Uncertain calibration of urban drainage models

A scientific approach to solve practical problems

Dipl.-Ing. Manfred Kleidorfer

Arbeitsbereich Umwelttechnik

Leopold Franzens Universität Innsbruck

Dissertation

eingereicht zur Erlangung des Grades

Doktor der technischen Wissenschaften

Innsbruck, November 2009

1. Reviewer:
Univ.-Prof. Dipl.-Ing. Dr. Wolfgang Rauch
Institut für Infrastruktur
Arbeitsbereich Umwelttechnik
Universität Innsbruck - Austria

IUT

2. Reviewer:
Univ.-Prof. Dr. Ana Deletic
Department of Civil Engineering
Institute for Sustainable Water Resources
Monash University Melbourne - Australia

ANCORA · IMPARO

Kurzfassung

In der Siedlungsentwässerung unterstützen numerische Simulationsmodelle Planung und Betrieb von Infrastruktureinrichtungen wie beispielsweise von Kanalisationssystemen, Kläranlagen, Infiltrationsanlagen oder Regenwasserbehandlungsanlagen. Sie sind dabei ein geeignetes Hilfsmittel zur Entscheidungsfindung wie umfangreiche Investitionsmittel am besten eingesetzt werden können. Alleine in Österreich wurden im Durchschnitt der Jahre 1993 bis 2000 jährlich ca. 780 Mio € in die Abwasserentsorgung investiert.

In den letzten Jahrzehnten ist mit steigender Rechenkapazität moderner Computer die Komplexizität der verwendeten Modelle rasant angestiegen und immer mehr Prozesse werden immer genauer abgebildet. Um jedoch realitätsnahe Simulationsergebnisse erzielen zu können, müssen derartige Simulationen durch Vergleich der Ergebnisse mit Messdaten kalibriert werden. Der Grund dafür liegt zum einen in der Abstraktion und Simplifikation der Realität zu einem vereinfachten Abbild bei der Modellbildung und zum anderen in den unumgänglichen Unsicherheiten in den Eingangsdaten im Zuge der Datenerhebung. Bei der Kalibrierung werden die relevanten Modell- bzw. Kalibrierungsparameter derart adaptiert, dass eine möglichst gute Übereinstimmung von Simulationsergebnissen und Messungen erzielt werden kann.

Mit steigender Komplexizität der Modelle steigt aber auch die Anzahl der Modellparameter und die Gefahr, im Zuge der Kalibrierung nicht das globale Optimum zu erreichen. So kann eine Kalibrierung zu unterschiedlichen Parametersätzen führen, die alle eine ähnlich gute Übereinstimmung zwischen Messung und Simulation für den Kalibrierungszeitraum erzeugen, wobei dann aber die Ergebnisse im Vorhersagezeitraum deutlich abweichen. Trotz

scheinbar erfolgreicher Kalibrierung kann ein Modell daher für eine Vorhersage und damit für eine Planung ungeeignet sein. Die für eine vorausschauende Planung (die Lebensspanne der Bauwerke liegt bei mehreren Jahrzehnten) notwendigen Prognosen (z.B. Änderung der Landnutzung, Klimawandel, Bevölkerungsentwicklung) kommen dann noch als zusätzliche Quelle von Unsicherheiten hinzu.

In dieser Dissertation werden unterschiedliche Quellen von Unsicherheiten (Unsicherheiten in den Eingangsdaten, in den Kalibrierungsdaten und in der Modellstruktur) und deren Einfluss auf die Ergebnisse unterschiedlicher Modelle analysiert. Weiters werden unterschiedliche Verfahren zur Analyse von Modellunsicherheiten aus verwandten Bereichen dargestellt und gezeigt, wie diese auf Anwendungen der Siedlungsentwässerung übertragen werden können. Der Fokus liegt dabei auf Bayesschen Methoden und deren numerische Lösung mittels Markov Ketten Monte Carlo Simulation.

Es kann gezeigt werden, dass die Verfügbarkeit der Kalibrierungsdaten (räumliche Auflösung und Länge der Datenreihe) die Güte der Modellergebnisse deutlich beeinflusst. Bei zu geringer Datendichte kann ein scheinbar kalibriertes Modell seine Fähigkeit, außerhalb des Kalibrierungszeitraumes zutreffende Prognosen zu machen, völlig verlieren. Dies hängt stark mit Unsicherheiten in den Eingangsdaten hervorgerufen durch die räumliche Regenverteilung zusammen.

Modellstrukturunsicherheiten sind von geringerer Bedeutung, allerdings nur wenn ein Modell überhaupt kalibriert werden kann. Wichtig dabei ist jedoch, dass ein Modellparameter immer spezifisch für dieses Modell bestimmt wird und dabei gewisse Unsicherheiten in der Modellstruktur kompensiert. Auch wenn ein Parameter in zwei Modellen denselben physikalischen Hintergrund hat (z.B. die abflusswirksame befestigte Fläche eines Einzugsgebietes) kann er nicht ohne weiteres von einem Modell in das andere übertragen werden.

Abschließend wird an den Beispielen der Kanalsysteme von Innsbruck und Linz (beide Österreich) gezeigt, wie Modelle der Siedlungsentwässerung kalibriert und verwendet werden, um die Erfordernisse für die Leistungsfähigkeit von Kanälen (ÖWAV Regelblatt 11) und Mischwasserbehandlung (ÖWAV Regelblatt 19) nachzuweisen.

Abstract

Urban drainage simulation models are state of the art instruments for planners, consultants and scientist working in the field of urban hydrology and they are used for design and operation of different infrastructural facilities such as sewer systems, wastewater treatment plants or stormwater treatment facilities. They support decision–making how money is best invested. In Austria alone 780 Mio € per year (in average from 1993 to 2000) were invested in urban drainage.

In the last decades development of more and more sophisticated approaches has been proceeding in urban drainage modelling steadily. With increasing complexity of models data requirements for model building and model calibration rise and it becomes more and more difficult to analyse model structure with taking into account uncertainties of input and calibration data. By now the accuracy of urban drainage model results in regard to model uncertainties is questioned. To obtain realistic modelling results accurate calibration is necessary. Model parameters have to be determined by manual or automatic calibration procedures, by comparing simulation data to observations. Such calibration can lead to different parameter sets which all reach a similarly good fit between measured and simulated data in the calibration period while the model results diverge outside the calibration period.

In this thesis several aspects of uncertainties and calibration of urban drainage models and their impact on model results are discussed. Furthermore different methods for uncertainty analysis are demonstrated and it is shown how they can be adapted to be used with urban drainage models. Thereby this thesis is focusing on Bayesian methods and their numerical solution in Markov Chain Monte Carlo Simulation.

Consequently it is shown that the impact of uncertainties is significant. Special attention has to be payed to temporal and spatial availability of calibration data. These uncertainties are strongly related to uncertainties of input–data respectively to uncertainties due to spatial rainfall distribution.

Analysis of model–structure uncertainties indicates that this source of uncertainties has minor impact on model results if the model can be calibrated sufficiently (which is not always the case). An important point regarding that topic is that model parameters cannot be transferred from one model to another even if they represent the same physical background (e.g. the effective impervious area). Model parameters mostly compensate uncertainties due to model structure and hence always are model specific.

Finally practical model applications for evaluating the requirements of the Austrian guidelines for proofing sewer capacity (ÖWAV Regelblatt 11) and treatment of combined sewage (ÖWAV Regelblatt 19) including model calibration are shown.

Acknowledgements

I would like to thank all those people who supported me during the last years.

First of all I want to thank my supervisor Prof. Dr. Wolfgang Rauch who guided this work. He gave me the chance to participate in several interesting research projects, to attend different international conferences and took always time for questions and discussions.

I also want to thank Prof. Dr. Ana Deletic for many suggestions during my research visit in Melbourne, for agreeing to review this thesis and especially for coming all the long way from Australia to Austria for my final exam.

Finally I would like to thank all my colleagues at the institute for many profitable discussions and a my family and friends for reminding me that there is a life outside the university.

Contents

List of Figures

List of Tables

List of Papers

The following papers are an integral part of this thesis. They are annexed to this thesis starting with page 141 in the printed version, but not in the online version due to copyright issues. Copies of the papers may be obtained from journal publishers.

1. Optimization of measurement campaigns for calibration of a conceptual sewer model
   M. Kleidorfer, M. Möderl, S. Fach and W. Rauch (2009)
   Published in Water Science and Technology **59** (8), 1523-1530,
   DOI:10.2166/wst.2009.154

2. Impact of input data uncertainties on urban stormwater model parameters
   M. Kleidorfer, A. Deletic, T.D. Fletcher and W. Rauch (2009)
   Published in Water Science and Technology **60** (6), 1545-1554,
   DOI:10.2166/wst.2009.493

3. A Bayesian approach for performance evaluation of stormwater models based on long-term high-resolution measurement data
   C.B.S. Dotto, **M. Kleidorfer**, A. Deletic, T.D. Fletcher, W. Rauch (2009)
   Submitted to Journal of Hydrology

4. A case independent approach on the impact of climate change effects on combined sewer system performance
   M. Kleidorfer, M. Möderl, R. Sitzenfrei, C. Urich and W. Rauch (2009)
   Published in Water Science and Technology **60** (6), 1555-1564,
   DOI:10.2166/wst.2009.520

5. CALIMERO: A model independent tool for autocalibration and uncertainty analysis
   M. Kleidorfer, G. Leonhart, M. Mair, H. Kinzel and W. Rauch (2009)
   Presented at the 8th International Conference on Urban Drainage Modelling, Tokyo, Japan, September 07. - September 11., 2009
   Published in Proceedings of the 8th International Conference on Urban Drainage Modelling, 2009
   Submitted to Water Practice and Technology

6. Hinweise zur Kalibrierung von hydrologischen Modellen für die Anwendung von ÖWAV Regelblatt 19 / neu (in German)
   M. Kleidorfer, S. Fach and W. Rauch (2008)
   Published in Wiener Mitteilungen Wasser-Abwasser-Gewässer 209, J1 - J24

7. Implementing real time control (RTC) in hydrological sewer models as practised in Linz (in German)
   M. Kleidorfer, S. Fach, M. Möderl and W. Rauch (2007)
   Published in Österreichische Wasser- und Abfallwirtschaft 9-10, 131 - 137,
   DOI:10.1007/s00506-007-0126-x

The following additional publications also resulted from the work during the PhD study. Although they are not part of this thesis, their content is also related to the topics presented here and could be interesting for readers. Copies of the papers may be obtained from journal publishers or conference organisers.

1. Defining Uncertainties in Modelling of Urban Drainage Systems
   A. Deletic, C.B.S. Dotto, D.T. McCarthy, **M. Kleidorfer**, G. Freni, G. Mannina, M. Uhl, T.D. Fletcher, W. Rauch, J.L. Bertrand-Krajewski, S. Tait (2009)
   Presented at the 8th International Conference on Urban Drainage Modelling, Tokyo, Japan, September 07. - September 11. 2009
   Published in Proceedings of the 8th International Conference on Urban Drainage Modelling, 2009
   Submitted to Water Science and Technology

2. Stormwater quality models: performance and sensitivity analysis
   C.B.S. Dotto, **M. Kleidorfer**, A. Deletic, T. D. Fletcher, D. T. McCarthy and W. Rauch (2009)
   Presented at the 8th International Conference on Urban Drainage Modelling, Tokyo, Japan, September 07. - September 11. 2009
   Published in Proceedings of the 8th International Conference on Urban Drainage Modelling, 2009
   Submitted to Water Science and Technology

3. Analysis of Sewer System Performance under Environmental change conditions
   M. Kleidorfer, S. De Toffol, M. Möderl and W. Rauch (2009)
   Proceedings of the 8th International Workshop on Precipitation in Urban Areas, St. Moritz, Switzerland, 2009

4. Kalibrierung und Bewertung von Simulationsergebnissen (in German)
   M. Kleidorfer and W. Rauch (2009)
   Published in Schriftenreihe zur Wasserwirtschaft. Optimierte Bemessung von Mischwasserentlastungsanlagen - Erfahrungen mit der Anwendung des neuen ÖWAV-Regelblattes 19. Graz: Verlag der Technischen Universität Graz (55), D1 - D32.

5. Identifying weak points of urban drainage systems by means of VulNetUD
   M. Möderl, **M. Kleidorfer**, R. Sitzenfrei and W. Rauch (2009)
   Presented at the 8th International Conference on Urban Drainage Modelling, Tokyo, Japan, September 07. - September 11. 2009
   Published in Proceedings of the 8th International Conference on Urban Drainage Modelling, 2009
   Submitted to Water Science and Technology

6. Abgleich von hydrologischem und hydrodynamischem Modell zur Verringerung der Unsicherheiten bei begrenzter Datengrundlage am Beispiel von Linz (in German)
   S. Fach, **M. Kleidorfer**, M. Möderl and W. Rauch (2008)
   Korrespondenz Abwasser 55/7, 756 - 765.
   DOI:10.3242/kae2008.07.002

7. Anforderung an Modelle und deren Kalibrierung (in German)
   S. Fach, **M. Kleidorfer** and W. Rauch (2008)
   innsbruck university press (IUP) (Series Forum Umwelttechnik und Wasserbau 1), 99 - 116

8. Reduktion von Mischwasseremissionen durch Optimierung eines Entwässerungssystems am Beispiel der Stadt Linz (in German)
   M. Möderl, **M. Kleidorfer**, S. Fach and W. Rauch (2007)
   Wiener Mitteilungen Wasser-Abwasser-Gewässer 203, K1 - K22

9. Aspects of calibration of hydrological models for the estimation of CSO performance
   M. Kleidorfer, S.S. Meyer and W. Rauch (2006)
   Presented at the 2nd International IWA Conference on Sewer Operation and Maintenance SOM 06, Vienna, Austria, October 26. - October 28. 2006
   Published in Proceedings of the 2nd International IWA Conference on Sewer Operation and Maintenance SOM 06. London: IWA Publishing, 353 - 360

10. Vergleich hydrodynamischer und hydrologischer Simulationsmodelle bei der Berechnung der Emissionen von Mischwasserbehandlungsanlagen. (in German)
    S. De Toffol, **M. Kleidorfer** and W. Rauch (2006)
    Wiener Mitteilungen Wasser-Abwasser-Gewässer 196, H1 - H20

Chapter 1

Introduction

The history of man is reflected
in the history of sewers.

Victor Hugo, Les Miserables

1.1 An introduction to urban drainage

Although the roots of the idea of draining populated areas can be traced back to the ancient Mesopotamian Empire, modern urban drainage concepts have their beginning in the 18th and 19th century as the large cities in Europe were threatened by epidemic plagues due to insufficient waste and wastewater disposal. Three main intentions caused the building of sewer systems and those three objectives are still the basic principles to be observed today (Butler and Davies, 2004):

- Protection from sanitary risks in urban areas by providing acceptable hygienic circumstances
- Flood protection
- Prevention from receiving water pollution

In addition to that three main objectives design and building of sewer systems of course also has to be sustainable and cost-effective.

Because urban areas and human activity have an impact on the natural water cycle two types of runoffs have to be handled by drainage systems: (1) wastewater – water that has been polluted for domestic or industrial use and (2) stormwater – water from rainfall (or another type of precipitation) that has been fallen on urban areas and is (at least partly) polluted due to its contact with the catchment surface and the human activity in the urban areas.

Two different concepts of sewer systems ensuring the drainage from urban areas are

- separated sewer systems (SSS)
  in which wastewater and stormwater is drained separated in different sewer pipes
- combined sewer systems (CSS)
  in which wastewater and stormwater is drained together in the same sewer pipes

In *separated sewer systems* wastewater is collected and drained to a wastewater treatment plant without any dilution due to storm events. Quantity and quality of the wastewater runoff changes in a diurnal and seasonal variation but is rather constant. Stormwater runoff has a very wide range of variation from zero (during dry period) to high peaks during storm events. Stormwater is usually drained to the receiving water directly without any treatment or after treatment in a stormwater treatment system. Although terminology differs by region the concepts are very similar: best management and practice (BMP) and low impact development (LID) in the United States, water sensitive urban design (WUSUD) in Australia and sustainable urban drainage systems (SUDS) in the United Kingdom. For a detailed description of stormwater treatment systems numerous publications are available (e.g. Blecken et al., 2009a,b; Bratieres et al., 2008; Burkhard et al., 2000; Coustumer et al., 2009; Deletic and Fletcher, 2006; Hatt et al., 2007, 2009; Li et al., 2007; Minton, 2002; Read et al., 2008; Siriwardene et al., 2007).

In Europe *combined sewer systems* are mainly used, in which wastewater and stormwater are drained together in the same sewer pipes to the wastewater treatment plant (WWTP). As the inflow to the WWTP (i.e. the outflow from the sewer system) is limited, such systems contain combined sewer overflows (CSOs) to protect the WWTP and urban areas from flooding. CSOs allow an overflow to a nearby river when the maximum sewer capacity is reached.

Additionally modified SSS (mainly SSS with local parts of CSS) or modified CSS (mainly CSS with local parts of SSS) are possible due to the historical development of the systems. Therein on–site infiltration facilities reduce the amount of stormwater runoff that reaches the sewer system.

A detailed and comprehensive introduction to urban drainage was published for example by Butler and Davies (2004) or Gujer (2002).

1.2 An introduction to urban drainage modelling

A model is a schematic (mathematical) description of physical coherences and it is always a simplification of reality neglecting known and unknown processes. Model results can only be an estimation of real occurrences. Quality and accuracy of that estimation depend on quality of input-data applied, model structure and model parameters, which have to be determined from reality by observation and/or by calibration. Usually different submodels are combined into one model and each submodel represents only a limited range of physical processes. Internally one submodel's results can display the input-data for another submodel (see figure 1.1). For example calculation of runoff from rainfall-data is often independent from calculation from pollutant concentration, but both results are necessary for calculating pollutant loads.

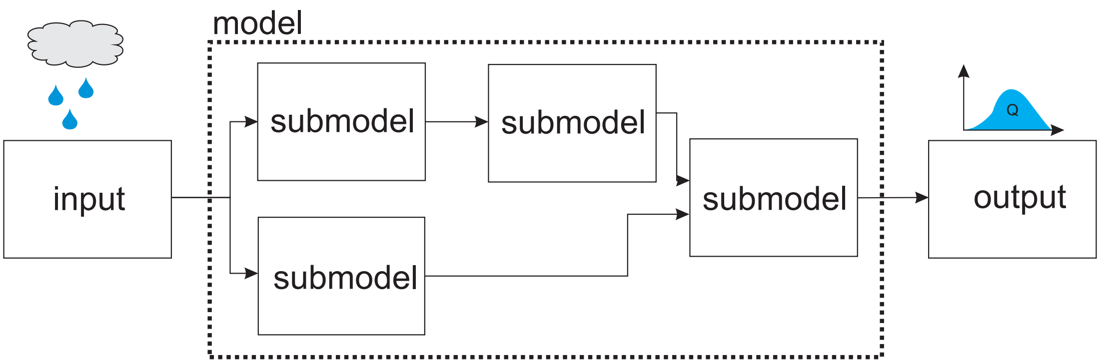

Figure 1.1: Model layout - Schematic representation of model composition

Since the 1970s and 1980s, when the first computer models for simulating urban drainage systems were introduced, the development of more sophisticated approaches for planning and evaluation proceeds steadily (Rauch et al., 2002). Today urban drainage simulation models are state of the art instruments for planners, consultants and scientists working in the field of urban hydrology and numerous commercial, freeware and open-source software products are available. Although rather simple design processes (e.g. the time-area method - developed in the 19th century and described by Kuichling (1889)) are still used in engineering practice they are continuously supplemented or replaced by modelling concepts in order to get more reliable results. More and more simulation models are embodied in modern guidelines and guiding rules (e.g. ÖWAV-RB 11, 2009; ÖWAV-RB 19, 2007).

The applications of urban drainage models are far-ranging from design, optimisation and evaluation of pipe networks, stormwater treatment facilities, combined sewer overflows and real time control strategies. They provide data for other simulation models as wastewater treatment plant simulation or river quality models (or are used in conjunction with such models in integrated urban drainage modelling) and they are used for predicting possible future scenarios (e.g. impact of climate change or urbanisation as shown for example by Arnbjerg-Nielsen (2008); Ashley et al. (2005); Butler et al. (2007); Mark et al. (2008); Semadeni-Davies et al. (2008)).

With increasing complexity of models, data requirements for model building and model calibration rise. Hence, it becomes more and more difficult to analyse model structure with taking into account impact of input and calibration data uncertainties. Today the accuracy of urban drainage model results with regard to model uncertainties is questioned. For example Hoppe and Gruening (2007) estimate that the impact of uncertainties in the input–data already exceeds the effect of measures for optimisation of sewer systems. Silberstein (2006) asks: “Hydrological models are so good, do we still need data?”

Numerous publications show the importance of consideration of different sources of uncertainties in environmental modelling including uncertainties in monitoring (e.g. Bertrand-Krajewski et al., 2003; Kleidorfer et al., 2009b; Overeem et al., 2008), hydrology of natural catchments (e.g. Beven, 2007; Beven and Binley, 1992; Beven and Freer, 2001; Carpenter and Georgakakos, 2004; Engeland et al., 2005; Kavetski et al., 2006a), stormwater quality modelling (e.g. Bertrand-Krajewski et al., 2002; Dotto et al., 2009; Haydon and Deletic, 2009; Kanso et al., 2005; Kleidorfer et al., 2009a; Lindblom et al., 2007), rainfall/runoff modelling (e.g. Lei, 1996; Lei and Schilling, 1996), integrated modelling (e.g. Freni et al., 2009a; Harremoës, 2003; Hoppe and Gruening, 2007; Mannina et al., 2006) and urban drainage modelling (e.g. Arnbjerg-Nielsen and Harremoës, 1996; Deletic et al., 2009; Kleidorfer et al., 2009a; Korving and Clemens, 2005; Rauch et al., 1998b; Thorndahl, 2008; Thorndahl et al., 2008).

To get reliable simulation results an accurate model calibration is indispensable. Therefore calibration parameters of a model have to be estimated by manual or automatic calibration procedures by comparing simulation data to observations. As shown by Gaume et al. (1998) such calibration can lead to different parameter sets, which all reach a similar good fit of measurement data and simulated data. This can happen due to nonlinear criteria functions which may have local minima and result in a failure of the calibration exercise. Hence, certain calibration algorithms may not find the global minimum (Kanso et al., 2003). Other important points when calibrating urban drainage models are to determine the required quantity and quality of calibration data as well as the choice of the objective function which is optimised during calibration. This topic is strongly related to uncertainty analysis.

1.3 Scope and structure of the thesis

1.3.1 Aim and scope

This thesis has two main purposes:
The first objective is the analysis of commonly used urban drainage models from different application areas with respect to their behaviour due to different sources of **uncertainties**. The methodologies presented should assist model users and model developers to gain a deeper insight to model performance, identify relevant calibration parameters, estimate possible model uncertainties and their impact on model results and to appraise a model's simulation results with taking into account model uncertainties.
The second objective is to demonstrate how model **calibration** in practical model applications can reduce model uncertainties. Additionally the question which data is necessary for reasonable model calibration is covered with taking into account impact of duration of measurement campaigns, location of measurement sites in a spatially

distributed model and choice of the objective function. The methodologies and results presented should assist model users to calibrate sewer models and to estimate the impact of data availability on model accuracy.

1.3.2 Structure of the dissertation

The first part of this thesis (chapter 2) starts with some general descriptions of modelling concepts in the field of urban drainage and also deals with data requirements (input-data, model-parameters, calibration-data) for urban drainage models. This is basically a literature review and a description of commonly used models. It provides the necessary background for the following chapters.

In chapter 3 the cases studies used in this thesis and the papers annexed are described. Those are the complete combined sewer systems of the cities Innsbruck and Linz (both Austria, Europe) and selected catchments of the separate sewer systems of Melbourne (Australia). The Australian data, a comprehensive stormwater data-set, was collected by Monash University Melbourne and described by Francey et al. (in press).

Chapter 4 introduces uncertainties in urban drainage models beginning with a definition of uncertainties (section 4.1) followed by several commonly used methods for parameter sensitivity and uncertainty analysis and propagation (section 4.3), which are used in the papers annexed.

Chapter 5 presents some applications and limitations of urban drainage models when working on practical projects with limited data availability and covers model calibration as part of the model building process. In Austria simulation models are defined as state–of–the–art method in the relevant guiding rules for proving flood protection (ÖWAV-RB 11, 2009) and receiving water quality protection (ÖWAV-RB 19, 2007) only since a few years. Thus the use of simulation models is still not very common and needs some assistance. The case studies presented demonstrate how urban drainage models are used to evaluate that requirements and show different steps in the modelling process (model choice, model building, model calibration). The papers annexed relating to this part of the thesis (**Paper VI** and **Paper VII**) are published in national Austrian and German journals in German to reach engineers working with sewer systems. The contents of the German article is summarised in English in this chapter.

1.3.3 Contents of the papers of this dissertation

The papers annexed are integral part of this thesis and cover different aspects of uncertainties in urban drainage models.

Paper I (Kleidorfer et al., 2009b) deals with uncertainties due to data availability for calibration of a conceptual sewer model. In addition to all sources of uncertainties in data collection due to the measurement methods itself, it is a key question which data has to be collected to calibrate a hydrological model, how long measurement campaigns should last and where that data has to be collected in a spatial distributed system as it is neither possible nor sensible to measure the complete system characteristics. In this paper we address this question by means of stochastic modelling. Using Monte Carlo Simulation different calibration strategies (selection of measurement sites, selection of rainfall-events) and different calibration parameters (overflow volume, number of overflows) are tested, in order to evaluate the influence on predicting the total overflow volume of the entire system. This methodology is applied in a case study with the aim to calculate the combined sewer overflow (CSO) efficiency. It can be shown that a distributed hydrological model can be calibrated sufficiently when calibration is done on 30 % of all existing CSOs based on long-term observation. Event based calibration is limited possible to a limited extend when calibration events are selected carefully as wrong selection of calibration events can result in a complete failure of the calibration exercise.

Paper II (Kleidorfer et al., 2009a) deals with impact of input–data uncertainties on urban stormwater models and with sensitivity of model parameters to input–data. It has been recognised that often more than one set of calibration parameters can achieve similar model accuracy. A probability distribution of model parameters should therefore be constructed to examine the model's sensitivity to its parameters. With increasing complexity of models, it also becomes important to analyse the model parameter sensitivity while taking into account uncertainties in input and calibration data. In this study a Bayesian approach was used to develop a framework for quantification of impacts of uncertainties in the model inputs on the parameters of a simple integrated stormwater model for calculating runoff, total suspended solids and total nitrogen loads. The framework was applied to two catchments in Australia. It was found that only systematic rainfall errors have significant impact on flow model parameters. The most sensitive flow parameter was the effective impervious area, which can be calibrated to completely compensate for the input data uncertainties. The pollution model parameters were influenced by both systematic and random rainfall errors. Additionally an impact of circumstances (e.g. catchment type, data availability) has been recognised.

Paper III (Dotto et al., submitted) presents a Bayesian approach for performance evaluation of stormwater models based on long-term high-resolution measurement data. Therein two rainfall / runoff models (MUSIC and KAREN) and two stormwater quality models (Regression model and Buildup-washoff model) with different level of complexity (number of processes covered, number of parameters involved in the simulation process) are compared. The models are tested with a comprehensive quantitative and qualitative dataset collected in urban stormwater systems at five sites of different land-use in Melbourne, Australia. Sensitivity analysis was carried out using a Bayesian approach. Essentially this investigation aimed to answer: (a) How do these models reproduce the measured data, i.e. the catchments' responses? (b) How sensitive are these models to their parameters? (c) What is the impact of catchment characteristics (land-use) on parameter sensitivity? By answering these questions this study provides insights on model choice, parameter significance and correlation. The rainfall/runoff models tested performed very similar suggesting that a simple model may be used for urban catchments without compromising the results. The effective impervious fraction is the most important parameter in both models and special attention should be paid to its value. Even with the robust calibration and parameter sensitivity approach used here, the water quality models tested, poorly represent reality and result in a high level of uncertainty. The study developed was very important to verify the efficiency of the calibration and sensitivity analysis approach. The method presented seems to be promising in terms of generating the posterior parameter distributions and also gives some valuable information on parameter interaction.

Paper IV (Kleidorfer et al., 2009c) deals with "long-term prediction uncertainties" which occur when trying to predict long-term environmental change effects (e.g. land-use, climate, population). Design and construction of urban drainage systems has to be done in a predictive way, as the average lifespan of such investments is several decades. The design engineer has to predict many influencing factors and scenarios for future development of a system (e.g. change in land use, population, water consumption and infiltration measures). Furthermore, climate change can cause increased rain intensities which leads to an additional impact on drainage systems. In this paper the behaviour of different performance indicators of combined sewer systems is compared when taking into account long-term environmental change effects (change in rainfall characteristics, change in impervious area and change in dry weather flow). By using 250 virtual case studies this approach is - in principle - a Monte Carlo Simulation in which not only parameter values are varied but the entire system structure and layout is changed in each run. Hence, results are more general and case-independent. For example the consideration of an increase of rainfall intensities by 20% has the same effect on the investigated performance criteria as an increase of impervious area of +40%. On the other hand such an increase of rainfall intensities could be compensated by a reduction of impervious area by 30% (e.g. by infiltration measures).

Paper V (Kleidorfer et al., submitted) presents the tool "CALIMERO" for generalised autocalibration and uncertainty analysis. The innovation of this tool is (a) the flexibility to work with any model which have input and output files in plain–text formats and can be started from the command line and (b) the possibility to consider a–priori knowledge on system behaviour. The algorithms for evaluating the objective function and the calibration algorithm itself are defined by the user in a scripting environment to provide best possible flexibility. A simple example shows the capabilities of the tool presented to adapt calibration algorithms depending on specific case study characteristics.

Paper VI (Kleidorfer et al., 2008) gives hints for calibration of a conceptual combined sewer model for evaluating the requirements of ÖWAV-RB 19 (2007) for receiving water protection. Therein the emissions from combined sewer overflows have to be evaluated in a long–term simulation of at least 10 years by means of hydrological models. In the paper the legal requirements and aspects of model calibration (e.g. data collection, parameter sensitivity) are discussed. Additionally a structured approach for model building and model calibration is suggested.

Paper VII (Kleidorfer et al., 2007a) deals with model–structure uncertainties exemplified on the implementation of a real time control in a conceptual sewer model for the case study Linz. The guideline ÖWAV-RB 19 (2007) requires the calculation of a ten–year average of combined sewer overflow (CSO) performance. Therefore usually hydrological simulation models are used. In such systems often real time control (RTC) is used to minimise emission of combined sewer overflow. It is common practise to use hydrodynamic models for the design and analysis of such a control. However, analysis of a RTC in a hydrological model needs to be abstracted to a certain degree. The RTC of the sewer system in the city of Linz is discussed to demonstrate how this can be done, and what control algorithms are impossible to reproduce in a hydrological model.

Chapter 2

Modelling Concepts

> I'm a model, you know what I mean.
>
> Right Said Fred

In this chapter some aspects of the concepts used in urban drainage modelling are overviewed. As a full description of all different approaches is not possible in this work, this part concentrates on the models used in the case studies of this thesis and in the papers annexed and provides the background for deeper analysis of model uncertainties and model parameter calibration. A more detailed introduction into urban drainage modelling can be found in literature (e.g. Achleitner, 2008; Akan and Houghtalen, 2003; Gujer, 2008; Rauch and Harremoës, 1998b).

2.1 Rainfall/runoff modelling

Rainfall data is the driving force in all urban drainage simulations. From measured (or in certain applications artificial) rainfall data the catchment runoff is calculated (see figure 2.1). For rainfall/runoff simulation usually runoff–production and runoff–concentration are calculated independently.

2.1.1 Runoff production

Runoff production is the process in which effective rainfall h_e, which contributes to surface runoff, is calculated from measured rainfall h_n. Rainfall intensity is thereby reduced by taking into account different sources of initial and continuing losses:

- initial losses
  - interception h_i
  - depression storage h_d
- continuing losses / permanent losses h_p
  - infiltration
  - evaporation

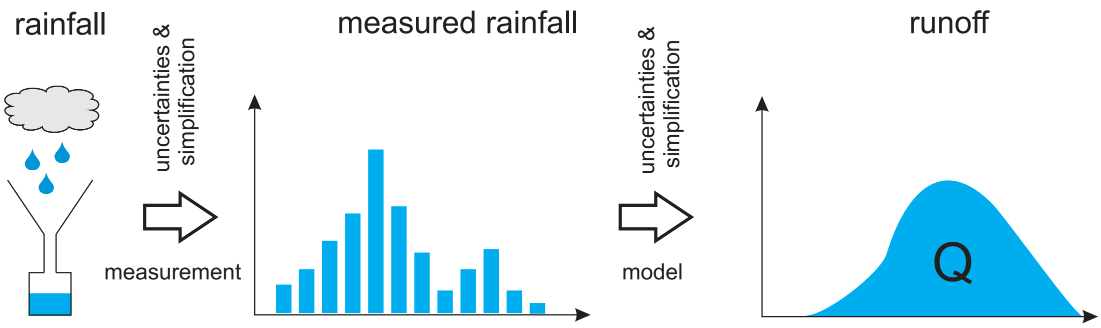

Figure 2.1: Rainfall / runoff modelling - Catchment runoff is calculated form measured rainfall data

Initial losses are model parameters. Although they have a physical background and can be estimated from catchment conditions their "real" value has to be determined during model calibration. Permanent losses can be model parameters or input data depending on model structure and data availability. For example some models require input of evaporation as timeseries similar to rainfall data, but in the majority of cases permanent losses are also represented by model parameters, which have to be determined during model calibration.

Commonly used models for runoff production are the threshold model, the percentage method and the limit value model described by Achleitner (2008). Figure 2.2 exemplifies the consideration of initial and permanent losses in the limited value method for calculation of effective rainfall height.

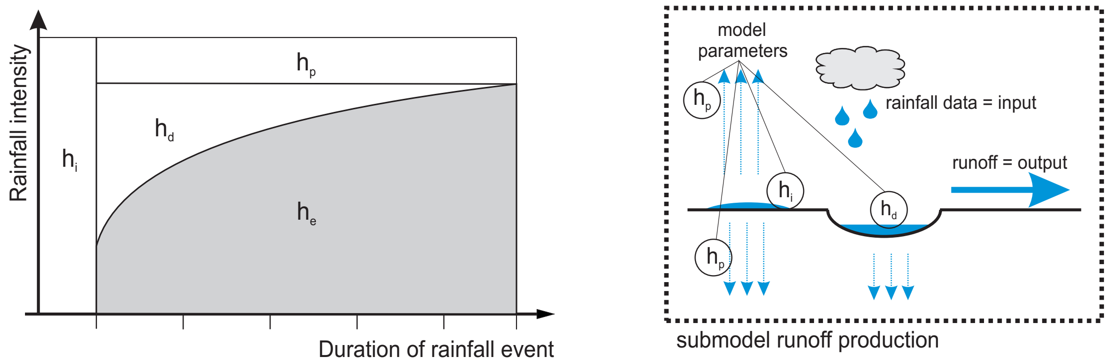

Figure 2.2: Limited value method - Calculation of the effective rainfall height via the limited value method

2.1.2 Runoff concentration

After effective rainfall intensity h_e has been calculated from h_n, surface runoff is calculated by surface routing. Here the aim is to determine the time which the rainwater takes to get to the manholes of the sewer pipes. In general two principle approaches can be distinguished, which are unit hydrograph methods or a kinematic wave model. The unit hydrograph method is based on the idea that a unique and time-invariant hydrograph results from effective rain falling over a particular catchment (Butler and Davies, 2004). Models based on the unit hydrograph method are the time-area method (isochrones method), reservoir models (single linear reservoir or linear reservoir in series) and the Muskingum method. The kinematic wave model is a simplification of the Saint Venant equations which are discussed in section 2.2.1. For a detailed description of the models see Beven (2001), Butler and Davies (2004) or Roberson et al. (1998).

2.2 Water transport / flow modelling

For simulation of flow in the sewer pipes two different concepts are possible which differ in the level of abstraction. Physical-mechanistic (hydrodynamic) models are based on basic physical principles as the continuity equations or the preservation of energy or momentum. Conceptual (hydrological) models describe physical relations and use simple descriptions of cause effect relations (Achleitner, 2008).

2.2.1 The Saint-Venant equations

For modelling of unsteady flow in open channels a pair of equations, the Saint-Venant equations, are commonly used. The Saint-Venant equations are – similar to the Euler equations – a simplification of the general Navier-Stokes equation for one–dimensional flow. The Saint-Venant equations are two equations, the continuity equation 2.1 and the dynamic (momentum) equation. The momentum equation can be formulated in terms of velocity (Eq. 2.2) or flow rate (Eq. 2.3).

$$v\frac{\partial A}{\partial x} + A\frac{\partial v}{\partial x} + b\frac{\partial y}{\partial t} = 0 \tag{2.1}$$

A . . . cross-sectional area
b . . . cross-sectional width
y . . . flow depth
v . . . velocity of flow
x . . . position of the section
t . . . time

$$\underbrace{\frac{\partial v}{\partial t}}_{\text{variation with time}} + \underbrace{g\frac{\partial y}{\partial x} + v\frac{\partial v}{\partial x}}_{\text{variation with distance}} - \underbrace{g(S_o - S_f)}_{\text{uniform steady conditions}} = 0 \tag{2.2}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right) + gA\frac{\partial y}{\partial x} - gA(S_o - S_f) = 0 \tag{2.3}$$

y . . . flow depth
v . . . flow velocity
Q . . . flow rate
x . . . distance
t . . . time
S_o . . . bed slope
S_f . . . friction slope
g . . . acceleration of gravity

The Saint-Venant equations are only valid under following conditions respectively when following assumptions are appropriate (Butler and Davies, 2004):

- pressure distribution is hydrostatic
- small bed slope
- uniform velocity distribution at channel cross-section
- prismatic channel

- friction losses estimated for steady flow are valid in unsteady flow
- lateral flow is negligible

Simplifications of the full Saint-Venant equations lead to two approximations. The diffusion wave (Eq. 2.4) neglects the variation with time and can be formulated as

$$\frac{\partial Q}{\partial t} + c\frac{\partial Q}{\partial t} = D\frac{\partial^2 Q}{\partial x^2} \tag{2.4}$$

in which the wave speed c and the diffusion coefficient D are regarded as constants. Additional neglecting of variation of flow rate with distance leads to the equation for kinematic wave (Eq. 2.5).

$$\frac{\partial Q}{\partial t} + c\frac{\partial Q}{\partial t} = 0 \tag{2.5}$$

Table 2.1 summarises which effects are reflected in different simplifications.

Table 2.1: Applications of simplified Saint-Venant equations after Butler and Davies (2004)

Accounts for	Kinematic wave	Diffusion wave	Dynamic wave
Wave translation	Yes	Yes	Yes
Backwater	No	Yes	Yes
Wave attenuation	No	Yes	Yes
Flow acceleration	No	No	Yes

2.3 Water quality modelling

2.3.1 Regression model

In urban drainage models often a simple regression model is used to estimate pollutant washoff from an impervious surface:

$$O = a \cdot Input^b \tag{2.6}$$

Here the model output O can either be pollutant concentrations C_t at a time t or pollutant loads (fluxes) L_t at a time t. Model input I can be measured rainfall data or routed runoff as result from a rainfall / runoff simulation. This model equation contains the parameters a and b that have to be calibrated.

This simple regression model is the most frequently used in practice in several stormwater models, such as XP-AQUALM (XP-Software, 1999), SWMM5 (Rossman, 2007) and P8-UCM (Palmstrom and Walker, 1990). In this thesis a regression model is used in **Paper II** for calculating pollutant loads from simulated flow and in **Paper III** for calculating pollutant concentrations from measured rainfall.

2.3.2 Exponential buildup / washoff function

The generation of pollutants in the runoff from an impervious surface is also very often described and modelled using the concepts of buildup and washoff. Buildup is the process in which pollutants accumulate on the surface during dry weather period. Washoff is the process of removing these accumulated pollution load by rainfall and incorporating it to the surface runoff. The attempt to model these two processes was proposed by Sartor et al. (1974). This algorithm is adopted in several of the stormwater software, such as SWMM (Rossman, 2007). Buildup during dry period is calculated after Sartor et al. (1974) and Deletic et al. (2000) using equation 2.7.

$$\frac{d\overline{M(t)}}{dt_d} = k_1 \cdot (M_0 - \overline{M(t)}) \tag{2.7}$$

Here $\overline{M}$ is the amount of solids available on the surface averaged over the area [$\mathrm{g\,m^{-2}}$], M_0 is the maximum amount of solids expected at the surface [$\mathrm{g\,m^{-2}}$] and k_1 is an accumulation constant [$\mathrm{day^{-1}}$]. Consequently the calibration parameters for buildup are M_0 and k_1.

Washoff during wet weather is calculated directly from rainfall intensity (not runoff) after the exponential function in equation 2.8.

$$\frac{d\overline{C_{t+tf}}}{dt} = k_2 \cdot \overline{M(t)} \cdot I(t)^{k_a} \cdot A_i \tag{2.8}$$

Here C is the washoff concentration [mg/l], $\overline{M}$ is the amount of solids available on the surface averaged over the area according to Eq. 2.7, I is the rainfall intensity [mm], A_i is the impervious area [$\mathrm{m^2}$], k_2 is the washoff coefficient and k_3 is the washoff exponent. The calculated concentration is shifted by the flowtime tf. Consequently the three calibration parameters for washoff simulation are k_2, k_3 and tf. In this thesis an exponential buildup / washoff model is used in **Paper III**.

2.4 Level of complexity

As mentioned urban drainage models are always simplifications of reality. The level of simplification (or level of complexity) expresses how accurate physical processes are described.

The simplest models are input / output relationships of the area of interest. Here model parameters usually don't have a physical background and hence they can only be estimated from model calibration. An example for such a kind of model is the simple regression model for estimating pollutant concentrations (section 2.3.1).

A conceptual model is a mathematical description of principal processes and their relations based on observations and on some degree of understanding of the physical processes. But it represents rather the concepts of one or more processes and not necessarily the detailed physical descriptions. Model parameters usually have a physical background and mostly can be estimated from collected data. The conceptual approaches used for modelling flow in sewer pipes are the same as for surface flow routing. Typical models are the time-area method, the single reservoir method, the linear reservoir in series and the Muskingum method.

The most complex models are based on fundamental equations of the physical processes in a highly detailed level (e.g. water transport modelling based on Saint-Venant equations). Often differential equations have to be solved which can cause intensive computing times. Results of such models are expected to be closest to reality but numerical instabilities can cause additional uncertainties.

The model choice is a key point in the modelling process highly depending on the aim of a specific study, the data availability and the computer resources available. A model has to be as simple as possible but as complex as necessary. A model which neglects important processes (important for a specific model application) might not be able to represent reality in a sufficient way. A model which is too complex might be difficult to calibrate because of the high number of calibration parameters especially when calibration parameters are correlated. Additionally the computing times rise with model complexity. Hence methods for uncertainty estimation which often require Monte-Carlo simulations with numerous simulation runs might become impossible to be performed.

2.5 Distributed urban drainage modelling

In practical applications urban drainage models are spatially distributed over a large area, for example a whole city. Although the modelling concepts as described above are the same as for a single catchment, this impacts calibration and uncertainty analysis because the number of calibration parameters increases. Like different submodels are combined to reach a specific modelling task, submodels in distributed urban drainage modelling are combined to predict behaviour of a whole drainage system. Figure 2.3 and Figure 2.4 show a distributed urban drainage model represented by the Software KAREN (Rauch and Kinzel, 2007) and CityDrain (Achleitner et al., 2007). For example in Figure 2.3 61 blocks are connected by links. Whereas one block has 7 parameters the model of the entire system has up to 427 parameters. Although in practical applications several parameters can kept fixed, one can clearly see that the number of parameters (and with it model–structure) increases significantly.

2.6 Integrated modelling

Integrated modelling is the combination of submodels of different systems (e.g. sewer system, waste water treatment plant, river water quality, groundwater) into one model in order to assess several aspects of the urban water system and interactions of different submodels for an integrated management of the system (Chocat et al., 2001; Harremoës and Rauch, 1996; Rauch et al., 1998a, 2002, 2005). Therefore additional systems have to be described apart from the models described above as for example river water quality models (Reichert et al., 2001; Shanahan et al., 2001; Vanrolleghem et al., 2001) or waste water treatment models (Henze et al., 2006). A schematic representation of

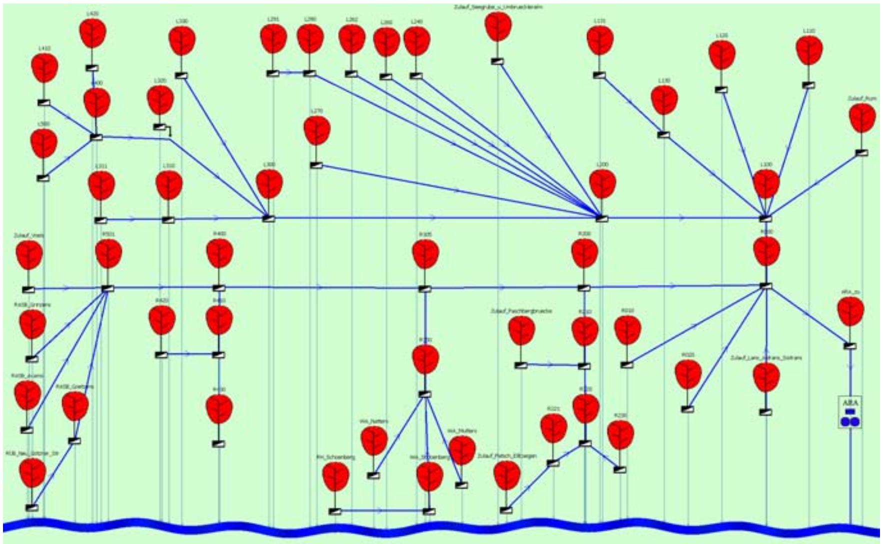

Figure 2.3: Distributed modelling - Sewer system of Innsbruck represented in the software KAREN

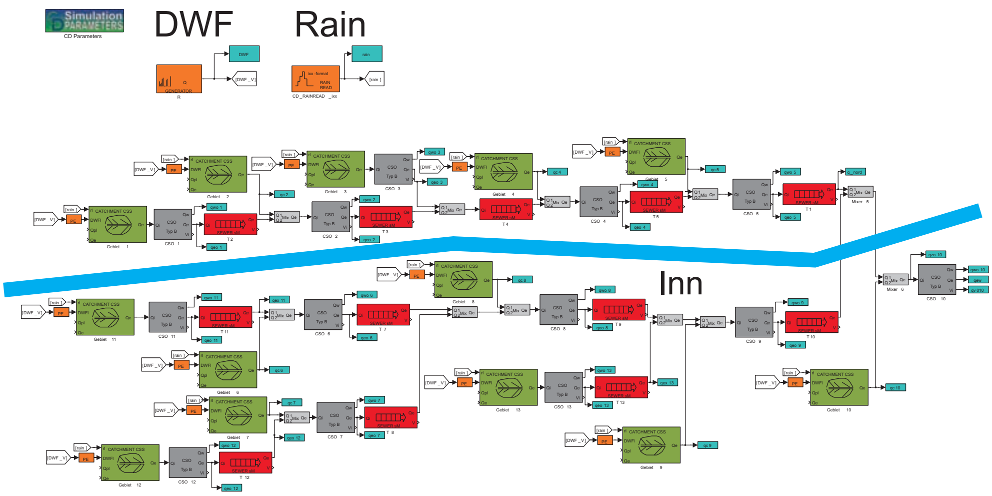

Figure 2.4: Distributed modelling - Sewer system of Innsbruck represented in the software CityDrain

different systems of a urban water system is shown in Figure 2.5.

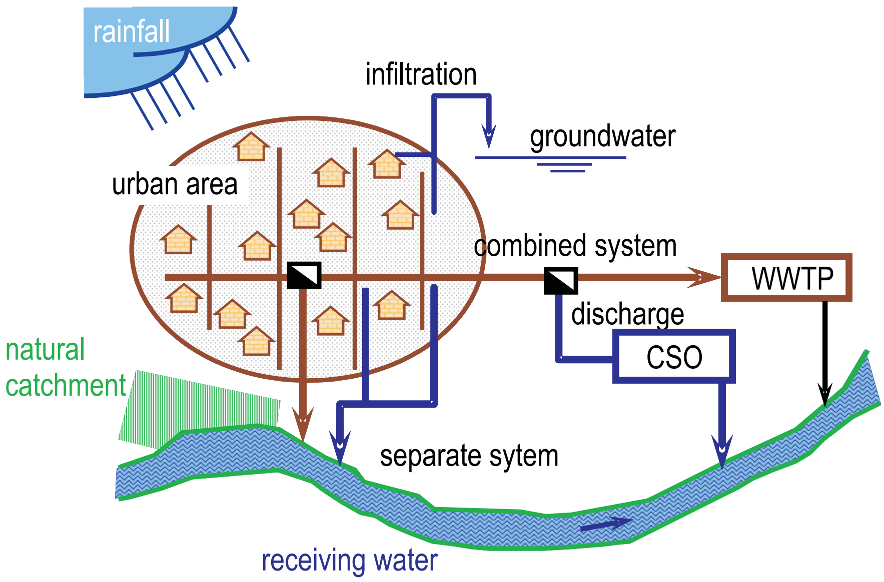

Figure 2.5: Integrated urban drainage modelling - Schematic representation of different systems

Rauch and Harremoës (1998a) evaluated the effect of combined sewer overflow reduction due to real–time control strategies in a combined sewer system on the oxygen concentration in the receiving water. Rauch and Harremoës (1996) analysed the impact of high hydraulic loads of waste water treatment plants on water pollution. De Toffol et al. (2006a) compared the impact of combined and separate sewer systems on ecological and economical performance indicators. Achleitner et al. (2005) and Achleitner and Rauch (2006) analysed how optimised hydro power gate operation can be used to reduce the impact of CSO discharge on receiving water pollution.

Chapter 3

Case Studies

I Reject Your Reality And
Substitute My Own.

Adam Savage

In this chapter the case studies used in the papers annexed are described. A detailed description is available in project reports from Kleidorfer et al. (2006c); Kleidorfer and Rauch (2006, 2007a,b,c,d) for Innsbruck and from Kleidorfer et al. (2007c, 2006b, 2007d); Möderl et al. (2007a,c) for Linz.

3.1 Case Study Innsbruck

Case study Innsbruck was used in **Paper I**, **Paper IV** and **Paper VI**.

3.1.1 Catchment characteristics

Innsbruck is a city located in Tyrol, Austria (see figure 3.1). The climate is alpine, so the region is characterised by cold winters and summers with intense rainfall. Innsbruck is drained in a combined sewer system with only a few very small parts of separate sewer systems. The total area drained is about 2 500 ha whereof 770 ha are impervious. 127 000 PE in Innsbruck and 38 000 PE from the surrounding communities are connected to the WWTP. The total network length (without private connection pipes) is approximately 240 km and contains about 5 500 manholes. An overview of the catchment characteristic is shown in table 3.1. The sewer system of Innsbruck as well as the setup of a hydrodynamic model was described in detail by Kleidorfer (2005) and summarised by Fach et al. (2007).

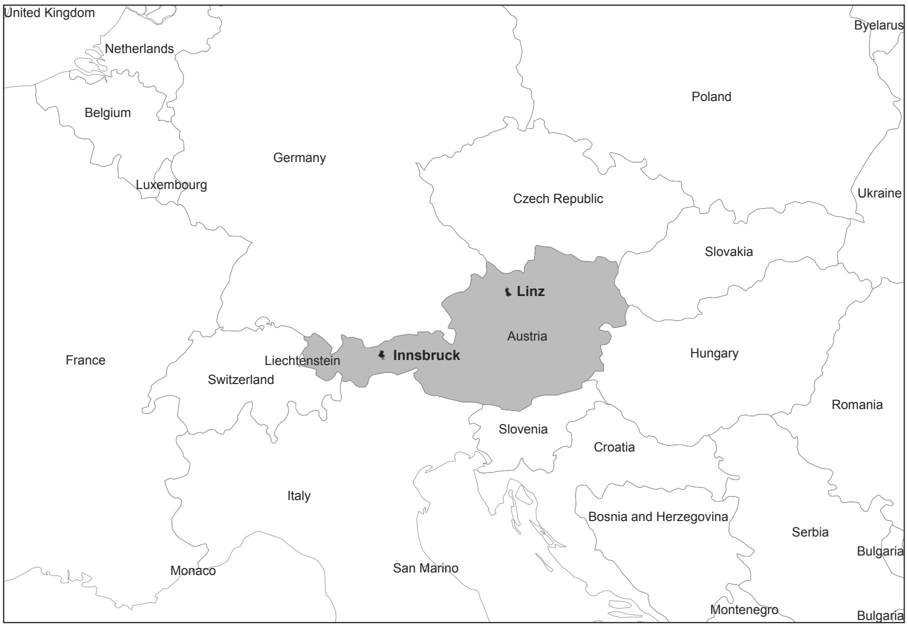

Figure 3.1: Case study description - Location of Innsbruck and Linz

Table 3.1: Catchment and sewer system characteristics of Innsbruck

total area [ha]	2 500
impervious area [ha]	774
inhabitants	127 000 + 38 000 [1]
storage volume [m^3]	5 100
manholes	5 500
network length [km]	240
number of CSOs	46

[1] surrounding communities

3.1.2 Data collection and evaluation

The available data for Innsbruck is:

- catchment area and total fraction imperviousness
- detailed sewer system data (pipes, basins, pumps, ...)
- rainfall data measurements at 4 rain gauges
- continuous water level measurements at 19 measurement sites
- flow measurements
- online total suspended solids (TSS) measurements

The data was provided by the operator of the sewer system "Innsbrucker Kommunalbetriebe" (IKB) or collected in cooperation with IKB. As since 2005 the monitoring network is continuously expanded for different studies and applications, different datasets were available. The location of the measurement sites for waterlevel and rainfall measurements as well as the sewer system of Innsbruck is shown in Figure 3.2.

3.1.2.1 Rainfall data

In Innsbruck rainfall data measurements from four rain gauges are available in sufficient temporal resolution of 15 minutes or higher.

Table 3.2: Rainfall measurements in Innsbruck

Starting from	Location
1987	University
1992	Airport
2006	WWTP
2006	Mühlau

Figure 3.2: Measurement sites Innsbruck - Location of rainfall and waterlevel measurements in Innsbruck

3.1.2.2 Flow measurements

Flow measurements with a temporal resolution of two minutes were arranged by IKB in measurement campaigns at four different measurement sites lasting two to four month (see Table 3.3).

Table 3.3: Flow measurements in the sewer system Innsbruck

Starting from	To	Location
09.06.2005	09.08.2005	Hoher Weg
07.12.2005	11.04.2006	Hallerstrasse
10.08.2005	10.10.2005	Mitterweg
04.10.2005	07.12.2005	Archenweg

From those flow measurements the dry weather flow can be evaluated. For that, only measurements at those days without rainfall during the last two days are evaluated. The evaluation of dry weather hydrographs for the measurement site "Archenweg" (comprising the northern part of Innsbruck which contains about 57 000 inhabitants) is shown in Figure 3.3 for weekdays and in Figure 3.4 for weekends. Therein the black line indicates the mean hydrograph.

Consequently the mean dry weather flow for Innsbruck can be calculated to 194 l/(inhabitant · day) on weekends and to 200 l/(inhabitant · day) on weekdays. By dividing the hydrograph by that mean value a normalised hydrograph can be calculated which is shown in Figure 3.5 for weekdays and in Figure 3.6 for weekends. The daily minimum is between 5 and 6 a.m. and the peak is at noon.

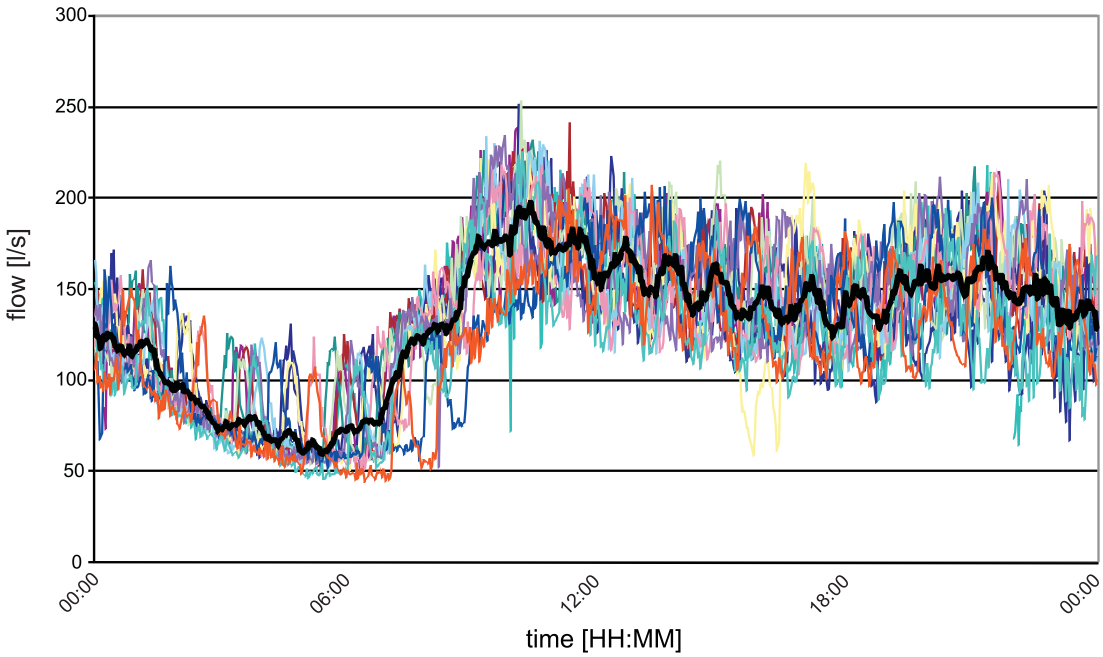

Figure 3.3: Dry weather flow Innsbruck on weekdays - Evaluation of flow measurements for determining dry-weather flow hydrograph for weekdays

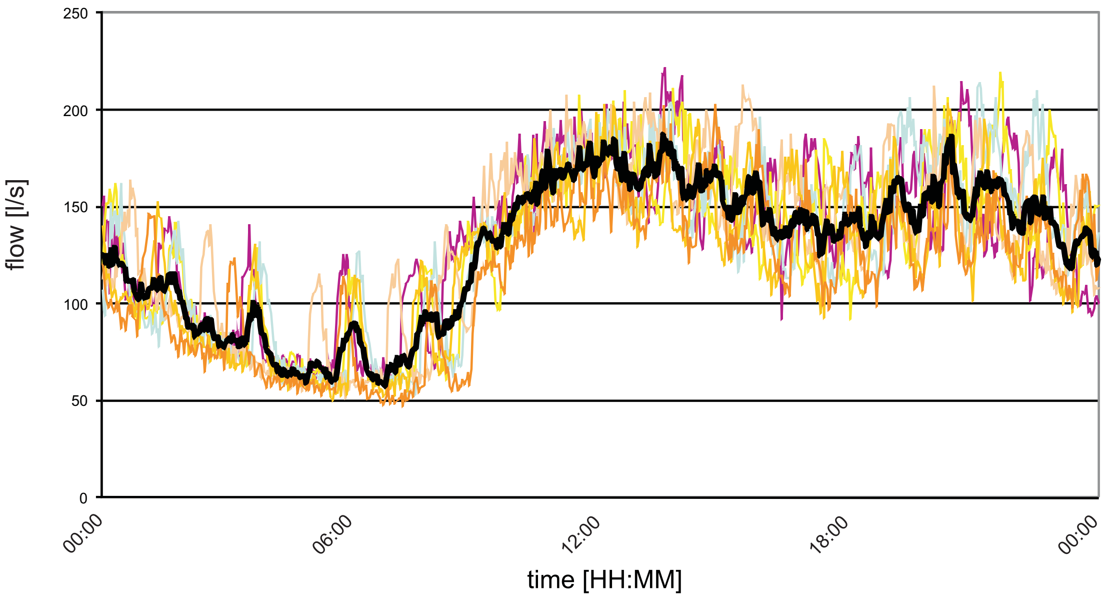

Figure 3.4: Dry weather flow Innsbruck on weekends - Evaluation of flow measurements for determining dry-weather flow hydrograph for weekends

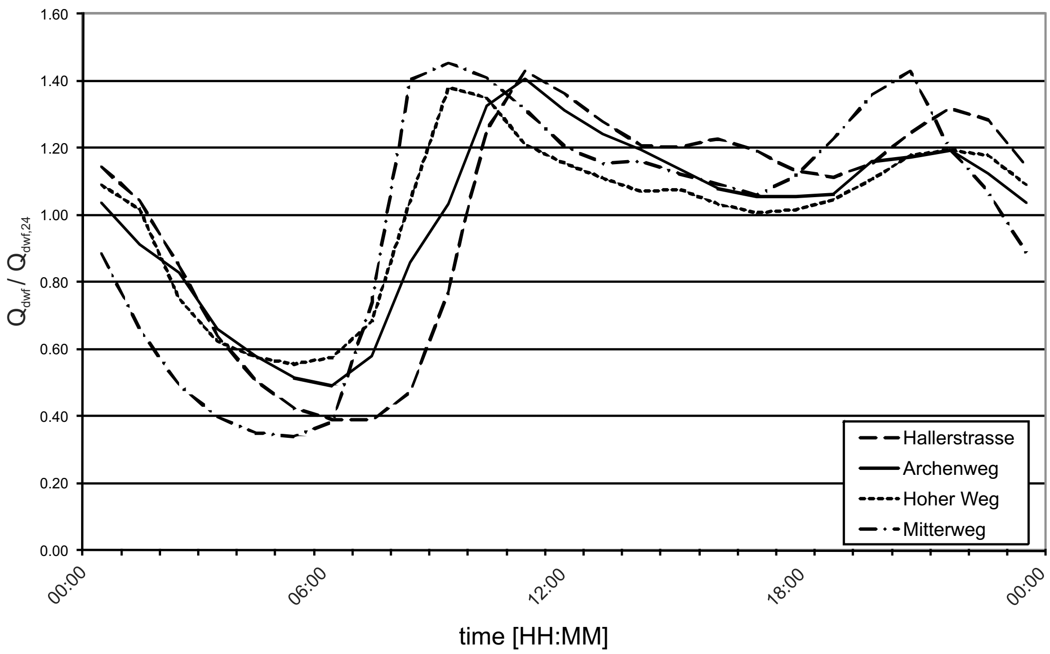

Figure 3.5: Normalised dry weather flow Innsbruck on weekdays - Evaluation of flow measurements for determining dry-weather flow hydrograph for weekdays

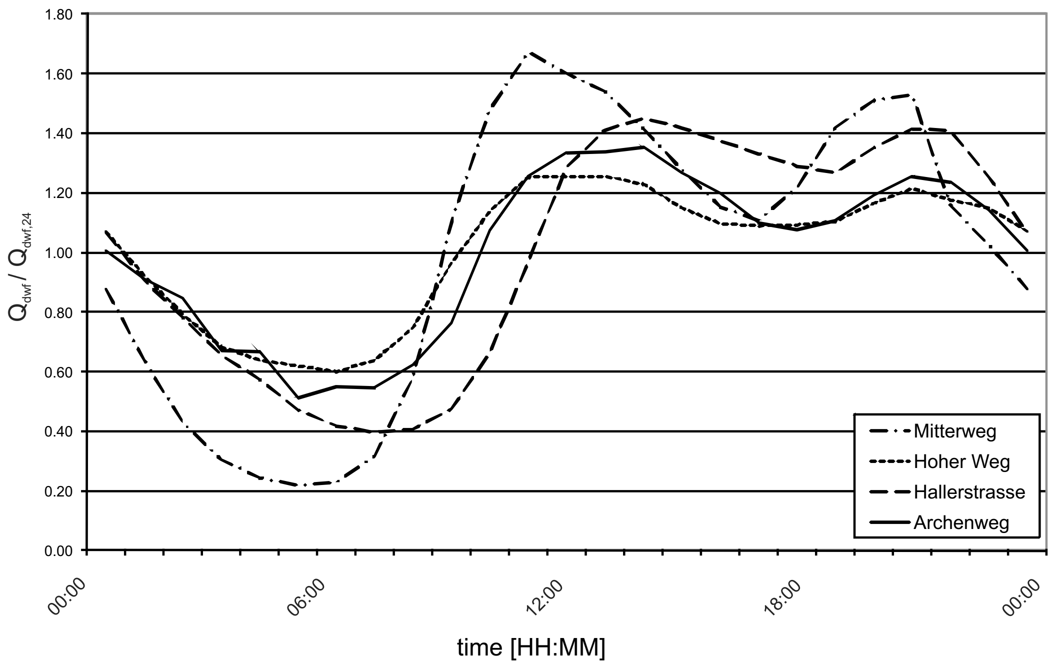

Figure 3.6: Normalised dry weather flow Innsbruck on weekends - Evaluation of flow measurements for determining dry-weather flow hydrograph for weekends

3.1.2.3 Water level measurements

Water level measurements by ultrasonic devices (see Figure 3.7) are available at 19 measurement sites (as of August 2009). Due to a successive extension of the monitoring network since the year 2006, the beginning of the data collection at the different sites varies. An overview of the water level measurements available is presented in Table 3.4. Those water level measurements are used for the calibration of hydrological and

Table 3.4: Water level measurements in the sewer system Innsbruck

Location	Starting from
Mariahilferstrasse	09.07.2006
Kugelfangweg Pumpwerk	09.07.2006
Vögelebichl	09.07.2006
Fuchsrain West	09.07.2006
Fuchsrain Ost	09.07.2006
Herrengasse	01.01.2006
Otto-Winter-Str. Süd	09.07.2006
Otto-Winter-Str. Nord	09.07.2006
Hallerstrasse-Feuerwache	09.07.2006
Innrain-Holzhammerbrücke	01.01.2006
Andreas-Hofer-Strasse / Schöpfstrasse	09.07.2006
Reichenauerstrasse - Rossbachstrasse	01.01.2007
Valiergasse RAGG	01.01.2007
Bürgerstrasse	01.04.2006
Gaswerk	01.07.2006
Sanatorium	01.01.2007
Dr. Sigismund Epp Weg	09.07.2006
Inndüker	09.07.2006
Vill - Iglerstrasse	29.05.2007

hydrodynamic urban drainage models, respectively for estimating CSO discharge from water level measurements (Fach et al., 2008b).

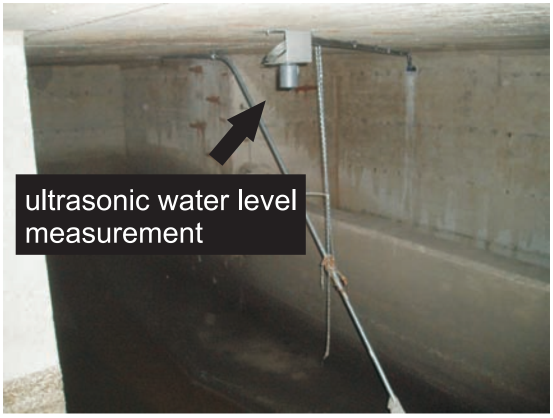

Figure 3.7: Water level measurement in Innsbruck - Water level measurement in Innsbruck via ultrasonic device

3.1.2.4 TSS concentration

TSS concentrations were measured using an UV-VIS spectrometer at the inflow to the WWTP (see figure 3.8) over a period of 8 month. The functionality of this type of measurement device is described for example by Hochedlinger (2005) or Gruber et al. (2006). The UV-VIS probe used for online measurements records light attenuation in the wavelength region between 200nm and 750nm. For TSS measurements the region between 350nm and 750nm is relevant. After a calibration (based on TSS analysis in the laboratory) TSS concentrations are calculated from light attenuation.

A comparison between results from TSS analysis in a laboratory and results from UV-VIS spectrometry is shown in Table 3.5, in Figure 3.9 for a dry weather day and in Figure 3.10 for a wet weather day.

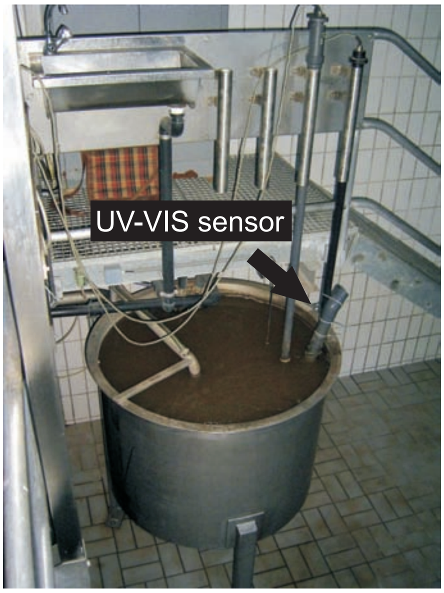

Figure 3.8: UV-VIS Sensor installation in Innsbruck - The UV-VIS sensor is installed to measure inflow to the WWTP

Table 3.5: Comparison of TSS measurements from laboratory and UV-VIS sensor

Date	TSS [mg/l] laboratory	TSS [mg/l] UV-VIS sensor
04.07.2006 09:00	267	276
04.07.2006 11:26	368	379
04.07.2006 13:11	271	324
04.07.2006 15:13	254	285
04.07.2006 16:17	292	293
25.07.2006 10:00	311	222
25.07.2006 11:48	279	236
25.07.2006 13:06	333	241

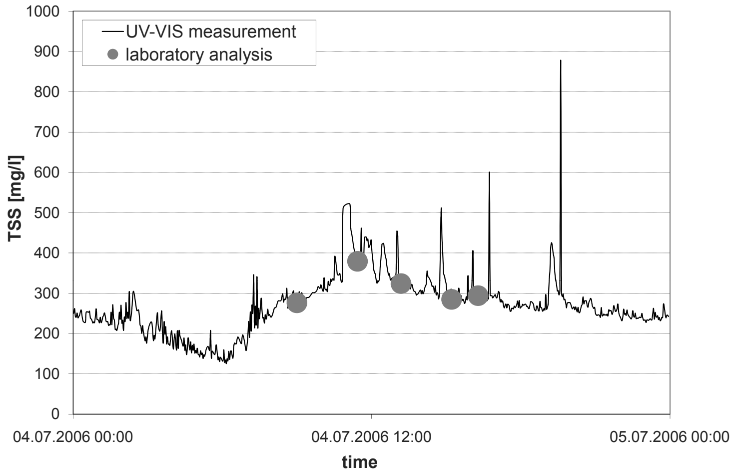

Figure 3.9: UV-VIS TSS measurement dry weather - Comparison between TSS analysis in laboratory and UV-VIS measurement

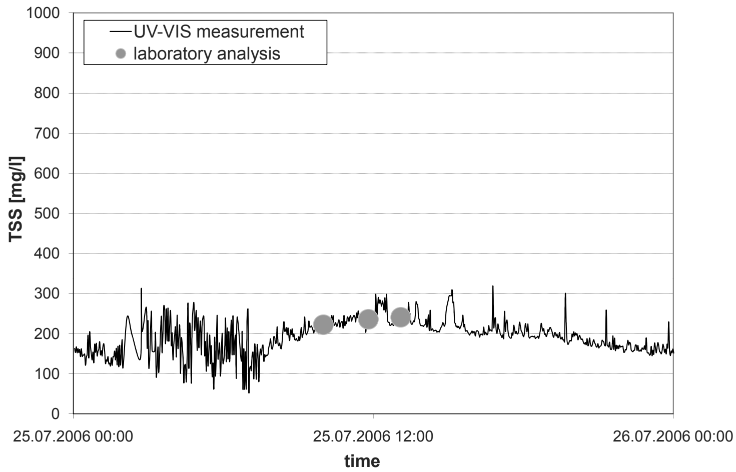

Figure 3.10: UV-VIS TSS measurement wet weather - Comparison between TSS analysis in laboratory and UV-VIS measurement

3.2 Case Study Linz

Case study Linz was used in **Paper VII**.

3.2.1 Catchment characteristics

The sewer system of Linz is basically combined with some parts of separate sewer systems and drains the city and 39 surrounding municipalities to the WWTP. The total catchment area drained is about 900 km^2 whereof about 13 000 ha are connected to the sewer system and 3 600 ha are impervious. Linz has about 190 000 inhabitants. Including 160 000 inhabitants from surrounding municipalities and several industrial plants consequently 350 000 PE are connected to the WWTP. The total network length is about 560 km in Linz and 280 km in surrounding municipalities. CSOs are discharged into the two rivers Donau and Traun. The main characteristics of the sewer system are shown in table 3.6.

Table 3.6: Catchment and sewer system characteristics of Linz

total area [ha]	13 000
impervious area [ha]	3 600
inhabitants	190 000 + 160 000 [1]
basin volume [m^3]	90 000
manholes	21 000
network length [km]	840

[1] surrounding communities

3.2.2 Data collection and evaluation

The available data for Linz is:

- catchment area and total fraction imperviousness
- detailed sewer system data (pipes, basins, pumps, ...)
- rainfall data measurements at 8 rain gauges
- continuous water level measurements at 5 measurement sites
- flow measurements during measurement campaigns

The data was provided by the operator of the sewer system “Linz AG” or collected in cooperation with Linz AG.

3.2.2.1 Rainfall data

In Linz and its surrounding area rainfall data measurements from eight rain gauges (see table 3.7) are available in a sufficient temporal resolution of 15 minutes or higher. Their location is shown in Figure 3.11.

Table 3.7: Rainfall measurements in Linz

Starting from	To	Location
1993	2005	Linz Stadt
1997	2004	Linz Urfahr
2000	2004	Goldwörth
1978	2004	Wels
1998	2005	Asten
04/2005	11/2005	Weikerlsee
04/2005	11/2005	Plesching
04/2005	11/2005	Lunzerstrasse

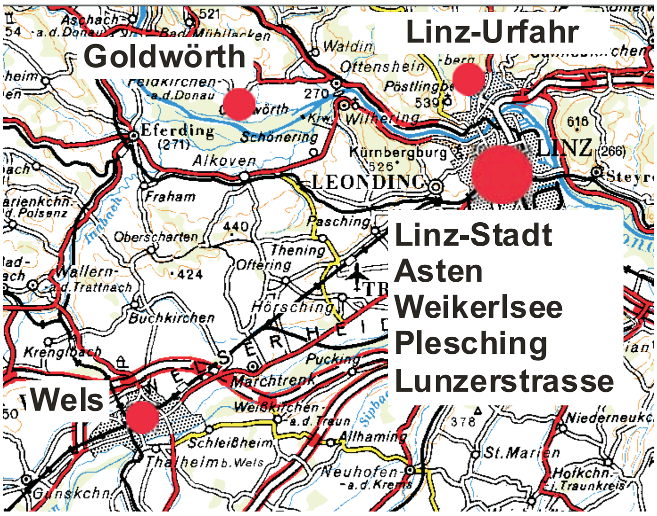

Figure 3.11: Rain gauges Linz - Location of available rain gauges in Linz and surrounding area

3.2.2.2 Water level measurements

Water level measurements are available at five measurement sites with a resolution of one minute. An overview of the water level measurements available is presented in table 3.8.

Table 3.8: Water level measurements in the sewer system Linz

Location	Starting from	To
Plesching	01/2004	04/2006
Hauptsammler Mitte	01/2004	04/2006
Füchselbachkanal	01/2004	04/2006
Lunzerstrasse	01/2004	04/2006
Weikerlsee	01/2004	04/2006

3.2.2.3 Flow measurements

Runoff measurements in Linz are available as measurement of the influent to WWTP from 07/2003 to 06/2006 in a temporal resolution of 15 minutes and at the interconnection points between surrounding municipalities and Linz from 02/2003 to 03/2003 in a temporal resolution of one minute. Thereof an average dry weather flow of about 1 600 l/s, a peak dry weather flow of about 2 400 l/s and a minimum dry weather flow of about 1 000 l/s were calculated.

3.3 Melbourne catchments

In **Paper II**, **Paper III** and **Paper V** a comprehensive stormwater dataset from catchments from Melbourne, Australia, was used for analysis. That data was collected and provided by Monash University, Melbourne and is described by Francey et al. (in press).

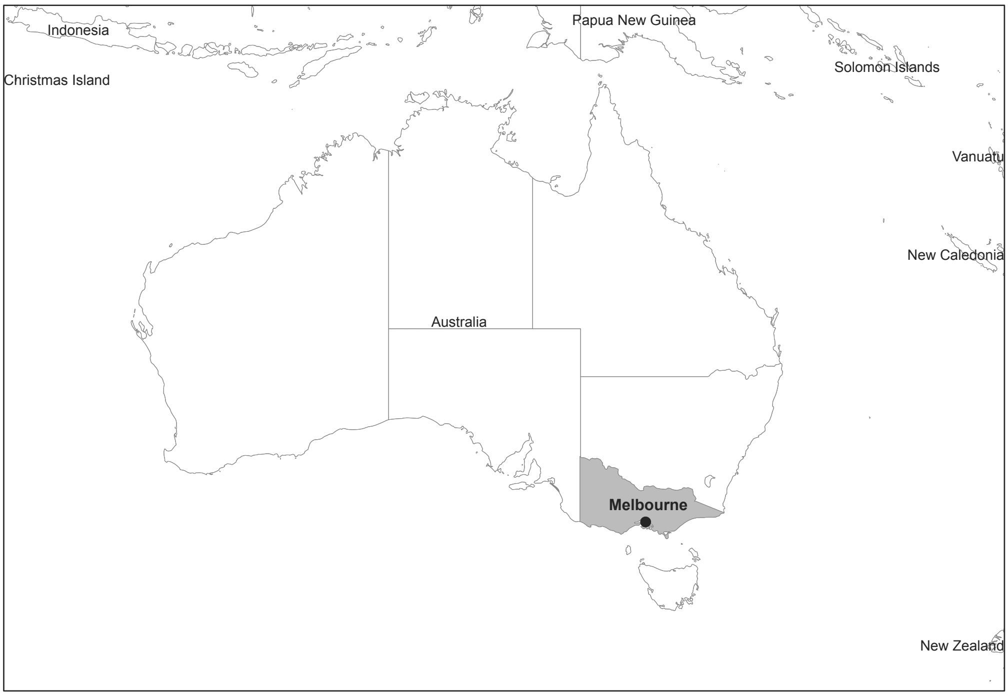

Figure 3.12: Melbourne - Map of Australia and Melbourne

3.3.1 Catchment characteristics

The dataset contains data on stormwater flows and pollution concentrations from 5 urban catchments of different land uses and sizes located in the eastern and south-eastern suburbs of Melbourne, Australia. The total fraction imperviousness TFI of the sites range from 0.2 to 0.8 and catchment areas ranged from just 10 to over 100ha. A summary of the catchments is presented in Table 3.9.

Table 3.9: Summary of catchment details Melbourne

Catchment	**Primary Land Use**	**Area [ha]**	**TFI**	**Distance to rain gauge [m]**
Gilby Rd, Mt. Waverly	Commercial	28.2	0.8	100
Madden Grove, Richmond	High Density Residential	89.1	0.74	600
Ruffeys Lake, Doncaster	Medium Density Residential	105.6	0.51	700
Shepards Bush, Glen Waverley	Medium Density Residential	38	0.45	550
Kilgerro Crt, Narre Warren Sth	Rural Residential	10.5	0.2	250

3.3.2 Data availability

Between 30 and 50 pollutographs for TSS and total nitrogen (TN) are available for each site and each event contains between 5 and 30 discrete samples. Rainfall data was monitored using rain gauges and 0.2mm tips were logged using one minute timesteps. The mean annual rainfall in these catchments ranges from 600 to over 800 millimetres per year. All catchments are serviced by separate stormwater and wastewater systems, but some cross-connections between systems are expected. Narre Warren is the only site in which on–site septic systems are adopted for sewage treatment.

Due to different catchment characteristics, different characteristics in runoff and pollution are to be expected as compared to the European (Austrian) case studies.

3.4 Didactic Example

To demonstrate the applicability of the methods presented in chapter 4 additional to the case studies presented above a simple example is used. This example is based on the regression model (see section 2.3.1) for estimating pollutant concentrations in which concentrations C are calculated from surface runoff Q (equation 3.1).

$$C = W \cdot Q^b \tag{3.1}$$

Equation 3.1 contains two calibration parameters W and b. To reduce computing time for this didactic example only one single event with a duration of 30 hours is used. The simulation time step is 6 minutes.

The runoff of that single event used as input was taken from the Richmond data set (see section 3.3). To avoid problems of real data uncertainties no measured pollutant concentrations are used in the analysis, but "real" data was generated by running the model with model parameters $W = 50$ and $b = 0.5$. This has the advantage that the best set of calibration parameters is known and one can clearly see how the methods for uncertainty analysis work. Figure 3.13 shows input data and synthetic calibration data used.

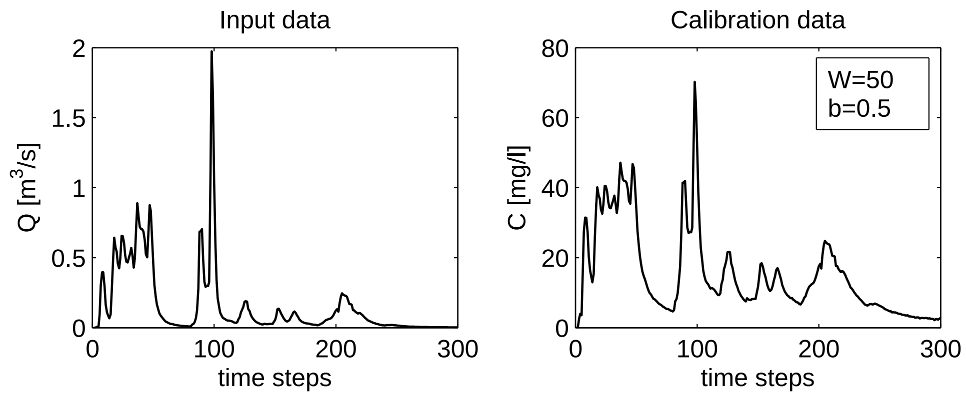

Figure 3.13: Didactic example: Input data (left) and calibration data (right) - Input data and synthetic calibration data created by running the regression water quality model with calibration parameters $W = 50$ and $b = 0.5$

Additionally artifical measurement uncertainties were applied on "real" data by sampling factors ϵ_i randomly from a uniform distibution U[0.8|1.2] and by multiplying that factors with the "real" data timeseries for each timestep i respectively. Consequently the timeseries used as measured data C_m is calculated from "real"data C_r after

$$C_m = \epsilon_i \cdot C_r. \tag{3.2}$$

Figure 3.14 shows the histogram of the residuals of "real" data and calibration data (i.e. the histogram of the deviation of the two timeseries) on the left hand side and a plot of these two timeseries on the right hand side. As one can clearly see the residuals are approximately Gaussian distributed with a maximum deviation of 10 mg/l.

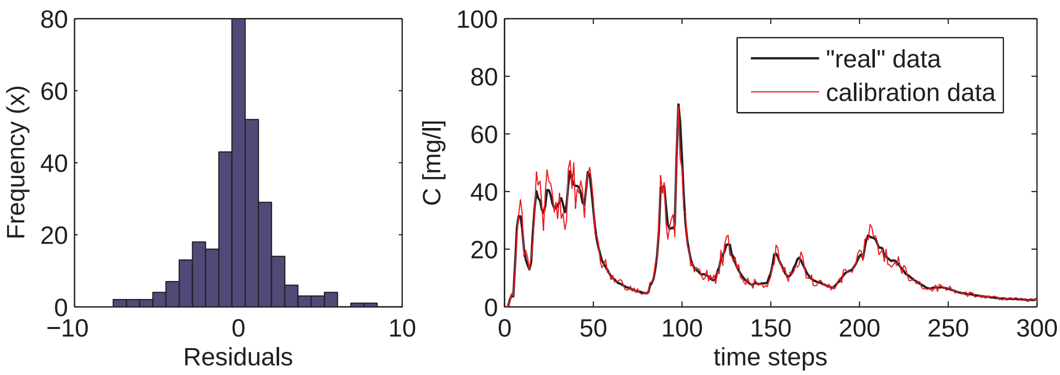

Figure 3.14: Didactic example: Comparison between real data and calibration data) - Comparison between generated "real" data and generated calibration as histogram of residuals (left) and timeseries (right)

Here no further analysis is undertaken; this data is used to demonstrate different methods for uncertainty analysis in the next chapter.

Chapter 4

Uncertainties in urban drainage modelling

> If you can not measure it, you can not improve it.
>
> Lord Kelvin

A model can never perfectly represent reality due to different sources of uncertainties in the modelling process. Hence model output is always uncertain and the impact of uncertainties on model results has to be estimated and interpreted in order to use such results in a reliable way for estimating behaviour of a real system. Much work has been done on analysis of uncertainties in environmental modelling (e.g. Beck, 1987; Beven, 2009; Thorndahl, 2008), but most studies mainly focus on uncertainties of modelling of large natural watersheds or concentrate on a specific source of uncertainty. Uncertainties in urban drainage modelling attracted increasing attention of scientist only in the last years and are usually still neglected in non-scientific practical projects.

This chapter deals with different sources of uncertainties in urban drainage modelling (section 4.1) and describes scientific methods of uncertainty analysis and propagation (section 4.3).

4.1 A general definition of uncertainties in urban drainage modelling

Often uncertainties are classified into two main groups (1) input-data uncertainties and (2) model structure uncertainties.

Here input data uncertainty is often understood as the problem that data collection is never accurate due to uncertainties of the measurement device and consists of random and systematic uncertainties. This type of uncertainties is described in the "ISO Guide to the Expression of Uncertainty in Measurement" (ISO, 2008) in which measurements are assumed to be Gaussian distributed. This Gaussian distribution is described by the mean value μ (the "true" measurement value) and the standard deviation σ. Furthermore following this idea of input-data uncertainties a strict definition of data-uncertainties assumes that data contains uncertainties but not errors. Hence data errors have to identified and removed prior to modelling.

While such an approach is applicable in rather well defined and simple systems with known boundary conditions, uncertainties in urban drainage models are mainly driven by questions of data availability. Additionally data rarely is "wrong" in the sense that an error can be eliminated during data processing, but often data is not representative for a specific modelling task. For example recorded precipitation data for a specific rainfall event can be measured in an accurate way, but still may not be suitable as input-data for a model when the spatial distribution is neglected.

For a detailed analysis of the impact of uncertainties on model results a more detailed classification is necessary. According to Beck (1991), Reichert (2009) divides the causes of uncertainties of model predictions into the following categories:

- non-deterministic behaviour of a system
- uncertainty of model parameter values
- uncertainty of the model structure
- uncertainty of external influence factors
- uncertainty of the numerical solution of the model equations

This is a general classification of uncertainties in environmental models not directly related to uncertainties in urban drainage modelling. Additional discussion is available for example from Beck (1983, 1987); O'Neill and Gardner (1979).

Non-deterministic behaviour of a system is the behaviour of a system which cannot be predicted. Although physicists would remark that according to current state of knowledge "true" randomness only exists in quantum mechanics such behaviour can reasonably be described by random model elements. Such a chaotic behaviour of a system has its origin in high sensitivity of a deterministic system to initial conditions (the initial state of a system can never be reproduced perfectly), in influence factors that are not measured and hence cannot be considered and in aggregation error due to the lack of spatial resolution.

Uncertainty of model parameter values is caused by uncertainties in the estimation of model parameter values during calibration.

Uncertainty of the model structure is caused by the point that a model is always a simplification of one or more real underlying processes. Structural model uncertainties consist of inadequate selection of model variables and processes (or inadequate formulation of processes) and of inadequate choice of the spatial and temporal resolution of a model. These uncertainties are very difficult to quantify and the only methodology to deal with that problem seems to compare model results of different models.

Uncertainty of external influence factors is what typically is called "input-data uncertainty". Hence this describes the impact of the environment on the model. This uncertainty can be estimated by deeper analysis of historical records of external influence factors in order to include external influence factors in the modelling process. For example uncertainties in rainfall measurements caused by wind could be modelled by including wind measurements to the model.

Uncertainty of the numerical solution of the model equations is uncertainty caused by the limited accuracy of numerical solutions. Although Reichert (2009) remarks that the other sources of uncertainty dominate uncertainty of the numerical solutions (and hence they usually can be neglected), Fach et al. (2007) report significant problems in the solution of the Saint Venant differential equations (see chapter 2) in commercial software, which can lead to a large error in mass balance.

Another type of uncertainty which is classified in this category occurs when Monte-Carlo simulations are used for calculating probability distributions. If the number of Monte-Carlo iterations is too small, this can lead to errors in the derivation of probability distributions.

Deletic et al. (2009) classify uncertainties related to urban drainage modelling in a bit different way as described below:

- Model input uncertainties related to
  - Measured input data
  - Estimated input data
  - Model parameters
- Calibration uncertainties related to
  - Measured calibration data uncertainties
  - Measured calibration data availability and choices
  - Calibration Algorithms
  - Criteria Functions
- Model structure uncertainties related to
  - Conceptualisation errors
  - Numerical methods and boundary conditions

4.1.1 Input-data uncertainties

Model inputs are required to run either a calibrated or a non-calibrated model and uncertainties in that input propagate to uncertainties in model output. Uncertainties in measured input data can often be characterised as (1) systematic uncertainties (e.g. due to an insufficient calibration of the measurement device) and/or (2) Gaussian distributed uncertainties (due to random effects).

As the main driving force for urban drainage simulation is rainfall data, it is expected that uncertainty in that input has significant impact on simulation results. Analysis of uncertainties in measurements of precipitation has a rather long history and a lot of publications dealing with estimation of those uncertainties are available (e.g. Sevruk, 1981, 1982, 1996). Sevruk (2002) presents results of a questionnaire of the World Meteorological Organization (WMO) and Einfalt et al. (2002) a survey of the Group on Urban Rainfall under the International Water Association (IWA) regarding different types of measurements. Both report tipping bucket gauges as the dominating method for obtaining high resolution rainfall data. Hoppe (2006) presents a very comprehensive overview of uncertainties of tipping bucket gauges based on literature review of numerous publications (La Barbera et al., 2002; Maksimovic et al., 1991; Rauch et al., 1998b; Schilling, 1991; Sevruk, 1996). Consequently he reports possible systematic errors of up to 30%.

In another study Zhu and Schilling (1996) analyse the impact of insufficient temporal rainfall resolution on simulation of annual combined sewer overflow.

As shown in **Paper II** a systematic error can be compensated during calibration of an rainfall / runoff model and a random error has nearly no effect as the rainfall / runoff model equalises random errors. Hence the measurement error of precipitation measurements is not dominating if the same rain gauge is used for calibration and estimation (which is not always the case).

Unfortunately systematic uncertainties are often highly influenced by scale issues of the model. Especially when large catchments are monitored by only a few rain gauges due to spatial rainfall distribution it might be possible that - to mention an extreme case - rainfall only occurs over the rain gauge or everywhere but over the rain gauge. Hence, measured rainfall might not be characteristic for real rainfall over the catchment (Arnaud et al., 2002; Chaplot et al., 2005; Krejci, 1996; Mikkelsen et al., 1998; Morin et al., 2006; Segond et al., 2007; Vaes et al., 2005). A possibility to consider spatial rainfall distribution are rainfall–radar measurements which are more and more used in urban drainage modelling (Einfalt et al., 2004). Nevertheless this radar measurements have to be calibrated on rain gauges which introduces a new source of uncertainties (Carpenter and Georgakakos, 2004; Villarini et al., 2007, 2008).

Estimated input data is model input data that is not directly measured but estimated from data collection. For example a parameter that is usually determined from orthophotos and terrain maps is the *total impervious area* ($A_{imp,T}$), but the parameter that is actually required for modelling is the *effective impervious area* ($A_{imp,eff}$). $A_{imp,eff}$ can be estimated to be 0.7 to $0.9 \cdot A_{imp,T}$ but its exact value has to be determined during calibration. Another example is the in–line sewer storage capacity in hydrological models of combined sewer systems. While the sewer storage capacity is considered in hydrodynamic models implicitly this parameter has to be estimated and calibrated in conceptual models. For more examples and a more detailed description see for example Kleidorfer et al. (2006a) or Kleidorfer et al. (2008).

Both (1) measured input data and (2) estimated input data can be affected by additional "long-term prediction uncertainties" which occur when trying to predict long-term environmental change effects (e.g. land-use change, climate change). Such predictions often contain substantial uncertainties. Not only estimation of uncertainties but also consideration of interaction of different parameters is important in such case. This topic is analysed in **Paper IV**.

Model parameters cannot be estimated from data collection, they have to determined during model calibration. Both – estimated input data and model parameter uncertainties – are strongly related to calibration uncertainties.

4.1.2 Calibration uncertainties

During model calibration certain parameters have to be estimated by comparing model output with corresponding measurements. Similar to uncertainty of input–data also measured calibration–data contains uncertainty depending on the measurement device used. Sometimes measured data can not directly be used for calibration because the measured data is no model output. In such a case calibration data has to be calculated from measured data. For example often combined sewer overflow discharge is calculated from water–level measurements (Fach et al., 2008b; Sitzenfrei et al., 2008), which causes additional uncertainty.

Additionally to uncertainties from the measurement device also uncertainties from spatial and temporal data availability occur. If calibration data are not representative, the calibrated model parameters will not be accurately estimated (e.g. calibrating a rainfall-runoff model during summer periods will produce model parameters which will not reflect summer period processes). For example, Mourad et al. (2005) used a random sampling methodology to understand the impact of data availability (i.e. number of events) on the calibration of several urban stormwater quality models. They found that, in order to adequately calibrate these models, it was often required to use the majority (between 60 to 100%) of the available data set during calibration.

Additionally, in the case of spatially distributed systems it is neither possible nor sensible to measure the complete system characteristics, and the question is raised how many measurement sites are necessary. This topic is regarded in **Paper I** where the impact of the number of measurement sites used for calibration of combined sewer systems is evaluated. The number of required sites is influenced by the time period used for calibration. For example, a similar calibration performance can be reached when using 30% of all available sites for calibration and a time period of one year, as compared to using 60% of all available sites with five single events.

Furthermore, the availability of calibration data impacts the uncertainty of a model's prediction outside the calibration period (McCarthy, 2008; Mourad et al., 2005) caused by the (input–data dependent) sensitivity of the model outputs to parameter changes (McCarthy, 2008).

Another point is the choice of the correct calibration variable that will best suit the model application (i.e. that model output that is compared during calibration). For example, there has been discussion on whether to calibrate load models using pollutant concentrations or fluxes, with fluxes most commonly used. McCarthy (2008) demonstrates that using concentrations produced more accurate predictions. In another example Kleidorfer et al. (2006a) demonstrated that not all outputs of a hydrological model of a combined sewer system as for example the runoff to the WWTP can be used for calibrating a model in order to estimate CSO efficiency.

Further uncertainties are caused by the choice of the calibration algorithm and by the objective function that is optimised (minimised or maximised). As urban drainage models are highly non–linear systems not all calibration algorithms can effectively find the global optimum of the objective function. And even if the global optimum can be found this does not necessarily mean that the objective function selected represents model output. For example the commonly used Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) is known to be mainly influenced by peaks in the timeseries and low values are under-represented. This is no problem when the modelling task is representing peak events (as it usually is) but calibration parameters found this way may probably not reproduce low events. For a comparison of the performance indicators Nash–Sutcliffe efficiency, Bias and Index of agreement see Achleitner (2008). **Paper V** presents the software tool CALIMERO where different calibration algorithms and different objective functions can be tested.

4.1.3 Model structure uncertainties

A model is always only a simplification of reality and model structure uncertainties are caused by processes which are not (or not sufficiently) represented by the model. This source of uncertainties is highly depending on the specific case study and the final aim of the modelling task. Hence the model user has to decide which processes are necessary to sufficiently represent reality. He has to choose the model based on such considerations and often model structure is only questioned if the model calibration fails. Usually they are not regarded explicitly but only in conjunction with other sources of uncertainties (for example model parameter uncertainties). To current state of knowledge an estimation of model structure uncertainties can only be done by comparing the performance of different models (Deletic et al., 2009).

4.2 The Likelihood function

The likelihood function (or often just called “likelihood”) in modelling is a measure how good simulation results and corresponding observations (calibration data) fit. Therefore the likelihood function describes the probability distribution of model outcomes as a function of model parameters and model input. Here it is important to note that in frequentist statistics (which is not the focus of this thesis) model parameters are not random variables of a specific distribution but fixed but unknown variables.

Reichert (2009) writes the likelihood function L of a model M in a generalised way as a function of model parameters Θ_M, model output for a specific measurement layout y^L and model input x (equation 4.1 left term) which is characterised by a probability density of the model results conditional on values of input x and model parameters Θ_M (equation 4.1 right term).

$$L_M(\Theta_M^L, y, x) = f_{Y_{M,obs}^L}(y^L|x, \Theta_M) \tag{4.1}$$

Here the different sources of uncertainties are implicitly considered in the probability density function of the model results in a generalised way.

Often is is assumed that the residuals follow a Gaussian distribution with mean μ and standard deviation σ with equation

$$f_{N(\mu,\sigma)}(z) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \, \mathrm{e}^{-\frac{1}{2}\frac{(z-\mu)^2}{\sigma^2}} \; . \tag{4.2}$$

For this special (but common) assumption of multi–normally distributed residuals about model output the likelihood function would be

$$L_M(\Theta_M^L, y, x) = \frac{1}{(2\pi)^{\frac{n}{2}} |C|^{\frac{1}{2}}} \cdot \mathrm{e}^{-\frac{1}{2}(y_{obs} - f(\Theta,x)^T C^{-1} (y_{obs} - f(\Theta,x))} \tag{4.3}$$

This is the case if the deviation between model output and calibration data can be attributed to a “measurement noise” or the inability of the model to reproduce system behaviour. Different types and sources of uncertainties can be combined in one distribution function or be distinguished as done for example in the Bayesian Total Error Framework of Kuczera et al. (2006). Here $y_{obs} - f(\Theta, x)$ is the vector of the deviation of observed data and model results (i.e. the residuals) and C is the residual covariance matrix. If the residuals do not cross–correlate C is a diagonal matrix and each element is the variance σ^2 of corresponding residuals. In such a case uncertainties are often expressed as weights w_i for each measurement point i with

$$w_i = \frac{1}{\sqrt{\sigma_i^2}}. \tag{4.4}$$

Because of computational difficulties which can occur especially when different probability density functions are combined or when n in equation 4.3 is a very large number (which is the case for example in long time–series) often the logarithm of the likelihood function is used. Then the power $n/2$ becomes a simple multiplication and is much easier to handle in the numerical solution.

In Bayesian and pseudo (or informal) Bayesian inference (see section 4.3) it is important if the likelihood function is "formal" or "informal". The formal definition follows from assumptions about the structure of the error (as above) and is directly connected to the probability density function of the observed random vector conditional to the knowledge of the parameter vector and the predictor vector (Mantovan and Todini, 2006). An informal (also "less-formal", "pseudo-formal") likelihood function is a more subjective rating of model output, also a binary rating ("good" result / "bad" result) is possible. Informal likelihood functions used in the literature are after Beven (2009) for example:

$$L_M(\Theta_M^L, y, x) = \frac{1}{C}\left(\frac{1}{\sigma^2}\right)^N \tag{4.5}$$

Here N is an empirically chosen shaping parameter, C is a scaling factor and σ^2 is the residual variance. Please note that in case of a perfect fit of model output and observed data $\sigma^2 = 0$ and $L_M(\Theta_M^L, y, x)$ would go to infinity.

Another commonly used measure of model output is the Nash-Sutcliffe criterion based on the Nash-Sutcliffe efficiency by Nash and Sutcliffe (1970)

$$L_M(\Theta_M^L, y, x) = \begin{cases} 1 - \frac{\sum_{t=1}^{T}(y_{obs}^t - f(\Theta,x)^t)^2}{\sum_{t=1}^{T}(y_{obs} - \bar{y_{obs}})^2} & \forall\Theta| \sum_{t=1}^{T}(y_{obs}^t - f(\Theta,x)^t)^2 \leq \sum_{t=1}^{T}(y_{obs} - \bar{y_{obs}})^2 \\ 0 & \forall\Theta| \sum_{t=1}^{T}(y_{obs}^t - f(\Theta,x)^t)^2 > \sum_{t=1}^{T}(y_{obs} - \bar{y_{obs}})^2 \end{cases} \tag{4.6}$$

Hence usually the Nash-Sutcliffe efficiency returns a value in the interval $]-\infty|1]$ (where 1 represents a perfect fit between model output and observed data) and the Likelihood criterion is constraint in the interval $[0|1]$. A likelihood value of 0 means that this model (i.e. parameter set) is rejected.

For more information about informal likelihood functions and more examples (e.g. Chiew-McMahon criterion, Normalised sum of squared errors, Index of agreement, Mean cumulative error, Normalised absolute error and maximum distance) please refer to Smith et al. (2008).

4.3 Methods for uncertainty analysis and propagation

4.3.1 Parameter sensitivity

Although parameter sensitivity analysis is not a method for uncertainty estimation of simulation models, knowledge about parameter sensitivity is crucial. This is the necessary background for any deeper analysis and helps to improve the understanding of the behaviour of the model and it also supports model calibration and data collection. Its goal is to explore the change in model output resulting from a change in model parameters or model inputs and to separate influential from non-influential parameters.

Usually parameter sensitivity analysis is undertaken during the development of models and often model users argue that they already know sensitivities of their parameters. While to some extend that may be true one has to keep in mind that parameter sensitivities often highly depend on the parameter range analysed and on the input data of a model. Hence a parameter might be sensitive for one application and insensitive for another one. Another point is, that model users *are* model developers when setting up a distributed drainage model in which different catchments are connected by sewer pipes (or different submodels are combined to one model). As already mentioned this just means that a final model consisting of i submodels (with n parameters for each submodel) has $n \cdot i$ parameters. By knowing sensitivities of model parameters the model user also can decide which data has to be collected very carefully and which one can be estimated without disturbing model output too much. For example Möderl et al. (2009) discuss aspects of spatial distributed parameter sensitivity.

4.3.1.1 Local sensitivity analysis

Local sensitivity analysis (also called “point sensitivity analysis”) investigates the sensitivity of a parameter with respect to the simulation results at a certain parameter value. As sensitivity depends on the value analysed, the parameter value should be chosen close to expected final value after model calibration. Typical choices are published “default” values or values gained trough preliminary analysis.

A measure for local sensitivity is the slope of the linearised function $f(\Theta_M)$ in dependency of the model parameters Θ_M (equation 4.7, see figure 4.1).

$$s_{i,j}(\Theta_M) = \frac{\partial f(\Theta_{M,j})}{\partial \Theta_{M,j}} \tag{4.7}$$

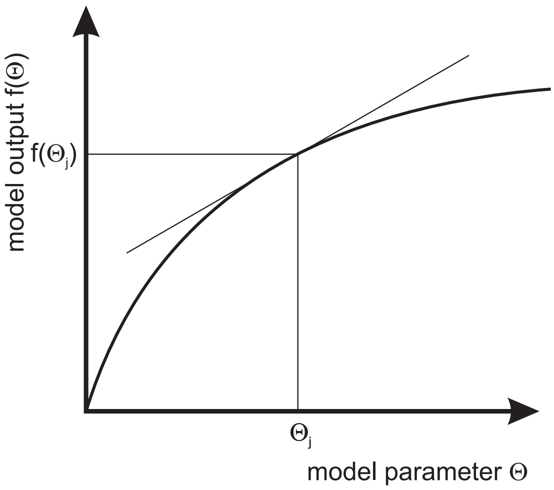

Figure 4.1: Local parameter sensitivity - function slope

According to the terminology of Reichert (2009) equation 4.7 is the absolute-absolute sensitivity. Here sensitivity depends on units of model results and model parameters and hence sensitivities can hardly be compared. For this reason it is often useful to normalise model results and model parameters to get sensitivities which can be compared. According to Reichert (2009) following definitions are possible:

- absolute-relative sensitivity describing absolute change of the results for a relative change of the parameter by 100%

$$s_{i,j}(\Theta_M) = \Theta_{M,j} \cdot \frac{\partial f(\Theta_{M,j})}{\partial \Theta_{M,j}} \tag{4.8}$$

- relative-absolute sensitivity describing relative change of the results for a absolute change of the parameter

$$s_{i,j}(\Theta_M) = \frac{1}{f(\Theta_{M,j})} \cdot \frac{\partial f(\Theta_{M,j})}{\partial \Theta_{M,j}} \tag{4.9}$$

- relative-relative sensitivity describing relative change of the results for a relative change of the parameter by 100%

$$s_{i,j}(\Theta_M) = \frac{\Theta_{M,j}}{f(\Theta_{M,j})} \cdot \frac{\partial f(\Theta_{M,j})}{\partial \Theta_{M,j}} \tag{4.10}$$

The gradient term $\frac{\partial f(\Theta_{M,j})}{\partial \Theta_{M,j}}$ in equations 4.7 to 4.10 is often difficult to evaluate analytically. In urban drainage modelling, when dynamic systems are modelled, it is nearly impossible as this would require differentiating the model equations. Hence usually this problem is solved numerically by using runs of the model with slightly different values of Θ_M. Then the gradient term can be approximated by

$$\frac{\partial f(\Theta_{M,j})}{\partial \Theta_{M,j}} = \frac{f(\Theta_{M,j} + \Delta\Theta_{M,j}) - f(\Theta_{M,j} - \Delta\Theta_{M,j})}{2\Delta\Theta_{M,j}} \tag{4.11}$$

and $\Delta\Theta_{M,j}$ is a small increment in the parameter value.

As already mentioned local sensitivity analysis is based on a linearization of the model. Its main advantage is that local sensitivity analysis is computationally relatively inexpensive and only requires one additional model run for each parameter analysed. Hence it can give useful information about model behaviour without much afford. In urban drainage modelling we cope with highly non-linear dynamical models and local sensitivity analysis can only give a rough estimation of impact of parameter changes on model output.

The evaluation of the sensitivity indices for the didactic example (section 3.4) at the parameter values $W = 50$ and $b = 0.5$ is shown in Figure 4.2 as function of model input. In the first row (Figure 4.2 (a)) the absolute change of model output for a relative change of the parameter according to equation (4.8) is presented. The second row (Figure 4.2 (b)) shows the relative change of model output (in %) for an absolute change of the parameter by one unit (equation (4.9)). The third row (Figure 4.2 (c)) shows the relative change of model output for a relative change of the parameter (equation(4.10)). Here one can clearly see that parameter sensitivity depends on model structure and is a function of model input. As local parameter sensitivity is based on a linearisation of the model, in Figure 4.2 (b) and (c) the linar relationship between model output and W (see equation 3.1) can be seen resulting in a constant sensitivity index.

4.3.1.2 Global sensitivity analysis

Global sensitivity analysis is an attempt to explore a wider range of parameter space. Although Reichert (2009) appropriately remarks that the sensitivity analysis never is

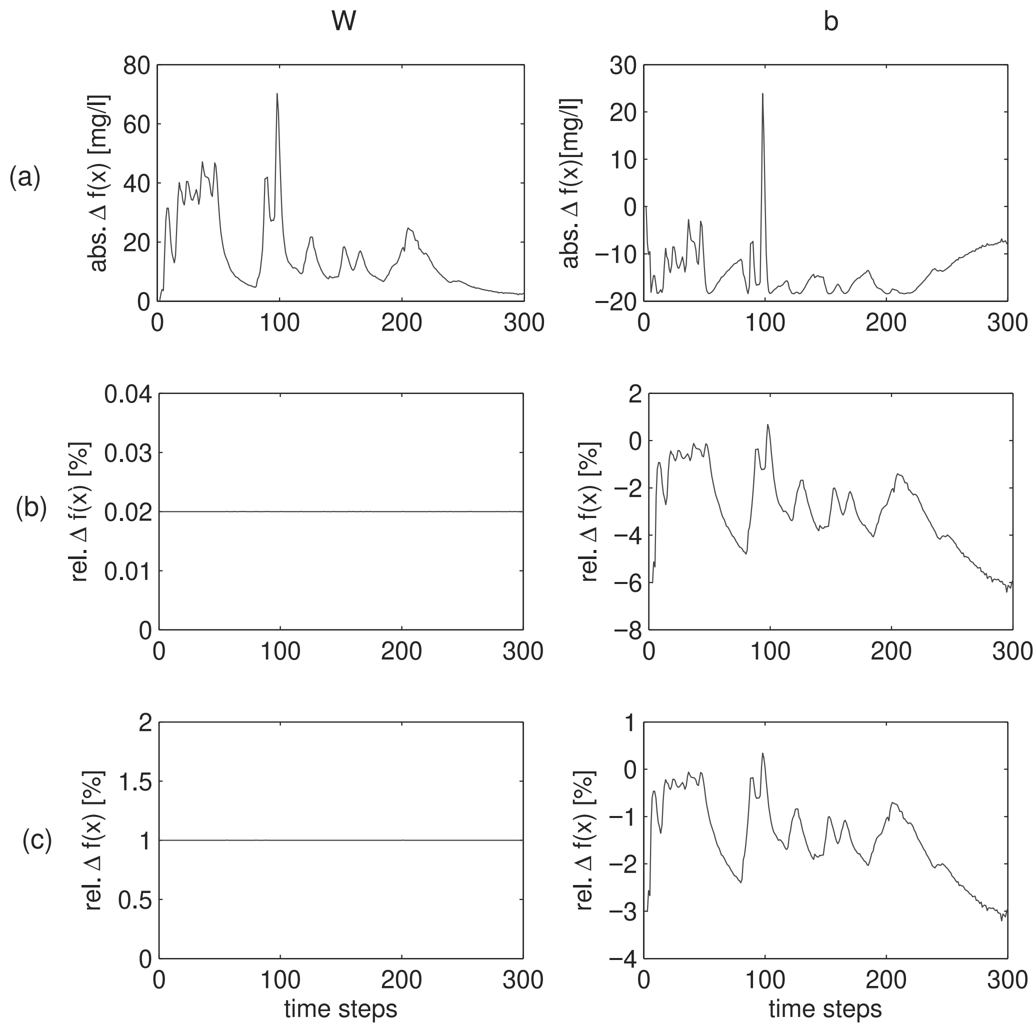

Figure 4.2: Local sensitivity indices of didactic example - Local sensitivity indices for the parameters W (left) and b (right) - (a) represents the absolute change of model output for a change of the parameter by 100%; (b) represents the relative change of model output for a change of the parameter by one unit; (c) represents the relative change of model output for a relative change of the parameter by 100%

"global" (i.e. always only a certain range can be analysed) and hence this should rather be called "regional" sensitivity analysis, the commonly used term "global sensitivity analysis" is also used here. As analytical methods for global sensitivity analysis usually cannot be used for evaluating urban drainage models (Reichert, 2009), here mainly numerical techniques are important. That's because urban drainage models are highly non-linear and their variables are not continuous (Arnbjerg-Nielsen and Harremoës, 1996). Therefore several methods exist. A review of different methods is available from Saltelli et al. (2006).

Contrary to local sensitivity analysis global sensitivity analysis does not start from a single point, but from a distribution of parameters and the parameters are random values sampled from that distribution. This is already very close to methods used for uncertainty analysis and if parameter distributions are chosen according to known uncertainties this can be interpreted as analysis of sensitivity of model result in respect to uncertainties of model parameters. The 'Hornberger-Spear-Young method' (Hornberger and Spear, 1980; Spear and Hornberger, 1980; Young, 1983) is mostly cited as method for sensitivity analysis (e.g. Beven, 2009; Thorndahl et al., 2008) but also sometimes as method for uncertainty estimation (e.g. Arnbjerg-Nielsen and Harremoës, 1996). In fact this is not a discrepancy as analysis of parameter sensitivities is part of uncertainty analysis (Deletic et al., 2009). Here it is important to note that several authors showed that the method used (especially the way how parameters are sampled from the distribution and how many iterations are undertaken) has significant impact on the apparent sensitivities (Beven, 2009). This is also shown by Tang et al. (2007) or Pappenberger et al. (2007).

4.3.1.3 Graphical Analysis

Cook and van Noortwijk (2000) present some methods for graphical sensitivity analysis. Out of those Reichert (2009) concentrates on scatter plot and scatter plot matrices which can be a convenient way for presenting impact of parameter changes on model results and for comparing sensitivity of different model parameters. Its advantage is that results are easy to visualise, understand and interpret. Therefore parameters are sampled from a distribution and model results for that parameters are plotted against parameter samples. If the model result is not sensitive to a parameter this leads to a wide spread of points around the mean of simulation results. Hence such plots represent correlations of parameter deviations and their impact on model output. Sensitivity of multiple parameters can be visualised in a scatter plot matrix.

Figure 4.3 and Figure 4.4 show a graphical sensitivity analysis to visualise sensitivities of the two parameters W and b in the didactic example (section 3.4). 1 000 random parameter samples were drawn from the uniform distribution U[25|75] for W from U[0.25|0.75] for b which is a sampling of $[-50\%| + 50\%]$ around their assumed value of $W = 50$ and $b = 0.5$.

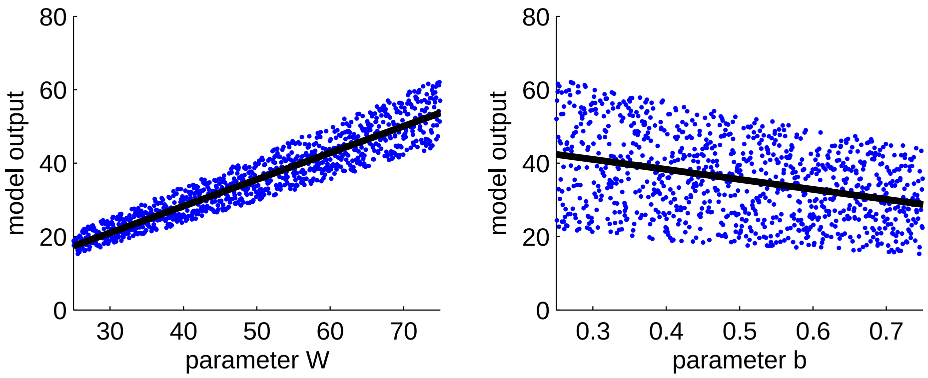

Figure 4.3: Graphical sensitivity analysis with scatter plots - Illustration of graphical analysis with scatter plots for regression model with $Q = 0.5$

In Figure 4.3 model output is plotted against W (left hand side) and b (right hand side) for assumed input $Q = 0.5$. The black line represents the mean of the model output with respect to the particular parameter. The wider spread on the right hand side shows that model output is more sensitive to a deviation of W than to a change in b.

Figure 4.4 shows results for the same analysis but with assumed input $Q = 1$. Here it is clear to see that b is completely insensitive (which is clear as 1^b is always 1) and model output is only influenced by W. This is also an example for the point that parameter sensitivity not only depends on the parameter range analysed but also on model input.

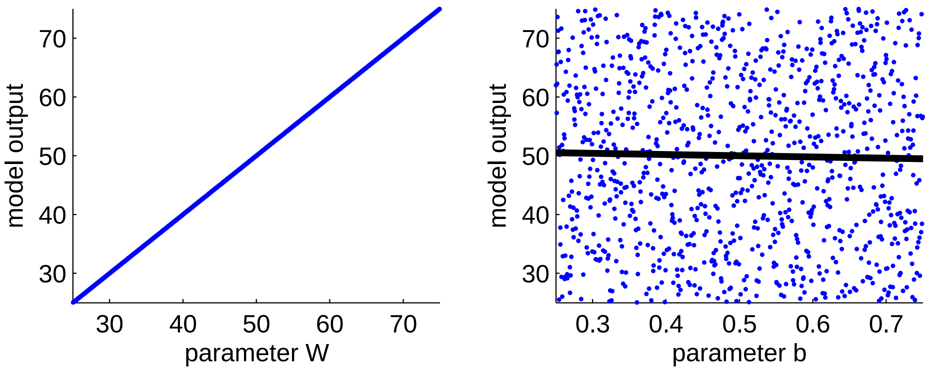

Figure 4.4: Graphical sensitivity analysis with scatter plots - Illustration of graphical analysis with scatter plots for regression model with $Q = 1.0$

4.3.1.4 Hornberger-Spear-Young (HSY)

The Hornberger-Spear-Young (HSY) method (Beck, 1983, 1987; Hornberger and Spear, 1981, 1980; Spear and Hornberger, 1980; Young, 1983) (also often cited as generalised sensitivity analysis (GSA)) was firstly used for estimating uncertainties of river quality models. They introduced the concept of a "problem-defining behaviour" of the system analysed. Therefore in a Monte-Carlo simulation parameters are sampled from distributions and model output of each iteration is classified to meet a specific behaviour B or not ($\neg B$). Subsequently two subsets of model results and their according model parameters are gained. One is showing the behaviour defined and one is not. By comparing cumulative distributions of the parameters of those two classes it is possible to estimate sensitivity of model output to changes of parameter values. If the cumulative distributions show no or little difference that parameter is interpreted as insensitive, otherwise it is not. This can easily be seen by plotting cumulative distributions or by applying a statistical measure like

$$d_{m,n} = \sup_x |S_n(x) - S_m(x)|. \tag{4.12}$$

Here S_n and S_m are the sample distribution functions estimating the cumulative distribution functions $F(\Theta_M|B)$ (representing behaviour B) and $F(\Theta_M|\neg B)$ (not representing behaviour B) for n behaviours and m non-behaviours. $d_{m,n}$ is used in the two-sample Kolmogorov Smirnov test and represents the maximum vertical distance between $F(\Theta_M|B)$ and $F(\Theta_M|\neg B)$. In the Kolmogorov Smirnov test the null hypothesis that the two sample distributions come from the same distribution is rejected at a level α if $\sqrt{\frac{nm}{n+m}}D_{n,m} > K_\alpha$. For ranking sensitivity of parameters $d_{m,n}$ can be used directly. Hence large values of $d_{m,n}$ indicate sensitive parameters and small values insensitive parameters, i.e. $d_{m,n} = 1$ is most sensitive and $d_{m,n} = 0$ is non–sensitive (Beven, 2009; Thorndahl, 2008). Of course this approach is not limited to two classes of model output and different “behaviours” could be introduced. For example Beven (2009) shows the application of HSY on a 4 parameter rainfall runoff model in which the simulations have been classified into ten subsets.

Numerous applications show the potential of this method. van Griensven et al. (2006) compare different methods (including HSY) for water quality estimation of two river in the Texas and Ohio. Sieber and Uhlenbrook (2005) use this method for sensitivity analysis of a distributed catchment model in Germany to verify the model structure and Guven and Howard (2007) identified critical parameters of a cyanobacterial growth and movement model.

The HSY method was applied on the didactic example (section 3.4) to demonstrate its functionality. Therefore 50 000 random parameter samples were drawn from the uniform distribution U[0|100] for W and from U[0|1] for b. Corresponding simulation results were compared to calibration data and the Nash–Sutcliffe efficiency E was evaluated. Simulation results with $E > 0$ were regarded to be “behavioural” and simulation results with $E < 0$ to be “non–behavioural”. Figure 4.5 shows the empirical cumulative distributions for both parameters W and b and for both groups. Additionally the maximum vertical distance d between the distributions of the two groups was evaluated. As one can clearly see the distributions for both groups are significantly different, hence model output is sensitive two both parameters.

4.3.2 Generalized Likelihood Uncertainty Estimation (GLUE)

Taking HSY generalised sensitivity analysis as basis Beven and Binley (1992) firstly used multiple behavioural models and presented a methodology called Generalized Likelihood Uncertainty Estimation (GLUE). GLUE is based on the idea that there is no “true” parameter set which can be found during model calibration because all

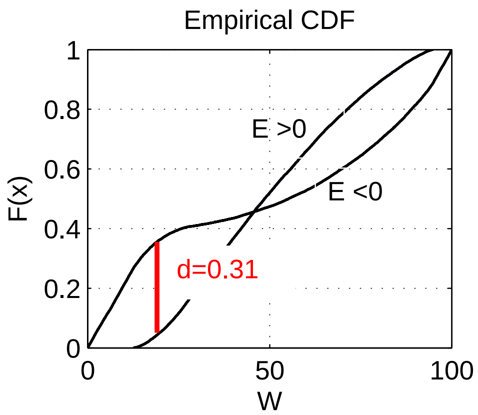

Figure 4.5: Hornberger–Spear–Young method - Empirical cumulative distributions of model parameters W (left) and b (right) for behavioural model results ($E > 0$) and non–behavioural model results ($E < 0$) with KolmogorovŰ-Smirnov distance d respectively

measurements (input and calibration data) as well as model parameters are to some extent uncertain. Hence, different parameter sets of a model can all lead to similar model results and to a similar fit with observation data. This is especially true for complicated models with a lot of calibration parameters and with correlations between some of them.

Consideration of multiple behaviours in GLUE is done by weighting model results according to a likelihood measure. This likelihood measure expresses model performance for example by comparing model results with calibration data, hence it can be a measure of the goodness of fit.

While in principle this is closely related to formal Bayesian methods (see section 4.3.3) the main difference is that the likelihood measure can be formal or informal (see section 4.2). In formal Bayesian methodology always a formal description of the likelihood function is required. This is extensively discussed in literature. For example Mantovan and Todini (2006) reported incoherencies of the GLUE methodology with Bayesian inference, Beven et al. (2007) replied that formal Bayesian inference is a special case of GLUE when a formal likelihood description is used and Mantovan et al. (2007) countered that "*Bayesian inference has been a well established and formalised theory for more than 60 years, whereas GLUE fails to benefit from the Bayesian inference results when improper likelihoods are used*". In a further response Beven et al.

(2008) show that one has to make very strong assumptions about the nature of the modelling error to formulate a formal likelihood function which is not always possible. As in real applications definition of uncertainties can be very complicated and the formulation of a formal likelihood function based on wrong or too simple error structure can lead to incorrect parameter distributions. Regardless of this discussion Freni et al. (2009b) also report that both methods perform similar when GLUE is based on the same assumptions as the Bayesian approach.

Nevertheless, several authors show applications of GLUE for different modelling tasks.

Freer et al. (1995) used GLUE to estimate predictive uncertainties of the rainfall / runoff model TOPMODEL in hydrology of watersheds. Page et al. (2003) investigated the uncertainty of historic and predicted acidic deposition of the model of acidification of groundwater in catchments (MAGIC). Rankinen et al. (2006) analysed the conceptual and parameter uncertainty of the semi-distributed INCA-N (Integrated Nutrients in Catchments-Nitrogen) model. McMichael et al. (2006) use GLUE for model predictive uncertainty estimation of the MIKE SHE hydrological model for estimating monthly streamflow in a semi-arid shrubland. Beven (2007) focuses on river water quality modelling with respect to the Water Framework Directive.

While all these papers describe applications from hydrology of natural catchments recently GLUE was also used for uncertainty estimation in urban drainage modelling. While in principle the differences of these fields seem to be marginal, different model structure and time-resolution can influence the (as mentioned) subjective choice of the likelihood function. In general urban drainage models are simpler regarding model structure (less processes, less parameters), but usually need a higher resolution of input data (5 - 15 min timesteps). This induces different characteristics of uncertainty (compare section 4.1). Mannina et al. (2006) and Freni et al. (2009a) analyse uncertainty of an integrated urban drainage system. Furthermore Freni et al. (2008b) analyse the impact of the acceptability threshold of the Nash-Sutcliffe criterion (i.e. the impact of the likelihood function) by varying the "accepted" Nash-Sutcliffe efficiency between 0 and 0.7. Freni et al. (2008a) use GLUE to test different sewer sediment erosion models. Thorndahl et al. (2008) apply GLUE on the commercial urban drainage model MOUSE based on six single events and Lindblom et al. (2007) compare GLUE and greybox modelling for predicting copper loads in stormwater systems.

Also the GLUE methodology was applied on the didactic example. Therefore 50 000 parameter samples were randomly drawn from an uniform distribution ($W \in \mathrm{U}[0|100]$ and $b \in \mathrm{U}[0|1]$). The likelihood function used was the Nash–Sutcliffe criterion (equation (4.6) in section 4.2) with an acceptability threshold of $E = 0$. From the 50 000 parameter samples 26 657 were accepted (i.e. E>0 was obtained), that is an acceptance ration of 53%.

Figure 4.6 shows scatter plots of parameter samples for W (left) and b (right) against E. This is also a visualisation of the response surface in a sectional view as usually done for multidimensional problems. But as in the relationship used here only two model parameters are included, the likelihood surface can also be visualised in a three–dimensional plot (see Figure 4.7).

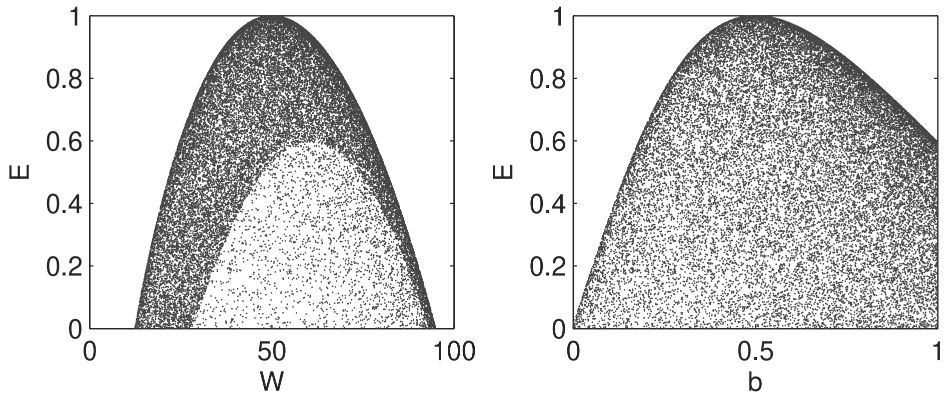

Figure 4.6: GLUE: Scatter plots - Scatter plot of model parameters W (left) and b (right) against corresponding E

From the evaluation of the likelihood surface a posterior distribution of model parameters can be constructed. Therefore the sampling range analysed for each model parameter (abscissas in Figure 4.6) is divided into 50 equally sized intervals. Consequently all likelihood values in each interval are summarised and these sums are presented in a bar–plot (see Figure 4.8) .

Figure 4.9 shows the posterior distribution of model parameters constructed for accepted model parameters, weighted after the value of the likelihood function obtained

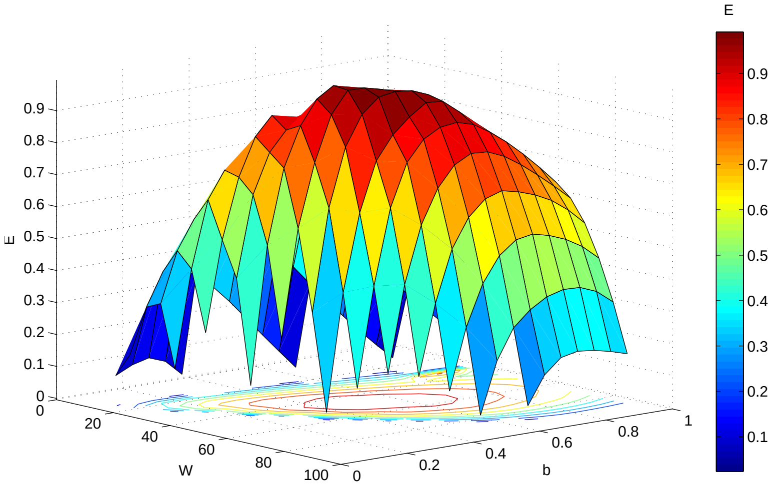

Figure 4.7: Visualisation of the response surface - Visualisation of the response surface of the regression water quality model with $E = f(W, b)$

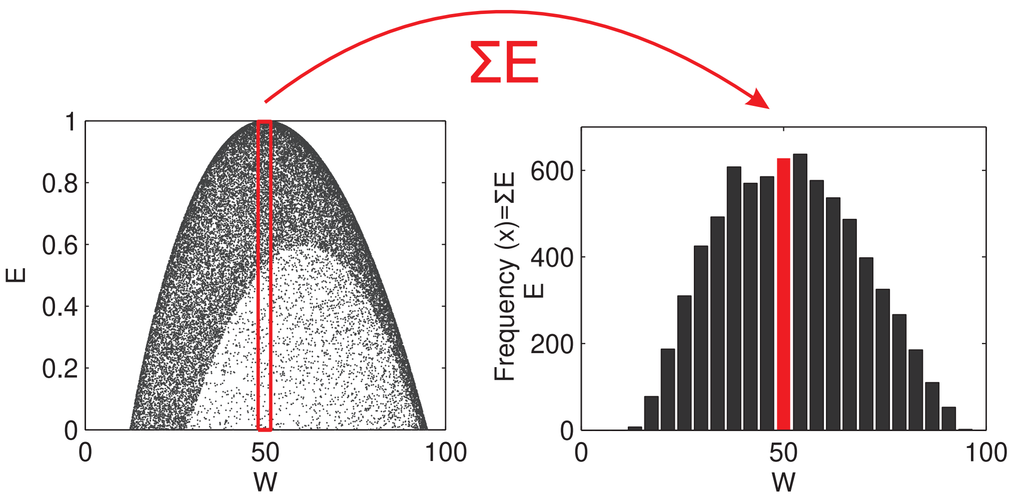

Figure 4.8: GLUE: Construction of posterior distribution - Construction of posterior distribution from likelihood function by summarising E for each interval

from the iterations. Figure 4.10 shows the corresponding empirical cumulative distributions of the model parameters W (left) and b (right) as well as the 25% percentile and the 75% percentile (red squares). The values of these percentiles is presented in table 4.1.

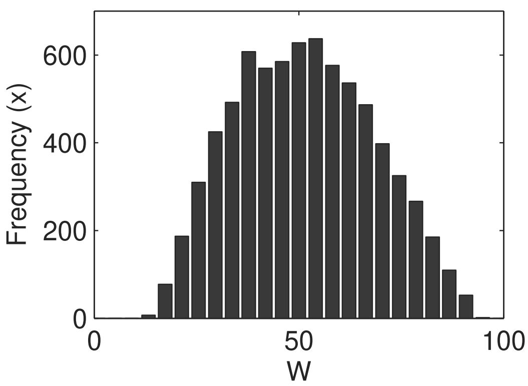

Figure 4.9: GLUE: Posterior distribution - Posterior frequency distribution of model parameters W (left) and b (right) weighted according Nash–Sutcliffe criterion

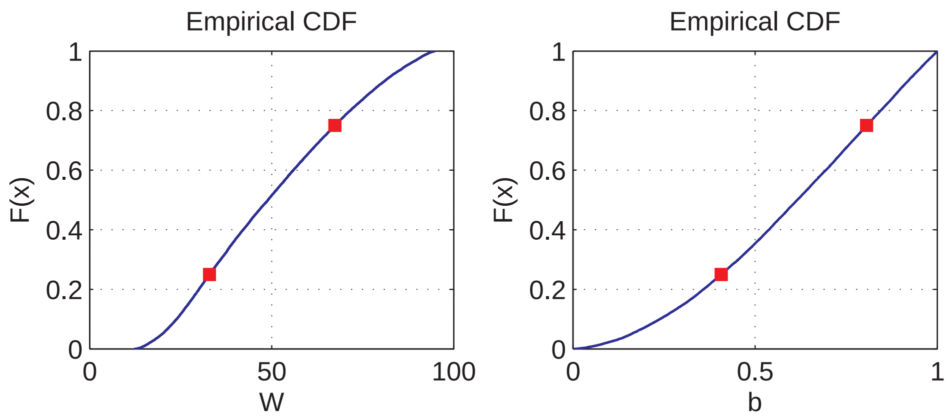

Figure 4.10: GLUE: Empirical cumulative distribution of model parameters - Empirical cumulative distribution of model parameters W (left) and b (right) as well as 25% and 75% percentiles (red squares)

Table 4.1: 25% and 75 percentile of the model parameters of the didactical example calculated with GLUE

Parameters	25% percentile	75% percentile
W	33	67
b	0.41	0.81

Consequently an uncertainty bound of model predictions can be evaluated. This bound showing again the 25% and the 75% percentile is presented in Figure 4.11 together with the measurement data. Here the measurement data perfectly lies in the uncertainty range for the whole event. This is not necessarily always the case when real measurement data is used instead of synthetic data.

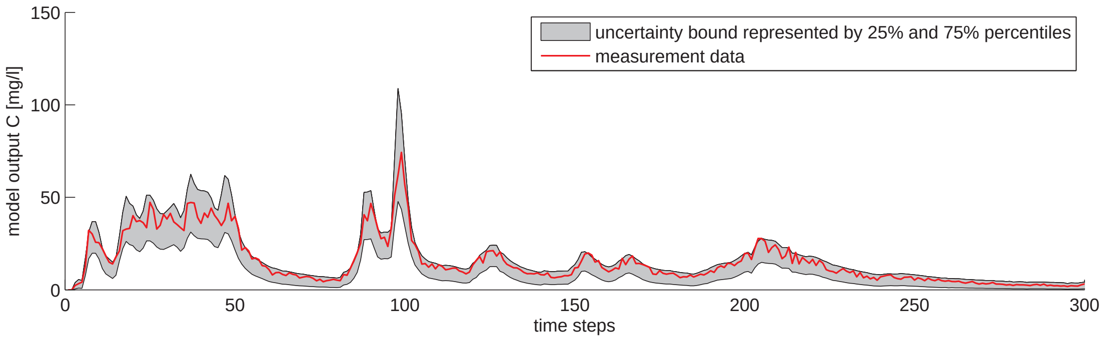

Figure 4.11: Uncertainty bound calculated with GLUE - Uncertainty bound (grey) calculated with GLUE and represented by the 25% and the 75% percentile of the accepted simulation results. The calibration data is represented by the red line.

4.3.3 Bayesian inference

Bayesian inference is a very popular approach in environmental modelling as by using this methodology a personal (subjective) "degree of belief" (described by a probability distribution) can be introduced in the modelling process. The fundamentals go back to Bayes' theorem (also often called Bayes' law) after Bayes and Price (1763) for calculation of conditional probabilities:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{4.13}$$

Here $P(A)$ is the "prior" known probability of occurrence of event A, $P(B)$ is the "prior" probability of occurrence of event B (here used as a normalising constant) and $P(B|A)$ is the conditional probability of the event B for given event A. Thereof $P(A|B)$, the conditional probability of A for given B is calculated. Hence Bayes' theorem describes how the personal belief $P(A)$ can be updated by observing B. A deviation of Bayes' theorem from conditional probability is for example available from Koch (1999).

Although there are many fields of applications in which prior distribution really represents a subjective belief of a single person, in environmental modelling prior distributions should represent more or less impartial knowledge. Hence prior distributions and likelihood functions should be formulated to express commonly accepted expert knowledge. Gillies (1991) suggests that subjective probabilities should be extended to an intersubjective interpretation of probabilities of groups of persons (commonly accepted expert knowledge of the scientific community).

In terms of modelling Bayes' theorem can be written as

$$P(\Theta|D) = \frac{P(D|\Theta) \cdot P(\Theta)}{P(D)} \tag{4.14}$$

with $P(\Theta)$ as prior distribution of a set of model parameters Θ, $P(D)$ as distribution of observations (data) and $P(D|\Theta)$ as conditional probability of observing data D for a given parameter–set Θ (i.e. the likelihood function - see section 4.2). Hence $P(\Theta|D)$ is the probability distribution for parameter–set Θ for given (observed) data D (often called posterior distribution). Furthermore $P(D|\Theta)$ is the updated parameter probability after the imposition of calibration constraints. $P(D)$ again is a normalisation constant that ensures that $P(\Theta|D)$ integrates to unity over parameter space as required for a probability distribution.

Using this formal learning strategy the prior distribution $P(\Theta)$ can be updated each time new data becomes available to finally approach the true joint probability distribution of the parameters. For example Lindley (2006) suggests that probability is the only way to deal with uncertainty in models. Therefore the Bayesian inference can be an efficient approach to deal with uncertainties in urban drainage modelling as long as the different components can be defined adequately (Beven, 2009).

A large number of publications in different fields show the application of Bayesian inference in environmental modelling (e.g. Ellison, 1996; Engeland et al., 2005; Guisan and Zimmermann, 2000; Kanso et al., 2003; Krzysztofowicz, 2002; Krzysztofowicz and Maranzano, 2004; Kuczera et al., 2006; Kuczera and Parent, 1998; Omlin and Reichert, 1999). Nevertheless Bayesian inference has two main point of criticism that it is (1) too casual because it allows subjective prior knowledge and (2) to strict because it requires a formal likelihood function.

4.3.3.1 How to define prior knowledge?

In Bayesian inference learning starts from prior knowledge (prior distribution) for all parameters that are considered to be uncertain. A common point criticism is that this introduces subjectivity to statistics. This is for example expressed in this joke about Bayesian statisticians:

> *A Bayesian is one who, vaguely expecting a horse, and catching a glimpse of a donkey, strongly believes he has seen a mule*

Although there is strong discussion among statisticians about objectivity and subjectivity of Bayesian methods and how to develop “objective Bayes” methods (see for example Berger, 2006; Fienberg, 2006; Goldstein, 2006; O’Hagan, 2006) most users of this method embrace the possibility to consider their expert knowledge (e.g. Beven, 2009; Kuczera et al., 2006). Furthermore the prior distribution is dominated by the likelihood function by repeated application of Bayes’ theorem.

Reichert (2009) demonstrates three possibilities do define prior knowledge:

- based on first principles or data
- information from experts
- vague prior knowledge

Prior knowledge based on *first principle data* is rarely available in urban drainage modelling as systems are too complex. An exception is the use of a posterior distribution from a previous study. As soon as more data gets available Bayesian learning can be continued by using such a posterior distribution as a prior distribution. However this would be equivalent as using the whole dataset at once. Another exception are distributions of parameters that have a physical meaning (e.g. estimated input data). In such a case a distributions derived from measurements can be used.

As already mentioned *expert knowledge* also can be used as prior knowledge in cases when no posterior distribution is available. The best way to use expert knowledge would be not to directly use subjective belief of a single person but to derive prior distributions from common knowledge of the scientific community. In this context Tversky and Kahneman (1974) analysed problems of eliciting distributions from personal beliefs when making judgements under uncertainty which are (1) *representativeness* (people often overestimate the probability of obvious events), (2) *availability* of instances when people are asked to assess the frequency of a certain incident and (3) *adjustment and anchoring* as people tend to conclude distributions from similar better known circumstances.

The third type is formulation of *vague knowledge* as prior distribution. The wider a prior distribution is, the less is its influence on the posterior distribution with an "non-informative" prior as extreme case. This would mean any value of parameter has the same probability. Reichert (2009) criticises that in environmental science often a uniform distribution within a certain interval is used as prior for three reasons:

- The probability density outside that interval is zero and hence the posterior distribution is also always zero, even when there is strong evidence from the data that a parameter should be a value from this range.
- It does not seem reasonable to assume a parameter's value to have the same probability close to the expected value and at the boundary of the interval.
- The characteristics of a prior distribution like mean, median and variance should not be sensitive to the choice of the boundaries of the interval.

On the other hand several authors argue that a uniform distribution within boundaries can be a reasonable choice if the user is aware that in this way only a certain parameter range is analysed (and the posterior distribution can not specify probabilities outside that range).

4.3.3.2 How to chose a likelihood function?

As already mentioned in section 4.2 and section 4.3.2 the definition of the likelihood function is a crucial point in Bayesian inference. Contrary to GLUE a formal definition of the likelihood function is required. In Bayesian inference $P(D|\Theta)$ has to formally represent the probability of predicting data D with given model parameters Θ. This also includes the assumption that the error models used are correct. For different methods of defining likelihood functions see 4.2 or (for more details) relevant literature (e.g. Beven, 2009; Koch, 1999).

4.3.4 Markov-Chain Monte-Carlo Simulation

Urban drainage models are non–linear complex systems. Usually it is not possible to use Bayesian inference analytically, which would require a multidimensional integration of the likelihood function. Hence numerical techniques for solving equation 4.14 are required. Therefore usually Monte Carlo simulations are used wherein samples from the posterior distribution are drawn and properties of the posterior are approximated by properties of the sample. A common method is Markov Chain Monte Carlo Simulation (MCMC) with the Metropolis-Hasting sampler (Hastings, 1970; Metropolis et al., 1953) or the Gibbs sampler. Here sampling after the Metropolis-Hastings algorithm is described as this method is used in **Paper II** and **Paper III** with the software tool

MICA (Doherty, 2003) and also in a row of other publications (e.g. Dotto et al., 2009; Gallagher and Doherty, 2007; Kanso et al., 2003, 2005; Kuczera and Parent, 1998).

A Markov chain is a stochastic process where all conditions necessary for future development of the process are contained in the present state. Hence future states are independent from past states. The idea of a Markov Chain Monte Carlo simulation in conjunction with Bayesian likelihood methods is to develop an effective method of integrating under the likelihood surface with emphasis on sampling model parameters mainly in areas of high likelihood (to be as effective as possible), but also to allow some sampling in areas of low likelihood (to avoid missing other optima). Therefore a Markov chain is a random walk, continuously learning from prior samples. For effective sampling usually multiple chains are simulated simultaneously. An advantage of Markov Chain Monte Carlo simulation is that no assumptions of model linearity or differentiability of model outputs with respect to parameter values are required (Gallagher and Doherty, 2007).

Starting from the current parameter set Θ a new parameter sets Θ^N is sampled from the posterior distribution. The new parameter set Θ^N is accepted if a random number β sampled from a uniform distribution in the interval $[0|1]$ meets the condition

$$\beta < \frac{P(\Theta^N|D) \cdot Q(\Theta^N|\Theta)}{P(\Theta|D) \cdot Q(\Theta, \Theta^N)} \tag{4.15}$$

whereas $Q(\Theta^N|\Theta)$ is the density function describing the transition from the current parameter set Θ to a new parameter set Θ^N. $P(\Theta|D)$ ($P(\Theta^N|D)$) is calculated by multiplying the likelihood function by the prior distribution of Θ (Θ^N) according to (4.14). In case of symmetrical transition distributions used the ratio $\frac{Q(\Theta^N|\Theta)}{Q(\Theta,\Theta^N)}$ is equal to unity. As here only relations of likelihoods are compared it is not required to calculate the denominator in equation (4.14). If the new parameter set is accepted the posterior distribution is updated, otherwise $\Theta^N = \Theta$.

It can be shown that this methods converges to the posterior distribution $P(\Theta|D)$ (Smith et al., 2008; Tierney, 1994). More details about the process is available fro example from Gallagher and Doherty (2007) or Doherty (2003).

Doherty (2003) suggests that for best performance the acceptance ratio should be around $20 - 70\%$. If it is too large it is possible that the range analysed is to narrow for a sufficient inverstigation of parameter space, if it is too small the algorithm is not effective. Therefore the variance of the posterior distribution can be adapted during simulation for best efficciency. For example Vrugt et al. (2003) combined the Metropolis-Hasting sampler with a shuffled complex evolution (SCE) method to get a shuffled complex evolution metropolis (SCEM-UA) algorithm in order to continuously update the proposal distribution.
The application of a formal Bayesian inference on the didactic example is very similar to the application of GLUE. Here the software tool MICA (Doherty, 2003) was used, which is the same as used in **Paper II** and **Paper III**. In MICA Gaussian distributed residuals are assumed in the likelihood function (equation (4.2) and (4.3)) and for parameter sampling the Metropolis–Hastings algorithm (equation (4.15)) is used. Hence the main difference to GLUE is that acceptance of parameter samples is less informal. Here 10 chains were simulated simultaneously starting from 10 different initial points to ensure that the wohle parameter space is investigated. The assumed prior distributions is are uniform distribution

Figure 4.12 shows the accepted parameter samples and the corresponding mean squared error (MSE). As one can clearly see all chains converge to the same point $W = 50.25$ and $b = 0.495$, which is very close to the assumed "true" point of $W = 50$ and $b = 0.5$ before the artificial errors have been applied on the measurement data. Furthermore the range of the parameter values accepted is very narrow compared to results of GLUE.

Figure 4.13 shows the histogram of the accepted parameter samples representing the shape of the posterior probability distribution, which is shown as cumulative distribution in Figure 4.14. Here again the 25% and the 75% percentile of the accepted parameters is shown (red squares).

Consequently again an uncertainty bound of model predictions can be evaluated. This bound showing is presented in Figure 4.15 together with the measurement data. Also here one can clearly see that the uncertainty bound is much narrower compared to the results of the GLUE methodology. This difference results from the difference in acceptance criterion. In the formal Bayesian inference used here the acceptance ratio is controlled by the Metropolis–Hasting sampler and only parameter samples with $E > 0.9$

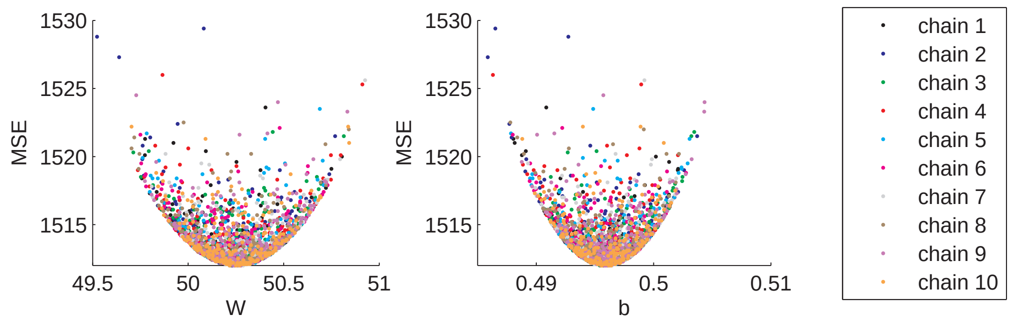

Figure 4.12: MCMC: Accepted parameter samples against MSE - Accepted parameter samples for W (left) and b (right) against mean squared error (MSE) for 10 different chains

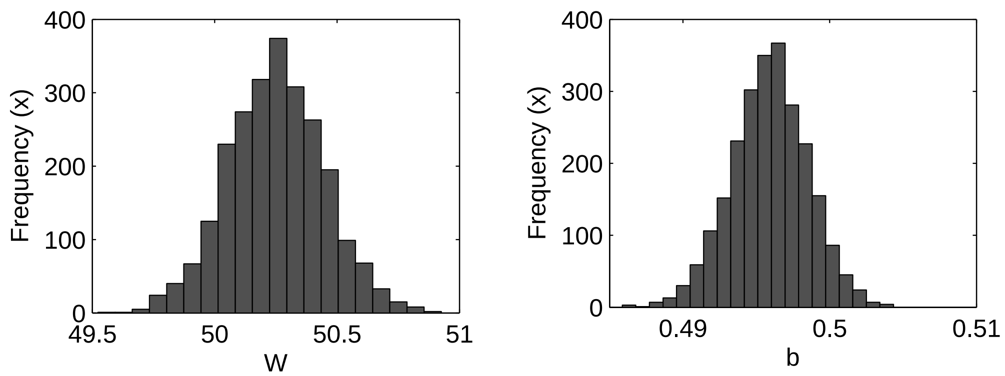

Figure 4.13: MCMC: Histogram of accepted parameter samples - Histogram of accepted parameter samples for W (left) and b (right)

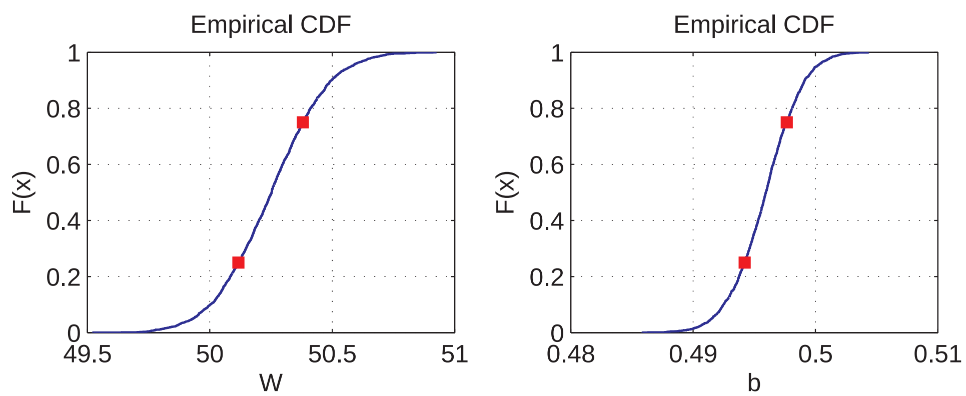

Figure 4.14: MCMC: Cumulative parameter distribution function - Empirical cumulative distribution function of accepted parameter samples for W (left) and b (right). The 25% and the 75% percentile of parameter samples is represented by red squares

were accepted. For GLUE the Nash–Sutcliffe criterion with an acceptability threshold of $E = 0$ was used, hence all parameter values with $E > 0$ were accepted. A further discussion about the impact of the acceptability threshold in GLUE is available from Freni et al. (2008b).

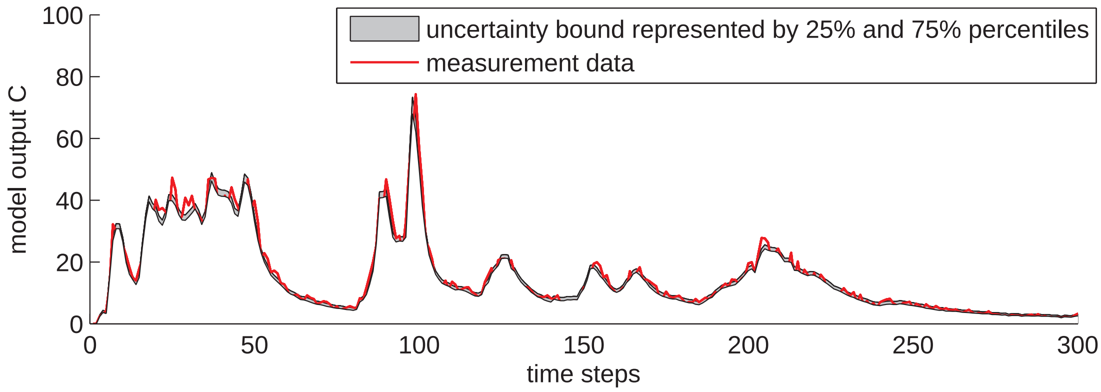

Figure 4.15: Uncertainty bound calculated with MCMC - Uncertainty bound (grey) calculated with MCMC and represented by the 25% and the 75% percentile of the accepted simulation results. The calibration data is represented by the red line.

4.3.5 Bayesian Total Error Estimation (BATEA)

The methods above describe how a posterior distribution $P(\Theta|D)$ of model parameters Θ for given input and calibration data D can be inferred. Therein all sources of uncertainties are implicitly expressed as intrinsic uncertainties of model parameters determined. This does not sufficiently meet requirements for a detailed consideration of different sources of uncertainties as defined in section 4.1. Hence Kavetski et al. (2003, 2006a,b); Kuczera et al. (2006) developed the Bayesian Total Error Analysis (BATEA) framework based on hierarchical Bayesian models in which each source of uncertainty is considered explicitly. Their aim was to develop a robust framework for rational assessment of predictive uncertainty that allows testing of different model hypothesis without confounding different sources of uncertainties (Kuczera et al., 2006).

Figure 4.16 shows a schematic representation of BATEA following a description of Thyer et al. (2009). Here model input errors, model structure errors and calibration errors are regarded to be independent. For example errors in rainfall measurements

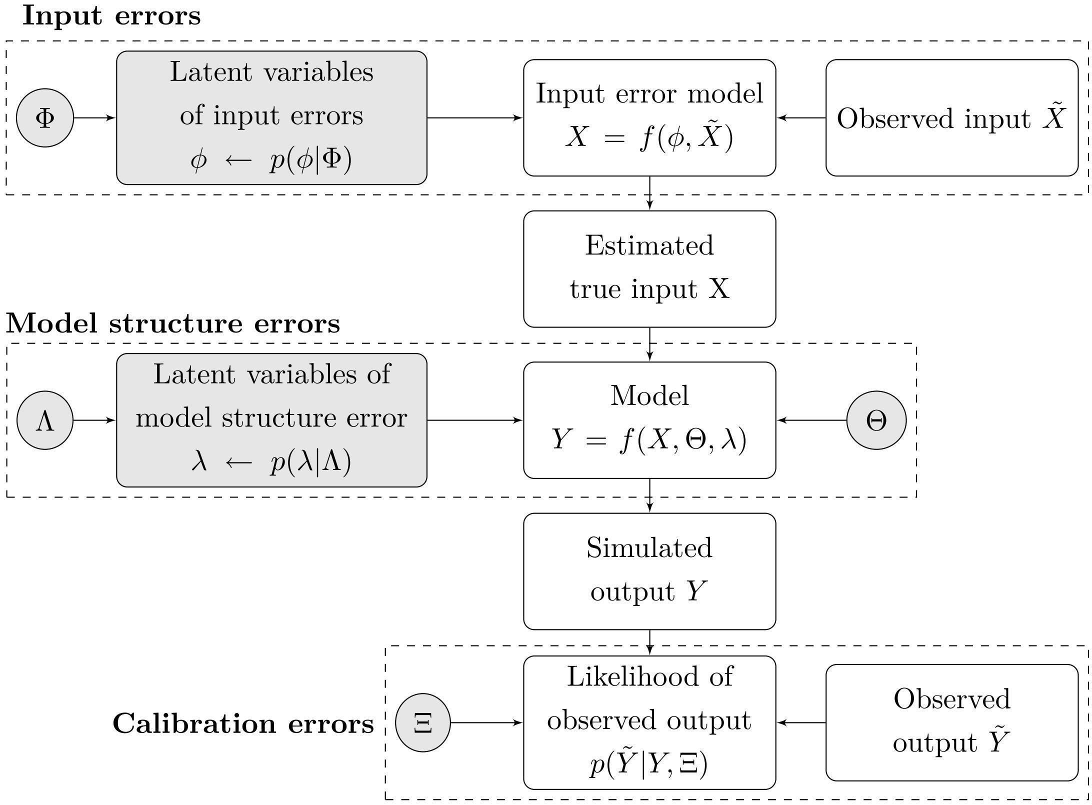

Figure 4.16: Schematic of BATEA

or calibration data should not be compensated by adapting model parameters. Therefore new (latent) variables have to be introduced. The variable ϕ in Figure 4.16 is such a parameter to estimate true input X from given observed input $\tilde{X}$ and an input error model $f(\phi, \tilde{X})$. In terms of hierarchical Bayesian modelling ϕ is a "latent" variable because true input can never be observed and hence ϕ has to be inferred. ϕ is sampled from the "hyperdistribution" $p(\phi, \Phi)$ and Φ are the "hyperparameters" describing the statistical properties of the input errors (for an example see below). For the other sources of uncertainties following this definition Λ describes the properties of model structure uncertainties and Ξ properties of calibration uncertainties, Θ again are model parameters. Consequently model output Y is a function of estimated true input X, model parameters Θ and latent variables of model structure uncertainties λ. In the likelihood function observed output $\tilde{Y}$ is compared with simulated output Y whereas Ξ describes uncertainties in measurement of calibration data.

Thyer et al. (2009), Kavetski et al. (2006b) and Kuczera et al. (2006) present an application of BATEA for a conceptual rainfall / runoff model to infer latent variables of input errors. They assume an input error model in which observed rainfall is disturbed by multiplicative errors, hence for each timestep t true input X_t can be calculated from observed input $\tilde{X}_t$ and a rainfall multiplier $\phi_{i(t)}$ after

$$X_t = \phi_{i(t)} \cdot \tilde{X}_t. \tag{4.16}$$

Furthermore they assume rainfall error follows a lognormal distribution which is characterised by its hyperparameters $\Phi = (\mu, \sigma)$

$$\log \phi_{i(t)} \leftarrow N(\mu, \sigma^2). \tag{4.17}$$

The hyperparameters can be assumed to be known (e.g. if the uncertainties of the measurement device are known), but most of the time they have to be inferred. This could be done independently from modelling process if they can be "calibrated" by comparing data from different sources (e.g. rain gauges and radar measurements) or together with other unknown parameters and hyperparameters in the Bayesian framework. In such a case one would not only get posterior distributions of model parameters, but also posterior distribution of latent variables (characterised by their hyperparameters). Consequently the full posterior distribution is proportional to a combination of the likelihood function, population distribution of latent variables and priors of model parameters and hyperparameters.

$$\underbrace{p(\Theta, \lambda, \Lambda, \phi, \Phi, \Xi | \tilde{Y}, \tilde{X})}_{\text{posterior}} \propto \underbrace{p(\tilde{Y} | \Theta, \lambda, \phi, \Xi, \tilde{X})}_{\text{likelihood}} \underbrace{p(\lambda|\Lambda) p(\phi|\Phi)}_{\text{population distributions of latent variables}} \underbrace{p(\Theta)p(\Lambda)p(\Phi)p(\Xi)}_{\text{priors}} \tag{4.18}$$

For a full deviation see the relevant literature (e.g. Kavetski et al., 2003, 2006a,b; Kuczera et al., 2006; Thyer et al., 2009). Strategies for sampling that high dimensional posterior distribution via Markov Chain Monte Carlo simulation are described for example by Renard et al. (2009).

Of course this methodology still suffers from problems of formulating correct error models for different sources of uncertainties and from problems of inferring parameters and hyperparameters. These problems are points of current research. Additionally a careful diagnostics of the posterior distribution is necessary to ensure that the posterior distribution is not ill posed (see e.g. Thyer et al., 2009). Nevertheless the main advantage is that uncertainties are considered where they happen and hence different parameters should not compensate for each other. This is necessary for a reliable estimation of uncertainties and for propagation of uncertainties on model output. Furthermore this methodology has significant advantages for estimation of predictive uncertainties when circumstances change. For example changes in the availability or quality of data (e.g. use of new measurement devices, installation of more rain gauges) can easily be implemented by adapting the specific error model without the requirement for a complete new calibration of the model. On the other side also the model can be changed, improved or extended without changing error models. This is especially interesting for urban drainage modelling where sewer systems continuously change because of building measures, pavement of areas, rehabilitation of sewers and so forth.

4.4 Uncertainty analysis or calibration (or both)?

The methods presented above aim on determining the most likely model parameters. Unlike during a point-calibration this is not the search for the best parameter set but the search for a distribution of likely model parameters.

This approach has the advantage, that the problems of finding the global optimum of the likelihood response surface (non-Ű uniqueness and nonŰ-identifiability) of an optimum) are avoided or at least mitigated. Non–uniqueness of an optimum is the problem of the lack of a single optimum or in other word there are multiple optima which all lead to a similar good fit between measured data and simulated data. NonŰ-identifiability of an optimum is the lack of a clear optimum or in other words the likelihood surface is rather flat so that the optimum cannot be found. Additionally it is known that the optimum found by search algorithms depends on different boundary conditions as for example the calibration dataset used, the model structure, the goodness–of–fit criterion, the calibration algorithm, and the initial parameter values. Furthermore these methods aim on analysing the shape of the response surface of a model and on the impact of different sources of uncertainties on that shape. Hence impact of uncertainties can be analysed by investigating the response surface as for example done in **Paper II**

. Consequently Bayesian statistics interprets model calibration in a way that the best parameter value does not exist and only is an imaginary value, some values are just more likely than others. This finally results in a distribution of model parameters.

The disadvantage is an enormous computational effort caused by the numerous (thousands) simulation runs required for constructing the posterior distribution of model parameters. Hence in practical applications when models are very complex this would result in an explosion of computing time and for practical model application we are still aiming on finding the best parameter set. Nevertheless uncertainties and their impact on model calibration also have to be regraded. Therefore findings from a deeper investigation of the response surface also can be used to improve point–calibration of a model. Additionally scientific research is also aiming to improve model calibration by working on better search algorithms, taking into account multiple search criteria and evaluating the Pareto optimality of a model (Gamerith et al., 2009; Muschalla et al., 2008). In the next chapter different aspects of model calibration in practical applications aiming on finding the best model parameters are discussed.
As "calibration" is meant as inverse problem of finding parameters with having observations of model input and model output available, both approaches (finding a distribution of parameters and finding the best value) is calibration. Nevertheless when the term "calibration" is used in the next chapter, a point-parameter search is meant.

Chapter 5

Practical model applications

Essentially, all models are
wrong, but some are useful.

George Box

In practical applications a comprehensive uncertainty analysis is often limited by data availability. In addition the methods for uncertainty estimation presented above are scientific methods and require numerous Monte Carlo simulations which are often not possible to be undertaken due to limited working time. Instead a sufficient model calibration is required (Kleidorfer et al., 2006a) depending on data availability.

In this chapter aspects of model calibration in practical model applications are discussed (section 5.1) aiming on finding the "best" parameter set (i.e. that parameter set, that leads to the best fit between model output and measurement data) instead of aiming to find a distribution of parameters. Furthermore the software tool CALIMERO for autocalibration is presented which was developed as part of this thesis (**Paper V**). Examples are shown how urban drainage models are calibrated and used in state–of–the–art guiding rules in Austria for protection from urban flooding and for limitation of CSO discharge for protection of receiving water quality (section 5.2 and 5.3, **Paper VI** and **Paper VII**).

5.1 Model calibration in practical applications

For model calibration measured input data and measured system behaviour has to be available for the same time period. By comparing model output and measured system behaviour the deviation can be calculated and model parameters can be adapted to reduce this deviation as far as possible. Thereby the deviation between measured and simulated system behaviour is calculated by means of an objective function which raises several questions as for example: Which data should be compared (e.g. flow or water levels, fluxes or concentrations)? How many measurement points are necessary for a sufficient model calibration in a spatial distributed system? How should the data be compared (i.e. which mathematical function should be used)?

5.1.1 Calibration data

It is evident that only data can be used for calibration that is also model output. For example a hydrological model cannot be calibrated on waterlevel measurements in the sewers as they are not estimated by the model. However, in some cases it is possible to convert measured data to be consistent with model output. For example Fach et al. (2008b) calculated CSO discharge from waterlevel measurements for calibration of a hydrological model. Therewith the amount of available data can be significantly increased at the expense of additional calibration data uncertainties. Nevertheless – as data availability is often the limiting factor for a sufficient model calibration – such an approach can be very useful.

In **Paper I** and Kleidorfer et al. (2006a) some aspects of the impact of data availability on calibration of a hydrological sewer model are discussed by means of stochastic modelling. For a detailed description see **Paper I**. When collecting data needed for calibration of a hydrological sewer model, it is crucial to pay attention to the selection of measurement sites and to the selection of rainfall events. In the case of spatially distributed systems it is neither possible nor sensible to measure the complete system characteristics and the question is raised about how many measurement sites are necessary. Evaluation of the impact of the number of measurement sites used for calibration shows that the number of required sites is influenced by the time period used for calibration. For example, a similar calibration performance can be reached when using 30% of all available sites for calibration and data over a time period of one year, as compared to using 60% of the sites with five single events. Furthermore a wrong selection of calibration events might result in a complete failure of the calibration exercise.

Mourad et al. (2005) evaluated the impact of calibration data availability on calibration of stormwater quality models and reports that the models are very sensitive to the number of available calibration events. When only few observations are used calibration might lead to a good fit during calibration but to a bad prediction. In order to adequately calibrate stormwater quality models it was required to use the majority (between 60 to 100%) of the available data set during calibration. In an other example McCarthy (2008) analysed differences when calibrating on pollutant concentrations or pollutant fluxes and reported more accurate predictions when calibrating on concentrations. This is thought to be caused by the fact that calibration of fluxes ($Q \cdot c$) is dominated by calibration of flow rates (Q) as parameters for calculating Q are more sensitive as parameters for calculating c.

As impact of calibration data availability is expected to be highly case specific, no general conclusion is possible which data is necessary for a sufficient calibration. Hence apart from model calibration, validation on data not used for calibrating is highly recommended as well as a sentitivity analysis of model parameters with respect to the expected range of input–data.

As sensitivity of model parameters highly depends on the input–data used, this also influences model calibration. It is only possible to calibrate that parameters, that are sensitive to model–output (i.e. insensitive parameters can have any value without much impact on model–output). Hence it is possible that – after model calibration – certain parameters are set to“wrong” values. Because of the minor impact on model results this does not influence predictive capability of a model if the input–data used for prediction is in the same range as input–data used for model calibration. But if the input–data used for calibration is not representative, it is possible that parameters which are insensitive during model calibration become highly sensitive afterwards. An obvious example is calibration of rainfall/runoff simulation of pervious parts of a catchment. Pervious areas only contribute to surface runoff during rainfall events with high rainfall intensity as soon as the infiltration capacity has been exceeded. Hence it is not possible to calibrate the corresponding model parameters during rainfall events with low intensity as that parameters have no impact on model–output. A rainfall/runoff model that is calibrated only on low rainfall events might consequently completely loose predictive capability.

The same effect is possible the other way around. For example parameters responsible for modelling of dry weather flow cannot be calibrated during rainfall events as the runoff is then dominated by surface runoff and the model parameters become insensitive. Hence a model calibrated that way might not be able to reproduce dry weather flow in an accurate way. For the most cases this will be of less relevance as usually the aim of a model is predict rain weather conditions. Nevertheless the user has to be aware of the limits of a model and its calibration.

For example this point is a particular problem for models which are used to predict extreme conditions or low–frequency events. Such events hardly can be calibrated and hence uncertainty of model prediction increases. Nevertheless a sensitivity analysis is highly recommended for any model applications. Additionally further research is required to analyse this topic with respect to different models.

5.1.2 Objective function

In equation 5.1 to 5.5 some commonly used measures for the "goodness–of–fit" between measured data points M_i (with their mean value $\bar{M}$) and simulated data points S_i (with their mean value $\bar{S}$) are presented.

The *Bias* B is one of the simplest ways to quantify the deviation between measurement and simulation. It is defined as the ratio of the mean of the simulation results to the mean of the measurement and ranges from $-\infty$ (if negative results are possible) to $+\infty$. It only accounts for a systematic error but can be useful because it indicates the quality in meeting the mass balance.

$$B = \frac{\bar{S}}{\bar{M}} \qquad]-\infty|+\infty[\qquad \rightarrow 1 \tag{5.1}$$

The *Coefficient of determination* R^2 is the square of the correlation coefficient and describes the total variance in the measured data that can be explained by the model and ranges from 0 to 1. It is the square of the Pearson correlation coefficient C (ranging from -1 to 1) and it only evaluates the linear relationship between measurement and simulation output. Hence R^2 is insensitive to additive and proportional errors (Achleitner, 2008; Legates and McCabe Jr, 1999). Additionally R^2 is known to overestimate impact of outliers compared to values near the observation mean (Legates and Davis, 1997).

$$R^2 = \left(\frac{\sum_{i=1}^{n}(M_i - \bar{M}) \cdot (S_i - \bar{S}))}{\sqrt{\sum_{i=1}^{n}(M - \bar{M})^2} \cdot \sqrt{\sum_{i=1}^{n}(S_i - \bar{S})^2}} \right)^2 \qquad [0|1] \qquad \rightarrow MAX \qquad (5.2)$$

The *Mean Squared Error* MSE incorporates bias and variance . It has the same unit as the square of the data being compared and ranges from 0 to $+\infty$. Hence it is difficult to evaluate from the value obtained how good the agreement between measurements and model output is.

$$MSE = \sum_{i=1}^{n}(M_i - S_i)^2 \qquad [0| + \infty[\qquad \rightarrow MIN \qquad (5.3)$$

Therefore several possibilities of normalising the MSE can be applied.

The *Nash–Sutcliffe–Efficiency* E (Nash and Sutcliffe, 1970) is also often called Coefficient of Efficiency and represents the ratio of MSE to the variance in the observed data, subtracted from unity and ranges from $-\infty$ to 1. Values $E < 0$ indicate that the mean of the observed data is a better prediction than model output. E is an improvement to d (see below) as it is sensitive to differences in the mean values of model output and measurement data. However, it still is more sensitive to outliers (peaks in the timeseries) as to values near the mean.

$$E = 1 - \frac{\sum_{i=1}^{n}(M_i - S_i)^2}{\sum_{i=1}^{n}(M_i - \bar{M})^2} \qquad] - \infty|1] \qquad \rightarrow MAX \qquad (5.4)$$

The *Index of agreement* d (Willmott, 1981; Willmott et al., 1985) varies from 0 to 1. Similar to Nash–Sutcliffe–Efficiency it is also very sensitive to outliers.

$$d = 1 - \frac{\sum_{i=1}^{n}(M_i - S_i)^2}{\sum_{i=1}^{n}(|M_i - \bar{M}| + |S_i - \bar{M}|)^2} \qquad [0|1] \qquad \rightarrow MAX \qquad (5.5)$$

Legates and McCabe Jr (1999) compare different measures for the “goodness–of–fit” and show that correlation–based indicators as the coefficient of determination should not be used for evaluating the agreement between model output and measurement data. For a further comparison of the different performance indicators see Achleitner et al. (2008) or Grecu and Krajewski (2000).

5.1.3 Calibration algorithms

During calibration the objective function is optimised (minimised or maximised) by adapting calibration parameters. While in manual calibration this is done by "trial and error", autocalibration tools contain search algorithms trying to find the global optimum. From the mathematical point of view calibration is an inverse optimisation problem of estimating optimal parameter values where numerous algorithms exist as for example genetic evolution algorithms (Rauch and Harremoës, 1998b), the maximum likelihood estimator, the generalised least squares estimator or the weighted least squares estimator just to listen a few.

A commonly used algorithm is the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) for nonlinear optimisation, which is also implemented in the auto-calibration tool PEST Doherty (1999) or CALIMERO (**Paper VII**). The Levenberg-Marquardt provides a numerical solution for minimising least square errors between measured data M and simulated data S over n time steps to find the best possible set of calibration parameters Θ:

$$Error(p) = \sum_{i=1}^{n}(M_i - S_i)^2 \rightarrow MIN \tag{5.6}$$

As simulated data can be expressed as a function of model input X and model parameters Θ, equation 5.6 can be written as

$$Error(p) = \sum_{i=1}^{n}(M_i - f(X_i, \Theta))^2 \rightarrow MIN. \tag{5.7}$$

A detailed description of the Levenberg-Marquardt algorithm is available from Moré (1978).

As the choice of the performance indicator (i.e. the objective function which is minimised during calibration) is essential, modern auto-calibration algorithms are based on multi-objective calibration algorithms (e.g. Madsen, 2000; Muschalla, 2006). Muschalla (2006) tested multi–objective evolution strategies for optimisation of water resources systems and developed an integrated optimisation and simulation tool for finding numerous solutions on the Pareto-optimum front. He tested this calibration algorithm on an integrated urban drainage model including natural and urban catchments, a sewer system, a WWTP and a water body and minimised investment costs and receiving water quality (dissolved oxygen concentration, ammonia concentration and frequency of high flood events) likewise. While this study was mainly focused on system optimisation the same technology can also be used for autocalibration as shown by Gamerith et al. (2009, 2008); Muschalla et al. (2008).

5.1.4 The Autocalibration Tool CALIMERO

CALIMERO is a software tool written in C++ using Qt libraries (Nokia, 2009) and it is designed to be used with nearly any computer model. The only requirements are that (a) the model can be run over command line without a graphical user interface and (b) that model input and output files are plaintext (i.e. not an encrypted or binary file format). The software architecture of CALIMERO with its interfaces to model and data is shown in Figure 5.1.

All relevant data for simulations can be imported into an internal database, including: model input data (e.g. rainfall data), calibration data (i.e. observed data such as flow measurements) and system data, including calibration parameters. Additional knowledge about system performance (e.g. information about measurement uncertainties and data collection) should also be considered during calibration and has to be described mathematically. Hence it is possible to add the information if, for example, certain datasets are highly reliable and have been collected carefully or if they are estimated roughly from old projects.

Fuzzy knowledge is rarely considered during model calibration, but should be added to the model calibration framework (being expressed mathematically) since this type of information is often essential to a successful model calibration. An example of fuzzy knowledge is the occurrence of combined sewer overflow discharges at a specific point in the system (i.e. “frequently” or “seldom”), often provided by sewer system operators.

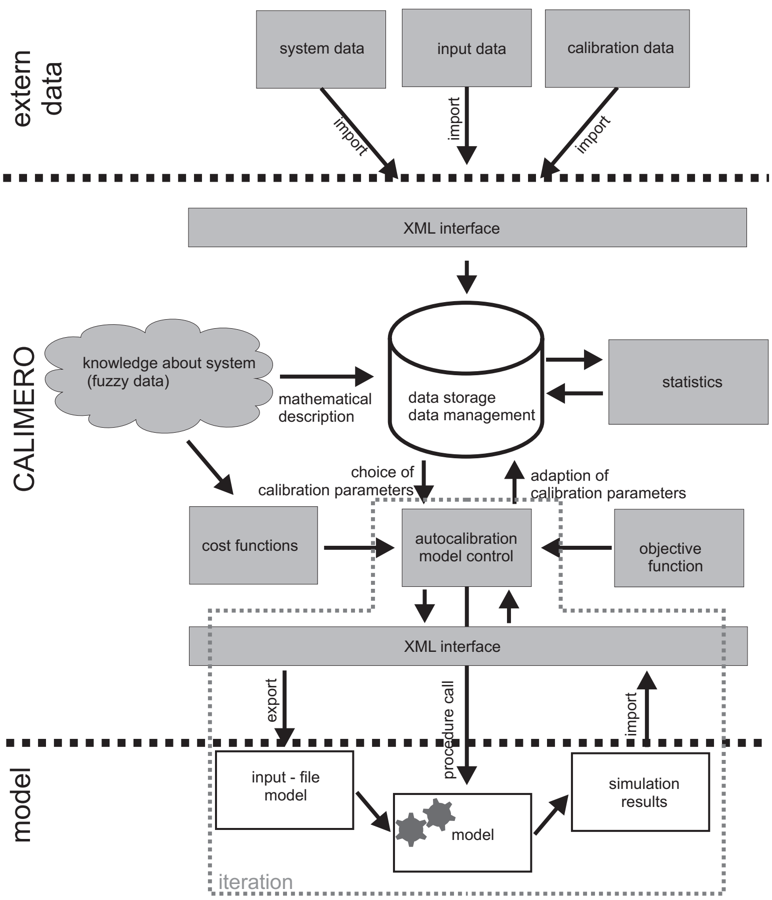

Figure 5.1: Concept of CALIMERO - Software architecture and interfaces

The objective function(s) (i.e. one or more values that are optimised during autocalibration) and the autocalibration algorithm itself can be defined via the script engine or selected from a predefined set.

5.1.4.1 Model and data interface

Import of model input data, system data and calibration data is possible via a predefined xml interface which has either to be prepared by the user prior to autocalibration or can be configured in CALIMERO. Therefore, the model input-files, a template of simulation results, calibration data and additional boundary conditions can be imported into CALIMERO in the same format as they are used by the model. Parameter names can be assigned to relevant values from the imported files for further use in the calibration script (see Figure 5.2). Simultaneously, templates for the model input-files are created the same way. If certain values from the model-input file are defined as calibration parameters, they are replaced during the calibration process prior to each iteration to test the new parameter values. Simulation results are defined in the same way: after assigning parameter names, these specified values are subsequently read from the simulation results in each iteration run and evaluated by the calibration scripts.

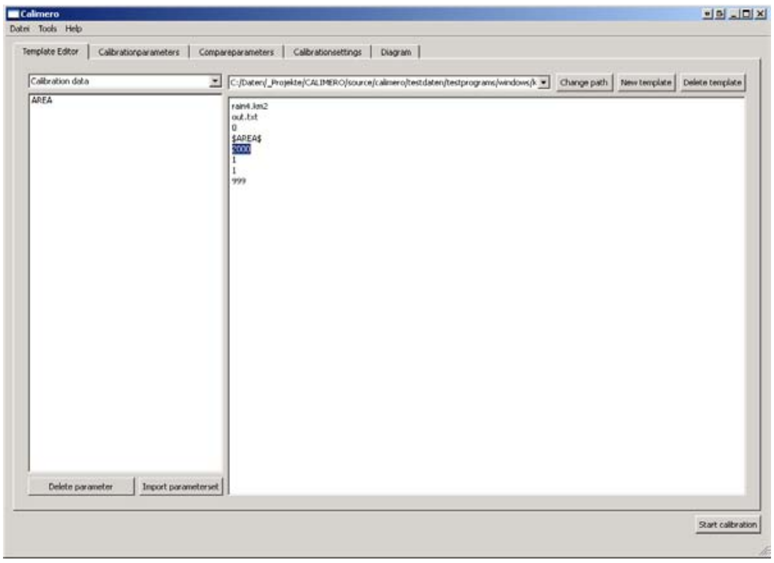

Figure 5.2: CALIMERO Screenshot - Definition of model parameters

5.1.4.2 Script integration

A drawback of many autocalibration tools is that they include one specific autocalibration algorithm and one specific objective function (e.g. most commonly minimisation of squared errors) which cannot be changed unless the user may change the source code.

In CALIMERO, objective functions and calibration algorithms are defined in a script engine to provide best possible flexibility. The scripting language follows ECMA / JavaScript specifications (ECMA-262, 1999) as this is a rather simple scripting language designed for non-programmers to work with. Due to its wide use in client side website programming, there are a lot of tutorials and manuals available. This standardised scripting language shall encourage the exchange of calibration scripts among different users. CALIMERO comes with a script editor and a script debugger for development and testing of algorithms (see Figure 3).

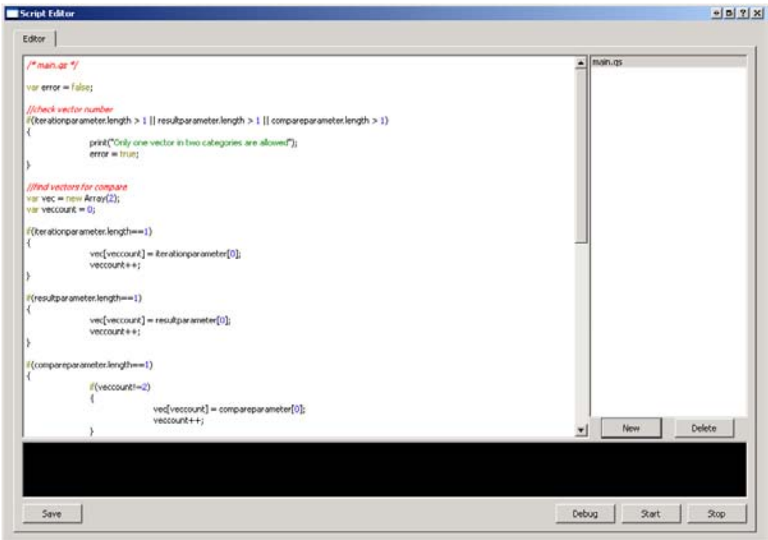

Figure 5.3: CALIMERO Screenshot - Script integration

5.1.4.3 Consideration of a-priori knowledge and boundary conditions

The term "a priori knowledge" does not completely correspond with the terminology of statistics and Bayesian inference in the sense that it describes an assumed, but mathematical exact, probability distribution. Here "a priori knowledge" is meant as additional, sometimes diffuse information about system behaviour and data accuracy.

Such information is often available from sewer system operators but hardly used in autocalibration (in contrast to manual calibration). For example when modelling a spatial distributed sewer system, data collection is mostly not homogeneous for the whole system. In certain areas it might have been carried out with, for example, a detailed, up-to-date examination of aerial photos and cadastral surveys for determining the fraction imperviousness in an accurate way. Data from other regions might come from former, and possibly outdated, investigations of vague origin.

Other examples are measurement devices which are known to record partly inaccurate data. A common practice is to completely exclude such doubtful data from calibration in order to not distort model calibration. But even such information can improve calibration, especially when working with badly defined systems under limited data availability (Kleidorfer et al., 2008). Additionally, measurement devices are calibrated for a specific data range (e.g. high waterlevels) and measurement uncertainties increase when recording data-points outside that range (e.g. very low waterlevels). In order to not consider less reliable data points, often a manual data processing is necessary. An exclusion of this less-reliable data points directly in the calibration algorithm itself helps the model user and reduces the effort for model calibration, especially when testing different calibration strategies.

By adapting algorithms for calibration and objective function evaluation, different data sources can be considered with different weights. Hence, all available data can be taken into account whereas reliable data-sets (or reliable ranges of measurement points) dominate autocalibration and less reliable data-sets are considered as additional information. Muschalla et al. (2008) present an application of multi-objective autocalibration, in which they conclude that multi-objective algorithms are highly sensitive to erroneous data. They expect an improvement in autocalibration by adapting the calibration algorithm to consider different objective functions. As the calibration algorithms are included in CALIMERO via script engine such adaptations are also possible for users who are not familiar with low level programming.

5.1.4.4 Calibration process and post–processing

During autocalibration the progress can be observed as all calibration parameters as well as the objective function evaluations are plotted (see Figure 5.4). The calibration is stopped if the deviation between measurement and calibration data (expressed in the objective function) falls below a user–defined threshold or is cancelled by the user. For post processing all parameter sets tested during autocalibration as well as the corresponding model results and evaluations of the objective function are stored in a database. From here any iteration can be exported to create the corresponding model–input file.

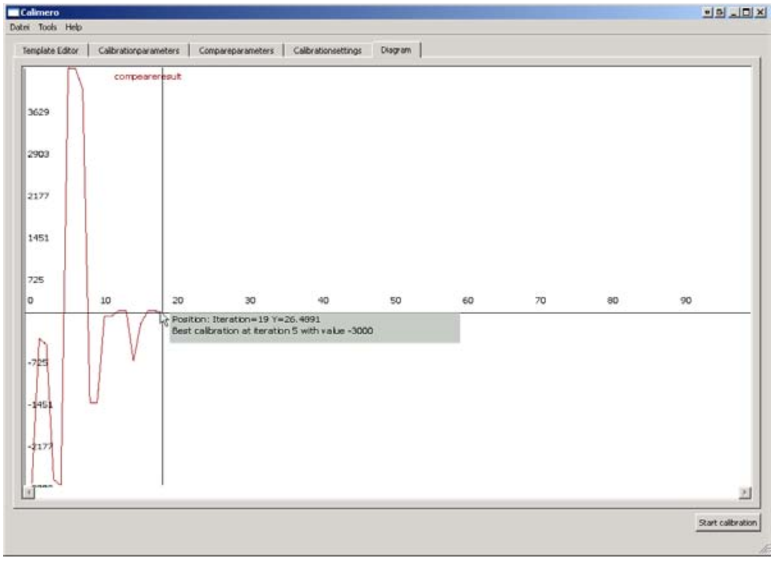

Figure 5.4: CALIMERO Screenshot - Evaluation of calibration process

5.2 UDM for flood protection

5.2.1 Legal requirements in Austria

The Austrian guiding rule ÖWAV-RB 11 (2009) is the national implementation of the standard ÖNORM EN 752 (1997) and describes the hydraulic calculation of sewer systems to prevent flooding in urban areas due to capacity overload. The main difference of that two guidelines is that ÖNORM EN 752 (1997) requires flooding proof while ÖWAV-RB 11 (2009) requires surcharge proof, which is easier to model. Surcharge happens when the water level in the sewers reaches the soil level and water starts flowing out of the system. Flooding is associated with any damage caused by this surcharge. This difference is reflected in different return periods used for the assessment (see table 5.1).

Table 5.1: Required design frequency in EN 752 and Regelblatt 11

Urbanisation category	Return period EN 752	Return period RB 11
rural areas	10	2
residential areas	20	3
industrial areas	30	5
underground transport facilities	50	10

For all but very simple systems a sufficient design of sewers is documented by means of hydrodynamic simulation either by using design storm events or historical time series. More detailed information on ÖWAV-RB 11 (2009) is available from De Toffol (2009).

In section 5.2.2 and 5.2.3 the application of ÖWAV-RB 11 (2009) is shown for the combined systems of the two cities Innsbruck and Linz.

5.2.2 Case study Innsbruck

5.2.2.1 System description

For Innsbruck (see section 3.1) a hydrodynamic model was set up using the software Hystem-Extran (ITWH, 2002). Figure 5.5 shows the system layout of simplified system and Table 5.2 presents some system characteristics. For a detailed system description see Kleidorfer (2005).

Table 5.2: System characteristics Hydrodynamic Model Innsbruck

Parameter	Value
pipe length	74 km
nodes	296
subcatchments	200
total area	2500 ha
impervious area	774 ha
inflow to WWTP	2.2 m^3/s

Figure 5.5: Hydrodynamic model Innsbruck - Hydrodynamic sewer model of Innsbruck

5.2.2.2 Model calibration

The model was calibrated by comparing measured and simulated water levels at selected measurements sites. Figure 5.6 shows the comparison of measured and simulated water levels for the rainfall event between 2006/08/03 and 2006/08/05 at the measurement sites “Vögelebichl” and “Mariahilferstrasse” as an example. Although this comparison shows a rather good agreement it is important to note, that there are significant differences in calibration performance for the same model when regarding different rainfall events. An example in which measured and simulated water levels do not agree at all is shown in Figure 5.7. Here waterlevels are plotted for the same measurement sites but for the rainfall event from 2006/03/05 until 2006/05/06. It is expected that in this case the difference origins from the fact that measured precipitation was snowfall and hence no surface runoff occurred.

This is an example for both (a) input–data uncertainties (because snowfall is not reflected in precipitation measurement) and (b) model–structure uncertainties (because no snow accumulation and melt model is implemented). In such a case the model user has to decide whether such discrepancies are acceptable or not depending on the aim of the study. In this case the hydrodynamic model was set up in order to evaluate the sewer system’s capacity. As in Alpine regions the relevant rainfall events with high intensities occur mainly as thunderstorms during the summer period a correct representation of precipitation events in winter is of minor interest.

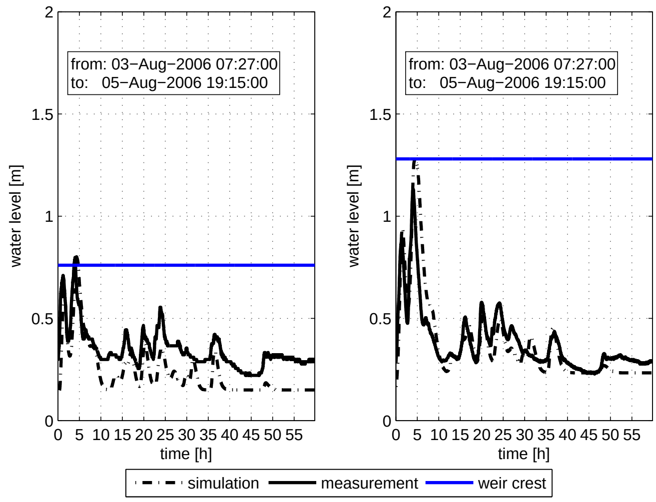

Figure 5.6: Hydrodynamic model Innsbruck: Calibration - Comparison between measured (solid line) and simulated (dashed line) water levels for the rainfall event 03. Aug - 05. Aug. 2006

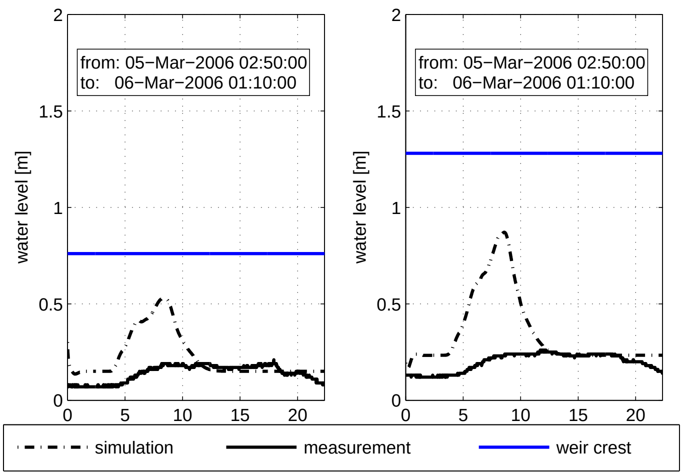

Figure 5.7: Hydrodynamic model Innsbruck: Calibration - Comparison between measured (solid line) and simulated (dashed line) water levels for the rainfall event 05. Mar - 06. Mar. 2006

But the reason for discrepancies between measured data and model output is not always so easy to find. Figure 5.8 shows the comparison between measured and simulated water levels for the rainfall event from 2006/09/16 to 2006/06/17. As one can clearly see measured water levels are much higher than simulated ones. Insufficient measurement of snowfall cannot be the reason in September. It is expected that here spatial rainfall distribution causes that difference.

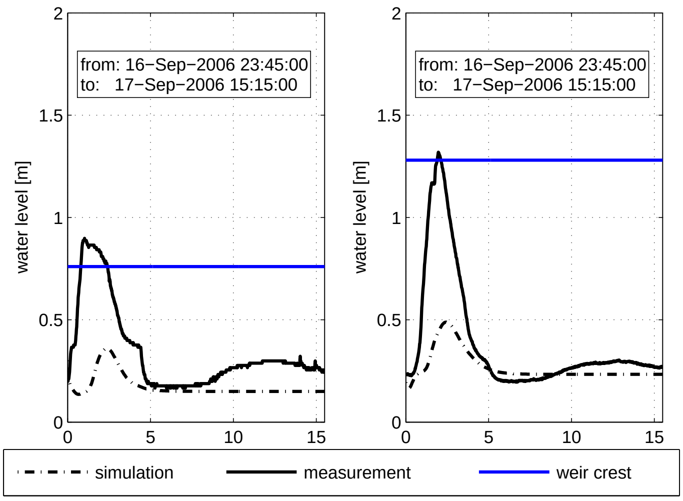

Figure 5.8: Hydrodynamic model Innsbruck: Calibration - Comparison between measured (solid line) and simulated (dashed line) water levels for the rainfall event 16. Sept - 17. Sept. 2006

5.2.2.3 Model results

After model calibration simulations for evaluating surcharge for different return periods with measured rainfall data and with design storms EulerII are undertaken. Figure 5.9 shows for example evaluated simulation results (areas where surcharge occurs) for a return period of r=5 [1/a] from the year 2006 and the urbanisation category according to Table 5.1.

As one can clearly see in the urbanisation category "city centre" (yellow) the requirements of ÖWAV-RB 11 (2009) are not met. Taking these results as starting point

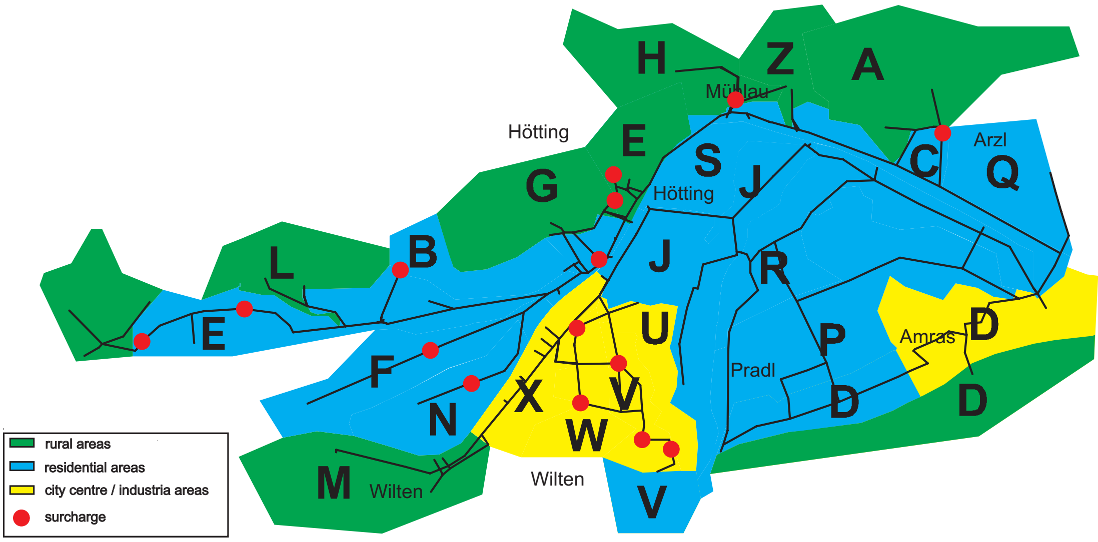

Figure 5.9: Hydrodynamic model Innsbruck: Simulation results - Surcharge

more detailed calculations by using more detailed hydrodynamic models can be done for areas where surcharge occurs.

5.2.3 Case study Linz

5.2.3.1 System description

The hydrodynamic model used for evaluating sewer system performance in Linz (see section 3.2) was set up using the software Mike-Urban (DHI). Figure 5.5 shows the system layout of simplified system and Table 5.2 presents some system characteristics. A system description is also available in **Paper VII** or from Möderl (2009) or Fach et al. (2008a).

Table 5.3: System characteristics Hydrodynamic Model Innsbruck

Parameter	Value
pipe length	378 km
nodes	397
subcatchments	192
total area	12709 ha
impervious area	3608 ha
inflow to WWTP	4.7 m^3/s

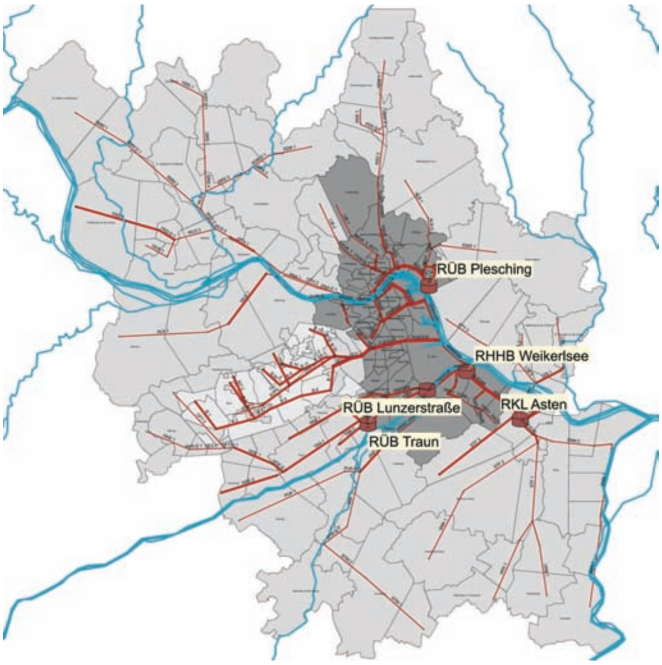

Figure 5.10: Hydrodynamic model Linz - Hydrodynamic sewer model of Linz

5.2.3.2 Model calibration

The hydrodynamic model of Linz is also calibrated by comparing measured and simulated waterlevels at the five measurement sites available. Figure 5.11 shows the comparison of measured (grey) and simulated (black) waterlevels for the rainfall events from 2004/08/17 to 2004/08/18, where a good agreement can be seen.

But also for Linz not all events could be calibrated sufficiently. Figure 5.12 shows the comparison of measured and simulated water levels for the rainfall event from 2005/01/20 to 2005/01/23. Here one can clearly see that the measured surface runoff is much higher (i.e. the runoff wave last longer) as estimated by simulation. It is expected that this is caused by snow melting processes which are not implemented in the software. When rainfall occurs while snow has accumulated on the surface the rather warm rainfall causes snow melting and an increased surface runoff. Hence that are model–structure uncertainties.

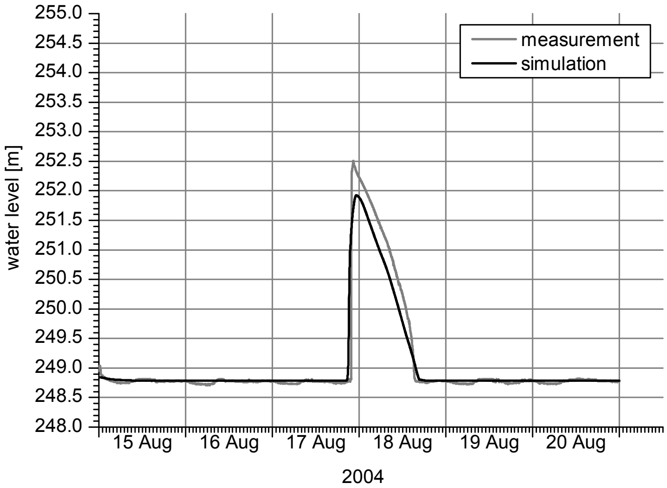

Figure 5.11: Hydrodynamic model Linz: Calibration - Comparison between measured (grey line) and simulated (black line) water levels for the rainfall event 17. Aug - 18. Aug. 2004

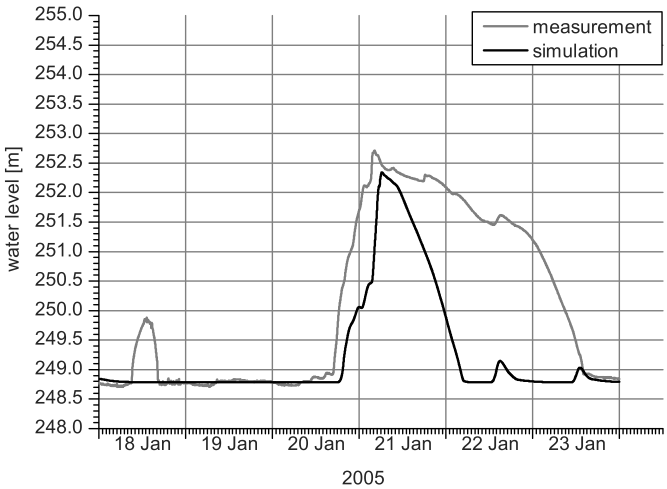

Figure 5.12: Hydrodynamic model Linz: Calibration - Comparison between measured (grey line) and simulated (black line) water levels for the rainfall event 20. Jan - 23. Jan. 2005

5.2.3.3 Model results

Simulation results were evaluated statistically to identify areas where surcharge occurs during strong rainfall events. Figure 5.13 shows results for return periods n=50a (green), n=50a (yellow) and n=10a (red). Hence green coloured catchments are safest; here surcharge happens statistically only every 50 years whereas surcharge at red coloured catchments happens statistically every 10 years. Of course these results also contain uncertainties, which are expected to be higher for estimating such high return periods. For a discussion about return period uncertainties refer for example to Thorndahl (2008).

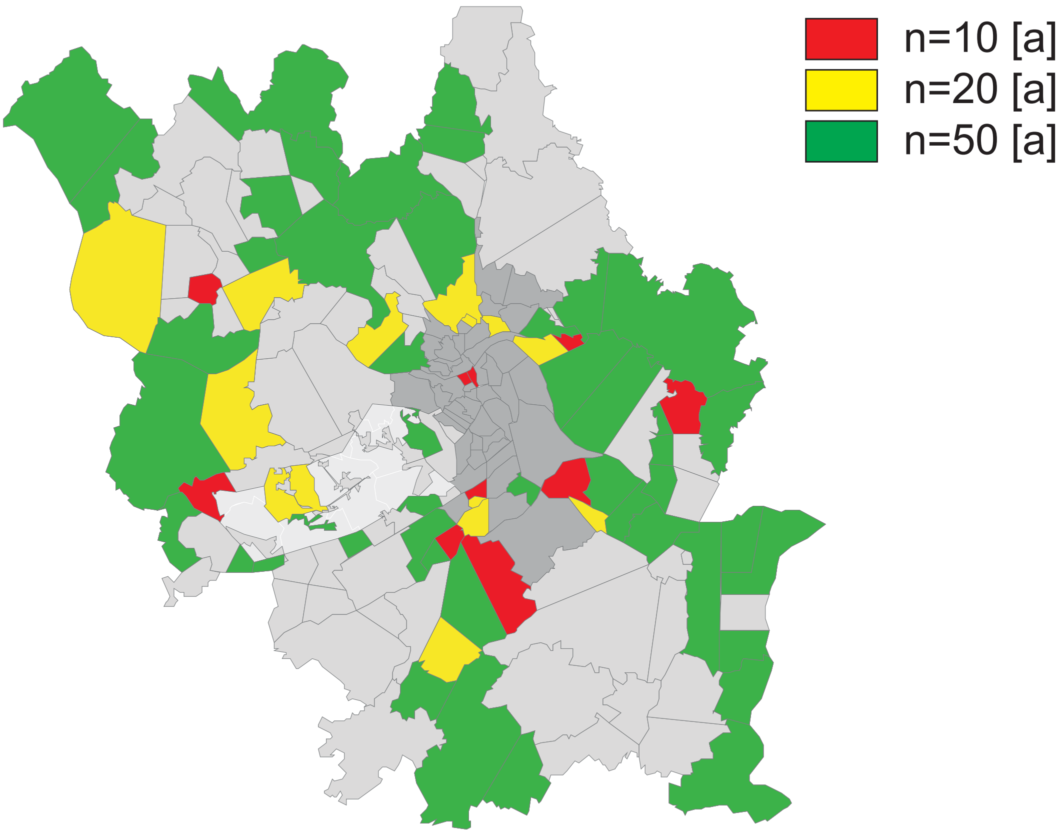

Figure 5.13: Hydrodynamic model Linz: Simulation results - Surcharge for return periods n=10a (red), n=20a (yellow) and n=50a (green)

5.3 UDM for prevention of receiving water pollution

5.3.1 Legal requirements in Austria

In Austria legal requirements for preventing receiving water pollution from combined sewer systems are defined in the Austrian standard ÖWAV-RB 19 (2007) and regulate design of CSO detention basins. Therein emissions from all CSOs of the entire sewer system connected to one WWTP are regarded together. A closer observation of discharge from single CSOs is of minor interest. ÖWAV-RB 19 (2007) introduces the efficiency of combined sewer overflows (CSO efficiency η) for dissolved pollutants η_d and particulate pollutants η_p as an indicator for CSO pollution. Thereby η is the part of the surface runoff treated at the WWTP and it is expressed in percentage. Consequently the CSO efficiency can be calculated after equation 5.8.

$$\eta = \frac{(VQ_c - VQ_d) \cdot c_c - VQO \cdot c_o}{(VQ_c - VQ_d) \cdot c_c} \cdot 100 = \frac{VQR \cdot c_c - VQO \cdot c_o}{VQR \cdot c_c} \cdot 100 \tag{5.8}$$

with

η	CSO efficiency	[%]
VQ_c	Total volume of the combined sewage	[$\mathrm{m^3\,a^{-1}}$]
VQ_d	Total volume of dry weather flow	[$\mathrm{m^3\,a^{-1}}$]
VQR	Total volume of surface runoff	[$\mathrm{m^3\,a^{-1}}$]
VQO	Total volume of overflow discharge	[$\mathrm{m^3\,a^{-1}}$]
c_c	Pollutant concentration in combined sewage	[$\mathrm{mg\,l^{-1}}$]
c_o	Pollutant concentration in overflow discharge	[$\mathrm{mg\,l^{-1}}$]

The required values for η_d and η_p depend on the design basis of the WWTP in population equivalents (PE) and the statistical rainfall intensity with a duration of 12 hours and return period once per year ($r_{720,1}$) for the investigated catchment. The requirements for η_d are shown in table 5.4 and the requirements for η_p are shown in table 5.5. For values not mentioned in the table interpolation should be used.

Table 5.4: Required CSO efficiency for dissolved pollutants

required η_d	**design basis of the WWTP (PE)**	
rainfall intensity	≤ 5000	≥ 50000
$r_{720,1} \leq 30mm/12h$	50	60
$r_{720,1} \geq 30mm/12h$	40	50

Table 5.5: Required CSO efficiency for particulate pollutants

required η_p	**design basis of the WWTP (PE)**	
rainfall intensity	≤ 5000	≥ 50000
$r_{720,1} \leq 30m$m/12h	65	75
$r_{720,1} \geq 30m$m/12h	55	65

The calculation of η requires long-term simulation (duration of at least 10 years) by means of either hydrological models or hydrodynamic models with a temporal resolution of rainfall data of 10 minutes or higher. The calculation is based on the assumption of constant pollutant concentration in time and along the sewer system. Furthermore for calculation of η_d a perfect mixture of wastewater and stormwater is assumed. Hence the concentrations c_c and c_o from equation 5.8 are the same and η_d can be calculated as

$$\eta_d = \frac{VQR - VQO}{VQR} \cdot 100. \tag{5.9}$$

For calculation of η_p the removal of sediments is expressed by the mean sedimentation efficiency η_{sed} which can be calculated after

$$\eta_{sed} = \frac{c_{c,CSO} - c_o}{c_{c,CSO}}. \tag{5.10}$$

and consequently η_d is calculated from η_d, VQO and η_{sed} for each CSO j after

$$\eta_p = \eta_d + \frac{\sum_j VQO_j \cdot \eta_{sed,j}}{VQR}. \tag{5.11}$$

As η_{sed} is difficult to determine ÖWAV-RB 19 (2007) presents typical values for sedimentation efficiency depending on the specific volume of the CSO structure (see Table 5.6).

Table 5.6: Sedimentation efficiency after ÖWAV-RB 19 (2007)

specific volume [$m^3\,ha^{-1}$]			
Hydrodynamic separator	Basin	In pipe storage with overflow downstream	η_{sed} [%]
0	0	0	0
3	5	10	20
7	10	20	35
>10	>15	>30	50

Apart from that emission based requirements in ÖWAV-RB 19 (2007) also criteria for the ambient water quality are defined, which comprehend six kinds of impacts:

- *Hydraulic impact.* To prevent impact on the biocoenosis the maximum CSO discharge with a return period of one year (Q_1) should be smaller than 10% to 50% of the maximum water discharge in the river with return period once per year (HQ_1)

$$Q_1 \geq 0.1 \text{ to } 0.5 \cdot HQ_1 \tag{5.12}$$

- *Acute ammonia toxicity.* The ammonia (NH_3) concentration depends on the ammonium (NH_4) concentration and on the dissociation equilibrium between NH_3 and NH_4 (which is influenced by temperature and pH–value). For salmonid water courses the NH_4 concentration calculated for one hour duration should not be higher than 2.5 mg/l (equivalent to 0.1 mg/l NH_3 with a pH-value of 8 and a temperature of 20 °C) and not higher than 5.0 mg/l (equivalent to 0.2 mg/l NH_3) for cyprinid water courses. The NH_4 concentration can be calculated after

$$c_{r,d} = \frac{Q_{r,u} + Q_d \cdot c_d \cdot \frac{Q_o}{Q_i+Q_o} + (Q_i + Q_o - Q_d) \cdot c_r \cdot \frac{Q_o}{Q_t+Q_o}}{Q_{r,u} + Q_o} \tag{5.13}$$

with

$c_{r,d}$	NH_4 conc. in the receiving water downstream the CSO discharge	[mg/l]
$c_{r,u}$	NH_4 conc. in the receiving water upstream the CSO discharge	[mg/l]
c_d	NH_4 conc. in the dry weather flow	[mg/l]
c_r	NH_4 conc. in the surface runoff	[mg/l]
$Q_{r,u}$	mean low water flow in the receiving water (MNQ)	[l/s]
Q_o	CSO discharge	[l/s]
Q_i	interceptor capacity	[l/s]
Q_d	dry weather flow	[l/s]

- *Oxygen concentration.* The oxygen concentration in the receiving water downstream of the CSO discharge should not be lower than 5 mg/l. As calculation of oxygen depletion can hardly be calculated based on point emissions from urban drainage modelling, according to ÖWAV-RB 19 (2007) no further investigation is necessary if no anaerobe conditions occur during dry weather flow and if the slope of the river is higher than 3–5 m/km. Otherwise deeper investigations including measurements in the water course are necessary.

- *Solids.* For total suspended solids in the receiving water values should not be higher than 50 mg/l. According to ATV-AGă2.1.1 (1993) this limit is expected to be observed if the ratio between population and mean low water flow in the receiving water (MNQ) is less than 25 PE/(l/s)

- *Hygienic impact.* The European Bathing Water Directive 2006/7/EC (2006) restricts faecal coliforms. For inland waters a limiting value of 500 CFU/100ml for *Escherichia coli* and 500 CFU/100ml for Intestinal enterococci (both based upon a 95–percentile evaluation) is required to reach the classification as "excellent quality". As the *Escherichia coli* concentration in combined sewage is much higher (10^5 to 10^7 CFU/100ml), in case of CSO discharge this limit cannot be kept and degradation takes several days. Hence for bathing water CSO discharge should be limited as far as possible.

- *Aesthetics.* At receiving water bodies, which are sensible to aesthetic impacts, rejects should be avoided by adding filters and racks at CSOs.

Further description of the requirements of ÖWAV-RB 19 (2007) is available from De Toffol (2009), Kleidorfer et al. (2006a) or Kleidorfer et al. (2008), where De Toffol (2009) also compares legislation and technical guidelines of 17 different counties.

5.3.2 Case study Innsbruck

5.3.2.1 System description

As described by De Toffol et al. (2006b) the calculation of CSO efficiency η requires long term simulations (duration of at least 10 years) by means of either hydrological models or hydrodynamic models. Taking the complexity of hydrodynamic calculations into account it is evident that computing time can easily reach extreme amounts (for a given test system several weeks). In contrast, a hydrological model of the same sewer system was calculated in some minutes (De Toffol et al., 2006b), which is due to the more simplified calculation method (Rauch et al., 2002). Therefore, the use of a hydrological model is an obvious choice for determining CSO pollution, if it can be set up in short time and can be calibrated accurately.

Figure 5.14 shows the hydrological model of Innsbruck represented by the software KAREN which is described with much detail by Rauch and Kinzel (2007). A description of the rainfall / runoff model used here is also available in **Paper II** and **Paper III**. In Table 5.7 the system characteristics are described.

Table 5.7: System characteristics Hydrological Model Innsbruck

Parameter	Value
subcatchments	35
CSOs	35
impervious area	781 ha
storage volume basins	5100 m^3
storage volume sewers	27360 m^3
inflow to WWTP	2.2 m^3/s

5.3.2.2 Model calibration

Kleidorfer et al. (2006a) discussed some aspects of calibration of hydrological models for the estimation of CSO performance and concluded that the overflow volume (V_{CSO}) and the number of overflow events (n_{CSO}) are good calibration parameters for estimating η. The duration of overflows (t_{CSO}) and especially the runoff to the WWTP (V_{WWTP})

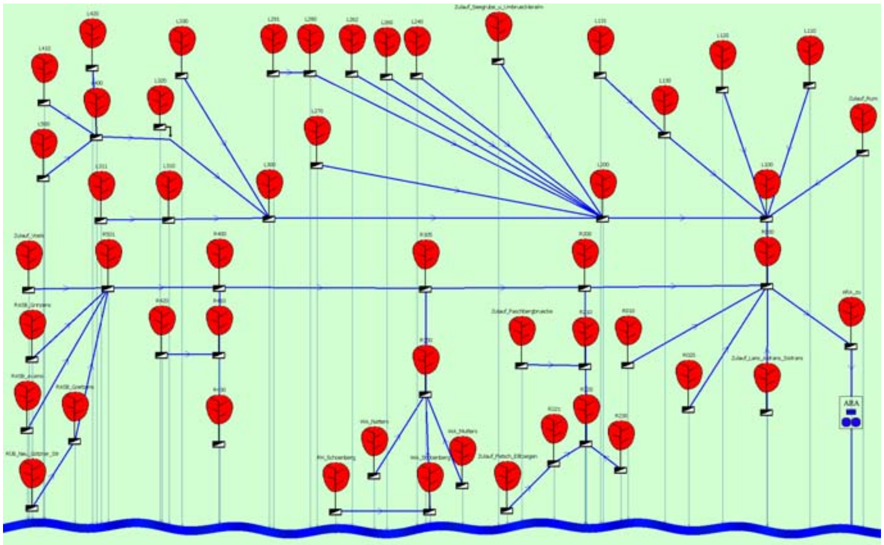

Figure 5.14: Hydrological model Innsbruck - Sewer system of Innsbruck represented in the software KAREN

are not adequate descriptors of system behaviour. V_{WWTP} is mainly influenced by the dry weather flow and barely changes with efficiency.

Of all analysed parameters which can be used for calibration the effective impervious area is the most important parameter and influences the CSO performance in a significant way. As hydrological models in contrast to hydrodynamic models do not represent inline sewer storage volume the virtual CSO storage volume (derived from real CSO storage volume + dynamically allocated inline storage of the main sewers of the catchment) and the virtual interceptor capacity (derived from real interceptor capacity and the dynamics of throttled flow) have to be estimated during calibration. That are model structure uncertainties which can be only estimated by comparing different models. This is discussed for example by De Toffol et al. (2006b), who compared hydrodynamic and hydrological models. All other model parameters can be neglected in the calibration process (i.e. reliable results can be reached with default values). This corresponds with findings in **Paper II** and **Paper III**, where the sensitivity of simulation results of KAREN are analysed with respect to the model parameters.

Additionally Kleidorfer et al. (2006a) showed that it is very difficult to predict the CSO performance when calibrating on single events because the choice of the event, which is used for calibration, impacts the result. Using several events for calibration improves the process only marginally as the calibration gets increasingly complex with the number of the events. Hence it is advisable to calibrate a hydrological model for long time series if the essential data is available. This is also discussed in **Paper I** where it was showed that random selection of rainfall events for calibration can lead to a complete failure of the calibration process. Additionally the impact of the number and location of measurement sites used for calibration was analysed in **Paper I**. Here measurements of 30% to 50% of all CSOs are required to reach a sufficient calibration. Furthermore, an exceeding number of measurement sites does not improve calibration performance anymore. These results and a suggestion for a calibration procedure are published in German in **Paper VI** as an assistance for engineers, who commence using urban drainage models after the release of ÖWAV-RB 19 (2007).

For Innsbruck no explicit CSO discharge measurements but only water level measurements are available. As in hydrological models water levels are no model output they cannot be used for calibration. Instead CSO discharge is calculated from water level measurements by means of overflow equations (e.g. Poleni equation). Although this causes calibration–data uncertainties (additional to uncertainties of water level measurements), Sitzenfrei et al. (2008) and Fach et al. (2008b) arranged simulations with a computational fluid dynamic model and showed that this procedure leads to sufficient results when the coefficients of discharge are set according to guiding rules (e.g. ATV-Aă111ăE, 1994). As suggested in **Paper I** the model is calibrated on long–time performance, i.e. on the sum of CSO discharge over the period of approximately two years. This was the entire time–period for which calibration–data was available for most measurement sites to the date of this study.

Figure 5.15 shows a comparison of measured (black bars) and simulated (grey bars) CSO discharge for 15 measurements sites.

5.3.2.3 Model results

After model calibration building measures intended for future developments are implemented to the model to evaluate the improvement caused by those measures. As for Innsbruck two long–time rainfall series (> 10 years) are available for estimation of CSO efficiency, both series are used for simulation. Table 5.8 shows required and estimated CSO efficiencies for future conditions with state 2007. Currently – in a new study – the

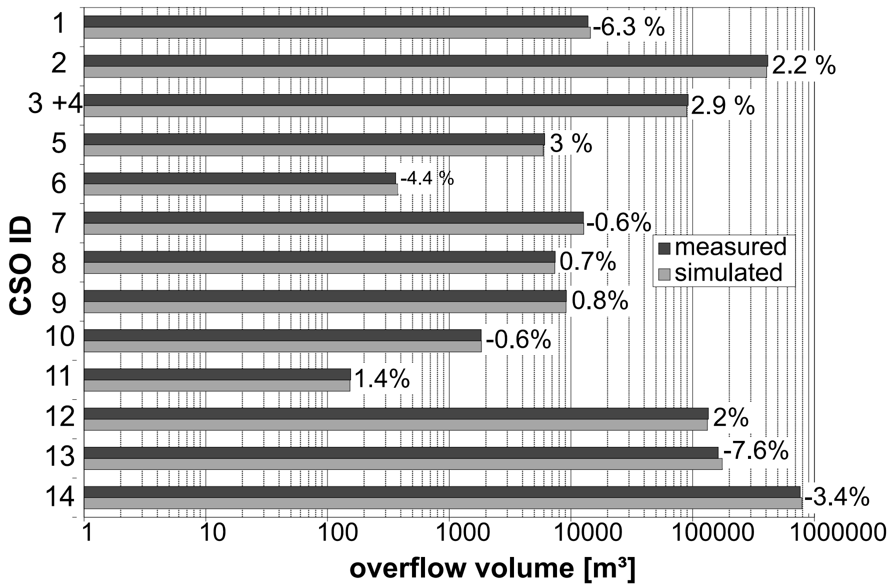

Figure 5.15: Hydrological model Innsbruck: Calibration - Comparison of simulated and measured CSO discharge volumes

effect of the building measures already realised is analysed by comparing model–results with new measurement data (2007 until now).

Table 5.8: Required and estimated CSO efficiency Innsbruck

Rain Gauge Innsbruck University 1992/01/01 – 2006/12/31		
	required	estimated
η_d [%]	58.5	68.2
η_p [%]	73.5	76.4
Rain Gauge Innsbruck Airport 1987/01/01 – 2005/12/31		
	required	estimated
η_d [%]	56.0	66.2
η_p [%]	71.0	75.2

Results from analysis of the impact for ambient water quality (see section 5.3.1 which are based on urban drainage modelling (hydraulic impact and acute ammonia toxicity) are given in Table 5.9 for the two rivers Inn and Sill.

Table 5.9: Impact for ambient water quality Inn and Sill

Impact	Requirement	University	Airport
Hydraulic impact Inn	$Q_1 < 0.5HQ_1$ =196 $m^3\,s^{-1}$	52 $m^3\,s^{-1}$	75 $m^3\,s^{-1}$
Hydraulic impact Sill	$Q_1 < 0.5HQ_1$ =47.4 $m^3\,s^{-1}$	3 $m^3\,s^{-1}$	4 $m^3\,s^{-1}$
Acute NH_4 toxicity Inn	$c_{r,d}$ <2.5 $mg\,l^{-1}$	0.34 $mg\,l^{-1}$	0.35 $mg\,l^{-1}$
Acute NH_4 toxicity Sill	$c_{r,d}$ <2.5 $mg\,l^{-1}$	0.25 $mg\,l^{-1}$	0.25 $mg\,l^{-1}$

From this requirements the sewer system Innsbruck also meets the criteria of ÖWAV-RB 19 (2007).

5.3.3 Case study Linz

5.3.3.1 System description

The hydrological model for Linz was set up in the software City Drain (Achleitner et al., 2007) and is illustrated in Figure 5.16. A detailed description is available from Fach et al. (2008a); Möderl et al. (2007a,b). The system characteristics are summarised in Table 5.10.

Table 5.10: System characteristics Hydrological Model Linz

Parameter	Value
subcatchments	30
CSOs	30
impervious area	2494 ha
storage volume basins	89995 m^3
storage volume sewers	392129 m^3
inflow to WWTP	4.7 m^3/s

In the combined sewer system of Linz additionally a real time control (RTC) has to be considered. The RTC aims to prevent discharge from the specific CSO “Weikerlsee”, or to allow discharge only if flooding would occur instead, because this CSO is located at a swimming lake at local recreation area. At the date of this study this RTC was still in a planning phase and hence no measurement data was available. It is common practise to use hydrodynamic models for the design and analysis of such a control. Hence, analysis of a RTC in a hydrological model needs to be abstracted to a certain degree which leads to model–structure uncertainties. The implementation of a RTC in the hydrological sewer model of Linz is discussed in **Paper VII**.

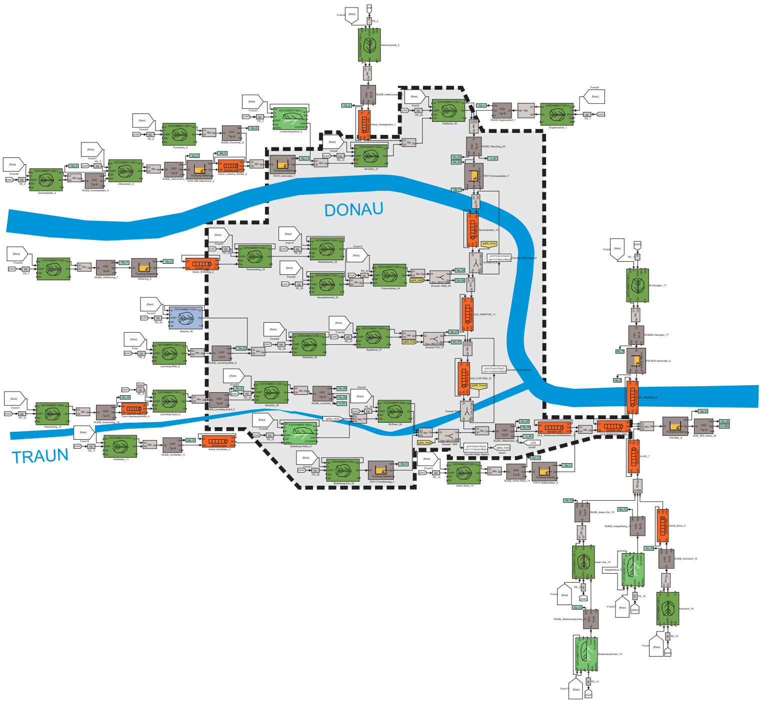

Figure 5.16: Hydrological model Linz - Sewer system of Linz represented in the software CityDrain from Kleidorfer et al. (2007b)

5.3.3.2 Model calibration

The model is calibrated on discharge measurements at the five CSOs "Plesching", "HSM", "Lunzerstrasse", "Füchselbach" and "Weikerlsee". The comparison of simulated and measured CSO discharge volumes is shown in Figure 5.17.

As the available data for calibration is limited to only a few rainfall events with discharge measurements, additionally simulation results from the hydrological model are compared to results from the hydrodynamic model (which can be calibrated on water level measurements). This synchronisation of hydrological and hydrodynamic models to reduce uncertainties in case of a limited data basis is described by Fach et al. (2008a). Both calculation methods can be used to keep the uncertainties of simulation results as low as possible. For example Figure 5.18 illustrates the comparison of surface runoff (left) and CSO discharge (right) which shows a good agreement.

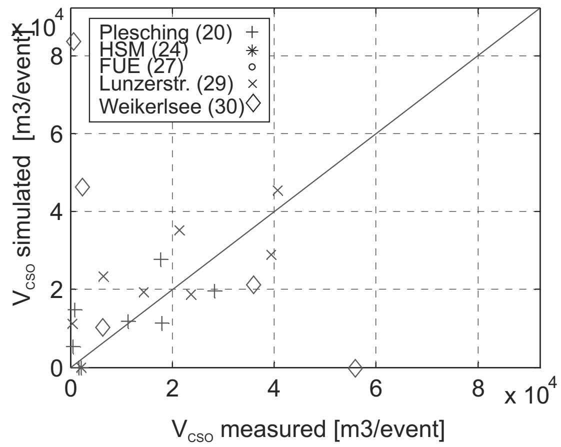

Figure 5.17: Hydrological model Linz: Calibration - Comparison of simulated and measured CSO discharge volumes

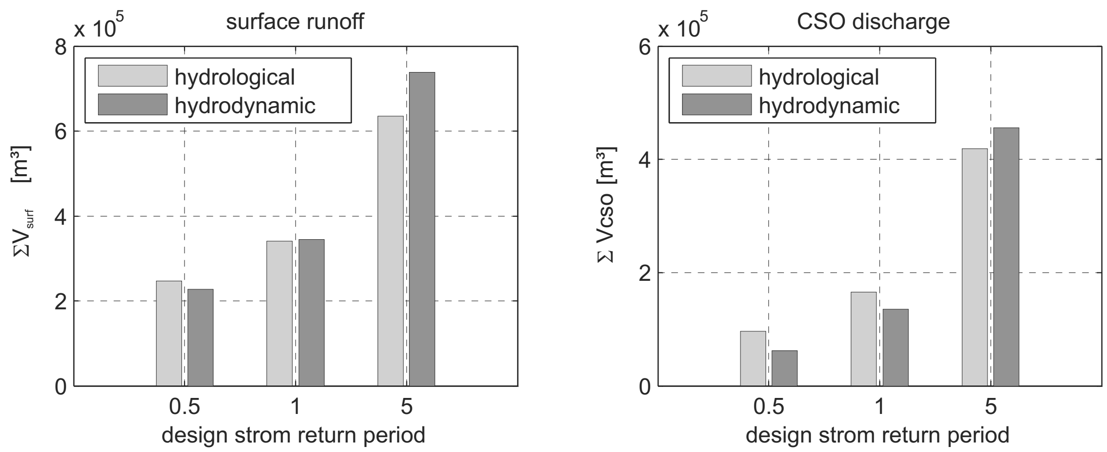

Figure 5.18: Hydrological model Linz: Calibration - Comparison of surface runoff (left) and CSO discharge (right) from hydrological and hydrodynamic simulation

5.3.3.3 Model results

The calculated CSO efficiencies η_d and η_p are shown in Figure 5.21 for nine different scenarios of future development of the sewer system. Therein the black line marks the value that has to be met according to ÖWAV-RB 19 (2007). As one can clearly see calculated values are higher as required ones for all scenarios.

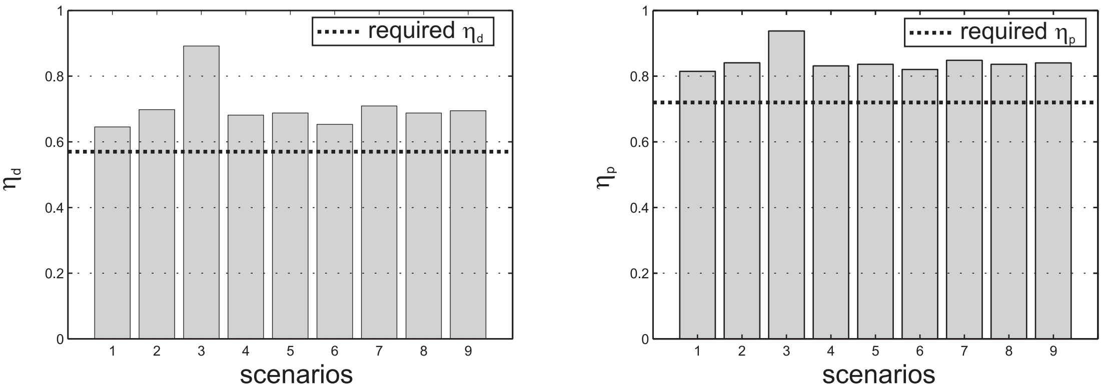

Figure 5.19: Hydrological model Linz: Results - Calculated η_d (left) and η_p (right) for different scenarios

To explain the operation of the RTC in Linz Figure 5.20 shows a schematic description of the implementation of the RTC in the hydrological model of Linz. The five throttle flaps “ULK”, “Traun”, “HSM”, “FUE” and “HSS” are controlled using the fill level in the basin “Weikerlsee” as sensor value. Figure 5.19 shows the operation of the RTC exemplified on one rainfall event. Figure 5.19 (a) shows the basin volume filled, (b) the runoff upstream the basin, (c) the discharge at the CSO “HSM” and (d) the runoff downstream the basin. The location of this hydrograph is also shown in Figure 5.20.

As soon as the maximum interceptor capacity downstream the CSO “Weikerlsee” is reached, the basin begins to fill (1) until a fill level of 38 000 m^3 is reached (2). This is the starting point of the RTC and the throttle flaps begin to close to restrict the runoff to the basin. This results for example in a discharge at CSO “HSM” instead of a discharge at the CSO “Weikerlsee” (Figure 5.19 (d)) beginning with (3). After 12 hours the basin has emptied (4).

As mentioned the goal of this RTC is not to reduce CSO discharge, but to prevent CSO discharge at the basin “Weikerlsee”. The effective reduction of total CSO discharge

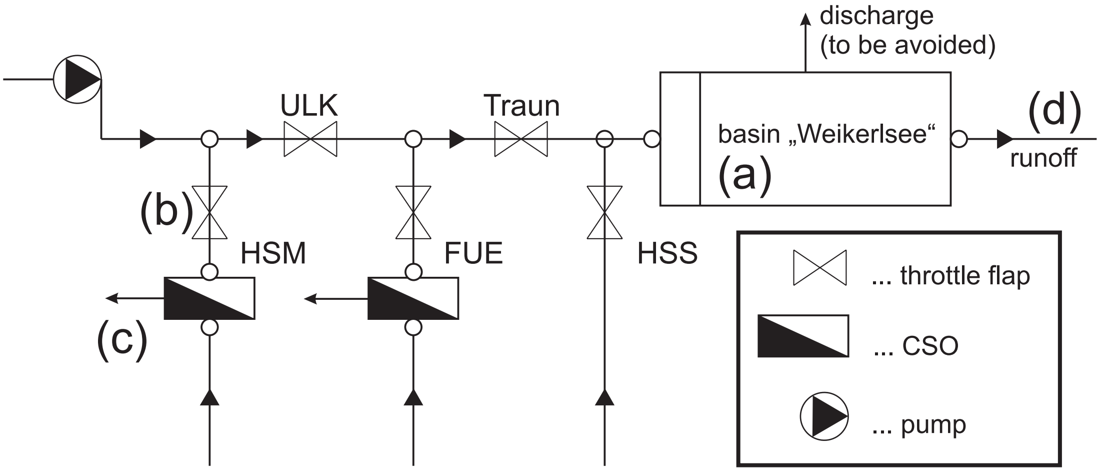

Figure 5.20: RTC in the hydrological model of Linz - Schematic description of the RTC in Linz

is only about 1 %. Hence, CSO discharge is shifted from CSO "Weikerlsee" to other ones. Figure 5.22 shows the distribution of CSO discharge with (black bars) or without RTC (grey bars) for the different CSOs for long–time simulation. Discharge is avoided at the CSO "Weikerlsee" and shifted mainly to "HSM" and "FUE".

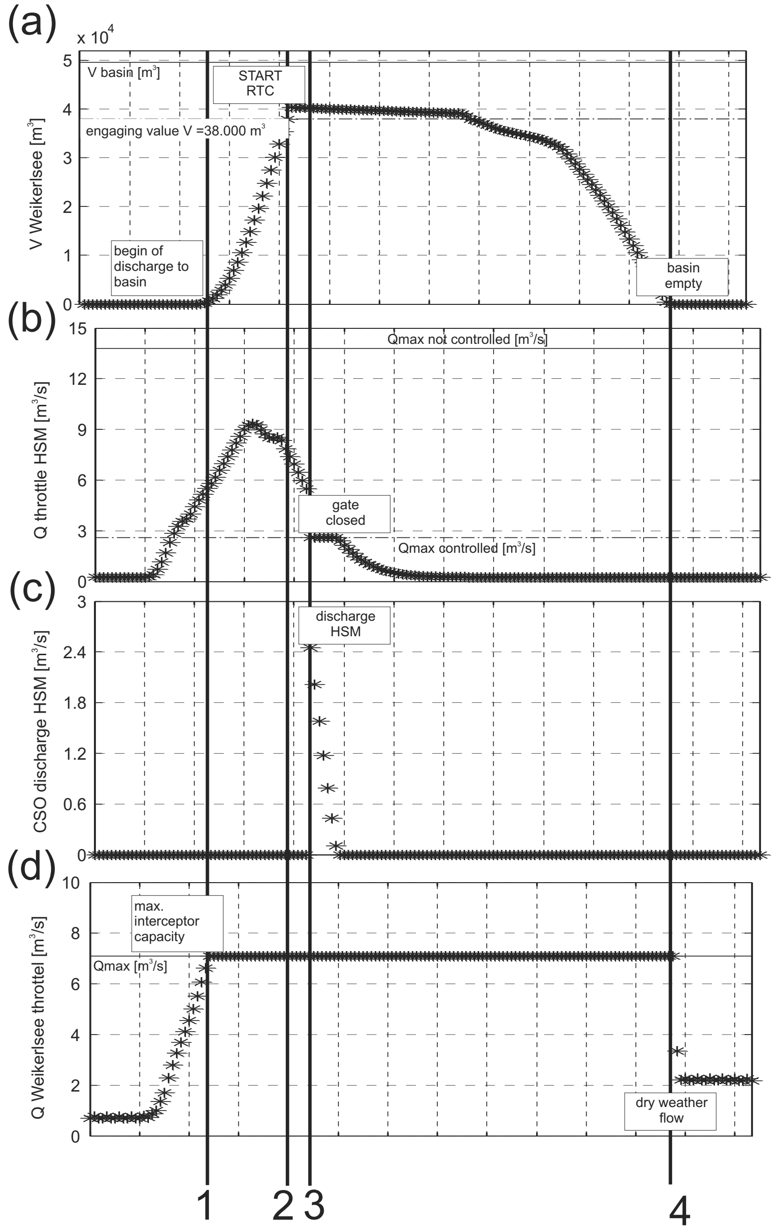

Figure 5.21: Operation scheme of the RTC in the hydrological model of Linz - Operation of the RTC exemplified on one rainfall event: (a) basin volume filled, (b) runoff upstream the basin, (c) discharge at the CSO "HSM", (d) runoff downstream the basin

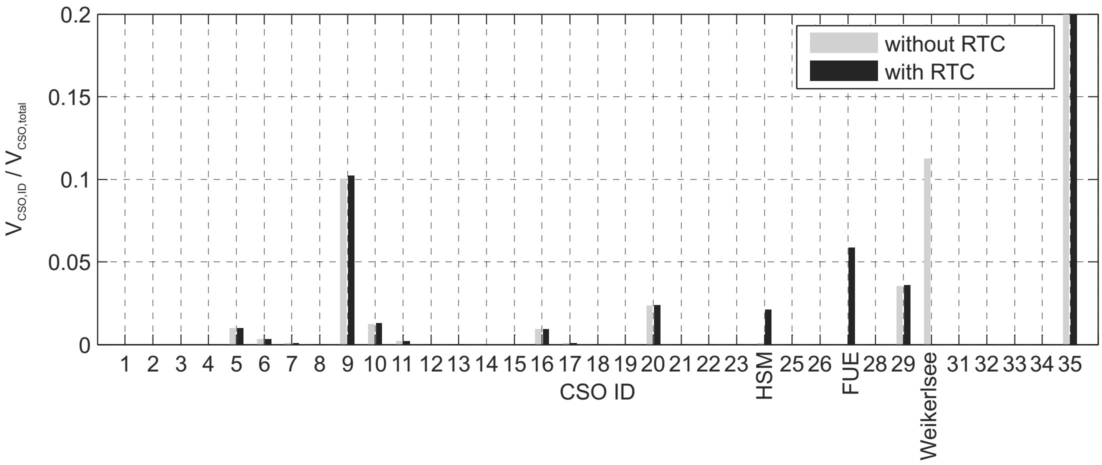

Figure 5.22: Hydrological model Linz: Results RTC - Distribution of CSO discharge with RTC and without RTC

Chapter 6

Conclusions, discussion & outlook

> The outcome of any serious research can only be to make two questions grow where only one grew before.
>
> Thorstein Veblen

6.1 Summary

In this thesis several aspects of uncertainty and calibration of urban drainage models are discussed. Chapter 2 describes the modelling concepts and chapter 3 the case studies used in the papers annexed. Chapter 4 starts with a classification of sources of uncertainties (section 4.1) and describes different methods for sensitivity and uncertainty analysis (section 4.3). **Paper I** focuses on impact of data availability and **Paper IV** on uncertainties due to long–time prediction. Of the methods described, the Markov Chain Monte Carlo simulation based on Bayesian inference (section 4.3.4), was tested in **Paper II** and **Paper III**. This is the preliminary stage of the Bayesian Total Error Analysis (BATEA) framework (section 4.3.5). In chapter 5 aspects of model calibration are discussed and examples for model applications are shown. This chapter mainly describes the contents of **Paper VI** and **Paper VII**.

In **Paper I** the impact of calibration data availability is analysed. Although long-term calibration is preferred to calibration on single events, single events can be used when increasing the number of measurement sites to reach a similar accuracy. The selection of rainfall events plays a decisive role. While a maximum deviation of less than 15 % could be reached when calibrating on the five rainfall events with highest peak or on the five rainfall events with the longest duration a pure random selection can result in a complete false estimation of overflow volume with a deviation of more than 75 %. Although all the results shown are case specific and any additional sources of uncertainties (e.g. measurement uncertainties, spatial rainfall distribution) were disregarded in this study, this investigation outlines relevant considerations when arranging measurement campaigns. The application of the methodology described in this paper on the sewer system of Innsbruck shows how those considerations can be implemented. In this case it was possible to calibrate a hydrological model sufficiently when calibrating on 30% of all existing CSOs.

In **Paper II** a Bayesian approach was used to develop a framework for the quantification of the impact of uncertainties in the model inputs on the parameters of a simple integrated stormwater model for calculating runoff, total suspended solids and total nitrogen loads. The framework was applied to two catchments in Australia. This study shows how the Metropolis algorithm can be adapted for evaluating sensitivity of calibration parameters of urban drainage models. This was done by using the software tool MICA. Such parameter sensitivity analysis takes into account all sources of modelling uncertainties and boundary conditions, such as availability of data, measurement uncertainties or different catchment characteristics, but it is not possible to distinguish between these different sources as in the full BATEA framework. Calibration parameters can compensate for all uncertainties, in order to achieve simulation results with the best possible fit to measurement data. The main advantage of this method is, that not only one “best calibration parameter set” is gained, but also a distribution of the most likely values of the model parameters.

In **Paper III** the same method is used for a comparison of two rainfall / runoff models (MUSIC and KAREN) and two stormwater quality models (Regression model and Buildup-washoff model) with different level of complexity. The rainfall/runoff models tested performed very similar suggesting that a simple model may be used for urban catchments without compromising the results. The effective impervious fraction is the most important parameter in both models and special attention should be paid to its value. Even with the robust calibration and parameter sensitivity approach used here, the water quality models tested poorly represent reality and result in a high level of uncertainty. The study developed was very important to verify the efficiency of the calibration and sensitivity analysis approach. The method presented seems to be promising in terms of generating the posterior parameter distributions and also gives some valuable information on parameter interaction.

In **Paper IV** the impact of environmental change effects (climate, land–use, population) on different performance indicators of combined sewer systems is analysed. In this paper some general coherences of sewer system behaviour under future development scenarios and climate change scenarios are presented. This is a contribution to enhanced system understanding taking into account long-term environmental change effects. For example the increase of rainfall intensities by the factor of 1.2 has the same effect as an increase of impervious area of +40%. Such an increase could be compensated by infiltration measures in current systems which lead to a reduction of impervious area by 30%. Another finding was that performance indicators representing CSO emissions show a negative correlation with performance indicators representing flooding. Increased conduit diameters improve the system capacity leading to a reduction of flooding, but on the other hand more combined flow can be conveyed downstream resulting in an increase of CSO discharge.

In **Paper V** the software tool CALIMERO for generalised autocalibration is presented. The novelty of that tool lies in the flexibility to work with any model which's input and output files are plaintext and which can be started from command line. The algorithms for evaluating the objective function and the calibration algorithm itself are defined in the scripting language ECMA /JavaScript via built-in script editor to provide best possible calibration results under consideration of additional knowledge about system behaviour. A simple example shows the capabilities of CALIMERO to adapt calibration algorithms depending on specific case study characteristics. Due to the modular design CALIMERO can also be used for automated uncertainty analysis (e.g. Monte Carlo simulation with subsequent results evaluation) with only a few adaptations, which will be the next step in development. The scripting language used is rather simple designed for non-programmers to work with. Due to its wide use in client side website programming there are a lot of tutorials and manuals available and shall encourage the exchange of calibration scripts among different users and consequently propagate more sophisticated uncertainty analysis even in non–scientific applications.

In **Paper VI** and **Paper VII** practical aspects of model calibration are discussed. **Paper VI** is intended to guide engineers working with numerical models for evaluating the requirements of ÖWAV-RB 19 (2007) and is focused on the Software KAREN (Rauch and Kinzel, 2007). Hence the model parameters and their sensitivities are described as well as sensitivity of input data to guide model user which data has to be collected very carefully and which data can be estimated from literature. Of all analysed parameters which can be used for calibration the effective impervious area has highest impact on calculation of CSO efficiency. It is clear that this parameter is the first choice for calibration and hence must be determined with great accuracy. Other important parameters are the virtual interceptor capacity (derived from real interceptor capacity and the dynamics of throttled flow) and the virtual CSO storage volume (derived from real CSO storage volume + dynamically allocated inline storage of the main sewers of the catchment). All other model parameters have minor impact on simulation results or can be estimated from literature and hence can usually be neglected in the calibration process.

Paper VII shows how a real time control (RTC) can be implemented in a hydrological model of a combined sewer system. It is common practise to use hydrodynamic models for the design and analysis of such a control. Hence, analysis of a RTC in a hydrological model needs to be abstracted to a certain degree. The RTC of the sewer system in the city of Linz is discussed in this paper to demonstrate how this can be done, and what control algorithms are impossible to be reproduced in a hydrological model. Therefore a hydrodynamic model was used for comparison. While waterlevel measurements in the sewers can be used in a hydrodynamic model they are not represented in hydrological models and cannot be used for control. The advantage of a RTC in a hydrological model is that this way such an RTC can be considered when continuous long–time simulations are undertaken. By comparing model of different complexity (hydrological and hydrodynamic models) this paper approaches analysis of model structure uncertainties.

6.2 Conclusions

In this thesis was shown that impact of uncertainties is significant. Special attention has to be payed to temporal and spatial availability of calibration data. It was shown that – as worst case – a model could seem to be calibrated sufficiently on few single events or on few measurement sites, but completely fails in predicting outside the calibration period. But these uncertainties are strongly related to uncertainties of input–data or to uncertainties due to spatial rainfall distribution, respectively. While it can clearly be seen that random rainfall errors related to measurement uncertainties of the rain gauge have only minor impact on model results and systematic rainfall errors are compensated during calibration, in practical model applications there are numerous rainfall events where the big discrepancy between measured data and simulated data can hardly be explained. The most plausible way of explaining that uncertainties is the influence of spatial rainfall distribution where the rainfall measured by the rain gauge is not characteristic for the rainfall of the entire catchment. As more measurement sites are used and as longer the time period used for calibration is, the bigger is the chance to "compensate" that uncertainties. That means, that still multiple (single) rainfall-events remain where measured data is not consistent with simulated data but in long–time average the spatial rainfall distribution is balanced. Hence if the limiting values to be observed are related to a long–time average that uncertainties are acceptable. Of course the question which data availability is required for a sufficient calibration remains and this case specific problem can only be solved by ensuring that model predictions are also reliable (not for all but for most) rainfall events in a validation period. Nevertheless further research especially on how to describe spatial rainfall distribution in an input–data error model is required.

Analysis of model–structure uncertainties by comparing hydrological models of different complexity and by comparing hydrological and hydrodynamic models indicates that this source of uncertainty has minor impact on model results. Of course this conclusion is only true *if* a model can be calibrated sufficiently. In case of the pollutant models analysed, which were not able to predict measured data, model–structure uncertainties are expected to be substantial. But an important point regarding that topic is that model parameters cannot be transferred from one model to another even if they represent the same physical background. Model parameters mostly compensate uncertainties due to model structure and hence are always model specific. For example calibrated values for EIF in MUSIC are different from calibrated values for EIF in KAREN or interceptor capacity from a hydrodynamic model is different from interceptor capacity in a hydrological model. Hence each model has to be calibrated on its own. The only possibility to combine different model structures is to calibrate one model on the simulation result of another model. This could be useful to align models in case of limited data availability.

Unlike the rainfall / runoff models the water quality models analysed could not be calibrated sufficiently. Hence here model structure uncertainties seem to be significant. The failure of these models is especially interesting as several publications are available which show a good agreement between measured data and simulated data. But in most studies which show good model performance, they are only calibrated on a few single events, while here a very comprehensive data set containing around 300 wet weather events was used. That also indicates that the impact of availability of calibration data is significant.

Impact of the choice of calibration algorithms and criteria functions was not analysed in this thesis. Nevertheless the software tool CALIMERO was presented which should support such an investigation in further research.

6.3 Discussion and outlook

During the last decades, the use of numerical models and software in the field of environmental engineering has grown steadily. Today numerical models are the state–of–the–art instrument for design, optimisation and evaluation of urban drainage infrastructural facilities and underpin decision making, operation and management. Hence since 1983 when Beck writes

> "*Sanitary engineers do not generally present their knowledge and hypotheses in a mathematical format.*"

a significant technological advance is visible. Despite that development uncertainties in urban drainage modelling due to lack of data, lack of process–understanding and spatial and temporal variability is still substantial. Hence another statement of Beck (1983)

> "*Relatively little attention, however, has been given to the problems of uncertainty and errors in the field data, of inadequate numbers of data, of uncertainty in the relationships between important system variables, and of uncertainty in the model parameter estimates.*"

is still true today in regard to urban drainage modelling. Additionally there is a gap in consideration of uncertainties between scientific and practical model applications. While at least some scientific studies dealing with uncertainties of urban drainage models are available in practical projects uncertainty analysis is usually not covered. This is expected to be true for three reasons: (a) Engineers have to cope with problems related to data availability. Hence a comprehensive uncertainty analysis is not possible due to insufficient input–data and calibration–data availability. (b) Despite having different methods for uncertainty analysis available, there is still a lack of a robust conceptual framework for describing and estimating uncertainties. Hence uncertainty analysis is highly case specific and remains a complex task which is usually neither required nor payed by the ordering party. (c) Guiding rules require the observation of limiting values but no estimation of uncertainties of model results. Hence it is difficult to communicate uncertainties to decision makers and often it is counterproductive to argue a certain value is met with – for example – a probability of "only" 95%. Additionally the effort for a comprehensive data collection for minimising uncertainties is not appreciated and reflected in the limiting values to be observed. Therefore sewer system operators are not encouraged to set up a comprehensive monitoring network.

One of the key points of future studies is the development of a robust uncertainties framework. The BATEA framework already tested for hydrology of natural catchments seems to be a promising approach having the main advantage that each source of uncertainty is considered explicitly at its point of origin. For further research the development of robust error models (especially error models for spatial rainfall distribution) adapted for urban drainage models is necessary. Additionally the scientific community has to formulate their expert knowledge to be used as prior knowledge in the Bayesian inference. The International Working Group on Data and Models, which works under the IWA/IAHR Joint Committee on Urban Drainage, already approaches this topic by organising workshops on uncertainty methodologies and by publishing common knowledge (e.g. Deletic et al., 2009).

For testing new methodologies as e.g. the BATEA framework case studies with a comprehensive dataset (input–data and calibration–data) are necessary. As this data is often limited, new strategies are usually tested on a few case studies only. Sitzenfrei et al. (2009) and Urich et al. (2009) developed the software tool VIBe (Virtual Infrastructure Benchmarking) which is an advancement of the Case Study Generator (Möderl et al., 2009) for generating virtual case studies of urban water systems. By using this software tool methodologies for uncertainty analysis can be tested more comprehensively with not data availability as limiting factor but only computing time. A further advancement are parallelised algorithms to reduce computing time as shown e.g. for the case of conceptual sewer system models by Burger et al. (2009). In this context also further investigation towards the impact of the objective function used in the calibration algorithms should be done.

As the BATEA framework considers different sources of uncertainties explicitly at their point of origin, this method would be particularly interesting to be applied on an integrated urban drainage model in which different submodels for different processes (e.g. sewer system, waste water treatment plant, river water quality, groundwater) are combined. By this approach uncertainties of the individual submodels could be analysed without having the problem that different calibration parameters compensate for each other.

Further research is also required to improve today's guidelines with respect to uncertainties in urban drainage models. In structural engineering the semi–probabilistic safety concept in which each parameter is multiplied by a partial safety coefficient is state–of–the–art. A similar concept would be adequate for urban drainage models and partial safety coefficients could depend on data availability (e.g. number of measurement sites, length of timeseries) and accuracy of data available (depending on the measurement method). Such an approach would not only encourage sewer system operators to improve their monitoring network (to reduce the required safety factor and consequently to reduce investment costs), but also help to communicate uncertainties of models to decision makers.

References

2006/7/EC, 2006. Directive 2006/7/EC of 15 February 2006 concerning the management of bathing water quality and repealing directive 76/160/EEC. 98

Achleitner, S., 2008. Modular conceptual modelling in urban drainage development and application of City Drain (PhD thesis). innsbruck university press. 11, 12, 13, 46, 79

Achleitner, S., De Toffol, S., Engelhard, C., Rauch, W., 2005. Model based hydropower gate operation for mitigation of CSO impacts by means of river base flow increase. Water Science & Technology 52 (5), 87–94. 20

Achleitner, S., Möderl, M., Rauch, W., 2007. City Drain - an open source approach for simulation of integrated urban drainage systems. Environmental Modelling & Software 22 (8), 1184 – 1195. 18, 103

Achleitner, S., Möderl, M., Rauch, W., 2008. Conceptual Modelling with City Drain Software in the Field of Sanitary Engineering. Korrespondenz Abwasser 55 (7), 746–755. 80

Achleitner, S., Rauch, W., 2006. Increase of river base flow by hydropower gate operation for mitigation of CSO impacts Ű- potential and limitations. Water Resources Management 21(9), 1478 – 1503. 20

Akan, A. O., Houghtalen, R. J., 2003. Urban hydrology, hydraulics, and stormwater quality: engineering applications and computer modeling. John Wiley & Sons, Inc., Hoboken, New Jersey. 11

Arnaud, P., Bouvier, C., Cisneros, L., Dominguez, R., 2002. Influence of rainfall spatial variability on flood prediction. Journal of Hydrology 260 (1-4), 216 – 230. 43

Arnbjerg-Nielsen, 2008. Quantification of climate change impacts on extreme precipitation used for design of sewer systems. In: 11th International Conference on Urban Drainage. Edinburgh, Scotland, UK. 4

Arnbjerg-Nielsen, K., Harremoës, P., 1996. The importance of inherent uncertainties in state-of-the-art urban storm drainage modelling for ungauged small catchments. Journal of Hydrology 179 (1-4), 305 – 319. 5, 53

Ashley, R. M., Balmforth, A., Saul, A. J., Blanskby, J., 2005. Flooding in the future Ű predicting climate change, risks and responses in urban areas. Water Science & Technology 52 (5), 265–273. 4

ATV-Aă111ăE, 1994. Standards for the hydraulic dimensioning and the performance verification of stormwater overflow installations in sewers and drains. 101

ATV-AGă2.1.1, 1993. Weitergehende Anforderungen an Mischwasserentlastungen. Korrespondenz Abwasser 35 (5). 98

Bayes, M., Price, M., 1763. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. Philosophical Transactions 53, 370–418. 63

Beck, M. B., 1983. Uncertainty and Forecasting of Water Quality. Springer, Berlin, Ch. Uncertainty, system identification and the prediction of water quality, pp. 3 – 68. 41, 55, 116

Beck, M. B., 1987. Water quality modelling: a review of the analysis of uncertainty. Water Resources Research 23 (8), 1393Ű1442. 39, 41, 55

Beck, M. B., 1991. Principles of modelling. Water Science & Technology 24(6), 1 – 8. 40

Berger, J., 2006. The case for objective Bayesian analysis. Bayesian Analysis 1(3), 385 – 402. 64

Bertrand-Krajewski, J. L., Bardin, J.-P., Mourad, M., Beranger, Y., 2003. Accounting for sensor calibration, data validation, measurement and sampling uncertainties in monitoring urban drainage systems. Water Science & Technology 47 (2), 95 – 102. 4

Bertrand-Krajewski, J. L., Barraud, S., Bardin, J.-P., 2002. Uncertainties, performance indicators and decision aid applied to stormwater facilities. Urban Water 4, 163–179. 4

Beven, K., 2001. Rainfall-Runoff Modelling: the Primer. John Wiley and Sons, Chichester, UK. 13

Beven, K., 2007. Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. Hydrology and Earth System Sciences 11 (1), 460–467. 4, 58

Beven, K., 2009. Environmental Modelling: An Uncertain Future? Routledge, Taylor and Francis Group. 39, 48, 53, 56, 64, 66

Beven, K., Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes 6 (3), 279–298. 4, 56

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

Beven, K., Smith, P., Freer, J., 2007. Comment on Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology by Pietro Mantovanand Ezio Todini. Journal of Hydrology 338 (3-4), 315 – 318. 57

Beven, K. J., Smith, P. J., Freer, J. E., 2008. So just why would a modeller choose to be incoherent? Journal of Hydrology 354 (1-4), 15 – 32. 57

Blecken, G.-T., Zinger, Y., Deletic, A., Fletcher, T. D., Viklander, M., 2009a. Impact of a submerged zone and a carbon source on heavy metal removal in stormwater biofilters. Ecological Engineering 35 (5), 769 – 778. 2

Blecken, G.-T., Zinger, Y., Deletic, A., Fletcher, T. D., Viklander, M., 2009b. Influence of intermittent wetting and drying conditions on heavy metal removal by stormwater biofilters. Water Research In Press, Accepted Manuscript, –. 2

Bratieres, K., Fletcher, T., Deletic, A., Zinger, Y., 2008. Nutrient and sediment removal by stormwater biofilters: A large-scale design optimisation study. Water Research 42 (14), 3930 – 3940. 2

Burger, G., Fach, S., Kinzel, H., Rauch, W., 2009. Parallel computing in integrated urban drainage simulations. In: Proceedings of the 8th International Conference on Urban Drainage Modelling. 118

Burkhard, R., Deletic, A., Craig, A., 2000. Techniques for water and wastewater management: a review of techniques and their integration in planning. Urban Water 2 (3), 197 – 221. 2

Butler, D., Davies, J., 2004. Urban Drainage, 2nd Edition. Spon Press, London. xxi, 1, 3, 13, 14, 15

Butler, D., McEntee, B., Onof, C., Hagger, A., 2007. Impact of climate change on sewer storage tank performance. Water Science & Technology 56 (12), 29–35. 4

Carpenter, T. M., Georgakakos, K. P., 2004. Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrological model. Journal of Hydrology 298, 202 – 221. 4, 43

Chaplot, V., Saleh, A., Jaynes, D., 2005. Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO3-N loads at the watershed level. Journal of Hydrology 312 (1-4), 223 – 234. 43

Chocat, B., Krebs, P., Marsalek, J., Rauch, W., Schilling, W., 2001. Urban drainage redefined: from stormwater removal to integrated management. Water Science & Technology 43(5), 64 – 68. 18

Cook, R. M., van Noortwijk, J. M., 2000. Sensitivity Analysis. John Wiley & Sons, Ch. Graphical methods, pp. 254 – 264. 53

Coustumer, S. L., Fletcher, T. D., Deletic, A., Barraud, S., Lewis, J. F., 2009. Hydraulic performance of biofilter systems for stormwater management: influences of design and operation. Journal of Hydrology In Press, Accepted Manuscript, –. 2

De Toffol, S., 2009. Sewer system performance assessment - an indicator based methodology (PhD thesis). Vol. 5 of Forum Umwelttechnik und Wasserbau. innsbruck university press, Innsbruck. 87, 98

De Toffol, S., Engelhard, C., Rauch, W., 2006a. Combined sewer system versus separate system - a comparison of ecological and economical performance indicators. Water Science & Technology 55 (4), 255–264. 20

De Toffol, S., Kleidorfer, M., Rauch, W., 2006b. Vergleich hydrodynamischer und hydrologischer Simulationsmodelle bei der Berechnung der Emissionen von Mischwasserbehandlungsanlagen. Wiener Mitteilungen 196, H1–H20. 99, 100

Deletic, A., Ashley, R., Rest, D., 2000. Modelling input of fine granular sediment into drainage systems via gully-plots. Water Research 34 (15), 3836–3844. 16

Deletic, A., Dotto, C. B. S., Fletcher, T. D., McCarthy, D. T., Bertrand-Krajewski, J. L., Rauch, W., Kleidorfer, M., Freni, G., Mannina, G., Tait, S., 2009. Defining uncertainties in modelling of urban drainage systems. In: Proceedings of the 8th International Conference on Urban Drainage Modelling, Tokyo. 5, 42, 46, 53, 118

Deletic, A., Fletcher, T. D., 2006. Performance of grass filters used for stormwater treatment–a field and modelling study. Journal of Hydrology 317 (3-4), 261 – 275. 2

Doherty, J., 1999. PEST Model-Independent Parameter Estimation User Manual 5th Edition. Watermark Numerical Computing. 80

Doherty, J., 2003. MICA - Model Independent Markov Chain Monte Carlo Analysis - User Manual. Watermark Numerical Computing, U.S. EPA. 67, 68

Dotto, C. B. S., Deletic, A., Fletcher, T. D., 2009. Analysis of parameter uncertainty of a flow and quality stormwater model. Water Science & Technology 60 (3), 717 – 725. 4, 67

Dotto, C. B. S., Kleidorfer, M., Deletic, A., Fletcher, T. D., Rauch, W., submitted. A bayesian approach for performance evaluation of stormwater models based on long-term high-resolution measurement data. Journal of Hydrology. 8

ECMA-262, 1999. Ecmascript language specification. Tech. rep., ECMA International, Genf. 84

Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirmbach, M., Vaes, G., Vieux, B., 2004. Towards a roadmap for use of radar rainfall data in urban drainage. Journal of Hydrology 299 (3-4), 186 – 202. 43

Einfalt, T., Arnbjerg-Nielsen, K., Spies, S., 2002. An enquiry into rainfall data measurement and processing for model use in urban hydrology. Water Science & Technology 45(2), 147 – 152. 43

Ellison, A. M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications 6 (4), 1036–1046. 64

Engeland, K., Xu, C.-Y., Gottschalk, L., 2005. Assessing uncertainties in a conceptual water balance model using Bayesian methodology. Hydrological Sciences Journal 50(1), 45 – 63. 4, 64

Fach, S., Kleidorfer, M., Möderl, Michael andRauch, W., 2008a. Synchronization of Hydrological and Hydrodynamic Models to Reduce Uncertainties in Case of a Limited Data Basis- The Example of the City of Linz. Korrespondenz Abwasser 55 (7). 91, 103, 104

Fach, S., Kleidorfer, M., Rauch, W., 2007. Anforderungen an Modelle und deren Kalibrierung. Forum Umwelttechnik und Wasserbau: Bemessung von Mischwasserentlastungen - Regelblattă19/neu 1, 99–116. 22, 42

Fach, S., Sitzenfrei, R., Rauch, W., 2008b. Assessing the relationship between water level and combined sewer overflow with computational fluid dynamics. In: 11th International Conference on Urban Drainage. Edinburgh, Scotland. 28, 44, 76, 101

Fienberg, S. E., 2006. When did Bayesian inference become “Bayesian"? Bayesian Analysis 1(1), 1 – 40. 64

Francey, M., Fletcher, T. D., Deletic, A., Duncan, H., in press. New insights into the quality of urban stormwater in south eastern Australia. Journal of Environmental engineering - ASCE. 6, 35

Freer, J., Beven, K., Ambroise, B., 1995. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. Water Resources Research 32(7), 2161 – 2173. 58

Freni, G., Mannina, G., Viviani, G., 2008a. Uncertainty assessment of sewer sediment erosion modelling. Urban Water Journal 5(1), 21 – 31. 58

Freni, G., Mannina, G., Viviani, G., 2008b. Uncertainty in urban stormwater quality modelling: The effect of acceptability threshold in the GLUE methodology. Water Research 42 (8-9), 2061 – 2072. 58, 70

Freni, G., Mannina, G., Viviani, G., 2009a. Uncertainty assessment of an integrated urban drainage model. Journal of Hydrology 373 (3-4), 392 – 404. 5, 58

Freni, G., Mannina, G., Viviani, G., 2009b. Urban runoff modelling uncertainty: Comparison among Bayesian and pseudo-Bayesian methods. Environmental Modelling & Software 24 (9), 1100 – 1111. 58

Gallagher, M., Doherty, J., 2007. Parameter estimation and uncertainty analysis for a watershed model. Environmental Modeling & Software 22, 1000 – 1020. 67

Gamerith, V., Muschalla, D., Koenemann, P., Gruber, G., 2009. Pollution load modelling in sewer systems: an approach of combining long term online sensor data with multi-objective auto-calibration schemes. Water Science & Technology 59 (1), 73–80. 74, 81

Gamerith, V., Muschalla, D., Schröter, K., Könemann, P., Gruber, G., 2008. Application of multi-objective auto-calibration schemes in pollution load modelling based on high-resolution calibration data sets. In: 11th International Conference on Urban Drainage. Edinburgh, Scotland. 81

Gaume, E., Villeneuve, J. P., Desbordes, M., 1998. Uncertainty assessment and analysis of the calibrated parameter values of an urban storm water quality model. Journal of Hydrology 210, 35–50. 5

Gillies, D., 1991. Intersubjective probability and confirmation theory. The British Journal for the Philosophy of Science 42 (4), 513–533. 63

Goldstein, M., 2006. Subjective Bayesian analysis: Principles and practice. Bayesian Analysis 1(3), 403 –420. 64

Grecu, M., Krajewski, W. F., 2000. A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting. Journal of Hydrology 239 (1-4), 69–84. 80

Gruber, G., Bertrand-Krajewski, J. L., de Bénédittis, J., Hochedlinger, M., Lettl, W., 2006. Practical aspects, experiences and strategies by using UV/VIS sensors for long-term sewer monitoring. Water Practice & Technology 1 (1). 29

Guisan, A., Zimmermann, N. E., 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135 (2-3), 147 – 186. 64

Gujer, W., 2002. Siedlungswasserwirtschaft, 2nd Edition. Springer-Verlag, Berlin. 3

Gujer, W., 2008. Systems Analysis for Water Technology. Springer-Verlag, Berlin. 11

Guven, B., Howard, A., 2007. Identifying the critical parameters of a cyanobacterial growth and movement model by using generalised sensitivity analysis. Ecological Modelling 207 (1), 11 – 21. 56

Harremoës, P., 2003. The role of uncertainty in application of integrated urban water modelling. In: IMUG Conference. Tilburg, Netherlands. 5

Harremoës, P., Rauch, W., 1996. Integrated design and analysis of drainage systems, including sewers, treatment plant and receiving waters. Journal of Hydraulic Research 34(6), 815 – 826. 18

Hastings, W., 1970. Monte Carlo sampling methods using markov chains and their applications. Biometrika 57, 97 – 109. 66

Hatt, B. E., Fletcher, T. D., Deletic, A., 2007. Treatment performance of gravel filter media: Implications for design and application of stormwater infiltration systems. Water Research 41 (12), 2513 – 2524. 2

Hatt, B. E., Fletcher, T. D., Deletic, A., 2009. Hydrologic and pollutant removal performance of stormwater biofiltration systems at the field scale. Journal of Hydrology 365 (3-4), 310 – 321. 2

Haydon, S., Deletic, A., 2009. Model output uncertainty of a coupled pathogen indicator-hydrologic catchment model due to input data uncertainty. Environmental Modelling & Software 24 (3), 322 – 328. 4

Henze, M., Gujer, W., Mino, T., van Loosedrecht, M., 2006. Activated sludge models ASM1, ASM2, ASM2d and ASM3. Water Intelligence Online 5. 18

Hochedlinger, M., 2005. Assessment of combined sever overflow emissions. Dissertation, Technische Universität Graz. 29

Hoppe, H., 2006. Unsicherheiten von Grundlagendaten im Rahmen integrierter Planungen urbaner Abwasserentsorgungssysteme. Schriftenreihe Siedlungswasserwirtschaft Bochum 51. 43

Hoppe, H., Gruening, H., 2007. Signifiance of uncertainties in the input data used in the integrated design of wastewater systems. In: 6th International Conference on Sustainable Techniques and Strategies in Urban Water Management - Novatech. Vol. 3. Lyon, France, pp. 1607 –1614. 4, 5

Hornberger, G. M., Spear, C. R., 1981. An approach to the preliminary analysis of environmental systems. Journal of Environmental Management 12, 7 – 18. 55

Hornberger, G. M., Spear, R. C., 1980. Eutrophication in peel inlet–I. the problem-defining behavior and a mathematical model for the phosphorus scenario. Water Research 14 (1), 29 – 42. 53, 55

ISO, 2008. ISO/IEC guide 98-3:2008: Uncertainty of measurement – part 3: Guide to the expression of uncertainty in measurement (GUM:1995). 40

ITWH, 2002. Hystem-Extran 6 Modellbeschreibung. Institut für technisch wissenschaftliche Hydrologie GmbH, Hannover, Germany. 87

Kanso, A., Gromaire, M.-C., Gaume, E., Tassin, B., Chebbo, G., 2003. Bayesian approach for the calibration of models: application to an urban stormwater pollution model. Water Science & Technology 47 (4), 77 – 84. 5, 64, 67

Kanso, A., Tassin, B., Chebbo, G., 2005. A benchmark methodology for managing uncertainties in urban runoff quality models. Water Science & Technology 51 (2), 163–170. 4, 67

Kavetski, D., Franks, S. W., Kuczera, G., 2003. Water Science & Applications Series 6. American Geophysical Union, Ch. Confronting input uncertainty in environmental modelling, pp. 49 – 68. 70, 72

Kavetski, D., Kuczera, G., Franks, S. W., 2006a. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. Water Resources Reseach 42 W03407. 4, 70, 72

Kavetski, D., Kuczera, G., Franks, S. W., 2006b. Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis. Journal of Hydrology 320, 187 – 201. 70, 72

Kleidorfer, M., 2005. Überrechnung und Bewertung eines Kanalsystems mittels hydrodynamischer Simulationen am Beispiel von Innsbruck. Master thesis, University of Innsbruck. 22, 87

Kleidorfer, M., Deletic, A., Fletcher, T. D., Rauch, W., 2009a. Impact of input data uncertainties on urban stormwater model parameters. Water Science & Technology. 4, 5, 7

Kleidorfer, M., Fach, S., Möderl, M., Rauch, W., 2007a. Implementing real time control (RTC) in hydrological sewer models as practised in Linz (in German). Österreichische Wasser- und Abfallwirtschaft 9-10, 131 – 137. 10

Kleidorfer, M., Fach, S., Möderl, M., Rauch, W., 2007b. Umsetzung von Kanalnetzsteuerungen in hydrologischen Modellen am Beispiel von Linz. Österreichische Wasser- und Abfallwirtschaft 59 (9-10), 131–137. 104

Kleidorfer, M., Fach, S., Rauch, W., 2008. Hinweise zur Kalibrierung von hydrologischen Modellen für die Anwendung von ÖWAV Regelblatt 19/neu. Wiener Mitteilungen 209. 9, 44, 85, 98

Kleidorfer, M., Leonhart, G., Mair, M., Kinzel, H., Rauch, W., submitted. Calimero: A model independent tool for autocalibration and uncertainty analysis. Water Practice and Technology. 9

Kleidorfer, M., Meyer, S., Möderl, M., Fach, S., Rauch, W., 2007c. Konzepterstellung für die Bewirtschaftung und Sanierung der Kanalisation Linz - Grobmodell. Tech. rep., IUT - Arbeitsbereich Umwelttechnik. 21

Kleidorfer, M., Meyer, S., Rauch, W., 2006a. Aspects of calibration of hydrological models for the estimation of CSO performance. In: 2nd International IWA Conference on Sewer Operation and Maintenance. Wien. 44, 45, 75, 76, 98, 99, 100

Kleidorfer, M., Meyer, S., Rauch, W., 2006b. Konzepterstellung für die Bewirtschaftung und Sanierung der Kanalisation Linz - Randbedingungen der Bewirtschaftung Hydrologische Grundlagen. Tech. rep., IUT - Arbeitsbereich Umwelttechnik. 21

Kleidorfer, M., Möderl, M., Fach, S., Rauch, W., 2007d. Konzepterstellung für die Bewirtschaftung und Sanierung der Kanalisation Linz - Teil 2. Tech. rep., IUT - Arbeitsbereich Umwelttechnik. 21

Kleidorfer, M., Möderl, M., Fach, S., Rauch, W., 2009b. Optimization of measurement campaigns for calibration of a conceptual sewer model. Water Science & Technology 58 (8), 1523–1530. 4, 7

Kleidorfer, M., Möderl, M., Rauch, W., 2006c. Kalibrierung von hydrodynamischen Kanalnetzberechnungen - Innsbruck Nord. Tech. rep., Institut für Infrastruktur, Arbeitsbereich Umwelttechnik, Universität Innsbruck. 21

Kleidorfer, M., Möderl, M., Sitzenfrei, R., Urich, C., Rauch, W., 2009c. A case independent approach on the impact of climate change effects on combined sewer system performance. Water Science & Technology. 8

Kleidorfer, M., Rauch, W., 2006. Konzeption eines Messprogramms für Feststoffe im Kanalisationssystems Innsbruck. Tech. rep., Institut für Infrastruktur, Arbeitsbereich Umwelttechnik, Universität Innsbruck. 21

Kleidorfer, M., Rauch, W., 2007a. Feststoffmessung mittels Multimeterspektrometrie. Tech. rep., Institut für Infrastruktur, Arbeitsbereich Umwelttechnik, Universität Innsbruck. 21

Kleidorfer, M., Rauch, W., 2007b. Kalibrierung von hydrodynamischen Kanalnetzberechnungen - Innsbruck Innenstadt. Tech. rep., Institut für Infrastruktur, Arbeitsbereich Umwelttechnik, Universität Innsbruck. 21

Kleidorfer, M., Rauch, W., 2007c. Risikomanagement Stadt Innsbruck - Hochwasser und Kanalisation. Tech. rep., Institut für Infrastruktir, Arbeitsbereich Umwelttechnik, Universität Innsbruck. 21

Kleidorfer, M., Rauch, W., 2007d. Wirkungsgradberechnungen nach ÖWAV Regelblatt 19 (2007) für das Kanalsystem Innsbruck. Tech. rep., Universität Innsbruck, Arbeitsbereich Umwelttechnik. 21

Koch, K. R., 1999. Introduction to Bayesian Statistics. Springer, Berlin. 63, 66

Korving, H., Clemens, F., 2005. Impact of dimension uncertainty and model calibration on sewer system assessment. Water Science & Technology 52 (5), 35–42. 5

Krejci, V., 1996. Integrated approach to the point-non point-pollution abatement in urban drainage. Water Science & Technology 33 (4-5), 9 – 15. 43

Krzysztofowicz, R., 2002. Bayesian system for probabilistic river stage forecasting. Journal of Hydrology 268 (1-4), 16 – 40. 64

Krzysztofowicz, R., Maranzano, C. J., 2004. Bayesian system for probabilistic stage transition forecasting. Journal of Hydrology 299 (1-2), 15 – 44. 64

Kuczera, G., Kavetski, D., Franks, S., Thyer, M., 2006. Towards a Bayesian total error analysis of coneptual rainfall-runoff models: Characterising model error using storm-dependent parameters. Journal of Hydrology 331, 161 – 177. 47, 64, 70, 72

Kuczera, G., Parent, E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. Journal of Hydrology 211, 69 – 85. 64, 67

Kuichling, E., 1889. The relation between the rainfall and the discharge of sewers in populous districts. Trans. Am. Soc. Civ. Eng. 20, 1–60. 4

La Barbera, P., Lanza, L. G., Stagi, L., 2002. Tipping bucket mechanical errors and their influence on rainfall statistics and extremes. Water Science & Technology 45(2), 1 – 9. 43

Legates, D., Davis, R., 1997. The continuing search for an anthropogenic climate change signal: Limitations of correlation-based approaches. Geophysical Research Letters 24, 18. 79

Legates, D., McCabe Jr, G., 1999. Evaluating the use of Şgoodness-of-fitŤ measures in hydrologic and hydroclimatic model validation. Water Resources Research 35 (1), 233–241. 79

Lei, J. H., 1996. Uncertainty analysis of urban rainfall - runoff modelling. Dissertation, Norwegian University of Science and Technology. 4

Lei, J. H., Schilling, W., 1996. Preliminary uncertainty analysis - a prerequisite for assessing the predictive uncertainty of hydrologic models. Water Science & Technology 33, 79–90. 4

Levenberg, K., 1944. A method for the solution of certain nonlinear problems in least squares. Quart. Appl. Math 2 (2), 164–168. 80

Li, Y., Deletic, A., Fletcher, T. D., 2007. Modelling wet weather sediment removal by stormwater constructed wetlands: Insights from a laboratory study. Journal of Hydrology 338 (3-4), 285 – 296. 2

Lindblom, E., Madsen, H., Mikkelsen, P., 2007. Comparative uncertainty analysis of copper loads in stormwater systems using GLUE and grey-box modelling. Water Science & Technology 56(6), 11 – 18. 4, 58

Lindley, D. V., 2006. Understanding Uncertainty. John Wiley & Sons. 64

Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. Journal of Hydrology 235 (3-4), 276–288. 81

Maksimovic, C., Buzek, L., Petrovic, J., 1991. Corrections of rainfall data obtained by tipping bucket rain gauge. Atmospheric Research 27, 45 – 53. 43

Mannina, G., Freni, G., Viviani, G., Saegro, S., Hafskjold, L., 2006. Integrated urban water modelling with uncertainty analysis. Water Science & Technology 54(6-7), 379 – 386. 5, 58

Mantovan, P., Todini, E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology 330 (1-2), 368 – 381. 48, 57

Mantovan, P., Todini, E., Martina, M. L., 2007. Reply to comment by Keith Beven, Paul Smith and Jim Freer on Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology 338 (3-4), 319 – 324. 57

Mark, O., Svensson, G., König, A., Linde, J., 2008. Analyses and adaptation of climate change impacts on urban drainage systems. In: 11th International Conference on Urban Drainage. Edinburgh, Scotland, UK. 4

Marquardt, D. W., 1963. An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics 11 (2), 431–441. 80

McCarthy, D., 2008. Modelling microorganisms in urban stormwater. Ph.D. thesis, Monash University, Melbourne. 45, 77

McMichael, C. E., Hope, A. S., Loaiciga, H. A., 2006. Distributed hydrological modelling in california semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. Journal of Hydrology 317 (3-4), 307 – 324. 58

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E., 1953. Equation of state calculations by fast computing machines. Journal of Chemical Physics 21 (6), 1087 – 1092. 66

Mikkelsen, P., Madsen, H., Arnbjerg-Nielsen, K., Jørgensen, H., Rosbjerg, D., Harremoës, P., 1998. A rationale for using local and regional point rainfall data for design and analysis of urban storm drainage systems. Water Science & Technology 37 (11), 7 – 14. 43

Minton, G., 2002. Stormwater Treatment. Resource Planning Associates. 2

Möderl, M., 2009. Modelltechnische Analyse non Netzerksystem der Siedlungswasserwirtschaft. Ph.D. thesis, University of Innsbruck. 91

Möderl, M., Butler, D., Rauch, W., 2009. A stochastic approach for automatic generation of urban drainagesystems. Water Science & Technology 59 (6), 1137 – 1143. 118

Möderl, M., Kleidorfer, M., Fach, S., Rauch, W., 2007a. Konzepterstellung für die Bewirtschaftung und Sanierung der Kanalisation Linz - Kanalbewirtschaftung Ű Teil 1. Tech. rep., IUT - Arbeitsbereich Umwelttechnik. 21, 103

Möderl, M., Kleidorfer, M., Fach, S., Rauch, W., 2007b. Reduktion von Mischwasseremissionen durch Optimierung eines Entwässerungssystems am Beispiel der Stadt Linz. Wiener Mitteilungen: Kanalmanagement 2007 - Unterirdische Kanalsanierung 203, K1–K22. 103

Möderl, M., Kleidorfer, M., Sitzenfrei, R., Rauch, W., 2009. Identifying weak points of urban drainage systems by means of vulnetud. In: Proceedings of the 8th International Conference on Urban Drainage Modelling. 49

Möderl, M., Meyer, S., Fach, S., Rauch, W., 2007c. Konzepterstellung für die Bewirtschaftung und Sanierung der Kanalisation Linz - Mischwasserbehandlung. Tech. rep., IUT - Arbeitsbereich Umwelttechnik. 21

Moré, J. J., 1978. The Levenberg-Marquardt algorithm: Implementation and theory. Springer. 80

Morin, E., Goodrich, D. C., Maddox, R. A., Gao, X., Gupta, H. V., Sorooshian, S., 2006. Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. Advances in Water Resources 29 (6), 843 – 860. 43

Mourad, M., Bertrand-Krajewski, J. L., Chebbo, G., 2005. Stormwater quality models: sensitivity to calibration data. Water Science & Technology 52 (5), 61–68. 45, 77

Muschalla, D., 2006. Evolutionäre multikriterielle otimierung komplexer wasserwirtschaftlicher systeme. Dissertation, Technische Universität Darmstadt. 81

Muschalla, D., Schneider, S., Schröter, K., Gamerith, V., Gruber, G., 2008. Sewer modelling based on highly distributed calibration data sets and multi-objective autocalibration schemes. Water Science & Technology 57 (10), 1547–1554. 74, 81, 85

Nash, J., Sutcliffe, J., 1970. River flow forecasting trough conceptual models part I - A discussion of principles. Journal of Hydrology 10 (3), 282–290. 46, 48, 79

Nokia, 2009. Qt-4.5 whitepaper. Nokia Corporation. 81

O'Hagan, A., 2006. Science, Subjectivity and Software (Comment on Articles by Berger and by Goldstein). Bayesian Analysis 1 (3), 445 – 450. 64

Omlin, M., Reichert, P., 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecological Modelling 115 (1), 45 – 59. 64

O'Neill, R. V., Gardner, R. H., 1979. Methodology in System Modelling and Simulation. Ch. Sources of uncertaintiey in ecological models, pp. 447 – 463. 41

ÖNORM EN 752, 1997. Entwässerungssysteme ausserhalb von gebäuden. 86

Overeem, A., Buishand, A., Holleman, I., 2008. Rainfall depth-duration-frequency curves and their uncertainties. Journal of Hydrology 348 (1-2), 124 – 134. 4

ÖWAV-RB 11, 2009. Richtlinie für die abwassertechnische Berechnung und Dimensionierung von Abwasserkanälen. 4, 6, 86, 87, 90

ÖWAV-RB 19, 2007. Richtlinie für die Bemessung von Mischwasserentlastungen. xxi, 4, 6, 9, 10, 95, 96, 97, 98, 101, 103, 106, 114

Page, T., Beven, K., Freer, J., A., J., 2003. Investigating the uncertainty in predicting responses to atmospheric deposition using the model of acidification of groundwater in catchments (MAGIC) within a generalised likelihood uncertainty estimation (GLUE) framework. Water, Air, & Soil Pollution 142(1-4), 71 – 94. 58

Palmstrom, N., Walker, W. W., 1990. P8 urban catchment model: UserŠs guide. Program documentation, and evaluation of existing models, design concepts, and Hunt-Potowomut data inventory. Tech. rep., The Narragansett Bay Project Report No. NBP-90-50. 16

Pappenberger, F., Beven, K., Frodsham, K., Romanowicz, R., Matgen, P., 2007. Grasping the unavoidable subjectivity in calibration of flood inundation models: A vulnerability weighted approach. Journal of Hydrology 333 (2-4), 275 – 287. 53

Rankinen, K., Karvonen, T., Butterfield, D., 2006. An application of the GLUE methodology for estimating the parameters of the INCA-N model. Science of The Total Environment 365 (1-3), 123 – 139. 58

Rauch, W., Aalderink, H., Krebs, P., Schilling, W., Vanrolleghem, P., 1998a. Requirements for integrated wastewater models – driven by receiving water objectives. Water Science & Technology 38 (11), 97 – 104. 18

Rauch, W., Bertrand-Krajewski, J. L., Krebs, P., Mark, O., Schilling, W., Schütze, M., Vanrolleghem, P. A., 2002. Deterministic modelling of integrated urban drainage systems. Water Science & Technology 45 (3), 81–94. 3, 18, 99

Rauch, W., Harremoës, P., 1996. The importance of the treatment plant performance during rain toacute water pollution. Water Science & Technology 34 (3-4), 1 – 8. 20

Rauch, W., Harremoës, P., 1998a. Correlation of combined sewer overflow reduction due to real-time control and resulting effect on the oxygen concentration in the river. Water Science & Technology 37 (12), 69 – 76. 20

Rauch, W., Harremoës, P., 1998b. On the application of evolution programs in urban drainage modelling. In: Fourth Conference on developments in Urban Drainage Modelling. London. 11, 80

Rauch, W., Kinzel, H., 2007. KAREN - User Manual. hydro-IT GmbH, Innsbruck. 18, 99, 114

Rauch, W., Seggelke, K., Brown, R., P., K., 2005. Integrated approaches in urban storm drainage: Where do we stand? Environmental Management 35(4), 396 – 409. 18

Rauch, W., Thurner, N., Harremoës, P., 1998b. Required accuracy of rainfall data for integrated urban drainage modeling. Water Science & Technology 37 (11), 81 – 89. 5, 43

Read, J., Wevill, T., Fletcher, T., Deletic, A., 2008. Variation among plant species in pollutant removal from stormwater in biofiltration systems. Water Research 42 (4-5), 893 – 902. 2

Reichert, P., 2009. Environmental System Analysis. Department of System Analysis, Integraated Assessment and Modelling, eawag: Swiss federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland. 40, 41, 47, 50, 51, 53, 65

Reichert, P., Borchardt, D., Henze, M., Rauch, W., Shanahan, P., Somlyody, L., Vanrolleghem, P., 2001. River water quality model no. 1 (RWQM1): II. Biochemical process equations. Water Science & Technology 43(5), 11 – 30. 18

Renard, B., Kavetski, D., Kuczera, G., 2009. Comment on ŚŚan integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic predictionŠŠ by Newsha K. Ajami et al. Water Resources Research 45, W03603. 72

Roberson, J. A., Cassidy, J. J., Chaudhry, M., 1998. Hydraulic Engineering. John Wiley and Sons Ltd. 13

Rossman, L. A., 2007. Storm Water Management Model User's Manual Version 5.0. EPA Unites Sates Environmental Protection Agency. 16

Saltelli, A., Ratto, M., Tarantola, S., Campolongo, F., 2006. Sensitivity analysis practices: Strategies for model-based inference. Reliability Engineering & System Safety 91 (10-11), 1109 – 1125, the Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004) - SAMO 2004. 53

Sartor, J. D., Boyd, G. B., Agardy, F. J., 1974. Water pollution aspects of street surface contaminants. Journal of Water Pollution Control Fed 46 (3), 458–467. 16

Schilling, W., 1991. Rainfall data for urban hydrology: what do we need? Atmospheric Research 27 (1-3), 5 – 21. 43

Segond, M.-L., Wheater, H. S., Onof, C., 2007. The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the lee catchment, uk. Journal of Hydrology 347 (1-2), 116 – 131. 43

Semadeni-Davies, A., Hernebring, C., Svensson, G., Gustafsson, L.-G., 2008. The impact of climate change and urbanisation on drainage in helsingborg, sweden: Combined sewer system. Journal of Hydrology 350 (1-2), 100–113. 4

Sevruk, B., 1981. Instruments and observing methods. Papers presented at the second WMO technical Conference on instruments and methods of observation 9. 43

Sevruk, B., 1982. Methods of correction for systematic error in point precipitation measurement for operational use. World Meteorological Operational Meteorology Rep 21(589). 43

Sevruk, B., 1996. Adjustment of tipping-bucket precipitation gauge measurement. Atmospheric Research 42, 237 – 246. 43

Sevruk, B., 2002. WMO questionnaire on recording precipitation gauges: state-of-the-art. Water Science & Technology 45(2), 139 – 145. 43

Shanahan, P., Borchardt, D., Henze, M., Rauch, W., Reichert, P., Somlybody, L., P., V., 2001. River water quality model no. 1 (RWQM1): I. Modelling approach. Water Science & Technology 43(5), 1 – 9. 18

Sieber, A., Uhlenbrook, S., 2005. Sensitivity analyses of a distributed catchment model to verify the model structure. Journal of Hydrology 310 (1-4), 216 – 235. 56

Silberstein, R., 2006. Hydrological models are so good, do we still need data? Environmental Modeling & Software 21 (9), 1340–1352. 4

Siriwardene, N., Deletic, A., Fletcher, T. D., 2007. Clogging of stormwater gravel infiltration systems and filters: Insights from a laboratory study. Water Research 41 (7), 1433 – 1440. 2

Sitzenfrei, R., Fach, S., Rauch, W., 2008. Auswirkung von Vereinfachungen bei der Bestimmung von Mischwasserentlastungsmengen - Quo vadis Poleni? Wiener Mitteilungen 209. 44, 101

Sitzenfrei, R., Kinzel, H., Rauch, W., 2009. A multi-layer cellular automata approach for algorithmic generation of virtual urban water systems - VIBe. In: Proceedings of the 8th International Conference on Urban Drainage Modelling. 118

Smith, P., Beven, K. J., Tawn, J. A., 2008. Informal likelihood measures in model assessment: Theoretic development and investigation. Advances in Water Resources 31 (8), 1087 – 1100. 48, 67

Spear, R. C., Hornberger, G. M., 1980. Eutrophication in peel inlet–II. identification of critical uncertainties via generalized sensitivity analysis. Water Research 14 (1), 43 – 49. 53, 55

Tang, Y., Reed, P., Wagener, T., van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. Hydrology and Earth System Sciences 11 (2), 793–817. 53

Thorndahl, S., 2008. Uncertainty assessment in long term urban drainage modelling. Phd thesis, Aalborg University. 5, 39, 56, 94

Thorndahl, S., Beven, K., Jensen, J., Schaarup-Jensen, K., 2008. Event based uncertainty assessment in urban drainage modelling, applying the GLUE methodology. Journal of Hydrology 357 (3-4), 421 – 437. 5, 53, 58

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., Srikanthan, S., 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resources Research 45 W00B14. 70, 72, 73

Tierney, L., 1994. Markov chains for exploring posterior distributions. The Annals of Statistics 22 (4), 1701–1728. 67

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. Science 185(4157), 1124 – 1131. 65

Urich, C., Sitzenfrei, R., Möderl, M., Rauch, W., 2009. An agent based approach for generating virtual sewer systems in the software VIBe. In: Proceedings of the 8th International Conference on Urban Drainage Modelling. 118

Vaes, G., Willems, P., Berlamont, J., 2005. Areal rainfall correction coefficients for small urban catchments. Atmospheric Research 77 (1-4), 48 – 59. 43

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., Srinivasan, R., 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. Journal of Hydrology 324 (1-4), 10 – 23. 56

Vanrolleghem, P., Borchardt, D., Henze, M., Rauch, W., Reichert, P., Shanahan, P., Somlyody, L., 2001. River water quality model no. 1 (RWQM1): III. Biochemical submodel selection. Water Science & Technology 43(5), 31 – 40. 18

Villarini, G., Lang, J. B., Lombardo, F., Napolitano, F., Russo, F., Krajewski, W. F., 2007. Impact of different regression frameworks on the estimation of the scaling properties of radar rainfall. Atmospheric Research 86 (3-4), 340 – 349. 43

Villarini, G., Serinaldi, F., Krajewski, W. F., 2008. Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. Advances in Water Resources 31 (12), 1674 – 1686. 43

Vrugt, J. A., Gupta, H. V., Bouten, W., Sorooshian, S., 2003. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resources Research 39(8). 68

Willmott, C., 1981. On the validation of models. Physical geography 2 (2), 184–194. 79

Willmott, C., Ackleson, S., Davis, R., Feddema, J., Klink, K., Legates, D., Rowe, C., 1985. Statistics for the evaluation and comparison of models. Journal of Geophysical Research 90 (C5), 8995–9005. 79

XP-Software, 1999. Xp-aqualm - technical description. Tech. rep., XP Software, Australia. 16

Young, P., 1983. Uncertainty and Forecasting of Water Quality. Springer, New York, Ch. The validity and credibility of models for badly defined systems, pp. 69 – 98. 53, 55

Zhu, H.-j., Schilling, W., 1996. Simulation errors due to insufficient temporal rainfall resolution – annual combined sewer overflow. Atmospheric Research 42 (1-4), 19 – 32. 43