir cascade considering either two or three reservoirs for runoff forecasting. This approach is conceptually simple, in fact, in most cases structural model deficiencies can be observed. However, it also allows for the identification of problems arising in parameter estimation and for modeling forecast uncertainties in the non-ideal case where

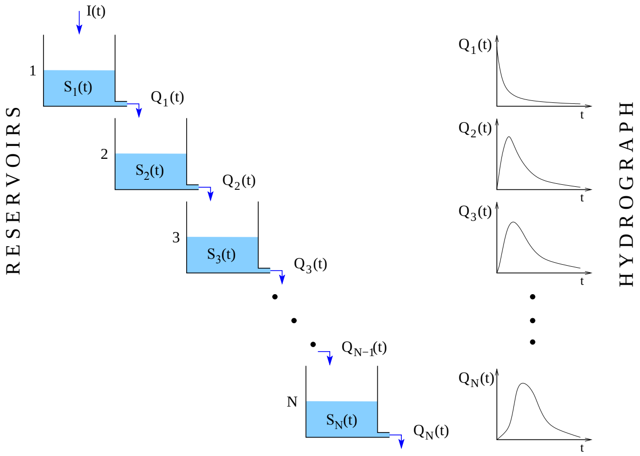


Figure 3.1: A system of N linear reservoirs. Corresponding hydrographs for the output $Q_n(t)$ are shown to the right (from [Tho11]).

the model does not perfectly describe the data. We consider this case relevant for practical applications.

A sketch of such a model layout is shown in Figure 3.1. We refer to [CMM88] for a detailed explanation of the reservoir cascade. The principal assumption is that runoff hydrographs can be described by routing the rainfall input through a series of virtual reservoirs where the outflow $Q_{i,t}$ at time t from a reservoir depends on the storage level $S_{i,t}$ in the reservoir and on a time constant K.

$$Q_{i,t} = \frac{1}{K} \cdot S_{i,t} \tag{3.3}$$

Rainfall input is considered only to the first reservoir and, considering an effective catchment area A and rainfall input P, we obtain a system of differential equations describing the runoff routing through the reservoir cascade:

$$d \begin{bmatrix} S_{1,t} \\ S_{2,t} \\ \vdots \\ S_{N,t} \end{bmatrix} = \begin{bmatrix} A \cdot P_t - \frac{1}{K} S_{1,t} \\ \frac{1}{K} S_{1,t} - \frac{1}{K} S_{2,t} \\ \vdots \\ \frac{1}{K} S_{N-1,t} - \frac{1}{K} S_{N,t} \end{bmatrix} dt \tag{3.4}$$

where the outflow hydrograph is found as

$$Q_t = \frac{1}{K} \cdot S_{N,t} \tag{3.5}$$

This model structure can be considered a grey-box model as it does not rely on first order principles but the model parameters effective catchment area A and time constant K can be interpreted physically.

3.2.2 The Linear Reservoir Cascade as Stochastic Grey-box Model

The model structure described above can easily be cast into a state-space layout and thus be applied as a stochastic grey-box model. This approach was adapted to modeling urban drainage systems by [BTM+11]. Considering $N = 2$ virtual reservoirs, we obtain the state equations

$$d\begin{bmatrix} S_{1,t} \\ S_{2,t} \end{bmatrix} = \begin{bmatrix} A \cdot P_t + a_0 - \frac{1}{K} S_{1,t} \\ \frac{1}{K} S_{1,t} - \frac{1}{K} S_{2,t} \end{bmatrix} dt + \begin{bmatrix} g_1(S,u,t,\theta) & 0 \\ 0 & g_2(S,u,t,\theta) \end{bmatrix} d\omega_t. \quad (3.6)$$

Here, the dry weather flow a_0 is considered as input to the first model states to avoid state values approaching 0 which can be problematic in the case of state dependent noise descriptions. g_1 and g_2 symbolize generic functions for scaling the variance of the diffusion.

The observation equation for time step k in hours h is found as

$$Y_k = Q_k = \frac{1}{K} S_{2,k} + D_k + e_k \quad (3.7)$$

where the dry weather variation D_k is described by a set of trigonometric functions

$$D_k = \sum_i^2 (s_i sin \frac{i2\pi k}{24h} + c_i cos \frac{i2\pi k}{24h}) \quad (3.8)$$

A transformation may be applied to the observations which will be discussed in Section 3.5. Depending on the catchment characteristics, more detailed model structures may be applied as discussed by [Bre12].

3.3 Model Structures for Uncertainty in Stochastic Grey-box models

Constant Diffusion
The diffusion term in equation 3.6 can take different structures. In the simpliest case, a constant diffusion term is applied:

$$g_i = \sigma_i \quad (3.9)$$

where the parameter σ_i for the i-th state is determined as part of the parameter estimation routine.

State Dependent Diffusion
[Bre12] and [TBM+12] have focused on diffusion structures that depend on the current state of the model. Suggested candidates were

1. A direct linear dependency on the state
$$g_i = \sigma_i \cdot S_i \tag{3.10}$$
2. An exponentially scaled dependency on the state with scaling parameter γ
$$g_i = \sigma_i \cdot S_i^\gamma \tag{3.11}$$
3. A logistic dependency on the state that reduces the variance if the state approaches an anticipated maximal level S_{max}. This approach corresponds to the case of sewer flows reaching the maximal pipe capacity.
$$g_i = \sigma_i \cdot S_i \cdot (S_{i,max} - S_i) \tag{3.12}$$

The reasoning behind these approaches is that the uncertainty of flow predictions is high during rain events and low in dry weather situations. Even the simple linear state dependency has proven to be a strong benchmark and hard to outperform by other diffusion structures. In addition, these approaches have the desirable property of scaling the diffusion to 0 if the states approach very small values so that negative forecasts are avoided. The state dependent diffusion term was therefore applied in several works (papers B,C and E).

Nevertheless, the state-dependent approaches are problematic for two reasons:

- Forecast uncertainty for dry weather periods and rain events is not addressed separately and lumped into the same parameters. This can lead to an underestimation of forecast uncertainties during rain periods if extended dry weather periods are included in the parameter estimation (papers B and E).
- Forecast uncertainty depends on the forecasted states themselves. This implies that, for example, an underestimation of observed flows by the forecast (in particular at the beginning of rain events) will also lead to an underestimation of forecast uncertainty (and vice versa). In addition, forecast uncertainty at the end of a rainfall event is typically small, although flows in the sewer system are still high. The state dependent diffusion structure will not be able to capture this behaviour.

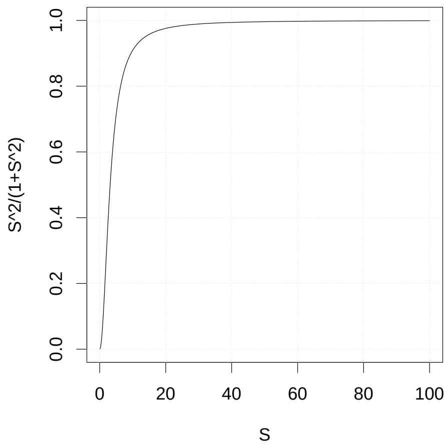


Figure 3.2: Fraction $\frac{S^2}{10+S^2}$ for different state values S. Realistic state values in the considered catchments are $> 300m^3$.

Input Dependent Diffusion
To overcome the above limitations, we propose a diffusion structure which describes forecast uncertainty as the sum of a constant dry weather uncertainty $\sigma_{i,1}$ and an external forcing F scaled by a constant parameter $\sigma_{i,2}$. It is desirable to avoid negative state forecasts to ensure stability of the model. Therefore, we introduce a rational state dependency as shown in equation 3.13.

$$g_i = (\sigma_{i,1} + \sigma_{i,2} \cdot F) \frac{S_i^2}{c + S_i^2} \tag{3.13}$$

The constant c can be selected by the modeller. The fraction in equation 3.13 should ideally be one for any realistically anticipated state values and approach 0 as S_i approaches 0 (Figure 3.2). We here use a value of $c = 1$.

The diffusion structure proposed in equation 3.13 resembles the formulation of model bias used to describe simulation uncertainty in [DHS$^+$13]. The time lagged rainfall input P_{k-l} is applied as external forcing F by the authors. The rainfall input, however, exhibits strong relative variations at irregular patterns. The diffusion scaling then becomes similarly irregular. Consequently, we apply a smoothed and time lagged rainfall input instead.

$$F_k = \lambda \cdot F_{k-1} + (1 - \lambda) P_{k-l} \tag{3.14}$$

The smoothing parameter $\lambda \in [0, 1]$ and the time lag $l \in [0, 1, \ldots, \infty)$ are identified as part of the parameter estimation procedure. Figure 3.3 depicts the effect of smoothening rainfall observations using different smoothing parameters.

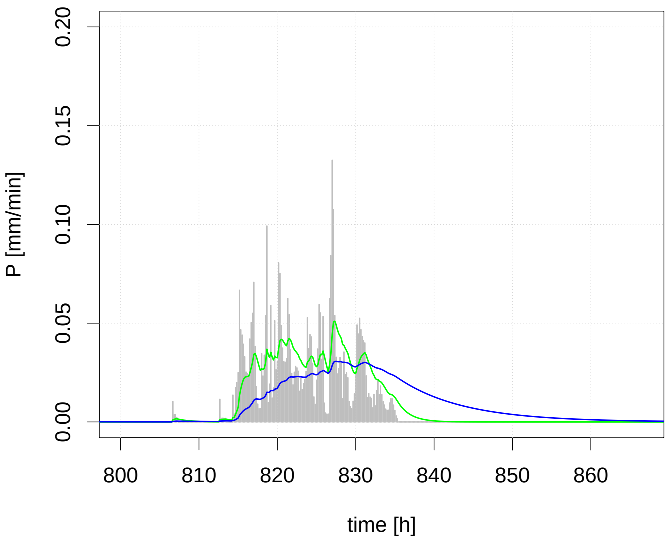


Figure 3.3: Rainfall observations in 10min resolution (grey) together with exponentially smoothened rainfall according to equation 3.14 with $l = 0$ and $\lambda = 0.9$ (green) and $\lambda = 0.98$ (blue)

Instead of the rainfall input, we can also use a smoothed version of the observed model errors as external forcing for the diffusion term in equation 3.13. If we consider innovations (or one-step-ahead prediction errors) ϵ_k for a forecast created at time step $k-1$, we obtain

$$F_k = \lambda \cdot F_{k-1} + (1-\lambda)\epsilon_k^2 \tag{3.15}$$

This may be interpreted as a model for generalized autoregressive conditional heteroskedasticity (GARCH) as applied in Econometrics ([Bol86]). This kind of external forcing for the diffusion term is expected to perform well in capturing the forecast uncertainty over short horizons. Over longer forecast horizons, the performance of this approach is questionable as the future model error is unknown at the time of forecast generation (paper G).

3.4 From State-Dependent Diffusion to Constant Diffusion - Lamperti Transformations

All structures of the diffusion term discussed in Section 3.3 include a dependency on the state value itself. Such structures are difficult to handle in *CTSM* as the extended Kalman Filter requires higher order terms ([Ves98]). We can avoid this problem by applying so-called Lamperti transformations that move the state dependency from the diffusion term of the SDE to the drift term. We obtain a set of SDE's with a more complicated (and typically highly non-linear drift term) and a state-independent (and ideally constant) diffusion term.

As discussed by [Møl10] and [Iac08], the application of such transformations is also advisable for the simulation of SDE's. State dependent diffusions can, together with the drift term, put restrictions on the state space. Considering for example equation 3.6 with a linearly state-dependent diffusion, limits the state-space to positive values for S_1 and S_2. When simulating the untransformed equation it is numerically not guaranteed that this limitation is obeyed. However, after an appropriate transformation the process lives on the entire real axis and numerical problems on the boundary of the domain are avoided ([Møl10]).

[Møl10] discusses the derivation of Lamperti transformations in detail. In this work, we give only the principle and summarize the results for the linear and the exponentially scaled state-dependent diffusion terms described in Section 3.3. In Section 3.3 we derive the transformation for the input dependent diffusion.

3.4.1 The General Lamperti Transformation

Consider the stochastic differential equation for the i-th state of a model with state vector X

$$dX_{i,t} = f_i(X_t, u_t, t, \theta)dt + \sigma_i(X_t, u_t, t, \theta)d\omega_t = f_i(\cdot)dt + \sigma_i(\cdot)d\omega_t. \tag{3.16}$$

Now consider a transformed state $Z_{i,t} = \phi(X_{i,t})$. Assuming a diffusion process which does not depend on other model states, we can derive a SDE for the transformed state using Itô's lemma ([Øks98])

$$dZ_i = \left(\frac{\partial\phi}{\partial t} + f_i(\cdot)\frac{\partial\phi}{\partial X_i} + \frac{\sigma_i^2(\cdot)}{2}\frac{\partial^2\phi}{\partial X_i^2}\right)dt + \sigma_i(\cdot)\frac{\partial\phi}{\partial X_i}d\omega_i \tag{3.17}$$

To obtain a transformed process with constant diffusion, the task is thus to define a transformation such that

$$\frac{1}{\sigma_i(\cdot)} = \frac{\partial\phi}{\partial X_i} \tag{3.18}$$

3.4.2 Lamperti Transformation for the Input Dependent Uncertainty Description

We consider the stochastic differential equation

$$dX_{i,t} = f_i(X_t, u_t, t, \theta)dt + \left(\varphi(F)\frac{X_{i,t}^2}{c + X_{i,t}^2}\right) d\omega_t \tag{3.19}$$

where φ is a function depending on an external forcing F (c.f. equation 3.13) and c is some constant.

As shown in equation 3.18 we find a transformed state $Z_i = \phi(X_i)$ from

$$\frac{X_i^2 + c}{\varphi(F) \cdot X_i^2} = \frac{\partial \phi}{\partial X_i}. \tag{3.20}$$

Integrating equation 3.20 results in

$$Z_i = \frac{1}{\varphi(F)} \left(X_i - \frac{c}{X_i}\right). \tag{3.21}$$

The corresponding backtransformation is found as

$$X_i = \frac{\varphi(F) \cdot Z_i}{2} + \sqrt{\frac{(\varphi(F) \cdot Z_i)^2}{4} + c}. \tag{3.22}$$

Note that we only consider the positive root of the quadratic equation as the untransformed state X_i must be positive.

We find

$$\frac{\partial \phi}{\partial t} = 0, \ \frac{\partial \phi}{\partial X_i} = \frac{X_i^2 + c}{\varphi(F) \cdot X_i^2}, \ \frac{\partial^2 \phi}{\partial X_i^2} = -\frac{2c}{\varphi(F) \cdot X_i^3} \tag{3.23}$$

and apply Itô's lemma ([Øks98]) to obtain the transformed state equation:

$$dZ_i = \left(0 + f_i(\cdot)\frac{c + X_i^2}{\varphi(F) \cdot X_i^2} - \frac{c \cdot \varphi(F) \cdot X_i}{(c + X_i^2)^2}\right) dt + 1 \cdot d\omega_i \tag{3.24}$$

We substitute X_i by equation 3.22 and use this new state equation for parameter estimation and simulation of the stochastic model in the transformed space. Also in the observation equation, we apply the backtransformation 3.22. Clearly, the backtransformation guarantees positive state values X_i.

The disadvantage of applying the Lamperti transformation is that a simple linear model equation can become strongly nonlinear in the transformed space and thus difficult to handle numerically.

3.4.3 Overview of Applied Lamperti Transformations

Table 3.1 gives an overview of the transformed state equations resulting for the diffusion structures applied in this work. The linear and exponentially scaled state dependencies are discussed in [Møl10] and [BTM$^+$11].

Table 3.1: Diffusion structures, state transformations and transformed state equations for selected diffusion types

Structure of diffusion term σ_i	Backtransformation from transformed state Z_i	Transformed state equation
$\sigma_1 \cdot X_i$	$X_i = e^{\sigma_1 \cdot Z_i}$	$dZ_{i,t} = \left(\frac{f_i(X_t, u_t, t, \theta)}{\sigma_1 \cdot X_{i,t}} - \frac{1}{2}\sigma_1 \right) dt + d\omega_{i,t}$
$\sigma_1 \cdot X_i^{\gamma}$	$X_i = (\sigma_1 \cdot (1-\gamma) \cdot Z_i)^{\frac{1}{1-\gamma}}$	$dZ_{i,t} = \left(\frac{f_i(X_t, u_t, t, \theta)}{\sigma_1 \cdot X_{i,t}^{\gamma}} - \frac{1}{2}\sigma\gamma X^{\gamma-1} \right) dt + d\omega_{i,t}$
$\varphi \cdot \frac{X_i^2}{c+X_i^2}$	$X_i = \frac{\varphi \cdot Z_i}{2} + \sqrt{\frac{(\varphi \cdot Z_i)^2}{4} + c}$	$dZ_i = \left(f_i(X_t, u_t, t, \theta) \frac{c+X_i^2}{\varphi \cdot X_i^2} - \frac{c \cdot \varphi \cdot X_i}{(c+X_i^2)^2} \right) dt + d\omega_{i,t}$

3.5 The Relation between Lamperti and Data Transformations

Several options for transforming the flow observations are discussed in [BTM+11] and [BMMM12]. In [BTM+11] a logarithmic transformation of the observations is suggested, based on the consideration that measurement uncertainty increases with the flow rate. The logarithmic transformation accounts for this behaviour as it leads to a multiplicative behaviour of the observation error.

We propose a different reasoning here. The parameter estimation procedure (c.f. Section 3.6.1) and the extended Kalman filter implemented in *CTSM* both assume that the innovations or one-step ahead predictions $\hat{Y}_{k|k-1}$ are normally distributed with covariance matrix $\sum_{k|k-1}^{yy}$. Assuming that, as a result of the Lamperti transform, the state prediction $\hat{Z}_{k|k-1}$ from equation 3.17 is normally distributed, the form of the observation equation (in our case linear, see equation 3.7) and the backtransformation (table 3.1) define which transformation should be applied to the observations. The exact same transformation will be applied also to the observation equation, and we need to obtain approximate normality for the resulting forecast $\hat{Y}_{k|k-1}$ (c.f Section 3.6.1).

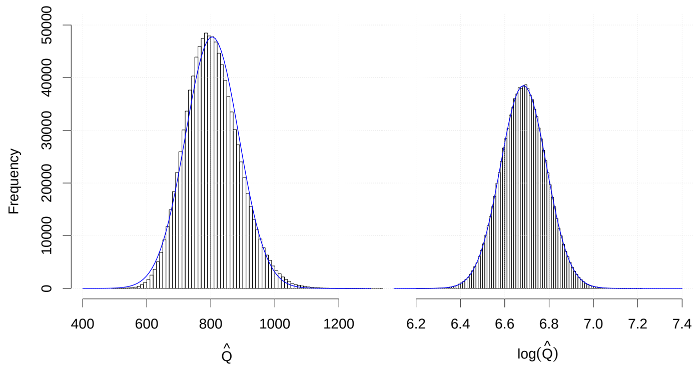
To illustrate this problem, we perform an experiment by sampling from an assumed normal distribution for Z. We select a time constant $K = 5\frac{1}{h}$ and mean and variance for the (imagined) state Z in such a way, that the flow prediction

$$\hat{Q} = \frac{1}{K} \cdot \hat{X} = \frac{1}{K} \cdot \phi^{-1}(\hat{Z}) \tag{3.25}$$

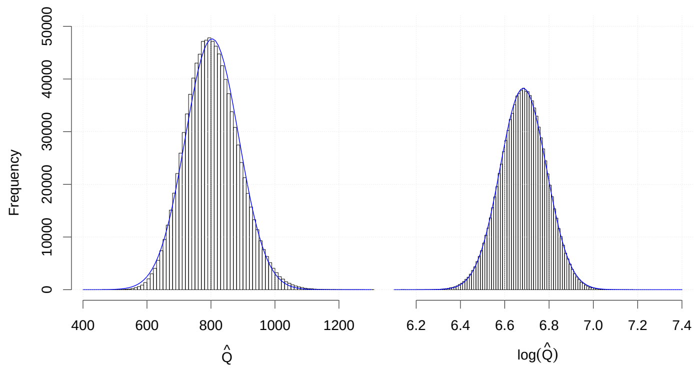
resulting from the backtransformed state X has a mean of $800m^3/h$ and a width of $320m^3/h$ of the 95% prediction interval.

We use a sample size of 10^6 values for Z. Considering the different variants of ϕ resulting from different Lamperti transformations, we assess whether $\hat{Q}$ and its log-transformed pendant $log(\hat{Q})$ can be assumed normally distributed. Figure 3.4 shows the corresponding histograms.

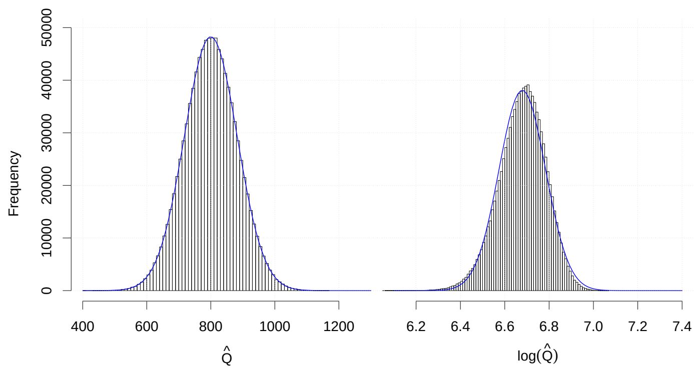
We see from figures 3.4a and 3.4b that for the linear state dependency and the exponentially scaled state dependency (we consider $\gamma = 0.8$), a log transform of the observation equation (and thus also the observations) will lead to a predictive distribution which is closer to normality. This is a shortcoming of papers B and E where no transformation was applied to the observations. We can, however, also notice in figures 3.4a and 3.4b that the distribution for $\hat{Q}$ is only slightly skewed, if the standard deviation is not too big. This is also the case in the articles B and E.



(a) Linear State Dependency



(b) Exponential State Dependency



(c) External Forcing

Figure 3.4: Histograms of flow predictions $\hat{Q}$ and log-transformed flow predictions $log(\hat{Q})$ considering different Lamperti transformations for the normally distributed state Z

For the new state dependency that was introduced for the consideration of external forcings (we consider $c = 1$ and $\varphi = 1$), we obtain the reverse picture (Figure 3.4c). The predictive distribution for $\hat{Q}$ is almost perfectly normal, while the distribution for $log(\hat{Q})$ is slightly skewed. Again the deviation is limited, but no transformation should be applied to the observation equation (and the observations) in combination with this Lamperti transformation.

We conclude that the choice of transformation for the observations should be based on considerations concerning the assumptions on normality for the Kalman filtering and parameter estimation procedures rather than temporal variations of the measurement error.

Finally, the choice of data transformation can also be affected by the quality of the observed data. Negative flow observations (observed in paper E, for instance) make it impossible to use a logarithmic transformation unless some smoothing is applied to the data which in turn introduces time lags into the observations.

3.6 Parameter Estimation in Stochastic Grey-box Models

3.6.1 Maximum Likelihood Estimation

The parameter estimation method implemented in *CTSM* is based on the maximum likelihood principle ([Mad08, Paw01]). This method is comprehensively documented in [KM03] and [KMJ04]. For time series data, the likelihood function is commonly expressed as a product of one step ahead conditional densities. Most commonly when working with *CTSM*, we use a frequentist approach where the only aim is to find the best performing parameter set. We do not consider parameter uncertainty when generating probabilistic forecasts. The underlying assumption is that commonly sufficient data is available to identify the correct set of parameters. However, it is also possible to consider prior information in the parameter estimation procedure (see Section 3.6.2).

The likelihood for a parameter set θ given a set of l-dimensional observations $y_i \in \mathbf{R}^l$

$$Y_n = [y_n, y_{n-1}, \ldots, y_1, y_0] \tag{3.26}$$

can be expressed as

$$L(\theta, Y_n) = \left(\prod_{k=1}^{n} p(y_k|Y_{k-1}, \theta) \right) p(y_0|\theta) \tag{3.27}$$

Assuming that the one-step ahead forecast errors (innovations) $\epsilon_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$ are iid and normally distributed with mean 0 and covariance matrix $R_{k|k-1} = V(y_k|Y_{k-1}, \theta)$, the likelihood function can be rewritten as

$$L(\theta, Y_n) = \left(\prod_{k=1}^{n} \frac{exp\left(-\frac{1}{2}\epsilon_k^T R_{k|k-1}^{-1} \epsilon_k\right)}{\sqrt{det(R_{k|k-1})}(\sqrt{2\pi})^l} \right) p(y_0|\theta) \tag{3.28}$$

Conditioning on y_0, equation 3.28 can be solved as an optimization problem. The one-step ahead forecasts and hence the innovations are obtained through extended Kalman filtering. This implies that during parameter estimation every forecast step is preceded by a state updating to the current observation.

The updating ensures that the model follows the realisation of the stochastic process described by the observations. We can avoid biased parameter estimates resulting from an adjustment of the simulation model to a single realisation of the (stochastic) reality and obtain a better description of the short term dynamics. We assume a normal distribution of the innovations which is usually a realistic assumption.

The principle is illustrated in Figure 3.5. The model is updated to the observations at every time step (each forecast starts closer to the observation than the previous forecasted value). In simple terms, the model parameters are identified by minimizing the average of the innovations shown in red.

The downside of this approach is that occasionally the state updating becomes to be dominant in the parameter estimation. The observation noise is then estimated excessively small and the model states are perfectly updated to the flow observations at every time step while the likelihood values become less sensitive to the physical model parameters which are then difficult to identify.

Furthermore, the mean of the one-step ahead predictions is generally obtained by evaluating only the drift term of the SDE as we assume a locally linear behaviour. This is not necessarily equal to the behaviour of the actual stochastic process.

Nevertheless, the Maximum likelihood based estimation has proven to provide robust results in a multitude of applications (for example [ARM13, BMPN00, BTM+11, Møl10, BM11]).

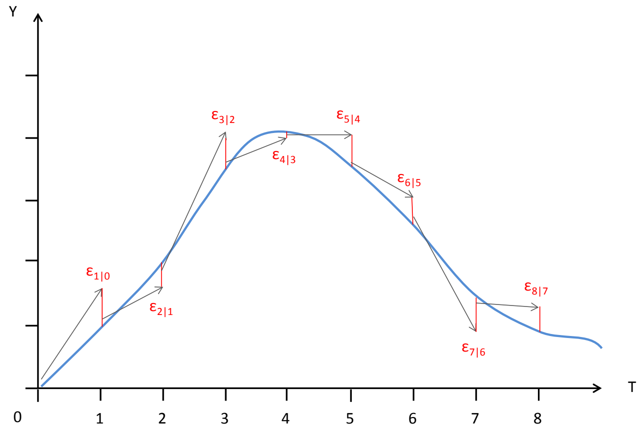


Figure 3.5: Estimation principle for the Maximum Likelihood approach. The true flow (with observations at time point $T = 1, \ldots, 8$) is shown blue, arrows symbolize one step-ahead model forecasts starting from an updated model state, red lines innovations ϵ.

3.6.2 Maximum A Posteriori Estimation

In the parameter estimation procedure provided by *CTSM* prior information can be considered by using the maximum a posteriori (MAP) estimate. The likelihood function is in this case extended with a penalty for the deviation of the parameter from its prior mean. Considering $\theta \in \mathbb{R}^p$, $\mu_\theta = E(\theta)$, $\Sigma_\theta = V(\theta)$ and $\epsilon_\theta = \theta - \mu_\theta$, we obtain

$$L(\theta, Y_n) = \left(\prod_{k=1}^{N} \frac{exp\left(-\frac{1}{2}\epsilon_k^T R_{k|k-1}^{-1} \epsilon_k\right)}{\sqrt{det(R_{k|k-1})}(\sqrt{2\pi})^l} \right) p(y_0|\theta) \frac{exp\left(-\frac{1}{2}\epsilon_\theta^T \Sigma_\theta \epsilon_\theta\right)}{\sqrt{det(\Sigma_\theta)}(\sqrt{2\pi})^p} \tag{3.29}$$

Again, conditioning on y_0, the MAP estimate is obtained as ([KM03, KMJ04])

$$\hat{\theta} = \arg\min(-ln(L(\theta|Y_n, y_0))). \tag{3.30}$$

Prior information can be assigned to selected parameters only by using non-informative priors with large standard deviations for the other parameters. This estimation approach can be used to avoid overly small estimates of the observation noise or if the data do not provide sufficient information to identify all the parameters.

3.6.3 Estimation Based on Multi-step Forecast Verification

The aim to apply the stochastic grey-box models for multi-step ahead forecasts motivated the search for a robust estimation method that allows us to identify a model which also captures the long-term dynamics of the system. This is not always guaranteed with the estimation approach described in Section 3.6.1. A first step in this direction was made in paper A and the estimation procedures were documented in paper C.

As we intend to use the runoff forecasting models for multi-step forecasts in a real-time control setting, the idea behind this approach is to estimate the model parameters based on the quality of multi-step forecasts rather than one-step ahead forecasts as described in Section 3.6.1. In this setting, multi-step forecasts are generated by repeated extended Kalman filtering without updating, as implemented in *CTSM* ([KM03]), and assumed normally distributed with variance $R_{i+j|i}$.

The objective function at time step i is found as a weighted average of the scores SC for flow prediction horizons $i+j$ up to the maximal horizon $i+k$

$$SC_i = \frac{1}{\sum_{j=1}^{k}(k-j+1)}\left(\sum_{j=1}^{k}(k-j+1)\cdot SC_{i,j}\right). \tag{3.31}$$

This is illustrated in Figure 3.6. Shorter horizons receive more weighting in the parameter estimation procedure, as the objective in practice is to generate forecasts of runoff volume for the different horizons $i+j$ to $i+k$. The runoff volume for a given horizon is an integral over the flow predictions for this horizon and all previous horizons. Flow predictions for shorter horizons thus affect the runoff volume forecast for more horizons than the flow predictions for longer horizons and should be given higher weighting.

We apply the *CRPS* (Section 2.2.3) as the score function SC for the flow predictions. The approach is documented as "‘model D"’ in [LMM14]. Advantages of the approach are a more robust parameter identification and that the model is estimated according to the forecasting objective, as the model is estimated based on multi-step forecasts rather than one-step ahead forecasts. A comparison in Section 3.6.5 shows that we obtain lower point forecast errors with this estimation approach.

However, the approach also has several disadvantages.

- *First*, although practically appealing, parameter estimation based on a

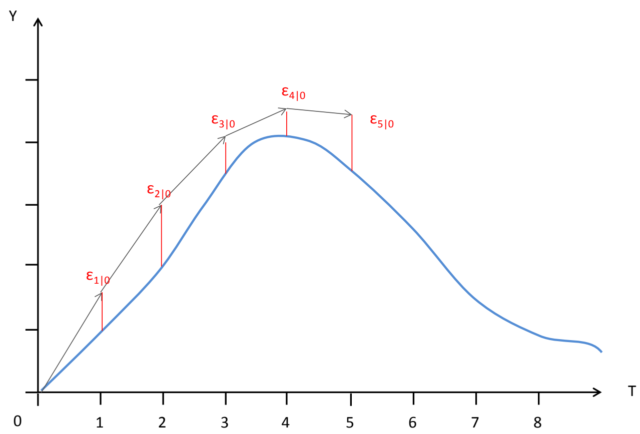


Figure 3.6: Estimation principle for the forecast based approach. The true flow (with observations at time point $T = 1, \ldots, 8$) is shown blue, arrows symbolize model forecasts generated at $t = 0$. In this example we consider a horizon $k = 5$. The score value for the first time step SC_0 is found by an evaluation of the probabilistic flow forecasts for every horizon $j = 1, \ldots, 5$

score function removes the theoretical support of the likelihood principle and thus provides a less well founded interpretability of prediction bounds, for example

- *Second*, the estimation approach described in [LMM14] relies on a normality assumption for multi-step predictions, which cannot necessarily be assumed on longer forecast horizons.

- *Third*, when generating multi-step forecasts using the EKF setup in *CTSM*, we do not actually simulate the stochastic process described by the SDE but only the drift term of the state equations ([KM03]). The expected value of the multi-step forecast used during parameter estimation and the expected value from an actual stochastic simulation of the SDE's (see Section 3.7) are therefore not necessarily the same.

The latter two issues can be avoided if multi-step ahead forecasts during parameter estimation are generated not by extended Kalman filtering but by an ensemble-based approach relying on a stochastic simulation of the SDE's as described in Section 3.7. Furthermore, as we observe an underestimation of forecast uncertainties by the models estimated using the forecast based approach (papers B, C, E), we can replace the $CRPS$ as a score function in equation 3.31 by another score function, for example, a Gaussian density.

3.6.4 Numerical Optimisation

All of the parameter estimation approaches described above rely on a numerical minimization of the objective function. We have in the first works applied Genetic algorithms ([Whi94], paper B) and the DDS algorithm ([TS07], papers B and E) because the methods are insensitive to missing values in the objective function evaluation. This behaviour was necessary because the previous version of $CTSM$ would frequently produce errors for "‘random"’ parameter combinations. We have applied these algorithms repeatedly starting from the optimum of the previous optimization run to ensure that a parameter set near the true optimum was identified.

As a result of improvements in $CTSM$ ([JKB+13]) the application of the $PORT$ algorithm ([Gay90]) became possible in later works (papers D and G). In paper G we combine this algorithm in series with the DDS algorithm if integer variables need to be optimized.

3.6.5 Comparing Estimation Approaches for a Sample Catchment

Extending the discussion in paper C, in this section we compare the forecast quality for three of the catchments considered in paper E (Amager East - EAm, Colloseum - COL, Kløvermarken - Klo). We consider a cascade of three reservoirs with time-invariant parameters as shown in paper E. We consider point and probabilistic forecast quality for models estimated using the Maximum Likelihood approach (ML, Section 3.6.1) and the forecast based approach using the $CRPS$ (Section 3.6.3). Forecast quality is compared on a 120min or 60 time step horizon for the four validation events described in paper E.

Figure 3.7 shows the point and probabilistic forecast skill for the different events and catchments while tables 3.2 and 3.3 depict the reliability Rel and $ARIL$ (c.f. 2.2.1) values for 90% prediction intervals, respectively. Point forecast are derived using the median of the probabilistic multi-step forecasts (c.f. 3.8).

We see that the model estimated using the forecast based approach (Section 3.6.3) provides better point (NSE, PI) and probabilistic forecast skill ($CRPS$) on the 60 time step horizon than the model estimated using the ML approach.

This is exactly the intended effect of performing parameter estimation based on forecasts. The forecast based estimation does, however, provide less reliable forecasts than the ML approach. The forecast uncertainty (table 3.3) is estimated smaller leading to smaller reliability values (Table 3.2). Also this tendency has been observed before (papers B, C, E).

Table 3.2: Reliability (Rel) of 90% prediction intervals on a 120min horizon for three catchments considered in paper E

Catchment	Estimation approach	Event 1	Event 2	Event 3	Event 4
EAm	ML	59%	76%	65%	49%
EAm	$CRPS$	60%	68%	72%	65%
COL	ML	26%	49%	76%	40%
COL	$CRPS$	40%	52%	88%	37%
Klo	ML	75%	79%	79%	65%
Klo	$CRPS$	87%	96%	89%	81%

Table 3.3: Average Interval Length ($ARIL$) of 90% prediction intervals on a 120min horizon for three catchments considered in paper E

Catchment	Estimation approach	Event 1	Event 2	Event 3	Event 4
EAm	ML	72%	73%	78%	57%
EAm	$CRPS$	118%	117%	117%	90%
COL	ML	55%	54%	83%	136%
COL	$CRPS$	84%	77%	117%	206%
Klo	ML	87%	78%	73%	82%
Klo	$CRPS$	251%	213%	182%	196%

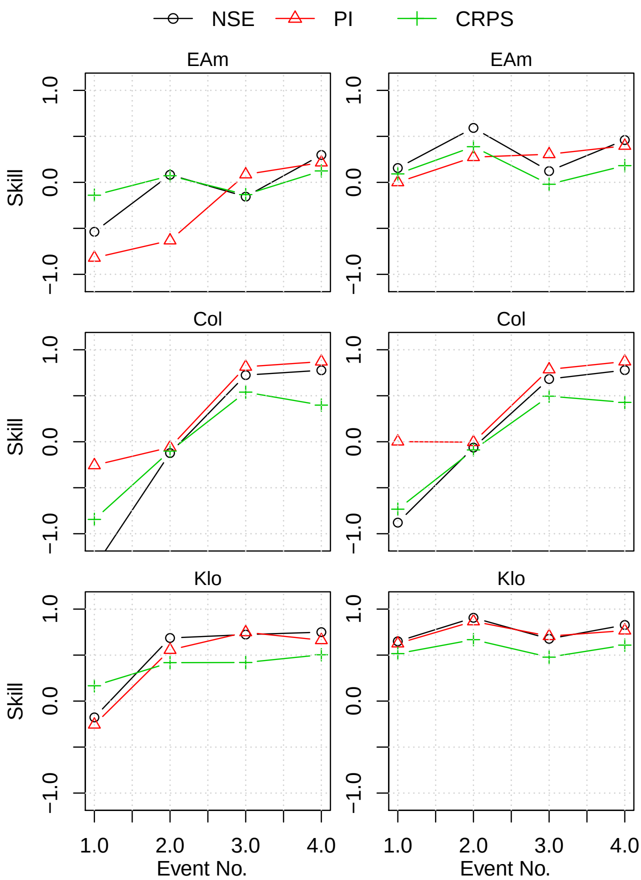


Figure 3.7: Forecast skill (NSE, Persistence Index PI, normalized CRPS) in the EAm, Col and Klo catchments for the validation events in paper E for models estimated using the ML based estimation approach (left, c.f. Section 3.6.1) and the forecast based estimation approach (right, c.f. Section 3.6.3)

3.7 Simulation of Stochastic Differential Equations

Simulations of stochastic differential equations distinguish themselves from those of ODE's in that we need to consider a coupled drift and diffusion term system. There is no single solution for a SDE but only a set of realisations which as a whole have statistical properties describing the stochastic process. Simulation methods for SDE's are described by [KP99] and [Iac08]. In this work, we mainly apply such stochastic simulations for the generation of ensemble-based probabilistic forecasts (see Section 3.8).

The simplest simulation method for SDE's is the Euler-Maruyama scheme ([KP99]). The considered time steps Δt are discretized into increments of equal length h. The simulated state for the i-th increment $\hat{X}_{t+h\cdot i|t}$ is then found as

$$\hat{X}_{t+h\cdot i|t} = \hat{X}_{t+h\cdot(i-1)|t} + f\left(\hat{X}_{t+h\cdot(i-1)|t}, u_{t+h\cdot i}, t+h\cdot i, \theta\right)\cdot h + \tag{3.32}$$

$$\sigma\left(\hat{X}_{t+h\cdot(i-1)|t}, u_{t+h\cdot i}, t+h\cdot i, \theta\right)\cdot \Delta W_{t+h\cdot(i-1)} \tag{3.33}$$

$\Delta W_{t+h\cdot(i-1)}$ is a realisation of an increment of the Wiener process which is equivalent to a random normal number with mean 0 and variance h. We apply the Euler-Maruyama scheme for the generation of probabilistic forecasts in papers E and F. A discretization step of $h = \Delta t/100$ is used for the simulations in these cases.

The Euler-Maruyama scheme is easily implemented but does have drawbacks in that very small increments h may be required to properly capture the physical and stochastic behaviour of the simulated process. We can avoid issues in modeling time-varying diffusions and make the state variables defined over the whole real domain by applying Lamperti transformations to our model (see Section 3.4). The resulting transformed state equations, however, are typically non-linear and stiff in some cases. Improved solution methods that are valid in such cases are implicit (such as the single step backward Euler method (SSBE, [HMS02]) - see paper G), account for the Jacobian of the drift term (such as the weak exponential scheme ([Mor05])) or apply predictor-corrector schemes ([BLP08]).

3.8 Generating Multi-step Probabilistic Forecasts from Stochastic Grey-box models

The eventual purpose of the considered stochastic grey-box models is to generate probabilistic forecasts of runoff volume for real-time control. The relevant forecast horizon depends on the current state of the system, in particular the current basin fillings. Flow forecasts must therefore be generated for a multitude of forecast horizons and runoff volumes must be derived from these.

In a probabilistic sense, we need to account for the correlation between the flow forecasts for different horizons. Furthermore, the runoff volume is a sum of the flow forecasts for different horizons. The distribution for this sum of random variables is not necessarily straight forward to derive. In paper C we have used a sampling approach while subsequently we have moved to an ensemble (or scenario) based approach (papers B, E and G) to derive the predicted distribution of runoff volume.

We aim to generate a probabilistic k-step ahead forecast of runoff volume $\hat{V}_{t+k|t}$ starting at time step t. We follow the scheme shown in Figure 3.8:

1. Assuming the last flow observation Q_t being available at t, we start with the updated model states $Z_{t,t}$ provided by the extended Kalman filter. We assume these to be multivariate normal with variance matrix $\Sigma_{t|t}$ (this assumption is also made in the Kalman filtering).
2. We generate N multivariate samples from this distribution (using the R-package ([RCT13]) *MASS* [VR02]). These serve as a starting point for the simulations.
3. We use the Euler-Maruyama scheme as described in 3.7 to generate N simulations of the Lamperti-transformed state equations up to $t + k$.
4. We integrate each flow simulation into a runoff volume and obtain a sample of N runoff volumes $\hat{V}_{t+k|t}$.
5. We empirically describe the forecasted distribution using quantiles from the sample of runoff volumes.

The observation noise in the proposed scheme can be disregarded as it is not relevant for decision making in real-time control. Correlation between the runoff forecasts for different horizons does not need to be considered explicitly as it is provided by the scenario simulations (other than if we would consider different models for different forecast horizons, for example). Moreover, we simulate the

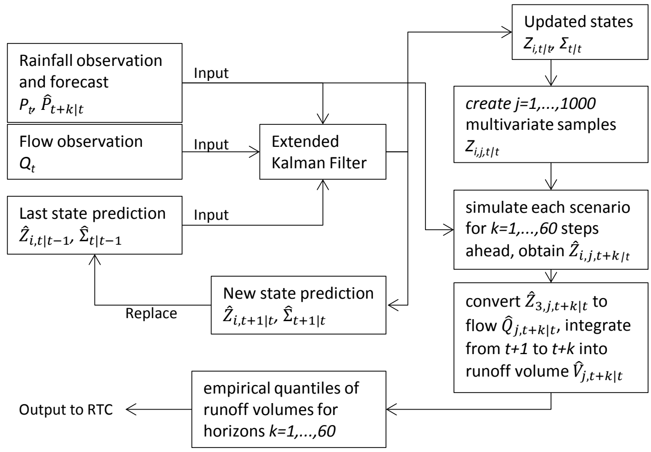


Figure 3.8: Ensemble approach for the generation of probabilistic forecasts of runoff volume $\hat{V}_{t+k|t}$ for different forecast horizons k

stochastic process described by the state equations 3.6 without any distributional assumptions.

However, the approach also has drawbacks. First, the computational effort is rather high as ensemble simulations need to be performed. Nevertheless, the approach is feasible and suitable for an on-line application as demonstrated in paper F. Second, the forecast approach is not consistent with the parameter estimation approaches described in Section 3.6. These rely on a forecast solely of the drift term of the model. We may consider this a drawback of the parameter estimation rather than the forecasting procedure which suggests further investigations in this direction.

3.9 Relation to Other Uncertainty Techniques Applied in Hydrology

A multitude of uncertainty techniques have been proposed in hydrology and other fields. These include Bayesian techniques, *GLUE* and other methods which are described below.

The grey-box modelling approach discussed in this work distinguishes itself in its very strong focus on on-line applications. It provides techniques for parameter estimation and data assimilation with very limited computational effort and the ability to generate probabilistic forecasts and account for model deficiencies during parameter estimation as an inherent feature. Drawbacks include the assumption of normality in the filtering and parameter estimation procedures and the limit on model complexity which is imposed by the EKF and the requirement to solve a multivariate system of SDE's.

Bayesian techniques are very widely applied in hydrology. These mainly distinguish themselves from the grey-box approach in the explicit consideration of parameter uncertainty during parameter estimation and simulation, while the grey-box approach 'lumps' parameter, structural and input uncertainty into the diffusion term. The use of prior information is common in the Bayesian school while it is an exception in the grey-box approach. Mostly, Bayesian approaches have a strong focus on uncertainty analysis in an off-line setting and structural model identification. Commonly, they require Monte Carlo simulations during parameter estimation and simulation which can be seen as their main drawback.

Uncertainty formulations used in conjunction with Bayesian techniques in the literature show a wide range of complexities. In the most simplistic case, errors are lumped into a single output error term e which is assumed iid (and typically normally distributed). The observations Y are then obtained depending on some model function g as

$$Y_t = g(X_t, u_t, t, \theta) + e \tag{3.34}$$

More advanced approaches account for input uncertainty using rainfall multipliers ([KKFT06],[RKK+10], [SBK13]) and structural uncertainty using time varying parameters ([RKK+10]). A large number of parameters is used in these approaches which makes them unattractive for on-line modelling purposes. This drawback is removed by adding a time varying bias term D to equation 3.34 in [RS12]. The bias captures structural model uncertainties and input uncertainties. A comparison with the grey-box approach in paper D indicates good simulation performances of this approach. The computational effort, however, is challenging even in this case.

The **GLUE** approach ([Bev93]) is widely used for uncertainty estimation in urban hydrology ([TBJSJ08, DMK+12, VMDM13]). It is methodologically similar to the Bayesian approaches but does not restrict itself to the theoretical likelihood framework. Due to the similarity to the Bayesian methods, we do not discuss the method further here.

Other approaches combine Bayesian estimation techniques and data assimilation using ensemble ([VDG+05]) or particle ([MDS12, VTDS13]) filtering approaches. The ensemble / particle filtering makes these approaches more computationally demanding than the grey-box approach. However, they are closer to an on-line application than the approaches above as the model can recursively be updated to new observations without considering the whole time series. In particular the approach presented by [VDG+05] is close to the grey-box approach in that it offers the possibility to use time-varying state noise descriptions and may be considered for the development of more complex conceptual model structures.

Other commonly applied uncertainty techniques focus on a **post-processing** of model results. The downside of these methods is that structural and input uncertainty is not explicitly accounted for during the identification of the actual model. On the other hand, these methods can perfectly be used in combination with complex physical models which is not true for the grey-box approach (unless a grey-box model is used as post-processing method rather than as forecasting model).

Among others, such approaches use time series models in combination with data assimilation in the physical model ([MS05]), Bayesian model averaging (BMA) ([HFZ13]) or non-parametric methods ([Pin07, PMN+09]). For the consideration of forecasts over multiple horizons either multivariate distributions (and suitable transformations of the model residuals) ([PMN+09, HFZ13]) or copulas ([TPM13]) must be considered to correctly capture correlation between different forecast horizons. As the correlation structure is likely to be time varying, it might be necessary to estimate it recursively as described in [PMN+09] and paper C. The feature of describing correlation between forecast horizons is inherent for approaches that generate probabilistic forecasts based on direct simulations of the model (such as the grey-box and the Bayesian approaches).

Chapter 4

Rainfall Inputs for On-line Runoff Forecasting

4.1 Requirements to Rainfall Input for On-line Runoff Forecasting

Runoff forecast models rely on rainfall observations and forecasts as input. This raises the issue of how good the rainfall input needs to be. Several criteria are important in this context:

- spatial resolution
- temporal resolution
- reliable measurement of the rainfall process
- reliable operation
- forecast quality

The criteria of spatial and temporal resolution have been intensely discussed in the literature. We have partly summarized this discussion in paper B. In addition, [Sch91] gives guidelines on the resolution rainfall measurements should have for design and operational applications in urban hydrology.

Reliability of rainfall measurements may be considered the key criterion in terms of a forecast model based operation of urban drainage systems. We consider reliability both in terms of operational availability and accuracy of the available measurements and forecasts. This matter is discussed further in Section 4.2. Identifying periods with an acceptable quality of rainfall data was the main issue during the work on paper E.

Finally, it seems obvious that a runoff forecast in an on-line setting profits from using a rainfall forecast. However, due to the reaction time of a catchment, the expected future runoff from a catchment up to a certain horizon will be well determined by the measured rainfall. The extent of this effect depends on the characteristics of the catchment (size, shape, degree of sealing and location of sealed areas).

In Section 4.3 we describe an experiment comparing runoff forecast quality on a 2 hour horizon with perfect rainfall forecast and without rainfall forecast and conclude that in some cases reasonable runoff forecasts can actually be obtained without rainfall forecast. On the contrary, [TR13] analyse the runoff forecast quality that can be obtained using radar rainfall forecasts in a 80ha urban catchment and find that the runoff forecast skill is very limited for horizons exceeding 60 minutes. We are not aware of systematic investigations into situations in which rainfall forecasts are beneficial for short term on-line runoff forecasting (in terms of catchment characteristics and forecast horizons) and what quality of rainfall forecasts is required. We consider this an important item of future research.

4.2 Raingauge vs. Radar Rainfall Input

4.2.1 Advantages and Disadvantages

In this thesis we have applied rainfall inputs from rain gauge an C-band radar measurements for runoff forecasting up to horizons of 120 minutes. Rainfall inputs from X-band local area weather radars (LAWR) were applied at the start of the project but found to not provide useful input for runoff forecasting. New processing methodologies have been developed recently which are likely to lead to more reliable rainfall observations with LAWR ([NTR12, NJR13]).

Below we discuss the advantages and disadvantages of rain gauge and radar rainfall measurements in terms of the criteria described in Section 4.1.

Spatial Resolution
Depending on the characteristics of the rain event, rainfall can vary strongly in space. An improved spatial resolution may generally be considered an advantage of radar rainfall measurements as compared to rain gauges. The former will provide an average of the rain intensities present within a radar pixel while the latter provide only a point measurement. Having only a point measurement may lead to false assumptions about the average rain intensity over a catchment and time displaced rainfall observations ([BGM13]).

In paper B we demonstrate that radar rainfall measurement as input to the stochastic runoff forecasting models has the potential to provide improved runoff forecasts in comparison to rain gauges. It cannot be clearly determined from the results, whether the improvement results from an improved rainfall forecast or an improved representation of the areal rainfall by the radar. Considering the results in Section 4.3 that show that reasonable runoff forecasts on a 120min horizon can be obtained without rainfall forecast in several cases in the bigger catchments (750ha) and the fact that the catchments considered in paper B are even bigger (1300 and 3000ha), we see an indication that spatial resolution may have an important role.

Figures 4.1 and 4.2 support this assumption. The plots show that the mean areal catchment rainfall derived from the time-statically adjusted radar data considered in paper B yields a similar or better correlation to the measured runoff from the catchment than the mean areal rainfall derived from rain gauges. Notably, the highest correlation is identified for lags of approximately 150 minutes between rainfall and runoff measurements. This is longer than the considered forecast horizon of 100 minutes. This result remains valid after prewhitening of the time series ([MS05]), the peak in the cross correlation is then identified for lags between 100 and 140 minutes.

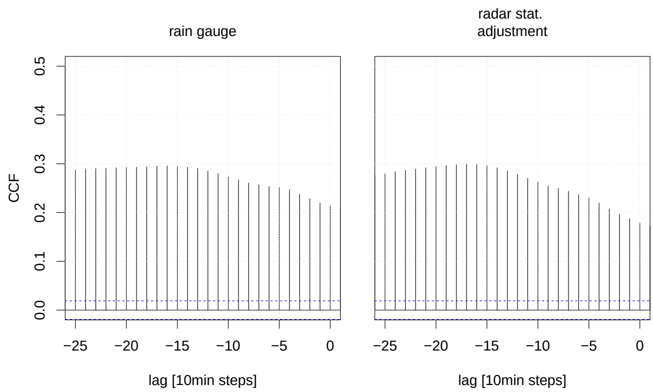


Figure 4.1: Cross correlation (CCF) between mean areal rainfall derived from rain gauges (left) and time-statically adjusted radar rainfall measurements (right) and runoff measurements in the Ballerup catchment (see paper B)

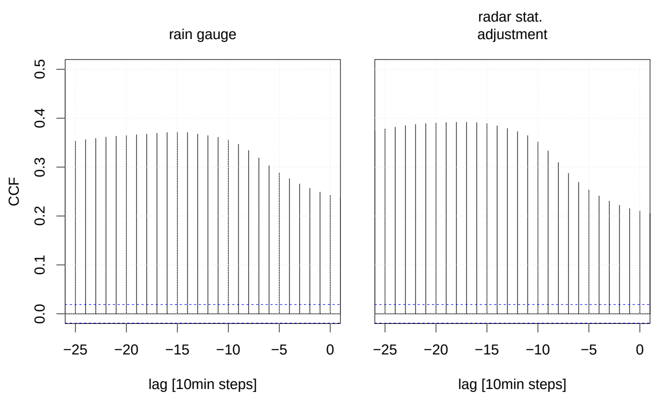


Figure 4.2: Cross correlation (CCF) between mean areal rainfall derived from rain gauges (left) and time-statically adjusted radar rainfall measurements (right) and runoff measurements in the Damhusåen catchment (see paper B)

Temporal Resolution

A clear advantage for one of the measurement methods cannot be identified in terms of temporal resolution. The available resolutions depend in both cases on the applied measurement principles, processing algorithms and the agreed protocols for on-line data transfer. In this work we were largely bound to a temporal resolution of 10 minutes for the rainfall measurements, corresponding to the available resolution of the C-band radar measurements.

In [Sch91] a temporal resolution Δt for the rainfall data that depends on the concentration time t_c of the catchment as

$$\Delta t = \frac{1}{3}t_c \cdots \frac{1}{5}t_c \tag{4.1}$$

is suggested. A coarse estimate for t_c in seconds can be found from the total catchment area A in m^2 as $t_c = \sqrt{A}$. Applying this to the Ballerup and Damhusåen catchments considered in papers A, B, C, D and G, we see that a 10 minute resolution of rainfall measurements is sufficient to model runoff from these catchments. This is, however, not necessarily true for the smaller catchments considered in papers E and F.

Reliable Measurement of the Rainfall Process

In a measuring sense, it is intuitive to assume the radar rainfall measurements less reliable for an on-line operation of drainage systems. There are issues in the conversion from the quantity measured by the radar (reflectivity for the C-band radar, drop counts for the LAWR) to rain intensities and in the attenuation of the signal (see for example [KVPT12, BK13, NJR13]).

However, as [BGM13] discusses, a time displacement of the rainfall measurement (which may, for example, result from measuring only at a point location) can have the same impact on runoff-forecasting models as a strong bias in the rainfall measurement. This issue needs to be considered when applying rain gauge observations as input to runoff forecasting models.

Moreover, in paper B we demonstrate that the rainfall input applied for runoff forecasting with automatically calibrated models does not need to resemble the observations on the ground. [TR13] evaluate the runoff forecast quality that can actually be obtained for a small catchment (80ha) using radar rainfall input and find reasonable runoff forecasts up to a horizon of 60 minutes.

Operational Reliability

Operational reliability needs to be ensured for both rain gauge and radar measurements. [ML13] show that a gauge network can be operated reliably for on-line purposes also over long periods. It may be a benefit that the maintenance of the rain gauges is within the responsibilities of the operator of the drainage network. This is quite often not the case for the weather radars. A careful processing of the radar rainfall measurements is required before using them as input for runoff forecasting models.

4.2.2 Merging Different Rainfall Inputs

Radar and rain gauges give different perspectives on the rainfall process. While the radar provides the spatial distribution of the rainfall, there are problems with obtaining accurate rain intensities due to spatially and temporally varying drop size distributions and signal attenuation ([BK13]). Rain gauges, on the other hand, provide very local measurements of the rainfall process, which, depending on the density of the gauge network, may not be spatially representative.

A multitude of publications is available on how to combine radar and rain gauge measurements into an improved observation of the 'true' rainfall process. We refer to the overview provided by [GD09] and the introduction of paper B. Common merging approaches focus on the adjustment of the radar measurements to make them more "'similar"' to rain gauge observations.

In paper B we demonstrate that this may not be necessary for on-line runoff forecasting models that are calibrated using radar rainfall input. On the contrary, if the adjustment is performed in an unsuitable way, for example, with strong time variations in the adjustment factors, the resulting rainfall measurement may yield worse results as input for runoff forecasting than the original radar measurement (paper B).

We have evaluated the performance of a simple statistical combination method presented by [GHL02] in paper A. In this approach, a simple autoregressive state space model

$$X_t = A \cdot X_{t-1} + e_t \tag{4.2}$$

$$Y_t = C \cdot X_t + s_t \tag{4.3}$$

is created for the rainfall process X_t with radar and rain gauge observations Y_t, parameter matrices A and C and assumed normal error vectors e_t and s_t (see [GHL02]). The resulting combined rainfall measurement corresponds to the rainfall state X_t, which in paper A is gridded with a resolution of 2x2km. The different rainfall measurements Y_t are merged with the rainfall state X_t using an ordinary Kalman filter and the model parameters are estimated by maximizing the likelihood of obtaining the radar and rain gauge measurements Y_t.

This approach is very appealing because physical information about the rainfall process can, in principle, be incorporated in the model A and because any kind of measurement providing information about rainfall can be included in the merging process given a suitable formulation of the observation equation in C.

The approach does, however, in the given form also have disadvantages in that we assume normality for the rainfall states X_t and observations Y_t. This assumption certainly does not hold for the rainfall process which is bounded at 0. Moreover, the Kalman filter algorithm cannot handle the consideration of areas corresponding to the full size of a C-band radar image. Such an image has an extent of 240x240 pixels and the resulting covariance matrix for e_t would be of dimension 57600x57600. Processing of the full radar image, however, is necessary to generate rainfall forecasts from the radar observations. This is also a disadvantage of the adjustment methodology presented by [Tod01] and [WOSo+13] which is based on the Kalman filter.

Finally, as a result of the discussion in paper B, we concluded that radar adjustment in particular, but also any procedure for merging different rainfall measurements in general, should focus on the final purpose of the resulting rainfall information. This may, for example, be the forecasting of runoff in an on-line setting. To avoid the above problems, we may apply geostatistical merging techniques as described by [BRH13, GD09]. We can then apply simple rainfall runoff models for one or several considered catchments and identify the parameters of the rainfall merging algorithm as a part of the overall parameter calibration for the rainfall runoff model by maximizing the likelihood of the the runoff observations. Such a setting could also be laid out with time varying parameters for the rainfall merging algorithm. These can be identified using a simple state space model layout and an (extended) Kalman filter (see, for example, [BDLY01]).

4.3 The Effect of Rainfall Forecasts on Runoff Forecast Quality - An Example

4.3.1 Problem Description

We have in Section 4.1 questioned whether rainfall forecasts always provide benefit for the generation of runoff forecasts for short horizons. This is underlined by the observation that runoff forecasts based on rain gauge measurements show a strong performance in paper B as compared to the forecasts using radar measurements and forecasts.

In this section we follow up on this discussion by analysing the performance of the stochastic runoff forecast models for the six subcatchments and 8 rain events considered in paper E. We use the same dataset as in paper E and refer the reader to this article for a description of the catchments. Rainfall information is available from C-band radar measurements in a 10min resolution, while a 2min time step is considered for the runoff forecast models and flow data. We consider forecasts of runoff volume for a forecast horizon of 120min (or 60 time steps) (c.f. Section 3.8). Rain events 1 to 4 are used for model calibration, while events 5 to 8 are used for model validation only.

We consider two sets of models:

(a) In the first set, future rainfall is assumed known during parameter estimation and forecast generation.

(b) In the second set, future rainfall is assumed unknown during parameter estimation and forecast generation.

In set (b), future rainfall is generated by extrapolating a local linear model. The linear model is fitted to the rainfall observations over the last 120min (in a resolution of 10min) and then extrapolated over the next 120min (in a resolution of 2min).

4.3.2 Results and Discussion

Figure 4.3 shows the point and probabilistic forecast skill for the different events and catchments while tables 4.1 and 4.2 depict the reliability *Rel* and *ARIL* (c.f. 2.2.1) values for 90% prediction intervals, respectively. Point forecasts are derived using the median of the probabilistic multi-step forecasts (c.f. 3.8).

Considering the score values in Figure 4.3, we can see a difference in the forecast skills of the models with known and unknown rainfall input in all catchments except Str. This catchment is a special case due to time varying dry weather flows. These lead to parameter estimates with very large time constants K and, as a result, to runoff forecasts that are hardly affected by the rainfall input. If we allow for a time-varying dry weather flow parameter a_0 in this model, the model using a known rainfall input performs clearly better than the one without (not shown).

In the other catchments we observe a strong loss of forecast skill in the EAm and WAm catchments when the future rainfall is considered unknown. In the Klo and Ler catchments (the biggest catchments considered), we observe similar forecast skill in the cases where rainfall input is known and unknown for some events. Evaluating the $ARIL$ values (Table 4.2), we see a clear increase of the forecast uncertainty in the EAm, COL and WAm catchments when the future rainfall is unknown, while the $ARIL$ values only slightly increase in the Klo and Ler catchments. This behaviour fits with the point forecast performance of the models for the calibration events 1-4.

4.3.3 Conclusion

We conclude that the quality of rainfall forecasts has an influence on the runoff forecast quality in the considered urban catchments. This effect is particularly pronounced in the smaller EAm, COL and WAm catchments. A similar result is found for an 80ha catchment in [TR13], where runoff forecast quality diminishes with the quality of the rainfall forecasts when exceeding a horizon of 60min.

However, in the larger catchments (Klo and Ler) the result is less clear. As discussed in Section 4.1, a rainfall forecast may consequently not be necessary for the generation of runoff forecasts for certain combinations of forecast horizon and catchment size. This behaviour can have practical relevance in cases where short-term on-line rainfall forecasts in catchments are not available due to failures or missing installations.

Table 4.1: Reliability (Rel) of 90% prediction intervals on a 120min horizon for the catchments considered in paper E

Catch-ment	Rain input	Rain Event 1	2	3	4	5	6	7	8	Mean
EAm	Known	80%	72%	82%	73%	59%	76%	65%	49%	70%
EAm	Unknown	74%	77%	84%	81%	65%	72%	71%	65%	74%
COL	Known	78%	78%	68%	64%	26%	49%	76%	40%	60%
COL	Unknown	82%	82%	70%	59%	43%	50%	78%	51%	65%
Klo	Known	77%	72%	65%	81%	75%	79%	79%	65%	74%
Klo	Unknown	75%	74%	65%	74%	74%	68%	80%	62%	71%
Ler	Known	67%	46%	21%	73%	66%	77%	78%	55%	60%
Ler	Unknown	69%	46%	21%	81%	81%	86%	78%	68%	66%
Str	Known	62%	76%	53%	72%	63%	78%	67%	62%	66%
Str	Unknown	89%	64%	45%	71%	73%	70%	12%	52%	60%
WAm	Known	51%	76%	79%	61%	81%	98%	78%	72%	74%
WAm	Unknown	63%	67%	72%	74%	72%	81%	78%	63%	71%

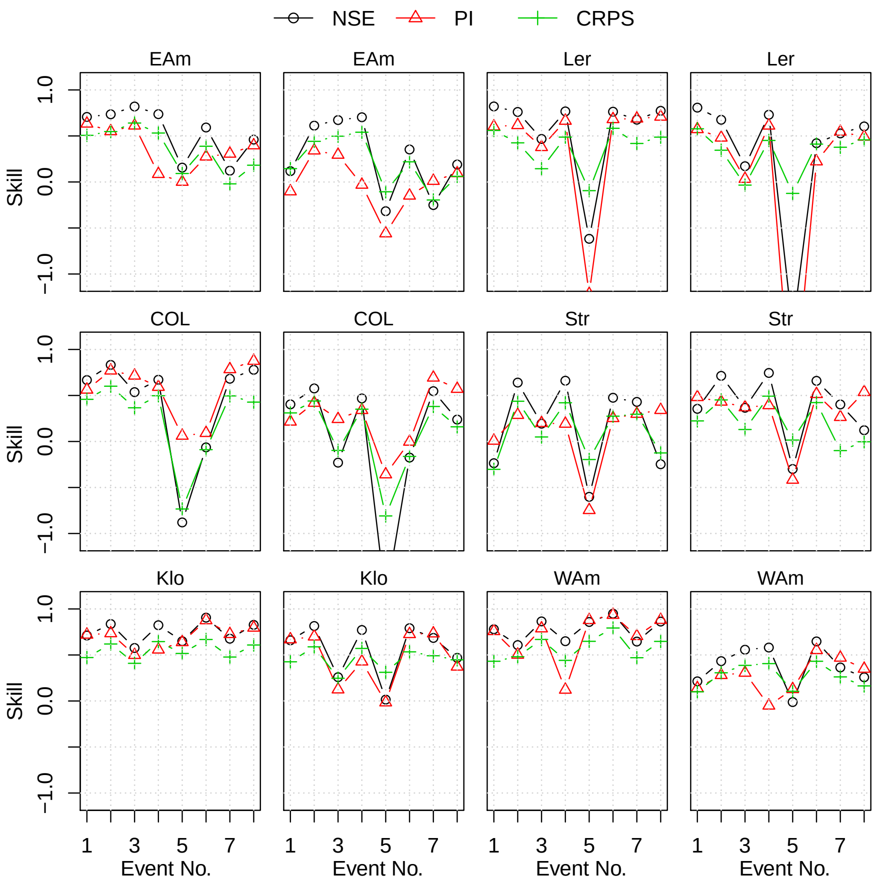


Figure 4.3: Forecast skill (NSE, Persistence Index PI, normalized CRPS) in the catchments considered in paper E. Events 1-4 were used for calibration, events 5-8 for validation only. We consider perfectly known future rain inputs (columns 1 and 3) and unknown future rain inputs (columns 2 and 4)

Table 4.2: Average interval length ($ARIL$) of 90% prediction intervals on a 120min horizon for the catchments considered in paper E

Catch-ment	Rain input	Rain Event 1	2	3	4	5	6	7	8	Mean
EAm	Known	75%	87%	79%	70%	72%	73%	78%	57%	74%
EAm	Unknown	109%	124%	114%	97%	105%	105%	115%	89%	107%
COL	Known	82%	103%	108%	128%	55%	54%	83%	136%	94%
COL	Unknown	107%	137%	150%	170%	75%	79%	113%	174%	125%
Klo	Known	84%	307%	110%	74%	87%	78%	73%	82%	112%
Klo	Unknown	88%	310%	115%	77%	89%	80%	75%	87%	115%
Ler	Known	269%	427%	235%	237%	139%	167%	200%	540%	277%
Ler	Unknown	279%	425%	283%	256%	146%	167%	222%	518%	287%
Str	Known	44%	51%	51%	47%	49%	47%	57%	49%	50%
Str	Unknown	46%	57%	59%	53%	50%	52%	74%	53%	55%
WAm	Known	47%	55%	54%	47%	52%	56%	64%	73%	56%
WAm	Unknown	64%	72%	71%	64%	69%	72%	86%	93%	74%

Chapter 5

Probabilistic Runoff Forecasts in Practice

5.1 Real-Time Control under Uncertainty - the DORA algorithm

The final purpose of generating probabilistic rainfall runoff forecasts is to improve decision making in real-time control. In this work we have considered the *DORA* (dynamic overflow risk assessment) algorithm ([VG14]) for decision making under uncertainty as it was created by the project partners in the *SWI* (Storm and Wastewater Informatics) project and allowed for the direct application of the probabilistic forecast models in a real world setup.

The *DORA* algorithm tries to reduce the impact of combined sewer overflows (CSO). It adjusts outflows from control points as the free variables in an optimization setting. Its objective function is to minimize the CSO cost in the whole catchment over the considered forecast horizon T. CSO cost C_F for a single structure is determined as a function of the forecasted runoff volume V_F to the control point

$$C_F = \int_0^\infty C(V_F) \cdot p(V_F) dV_F, \tag{5.1}$$

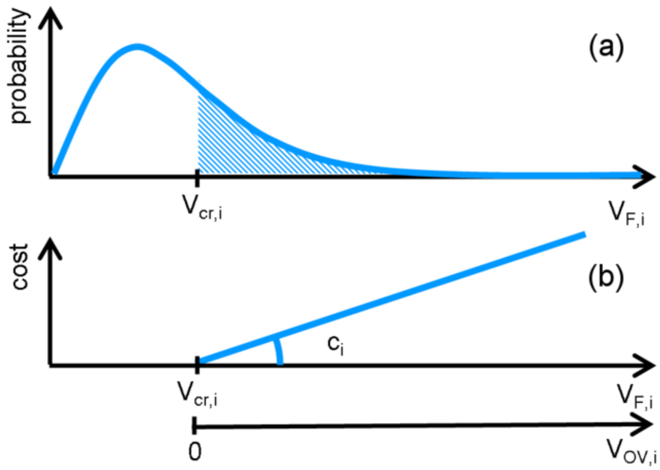


Figure 5.1: *Subfigure a*: Probabilistic runoff volume forecast for a horizon T at structure i with critical volume V_{cr} (marks the runoff volume where the basin is filled and CSO starts), *subfigure b*: overflow cost $c_i = C(V_{F,i})$ as a function of predicted volume $V_{F,i}$ and resulting overflow volume $V_{OV,i}$ (from [VG14])

where $p(V_F)$ corresponds to the forecasted probability that a certain runoff volume will occur and $C(V_F)$ to the overflow cost associated with this volume.

The cost function $C(V_F)$ must be defined by the stakeholders. In this work it is a piecewise linear function of V_F which is equal to 0 if V_F does not induce overflow (Figure 5.1). The slope of the cost function determines how much weight an overflow structure receives in the control setup.

DORA is attractive as a control algorithm, because it allows to account for forecast uncertainty without exploding dimensionality in the optimization setting. It can account for any distributional shapes of the probabilistic forecasts by evaluating the integral in equation 5.1 either parametrically or non-parametrically. In paper F we have implemented an empirical evaluation of equation 5.1 using quantiles of the scenario based probabilistic forecasts described in Section 3.8 in a 2% resolution. Other criteria such as pollutant loads, surface flooding or energy consumption as a result of pumping or water treatment can be implemented in the same framework.

A major drawback of the current implementation of the *DORA* algorithm is that only a single forecast horizon T is considered in the decision making. Consequently we need to define this horizon at every control time step and the resulting outflows will correspond to an average setting over this horizon rather than the series of decisions which will be implemented in reality.

5.2 Current Practical Implementation for Probabilistic Runoff Forecasting

The current practical implementation of *DORA* in Copenhagen uses runoff forecasts generated by conceptual models that are recalibrated in intervals of 10 minutes. This process is described in [LPB+14]. We have included this approach as a benchmark in papers E and F.

The runoff forecast uncertainty in this approach is described using a Gamma distribution. The mean of this distribution is found as

$$E[X] = k \cdot \theta, \tag{5.2}$$

where k is the shape and θ the scale parameter of the distribution. In the implementation, the mean of the distribution is assumed equal to the predicted runoff volume V_F for the considered horizon and we thus have $\theta = V_F/k$. The shape parameter is fixed to $k = 3$.

In paper C we demonstrate that this assumption of a fixed shape parameter leads to a strong overestimation of forecast uncertainty for a reasonably well performing forecast model.

5.3 Control Results Using Probabilistic Forecasts

An experiment in paper C shows that a strong overestimation of forecast uncertainties for runoff volume will also lead to a strong overestimation of overflow risk in accordance with equation 5.1.

Neglecting forecast uncertainty in the computation of overflow risk, on the other hand, hardly affects the computed overflow risk as compared to an evaluation with a proper quantification of forecast uncertainty. The reason for this behaviour is the simplistic layout considered in the experiment. During the evaluation of predicted overflow risk over the time series, we mainly stay in the linear range of the cost functions while only very few data points are located at the break point from zero cost to non-zero cost (c.f. Figure 5.1).

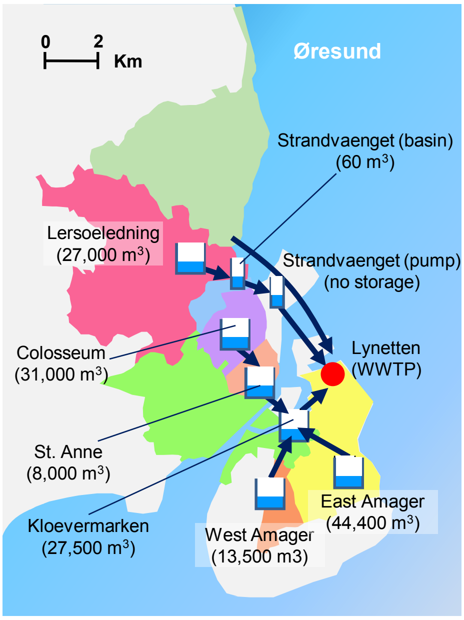


Figure 5.2: Control points in the Lynetten catchment with available storage volume (from paper F)

A different behaviour can be expected for non-linear cost functions and when using probabilistic forecasts in an actual optimization setting for real time control. Due to the variety of different basin outflows considered during the optimization, the critical runoff volume V_{cr} that leads to CSO will move around on the x-axis at every time step and a proper quantification of runoff forecast uncertainty will affect decision making.

This was tested for the Lynetten catchment in Copenhagen in paper F. We implemented grey-box based runoff forecast models for the subcatchments EAm, Col, Klo, Ler, Str and WAm (Figure 5.2). A forecast model is in this case applied only for the specific subcatchment of a control point, while inflow from upstream control points is considered as a model input as it is determined by the control algorithm.

The results in the article confirm our hypothesis. Runoff forecasts based on grey-box models that account for uncertainty lead to a clear reduction of CSO volume compared to both the baseline and the grey-box based runoff forecasts where forecast uncertainty is neglected in the decision making. Similar to the results in paper C, an overestimation of runoff forecast uncertainty with the approach based on the Gamma distribution leads to a significantly worse performance in decision making and increased CSO volumes.

Chapter 6

Conclusions

This thesis focuses on the development of methods for the generation of probabilistic runoff forecasts using stochastic grey-box models. We have made the step from considering this type of models for simulation studies towards an implementation that can actually be applied on-line. Forecast horizons of up to 120 minutes are considered in this work and the forecasts are applied in a real-time control context.

Using the questions posed in Section 1.3 as an outline, here we present the main conclusions from this work.

1. *What rainfall inputs should be used for short-term runoff forecasting and, in particular, do we benefit from using quantitative precipitation estimates (QPE) from weather radar?*

   The results from paper B suggest that we can obtain better runoff forecasts when using radar rainfall measurements rather than rain gauge observations. It is not clear whether the improvement results from a better measurement of the rainfall process or an improved rainfall forecast provided by the radar. However, catchments with reduced areas of 1300 and 3000ha were considered. For such large catchment sizes, the influence of the rainfall forecasts on the runoff forecasts may be small on a forecast horizon of 2 hours. In addition, a correlation analysis in Section 4.2.1 indicates a better description of the mean areal rainfall by the radar measurements.

The operational reliability of radar rainfall measurements is problematic. Rain gauges, on the other hand, provide a mature method for measuring rainfall. Operators of urban drainage networks, weather services and authorities are therefore experienced in the operation of hardware and data analysis ([JRMM98]). Rain gauges can thus be considered a "‘robust"’ means of measuring rainfall. Our experience from practical implementation in papers E and F suggests that this is not necessarily the case for radar rainfall measurements in Denmark.

2. *Do quantitative precipitation forecasts (QPF) provide benefit for short term runoff forecasts?*

   As discussed in Sections 4.1 and 4.3, the extent to which runoff forecasts can benefit from rainfall forecasts depends on the catchment characteristics and the considered forecast horizon. [TR13] demonstrate in an 80ha catchment that the benefits from radar rainfall forecasts on runoff forecasts diminish when exceeding horizons of 60 minutes. On the other hand, we show in Section 4.3 that in some cases we can obtain reasonable runoff forecasts without (or with very simple) rainfall forecasts.

   We are not aware of systematic investigations that evaluate which forecast horizons and catchment characteristics require rainfall forecasts to generate good on-line runoff forecasts and consider this an interesting item of future work.

3. *Do short term runoff forecasts benefit from a combined rainfall input making use of both rain gauge and radar rainfall measurements?*

   As discussed in paper B, numerous authors in the literature suggest that radar rainfall measurements should be adjusted to rain gauges before being used in (urban) hydrological applications. The validation of the adjustment methodology is, however, performed by comparing the adjusted radar data to rain gauge measurements or by using them as input for models that were calibrated using rain gauge measurements. The results in paper B indicate that if the adjustment is performed in an unsuitable way, the resulting rainfall measurement will yield worse results than using the non-adjusted radar rainfall data, even though the adjusted radar data have smaller bias when compared to the rain gauge observations.

   Nevertheless, it seems somewhat natural that the consideration of additional rainfall information from different sources should result in an improved information about the rainfall process. We emphasize that the adjustment (or merging) of radar and rain gauge measurements should focus on the runoff forecasting purpose. This means that the adjustment methodologies should be calibrated using an objective function that is based on the runoff forecast skill that can be obtained with the adjusted rainfall input.

The investigation presented in paper A does not yet account for the above considerations. Nevertheless, it demonstrates that a merging of radar and rain gauge measurements yields an improved runoff forecasting skill. In this context, it is important to mention the aspect of operational reliability, as the merging of two different types of rainfall information will introduce redundancy in the forecast system and thus increase operational reliability.

4. *How can forecast models and parameters be identified in the context of noisy data and provide forecasts over a multitude of horizons?*

   The probabilistic runoff forecasts generated by the stochastic grey-box models are required over a multitude of forecast horizons in the considered real time control applications. We suggest that the forecast models are estimated with this purpose in mind, minimizing the (weighted, probabilistic) forecast error of the multi-step predictions, rather than applying the (theoretically more sound) estimation method based on likelihood maximization which is implemented in the software framework *CTSM*.

   We develop an estimation methodology based on minimisation of the continuous ranked probability score ($CRPS$) for multi-step ahead forecasts of runoff volume in paper C. The methodology currently has drawbacks as we assume normal distribution for the multi-step ahead forecasts and the resulting models underestimate forecast uncertainty. Nevertheless, we demonstrate in Section 3.6.5 that the new estimation method improves runoff forecast skill on a 2 hour (60 step) horizon.

   An important conclusion from these considerations is that the model should ideally be estimated considering the intended application in mind. The state-of-the-art maximum likelihood approach that is usually applied for stochastic grey-box models will commonly identify models that perform well on short horizons. If we want to apply the models for multi-step predictions over longer horizons, different estimation approaches should be considered.

5. *How can dynamically changing forecast uncertainties be correctly captured in a probabilistic model structure?*

   The results in paper G suggest that it is suitable to model runoff forecast uncertainty as a combination of a constant dry weather uncertainty and a dynamic uncertainty for rain periods which is scaled by a smoothed version of the rainfall input.

   This type of model structure yields better results than the linear dependence of forecast uncertainty on the predicted model states which was suggested by [BTM^{+}11]. In particular, we can avoid the mix up of dry and wet weather uncertainty which, for the linear state dependence, leads to very large forecast uncertainties in some of the catchments in paper G. We are furthermore able to render forecast uncertainty independently from the forecasted runoff and can thus better capture forecast uncertainties occuring, for example, in the start of a rain event.

Nevertheless, the new approach cannot compensate for cases where the physical model structure is clearly insufficient for the data and we will still obtain unreliable forecasts in such cases.

6. *How can probabilistic forecasts be generated for decision making in real-time control?*

   The thesis builds on the assumption that simple models that can be tuned to minimize the forecast error need to be used for forecasting in an on-line context and that we need to account for forecast uncertainty in decision making. Following this line of thought, we use stochastic grey-box models to generate probabilistic on-line forecasts.

   We propose a scenario-based approach for the generation of multi-step probabilistic runoff forecasts. This approach is not bound to distributional assumptions. Moreover, it spares us the task of explicitly modelling the correlation between probabilistic runoff forecasts for different horizons, as it is inherent in the scenarios. Scenario simulations can be created from SDE's using the methods described in Section 3.7.

   Scenario (or ensemble) forecasts also provide flexibility for the use of the probabilistic information in the real-time control algorithm. We have implemented a real-time control scheme which makes use of the quantiles of the probabilistic forecasts (see Section 5.1 and paper F).

7. *What effect does the consideration of forecast uncertainty have on the efficiency of real-time control schemes?*

   We have evaluated the effect of runoff forecast uncertainty on the expected combined sewer overflow risk in paper C and on decision making in real-time control in paper F. From both applications we can conclude, that real-time control (in our case with respect to CSO) will benefit from a correct quantification of forecast uncertainty if non-linearities exist in the relation between the runoff forecast and the objective function of the control scheme.

   In this thesis, we apply piecewise linear cost functions (see Section 5.1). These cost functions are zero if a forecasted runoff volume will not lead to overflow and increase linearly with the runoff volume otherwise. When the uncertainty of the forecasted runoff volume is strongly overestimated, the forecasted overflow risk (the objective function) will be affected by this non-linearity and forecasted too big. This was demonstrated in paper C and leads to suboptimal decision making in paper F.

   When neglecting forecast uncertainty during decision making, the case is less obvious in paper C, because the estimation of overflow risk is only for very few time steps affected by the non-linearity in the cost function according to Figure 5.1.

However, considering an actual real-time control setting based on an optimization of basin outflows, the position of the non-linearity in the cost function (V_{cr} in Figure 5.1) with respect to the forecasted runoff volume changes depending on the basin outflows defined by the controller. A correct (or reasonable) quantification of forecast uncertainty then becomes important. We can see this from the strong performance of the control scheme when applying stochastic grey-box models in paper F.

Chapter 7

Outlook

As an outcome of this thesis, we have established a framework of stochastic, conceptual models that can be applied for probabilistic runoff forecasting in an on-line context in urban hydrology. We have suggestions for how different sources of observations of rainfall should be used for runoff forecasting and how forecast uncertainties can be quantified reliably. Furthermore, we have obtained results that suggest that an appropriate quantification of forecast uncertainty has a positive impact on decision making in real-time control, while a large overestimation of forecast uncertainty will negatively impact real-time control schemes.

As a result of new developments in the field and of remaining deficiencies in the applied modelling approach, we suggest further research in the following directions:

Optimal combination of rainfall input, rainfall forecast and rainfall-runoff model
For on-line runoff forecasting, rainfall input, rainfall forecast and runoff forecast model should be considered as a chain. First steps in this direction are taken by [TR13], for example.

The combination of different rainfall inputs (radar and rain gauge, for example) has the potential to improve runoff forecasts (see paper A and the introductory discussion in paper B). The combination of these inputs, however, should be performed in such a way that runoff forecast quality is optimized.

This implies setting up a framework where adjustment procedures and runoff forecast models are tuned in one and the same calibration procedure. We are not aware of any such framework than can be applied operationally. The approach presented by [GHL02] and applied in paper A follows this line of thought but has limitations in the extent of datasets that can be considered.

When do rainfall forecasts improve on-line runoff forecasts
The results in Section 4.3 suggest that we may in some cases be able to generate runoff forecasts without or with very limited information about the future rainfall. In the literature, discussions on the influence of the uncertainty of rainfall input on simulation quality are available (see the work by [BGM13], [LAP+14] and [SF86] for example). A structured discussion considering several catchments of different characteristics is missing in the literature on in urban hydrology. Such a discussion should consider these parameters

- forecast horizon,
- catchment size,
- catchment characteristics (in particular reduced area).

For generality, simulation studies based on theoretical catchments can be considered, similar to the approach in [SF86]. We would expect to find that rainfall forecasts are not required for applications using runoff forecasts on very short horizons. Similarly, for large catchments we may find the forecast uncertainty related to weather models acceptable in order to, for example, coarsely determine process settings on the wastewater treatment plant.

Optimal estimation of probabilistic rainfall-runoff models
If the stochastic grey-box models should be applied for generating runoff forecasts on longer forecast horizons, they should be calibrated in a way that reflects this application. We propose a corresponding estimation approach in paper C. This approach can be improved in the following ways:

- Parameter calibration is currently performed by generating multi-step predictions using the extended Kalman filtering approach implemented in *CTSM* and by minimizing the *CRPS* score averaged over the different forecast horizons. In several applications (papers B and E) we have observed that this approach leads to reasonable model performance in terms of point forecast quality. Runoff forecast uncertainties, however, are somewhat underestimated. We suggest to consider whether this problem can be solved by considering different score functions (such as likelihood inspired scores or a combination of interval scores focusing on different quantiles).

- During parameter calibration, forecasts are generated through the extended Kalman filter implemented in *CTSM*. We should investigate if it is possible to (in a computationally feasible way) base the parameter estimation procedure on direct simulations of the SDE's as described in Section 3.8. Such an approach would not rely on the assumption of normality of the forecasts and give a better representation of the actual stochastic processes. Particle filtering approaches as described by [MDS12] may also be relevant in this context.

In addition, there is also an issue of what time resolution should be selected for the runoff forecast models. When considering longer time steps (and thus shorter forecast horizons in terms of time steps), the forecast error from the very simple model structures may in some cases be smaller. This issue was not investigated in this work.

Structural complexity
The stochastic grey-box approach using *CTSM* has limitations in how complex the considered models can be. Discussions in the literature (for example [SF86, Bre12]) and the results in paper B suggest that somewhat more complex structures give benefit to short-term forecasts. In real-time control we currently apply a separate stochastic runoff forecast model for each control point. The forecast is then used as input for the decision making algorithm. It is very likely that the control schemes could benefit if we used a single stochastic model for the whole catchment which is then also used during decision making. Finally, new applications such as the development of on-line models for the capacity of the wastewater treatment plant also call for stochastic modelling techniques that allow for greater model complexity to represent the different processes.

The number of model states that can be considered in the stochastic grey-box approach is limited because an extended Kalman filter is applied. In this context we can consider different updating methods described, for example, by [Bor14] and [HVLKS11]. In particular, the combined Bayesian uncertainty description and data assimilation approach described by [VDG+05] is very related to the grey-box approach and should be tested with respect to the dynamic modelling of uncertainties and the generation of reliable forecasts on multi-step horizons.

In this context, Maximum a posteriori parameter estimation may be helpful for identifying parameters for more complex model structures.

Future applications
From a practical viewpoint, a foreseen item of future research is the development of libraries of stochastic rainfall-runoff models in combination with automated routines for residual analysis. Multiple models of different complexity can then be tested for a given catchment and the best performing model structure can be selected for on-line applications.

In a qualitative sense, we should investigate model structures that can be used for on-line forecasting of water quality and treatment capacity of the WWTP. Such structures will allow for an integrated operation of the urban drainage system with respect to pollutant loads and thus the actual stress on natural water bodies. First steps were taken by [BNMP99], but generally the model structures in this area are extremely complex and not suitable for on-line purposes.

Furthermore, we expect the integrated operation of urban energy and water systems (Smart Cities) to become a major driver of the on-line operation of urban drainage systems. Such a combined operation offers strong economic incentives for the operators of drainage systems and environmental benefits for society. This calls for the development of new control strategies that account for incentives from the energy market and uncertainties in both forecasts of electricity prices (for example [JPN$^+$13]) and loads from the drainage system, as well as model structures that can forecast the energy production and demand in the urban water cycle.

Bibliography

[ARM13] Philip Delff Andersen, Carsten Rode, and H Madsen. An arctic low-energy house as experimental setup for studies of heat dynamics of buildings. *Front. Archit. Res.*, 2(4):488–499, December 2013.

[ATCV13] Janelcy Alferes, Sovanna Tik, John Copp, and P A Vanrolleghem. Advanced monitoring of water systems using in situ measurement stations: data validation and fault detection. *Water Sci. Technol.*, 68(5):1022–1030, January 2013.

[BCG+13] Neil D Bennett, Barry F W Croke, Giorgio Guariso, Joseph H A Guillaume, Serena H Hamilton, Anthony J Jakeman, Stefano Marsili-Libelli, Lachlan T H Newham, J.P. Norton, Charles Perrin, Suzanne A Pierce, Barbara Robson, Ralf Seppelt, Alexey A Voinov, Brian D Fath, and Vazken Andreassian. Characterising performance of environmental models. *Environ. Model. Softw.*, 40:1–20, 2013.

[BDLY01] P E Brown, P J Diggle, M E Lord, and P C Young. Space–time calibration of radar rainfall data. *J. R. Stat. Soc. Ser. C - Appl. Stat.*, 50(2):221–241, 2001.

[Bev93] K J Beven. Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Water Resour.*, 16(1):41–51, 1993.

[BGM13] M Borup, M Grum, and P S Mikkelsen. Comparing the impact of time displaced and biased precipitation estimates for online updated urban runoff models. *Water Sci. Technol.*, 68(1):109–116, 2013.

[BK13] A Berne and W F Krajewski. Radar for hydrology: Unfulfilled promise or unrecognized potential? *Adv. Water Resour.*, 51:357–366, January 2013.

[BLP08] Nicola Bruti-Liberati and E Platen. Strong predictor-corrector Euler methods for stochastic differential equations. *Stochastics Dyn.*, 8(3):561–581, September 2008.

[BM11] P Bacher and H Madsen. Identifying suitable models for the heat dynamics of buildings. *Energy Build.*, 43(7):1511–1522, July 2011.

[BMMM12] A Breinholt, J K Møller, H Madsen, and P S Mikkelsen. A formal statistical approach to representing uncertainty in rainfall–runoff modelling with focus on residual analysis and probabilistic output evaluation – Distinguishing simulation and prediction. *J. Hydrol.*, 472–473:36–52, 2012.

[BMPN00] H Bechmann, H Madsen, N K Poulsen, and M K Nielsen. Grey box modeling of first flush and incoming wastewater at a wastewater treatment plant. *Environmetrics*, 11(1):1–12, 2000.

[BNMP99] H Bechmann, M K Nielsen, H Madsen, and N K Poulsen. Grey-box modelling of pollutant loads from a sewer system. *Urban Water*, 1(1):71–78, 1999.

[Bol86] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *J. Econom.*, 31(3):307–327, April 1986.

[Bor14] M Borup. *Real time updating in distributed urban rainfall runoff modelling*. Phd, Technical University of Denmark (DTU), 2014.

[Bre12] A Breinholt. *Uncertainty in prediction and simulation of flow in sewer systems*. Phd, Technical University of Denmark, 2012.

[BRH13] C. Berndt, E. Rabiei, and U. Haberlandt. Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *J. Hydrol.*, in press, October 2013.

[Brö12] J Bröcker. Evaluating raw ensembles with the continuous ranked probability score. *Q. J. R. Meteorol. Soc.*, 138(667):1611–1617, 2012.

[BS05] D Butler and M Schütze. Integrating simulation models with a view to optimal control of urban wastewater systems. *Vulnerability Water Qual. Intensively Dev. Urban Watersheds*, 20(4):415–426, 2005.

[BTM+11] A Breinholt, F O Thordarson, J K Møller, M Grum, P S Mikkelsen, and H Madsen. Grey-box modelling of flow in sewer systems with state-dependent diffusion. *Environmetrics*, 22(8):946–961, 2011.

[CMM88] V T Chow, D R Maidment, and L W Mays. *Applied hydrology*. McGraw-Hill, 1st edition, 1988.

[CNH96] J Carstensen, M K Nielsen, and P Harremoës. Predictive Control of Sewer Systems by Means of Grey-box Models. *Water Sci Technol*, 34(3-4):189–194, 1996.

[DDR13] D J Dürrenmatt, D Del Giudice, and J Rieckermann. Dynamic time warping improves sewer flow monitoring. *Water Res.*, 47:3803–3816, 2013.

[DHI03] DHI. MOUSE User Guide and Reference Manual. *Agern Alle*, 11, 2003.

[DHS+13] D Del Giudice, M Honti, A Scheidegger, C Albert, P Reichert, and J Rieckermann. Improving uncertainty estimation in urban hydrological modeling by statistically describing bias. *Hydrol. Earth Syst. Sci.*, 17:4209–4225, 2013.

[DMK+12] C.B.S. Dotto, Giorgio Mannina, M. Kleidorfer, L Vezzaro, Malte Henrichs, D.T. McCarthy, Gabriele Freni, W Rauch, and A. Deletic. Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling. *Water Res.*, 46(8):2545–58, May 2012.

[FB05] L Fuchs and T Beeneken. Development and implementation of a real-time control strategy for the sewer system of the city of Vienna. *Water Sci. Technol.*, 52(5):187–194, 2005.

[FPM11] O Fradet, M Pleau, and C Marcoux. Reducing CSOs and giving the river back to the public: innovative combined sewer overflow control and riverbanks restoration of the St. Charles River in Quebec City. *Water Sci. Technol.*, 63(2):331–338, 2011.

[Gay90] D M Gay. Usage Summary for Selected Optimization Routines. Technical report, AT&T Bell Laboratories, Murray Hill, NJ, United States, 1990.

[GD09] E Goudenhoofdt and L Delobbe. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrol. Earth Syst. Sci.*, 13(2):195–203, 2009.

[GHL02] M Grum, P Harremoës, and J J Linde. Assimilating a multitude of rainfall and runoff data using a stochastic state space modeling approach. In *9th Int. Conf. Urban Drain.*, volume Portland,, 2002.

[GR07] T Gneiting and A E Raftery. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.*, 102(477):359–378, 2007.

[GTC+11] M Grum, D Thornberg, M L Christensen, S A Shididi, and C Thirsing. Full-Scale Real Time Control Demonstration Project in Copenhagen's Largest Urban Drainage Catchments. In *Proc. 12th Int. Conf. Urban Drain.*, Porto Alegre, Brazil, 2011.

[HFZ13] S. Hemri, F. Fundel, and M. Zappa. Simultaneous calibration of ensemble river flow predictions over an entire range of lead times. *Water Resour. Res.*, 49:6744–6755, October 2013.

[HMS02] D J Higham, X Mao, and A M Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, 40(3):1041–1063, 2002.

[HVLKS11] CJ Hutton, CS Vamvakeridou-Lyroudia, Z Kapelan, and DA Savic. Real-time modelling and Data Assimilation techniques for improving the accuracy of model predictions. *Prep. 2011.010 Deliv. D3.6.2*, page 83, 2011.

[Iac08] S M Iacus. *Simulation and inference for stochastic differential equations : with R examples.* Springer Series in Statistics. Springer, Milan, Italy, 2008.

[JKB+13] R Juhl, N R Kristensen, P Bacher, J K Møller, and H Madsen. CTSM-R User Guide, 2013.

[JPN+13] Tryggvi Jónsson, Pierre Pinson, Henrik Aalborg Nielsen, Henrik Madsen, and Torben Skov Nielsen. Wind Power Predictions. *Ieee Trans. Sustain. Energy*, 4(1):210–218, 2013.

[JRMM98] H K Jørgensen, S Rosenørn, H Madsen, and P S Mikkelsen. Quality control of rain data used for urban runoff systems. *Water Sci. Technol.*, 37(11):113–120, 1998.

[JXZS10] X Jin, C Y Xu, Q Zhang, and V P Singh. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *J. Hydrol.*, 383(3-4):147–155, 2010.

[KKFT06] George Kuczera, Dmitri Kavetski, S Franks, and Mark Thyer. Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *J. Hydrol.*, 331(1):161–177, 2006.

[KM03] N R Kristensen and H Madsen. Continuous Time Stochastic Modelling CTSM 2.3 - Mathematics Guide. Technical report, Technical University of Denmark (DTU), Lyngby, Denmark, 2003.

[KMJ04] N R Kristensen, H Madsen, and S B Jørgensen. Parameter estimation in stochastic grey-box models. *Automatica*, 40(2):225–237, 2004.

[KP99] P E Kloeden and E Platen. *Numerical Solution of Stochastic Differential Equations*, volume 3rd. Springer, Berlin-Heidelberg-New York, 1999.

[KVPT12] S Kraemer, H R Verworn, A Pfister, and A Treis. Streckenintegrierte Regenmessung mit dem Mikrowellendämpfungsdifferenzverfahren und Quantifizierung von Radarsignaldämpfung. *Hydrol. und Wasserbewirtschaftung*, 56:59–77, 2012.

[LAP+14] F. Lobligeois, V. Andréassian, C. Perrin, P. Tabary, and C. Loumagne. When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrol. Earth Syst. Sci.*, 18(2):575–594, February 2014.

[Lju10] G M Ljung. Perspectives on system identification. *Annu. Rev. Control*, 34(1):1–12, April 2010.

[LMM14] R Löwe, P S Mikkelsen, and H Madsen. Stochastic rainfall-runoff forecasting: parameter estimation, multi-step prediction, and evaluation of overflow risk. *Stoch. Environ. Res. Risk Assess.*, 28:505–516, 2014.

[LPB+14] N S V Lund, J W Pedersen, M Borup, P S Mikkelsen, and M Grum. Auto-Calibration For Updating Of Linear Reservoir Models Used In Flow Forecasting Of Urban Runoff. *Hydrol. Process.*, (in preparation), 2014.

[Mad08] H Madsen. *Time series analysis.* Chapman and Hall/CRC, Boca Raton, FL, 2008.

[MB10] Christos K. Makropoulos and David Butler. Distributed Water Infrastructure for Sustainable Communities. *Water Resour. Manag.*, 24(11):2795–2816, January 2010.

[MDS12] Hamid Moradkhani, C M DeChant, and S Sorooshian. Evolution of ensemble data assimilation for uncertainty quantifcation using the particle filter markov chain monte carlo method. *Water Resour. Res.*, 48(12), 2012.

[ML13] F Männig and M Lindenberg. Betriebserfahrungen mit der Abflusssteuerung des Dresdner Mischwassernetzes. *KA Korrespondenz Abwasser, Abfall*, 2013(12):1036–1043, 2013.

[MMI05] M Maeda, H Mizushima, and K Ito. Development of the real-time control (RTC) system for Tokyo sewage system. *Water Sci. Technol.*, 51(2):213–220, 2005.

[Møl10] J K Møller. *Stochastic State Space Modelling of Nonlinear systems.* Imm-phd-2010-246, Technical University of Denmark (DTU), 2010.

[Mor05] C. M. Mora. Weak exponential schemes for stochastic differential equations with additive noise. *IMA J. Numer. Anal.*, 25(3):486–506, February 2005.

[MS05] H Madsen and C Skotner. Adaptive state updating in real-time river flow forecasting—a combined filtering and error forecasting procedure. *J. Hydrol.*, 308(1–4):302–312, December 2005.

[MW77] A H Murphy and R L Winkler. Reliability of Subjective Probability Forecasts of Precipitation and Temperature. *J. R. Stat. Soc. C (Applied Stat.*, 26(1):41–47, 1977.

[NJR13] J E Nielsen, N.E. Jensen, and M R Rasmussen. Calibrating LAWR weather radar using laser disdrometers. *Atmos. Res.*, 122:165–173, 2013.

[NM06] H A Nielsen and H Madsen. Modelling the heat consumption in district heating systems using a grey-box approach. *Energy Build.*, 38(1):63–71, 2006.

[NS70] J E Nash and J V Sutcliffe. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.*, 10(3):282–290, 1970.

[NTR12] J E Nielsen, S Thorndahl, and M R Rasmussen. Evaluering af ny processeringsmetode til LAWR. Technical report, Department of Civil Engineering, Aalborg University, Aalborg, Denmark, 2012.

[Øks98] B Øksendal. *Stochastic differential equations : An introduction with applications.* Springer, 6th edition, 1998.

[Paw01] Yudi. Pawitan. *In all likelihood : Statistical modelling and inference using likelihood.* Clarendon Press, Cork, Ireland, 2001.

[PCL+05] M Pleau, H Colas, P Lavalle, G Pelletier, and R Bonin. Global optimal real-time control of the Quebec urban drainage system. *Environ. Model. Softw.*, 20(4):401–413, 2005.

[PCR+09] V Puig, G Cembrano, J Romera, J Quevedo, B Aznar, G Ramón, and J Cabot. Predictive optimal control of sewer networks using CORAL tool: application to Riera Blanca catchment in Barcelona. *Water Sci. Technol.*, 60(4):869–878, 2009.

[Pin07] P Pinson. Non-parametric probabilistic forecasts of wind power: required properties and evaluation. *Wind Energy*, 10(6):497, 2007.

[Pin13] P Pinson. Wind Energy : Forecasting Challenges for its Operational Management. *Stat. Sci.*, 2013.

[PMN+09] P Pinson, H Madsen, H A Nielsen, G Papaefthymiou, and B Kloeckl. From Probabilistic Forecasts to Statistical Scenarios of Short-term Wind Power Production. *Wind Energy*, 12(1):51–62, 2009.

[RCT13] R-Core-Team. R: A Language and Environment for Statistical Computing, 2013.

[Ref96] J C Refsgaard. Terminology, Modelling Protocol And Classification of Hydrological Model Codes. In Michael B. Abbott and Jens Christian Refsgaard, editors, *Distrib. Hydrol. Model.*, volume 22 of *Water Science and Technology Library*, pages 17–39. Springer Netherlands, Dordrecht, Netherlands, 1996.

[RKK+10] B Renard, Dmitri Kavetski, George Kuczera, Mark Thyer, and S Franks. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resour. Res.*, 46(5):W05521, 2010.

[RS12] P Reichert and N Schuwirth. Linking statistical bias description to multiobjective model calibration. *Water Resour. Res.*, 48(9):W09543, 2012.

[RTSS13] Helena M. Ramos, Charlotte Teyssier, Irene Samora, and Anton J. Schleiss. Energy recovery in SUDS towards smart water grids: A case study. *Energy Policy*, in press, November 2013.

[SBB02] M Schütze, D Butler, and M B Beck. *Modelling, simulation and control of urban wastewater systems.* Springer, London, 2002.

[SBK13] Siao Sun and Jean-Luc Bertrand-Krajewski. Separately accounting for uncertainties in rainfall and runoff: Calibration of event-based conceptual hydrological models in small urban catchments using Bayesian method. *Water Resour. Res.*, 49(9):5381–5394, 2013.

[SCC+04] M Schütze, A Campisano, H Colas, W Schilling, and P A Vanrolleghem. Real time control of urban wastewater systems—where do we stand today? *J. Hydrol.*, 299(3):335–348, 2004.

[Sch91] W Schilling. Rainfall data for urban hydrology: what do we need? *Atmos. Res.*, 27(1):5–21, 1991.

[SF86] W Schilling and L Fuchs. Errors in stormwater modeling-a quantitative assessment. *J. Hydraul. Eng.*, 112(2):111–123, 1986.

[SGT+13] A K Sharma, T Guildal, H A R Thomsen, P S Mikkelsen, and B N Jacobsen. Aeration tank settling and real time control as a tool to improve the hydraulic capacity and treatment efficiency during wet weather: results from 7 years' full-scale operational data. *Water Sci. Technol.*, 67(10):2169–2176, 2013.

[SKZ+12] Michael Smith, Victor Koren, Ziya Zhang, Yu Zhang, Seann Reed, Zhengtao Cui, Fekadu Moreda, Brian a. Cosgrove, Naoki Mizukami, and Eric Anderson. Results of the DMIP 2 Oklahoma experiments. *J. Hydrol.*, 418-419:17–48, February 2012.

[SLBF13] K Seggelke, R Löwe, T Beeneken, and L Fuchs. Implementation of an integrated real-time control system of sewer system and waste water treatment plant in the city of Wilhelmshaven. *Urban Water J.*, 10(5):330–341, September 2013.

[TBJSJ08] S Thorndahl, K J Beven, J B Jensen, and Kjeld Schaarup-Jensen. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. *J. Hydrol.*, 357(3):421–437, 2008.

[TBM+12] F O Thordarson, A Breinholt, J K Møller, P S Mikkelsen, M Grum, and H Madsen. Uncertainty assessment of flow predictions in sewer systems using grey box models and skill score criterion. *Stoch. Environ. Res. Risk Assess.*, 26(8):1151–1162, 2012.

[Tho11] F O Thordarson. *Grey Box Modelling of Hydrological Systems*. Phd, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark, 2011.

[Tod01] E Todini. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. *Hydrol. Earth Syst. Sci.*, 5(2):187–199, 2001.

[TPM13] Julija Tastu, P Pinson, and H Madsen. Space-time scenarios of wind power generation produced using a Gaussian copula with parametrized precision matrix. Technical report, Technical University of Denmark (DTU), Kgs. Lyngby, Denmark, 2013.

[TR13] S Thorndahl and M R Rasmussen. Short-term forecasting of urban storm water runoff in real-time using extrapolated radar rainfall data. *J. Hydroinformatics*, 15(3):897–912, 2013.

[TS07] B A Tolson and C A Shoemaker. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.*, 43(1):W01413, 2007.

[VDG+05] J A Vrugt, C G H Diks, Hoshin V Gupta, W Bouten, and J M Verstraten. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resour. Res.*, 41(1):W01017, 2005.

[Ves98] M Vestergaard. *Nonlinear filtering in stochastic volatility models*. Master, Technical University of Denmark (DTU), Lyngby, Denmark, 1998.

[VG14] L Vezzaro and M Grum. A generalized Dynamic Overflow Risk Assessment (DORA) for urban drainage Real Time Control. *J. Hydrol.*, submitted, 2014.

[VMA+03] Henk Vanhooren, Jurgen Meirlaen, Youri Amerlinck, Filip Claeys, Hans Vangheluwe, and Peter A Vanrolleghem. WEST : modelling biological wastewater treatment. *J. Hydroinformatics*, 5(1):27–50, 2003.

[VMDM13] L Vezzaro, P S Mikkelsen, A. Deletic, and D.T. McCarthy. Urban drainage models – simplifying uncertainty analysis for practitioners. *Water Sci. Technol.*, 68(10):2136, November 2013.

[VR02] W N Venables and B D Ripley. *Modern Applied Statistics with S*, volume Fourth. Springer, New York, 2002.

[VTDS13] J A Vrugt, C J F Ter Braak, C G H Diks, and G Schoups. Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Adv. Water Resour.*, 51:457–478, January 2013.

[Whi94] D Whitley. A genetic algorithm tutorial. *Stat. Comput.*, 4(2):65–85, 1994.

[Wil11] D S Wilks. *Statistical methods in the atmospheric sciences*, volume 100. Academic press, Oxford, United Kingdom, 2011.

[WOSo+13] L Wang, Susana Ochoa, Nuno Simões, C Onof, and C Maksimović. Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. *Water Sci. Technol.*, 68(4):737–747, 2013.

Part II

Papers

Papers Included in the Thesis

A Roland Löwe, Peter Steen Mikkelsen, Michael Rasmussen, Henrik Madsen (2013). "'State-space adjustment of radar rainfall and skill score evaluation of stochastic volume forecasts in urban drainage systems"', *Water Science and Technology*, 68(3): 584–590, 2013.

B Roland Löwe, Søren Thorndahl, Peter Steen Mikkelsen, Michael Rasmussen, Henrik Madsen (2014). "'Probabilistic online runoff forecasting for urban catchments using inputs from rain gauges as well as statically and dynamically adjusted weather radar"', accepted by *Journal of Hydrology*, 2014.

C Roland Löwe, Peter Steen Mikkelsen, Henrik Madsen (2014). "'Stochastic rainfall-runoff forecasting: parameter estimation, multi-step prediction, and evaluation of overflow risk"', *Stochastic Environmental Research and Risk Assessment*, 28: 505–516, 2014.

D Dario Del Giudice, Roland Löwe, Henrik Madsen, Peter Steen Mikkelsen, Jörg Rieckermann (2014). "'Comparing two stochastic approaches to predict urban rainfall-runoff with explicit consideration of model bias"', in preparation for *Water Resources Research*, 2014.

E Roland Löwe, Luca Vezzaro, Peter Steen Mikkelsen, Morten Grum, Henrik Madsen (2014). "'Investigating the use of probabilistic forecasts for RTC of urban drainage systems - A Layout for Probabilistic Online Forecasting of Sewer Flows"', in preparation for *Environmental Modelling and Software*, 2014.

F Luca Vezzaro, Roland Löwe, Henrik Madsen, Morten Grum, Peter Steen Mikkelsen (2014). "'Investigating the use of probabilistic forecasts for RTC of urban drainage systems – Full scale testing in the city of Copenhagen, Denmark"', in preparation for *Environmental Modelling and Software*, 2014.

G Roland Löwe, Rune Juhl, Peter Steen Mikkelsen, Henrik Madsen (2014). "'Forecasting Operational Runoff Forecast Uncertainties - State, Rainfall and Error Dependencies"', in preparation for *Hydrology and Earth System Science Discussions*.

In this online version of the thesis, the articles are not included but can be obtained from electronic article databases e.g. via

`www.orbit.dtu.dk`

or upon request from the author (`rolo@dtu.dk`) or DTU Library:

DTU Library, Technical Information Center of Denmark
Anker Engelunds Vej 1, Building 101D
2800 Kgs. Lyngby, Denmark
service@dtic.dtu.dk