

Stochastic Modelling of Hydrologic Systems

Harpa Jonsdottir

Kongens Lyngby 2006
IMM-PHD-2006-150

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

IMM-PHD: ISSN 0909-3192

Summary

In this PhD project several stochastic modelling methods are studied and applied on various subjects in hydrology. The research was prepared at the Department of Informatics and Mathematical Modelling at the Technical University of Denmark.

The thesis is divided into two parts. The first part contains an introduction and an overview of the papers published. Then an introduction to basic concepts in hydrology along with a description of hydrological data is given. Finally an introduction to stochastic modelling is given.

The second part contains the research papers. In the research papers the stochastic methods are described, as at the time of publication these methods represent new contribution to hydrology. The second part also contains additional description of software used and a brief introduction to stiff systems. The system in one of the papers is stiff.

In Paper [A] a conditional parametric modelling method is tested. The data originate from a waste water treatment plant in Denmark, and consists of precipitation measurements and flow in a sewage system. The goal is to predict the flow and the predictions are to be used for automatic control in the waste water treatment plant. The conditional parametric modelling method is a black box method. The characteristic of such a model is that the model's parameters are not constants, but vary as a function of some external variables. In Paper [A] two types of conditional parameter models were tested; a conditional FIR model and a conditional ARX model. The parameter variation is modelled as a local regression and the results are significant improvements compared to the

traditional linear FIR and ARX models. The method of conditional parameter modelling is also good for sensitivity analysis since it can be used to investigate how parameters change when external variables/circumstances change. These investigations might then be used for further (more global or even more physical) modelling development.

In Paper [B] the grey box modelling approach is used, by using Stochastic Differential Equations. The parameter estimation is performed by use of the program CTSM (Continuous Time Stochastic Modelling). The field of study is the traditional rainfall-runoff relationship in a large watershed with one precipitation measurement station, one discharge measurement station and snow accumulation during winter. The rainfall-runoff relationship is thus both non-linear and non-stationary. Furthermore, the system is stiff, and advanced statistical and numerical methods must be used for parameter estimation. The model structure is kept simple in order to be able to identify all the model parameters. The case study is from a 1132 km² mountainous area in northern Iceland with altitude range of about 1000 m. The model performs well, despite of the fact that the input series is only one single series of temperature and one single series of precipitation, measured in the valley, close to the river mouth.

In Paper [C] the topic is a drought analysis in a reservoir related to a hydropower plant. A stochastic model is developed and the model is used to simulate a time series of discharge data which is long enough to achieve a stable estimate for risk assessment of water shortage. Since the available data are used to design the hydropower plant, it is demonstrated that the only way to estimate the risk of water shortage during a hydropower's lifetime is by using a stochastic simulation.

In Paper [D] the data originate from a small creek in Denmark with two measurement stations. The subject is flow routing where the upstream flow is used as an input for modelling the downstream flow. As in Paper [B], the grey box modelling approach is used, by using Stochastic Differential Equations and the parameter estimation is performed by use of the program CTSM. The model formulation is a linear reservoir model. However, the non-measured lateral inflow between the two measurement stations is modelled as a state variable and thus a dynamic estimate of the flow is achieved. This can be useful when modelling chemical processes in the water.

In general, the papers show the advantages of stochastic modelling for describing both non-linearities and non-stationaries in hydrological systems.

Resume

I dette Ph.D. projekt er forskellige stokastiske modelleringsmetoder studeret og afprøvet inden for forskellige områder i hydrologi. Forskningen har været udført ved Informatik og Matematisk Modellering, DTU.

Afhandlingen er delt i to dele. Den første del indeholder en indledning og en oversigt over fire artikler, skrevet som en del af forskningsarbejdet. Dernæst kommer en introduktion til grundlæggende begreber i hydrologi samt en beskrivelse af hydrologiske data. Til slut er der en introduktion til stokastisk modellering.

Anden del indeholder de 4 artikler, hvor de stokastiske metoder, og især hvordan disse metoder yder nye bidrag til den hydrologiske videnskab, er beskrevet. Anden del indeholder også en nærmere beskrivelse af software samt indledning til den matematiske analyse af stive systemer.

I Artikel [A] er betinget parametrisk modellering afprøvet. De data, som bruges, stammer fra et rensningsanlæg i Danmark. Disse data består af nedbør i et afstrømningsområde og afstrømningsmålinger i områdets kloaksystem.

Formålet er en forudsigelse af afstrømningen med de formål at bruge forudsigelserne i automatisk kontrol i rensningsanlæggets driftssystem. Den betingede parametriske modelleringsmetode er en black box metode. Kendetegnet ved denne type af modeller er, at modellens parametre ikke er konstante, men ændrer sig som funktioner af ydre forhold. I Artikel [A] er afprøvet to typer af betingede parametriske modeller: Betingede FIR modeller og betingede ARX modeller. Parametrenes dynamik er modelleret ved lokal regression, og resultaterne er en markant forbedring i forhold til de traditionelle lineære FIR og

ARX modeller. Den betingede parametriske modelleringsmetode er også nyttig i følsomhedsanalyser, da den kan bruges til at undersøge, hvordan parametrene ændrer sig efter ændringer i de ydre forhold. Denne type af undersøgelse kan bruges til videre modeludvikling, eventuelt til udvikling af mere fysisk udformede modeller.

I Artikel [B] er grey box modelleringsmetoden brugt, ved at bruge Stokastiske Differential Ligninger. Parameterestimationen er udført ved at bruge programmet CTSM (Continuous Time Stochastic Modelling). Sammenhængen mellem nedbør og afstrømning fra et stort opland er studeret. Om vinteren falder nedbør både som regn og sne i bjergene. Sneen samles op som vand i elven om foråret. Afhængigheden mellem nedbør og afstrømning er derfor ikke lineær og ikke stationær. Desuden er systemet stift, hvilket gør avancerede statistiske og numeriske metoder nødvendige. Modellens struktur er enkelt formuleret, således at alle modelles parametre kan identificeres. Data stammer fra en elv i det nordlige Island. Oplandet er 1132 km² med en højdeforskel på 1000 m. Selvom modellen kun bruger én nedbørsserie og én temperaturserie som input, virker modellen overordenlig tilfredsstillende.

Emnet i Arikel [C] er en analyse af risikoen for tømning af et vandmagasin i en flod i Island, som vil forårsage elsvigt fra det tilsluttede vandkraftværk. En stokastisk model er udviklet, og den er brugt for at simulere en afstrømningsdataserie, som er lang nok til at opnå et stabilt estimat for risiko for tømning af vandmagasinet. Det er vist, at den eneste tilfredsstillende måde til at estimere risikoen for tømning af vandmagasinet i vandkraftværkets økonomiske levetid, brugning af stokastisk simulation, da alle de eksisterende data er brugt til at designe vandkraftværket og magasinet. Populære ingeniørmæssige metoder som f. eks. sumkurvemetoden kan ikke håndtere hændelser med længere gentagelsesperioder end måleseriens længde.

I Artikel [D] bruges data, som stammer fra en lille å i Danmark med to målestationer i åen. Emnet er at forudsige vandhøjden ved nedstrømspunktet i systemet på basis af målinger ved opstrømspunktet i systemet samt nedbørsmålinger. Lige som i Artikel [B] er grey box modelleringsmetoden brugt ved at bruge Stokastiske Differentialligninger, og parameterne er estimeret i programmet CTSM. Modellens formulering er en lineær reservoir model, dog med den utraditionelle tilføjelse, at den ikke målte indstrømning imellem de to stationer er indført som en tilstandsvariabel. Dette medfører, at indstrømningen mellem de to stationer er estimeret dynamisk. Det har den fordel, at resultaterne kan bruges, når kemiske prosesser skal modelleres.

Generelt viser artiklerne fordele ved at bruge stokastisk modellering, der kan bruges til at analysere både ikke linearitet og ikke stationaritet i hydrologiske systemer.

Preface

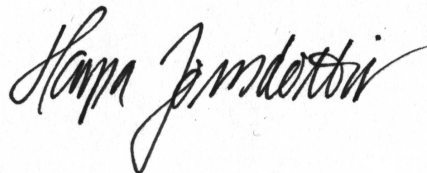
This thesis is a part of the fulfillment in completion of the PhD degree in engineering at the Department of Informatics and Mathematical Modelling (IMM) at the Technical University of Denmark. The projects were carried out at the IMM and at the National Energy Authority in Iceland.

Different stochastic models have been developed and tested on different hydrological problems. The main focus is on the modelling methodology, the parameter identification and the importance of stochastic modelling in general.

The thesis consists of a summary report, a short introduction to hydrology, as well as an introduction to stochastic modelling. Furthermore, a description of software and introduction to stiff system can be found in appendices, along with a collection of four research papers, already published or to be published.

Reykjavik,

June 2006

A handwritten signature in black ink, reading "Harpa Jonsdottir". The signature is written in a cursive style with a large, prominent initial 'H'.

Harpa Jonsdottir

Papers included in the thesis

- [A] Harpa Jonsdottir, Henrik Aalborg Nielsen, Henrik Madsen, Jonas Eliasson, Olafur Petur Palsson and Marinus K Nielsen. Conditional parametric models for storm sewer runoff *Water Resources Research*. Accepted.
- [B] Harpa Jonsdottir, Olafur P Palsson and Henrik Madsen. Parameter estimation in a stochastic rainfall-runoff model *Journal of hydrology*, 2006, Vol 326, p. 379-393.
- [C] Harpa Jonsdottir, Jonas Eliasson and Henrik Madsen. Assessment of serious water shortage in the Icelandic water resource system. *Physics and Chemistry of the Earth*, 2005, Vol 30, p. 420-425.
- [D] Harpa Jonsdottir, Judith L. Jacobsen and Henrik Madsen. A grey box model describing the hydraulic in a creek. *Environmetrics.*, 2001, p. 347-356.

Acknowledgements

This PhD project is carried out at the Department of Informatics and Mathematical Modelling (IMM) at the Technical University of Denmark, DTU, and at the Hydrological Service division at the National Energy Authority in Iceland. Thus, I have been in contact with many interesting and competent people while working on this project. I would like to thank them all, some for inspiring conversations, others for making me laugh.

First of all, I want to thank my supervisors, Professor Henrik Madsen at IMM DTU, Associate Professor Ólafur Pétur Pálsson and professor Jónas Elíasson, both at the Faculty of Engineering, at the University of Iceland (UI). They have been both supportive and patient and I have been very lucky to have had the opportunity to work with such experienced and qualified people, Elíasson in the field of hydrology and Madsen and Pálsson in the fields of statistics and mathematical modelling.

While working on the different projects I have cooperated with various people:

While working on my first paper, published in *Environmetrics*, PhD Judith J. Jacobsen PhD, was a great support.

The project of parameter estimation in a rainfall-runoff model, the results of which were published in *Journal of Hydrology*, turned out to be a challenging project. It involved a numerical task not quite seen on before. Professor Sven Sigurdsson, and Associate Professor Kristján Jónasson, both at UI, provided valuable advice related to numerical methods and stiff systems. They are both experts in numerical analysis. PhD Niels Rode Kristensen who implemented the

latest version of the program CTSM used for parameter estimation in stochastic differential equations also gave qualified advises.

The project concerning the waste water in a sewage system also involved many people. PhD Marinus K. Nielsen provided valuable information about the system and Associate Professor Henrik Aalborg Nielsen at IMM, provided assistance related to the LFLM software package in S-PLUS.

Furthermore, I want to thank the staff at the Hydrological Service for allowing me to be one of them, particularly, Stefanía G. Halldórsdóttir. I also want to thank the departments heads, Árni Snorrason, Páll Jónsson and Kristinn Einarsson for professional assistance and for providing me with such good facilities.

I cannot complete this acknowledgement without thanking my two daughters Katrin and Freyja who during the last months often had to put of with a tired and irritated mom, and I am very grateful to my mother and to my sister for all their babysitting.

To all others not mentioned: Thanks.

Contents

Summary	i
Resume	iii
Preface	v
Papers included in the thesis	vii
Acknowledgements	ix
1 The theme	1
1.1 Overview of papers included	1
1.2 Comparison of the models	9
1.3 Why stochastic modelling?	12
1.4 Conclusion and discussion	13
2 Introduction to hydrology	19

2.1	The history of water resources	19
2.2	The hydrological cycle	23
2.3	The storage effect	24
2.4	Surface water hydrology	25
2.5	Subsurface water and groundwater	32
3	Hydrological data	33
3.1	Precipitation	33
3.2	Discharge	35
3.3	River ice	39
4	The stochastic dynamic modelling	45
4.1	Model categorization	45
4.2	Non-linear models in general	46
4.3	Stochastic differential equations and parameter estimation	48
4.4	The family of linear stochastic models	53
4.5	Non-linear models in discrete time	57
4.6	Overview of the statistical methods	62
4.7	Motivation for using grey box modelling	65
A	Conditional parametric models for storm sewer runoff	69
B	Parameter estimation in a stochastic rainfall-runoff model	79
C	Assessment of serious water shortage	

in the Icelandic water resource system	97
D A grey box model describing the hydraulics in a creek	105
E The program CTSM	117
E.1 Introduction	118
E.2 Filtering methods	119
E.3 Optimization routine	121
F Stiff systems	123
F.1 Introduction	124
F.2 The stiffness of the non-linear system in Paper [B]	126

CHAPTER 1

The theme

This thesis is a compilation of a PhD project at the Institute of Informatics and Mathematical Modelling, at the Technical University of Denmark. The field of research is stochastic modelling in hydrology. New methods are tested and applied to different hydrological problems. A description of statistical/numerical methods and the results of the applications are found in the papers [A]-[D].

1.1 Overview of papers included

The hydrological subjects are on very different scales and with different aspects. The research is within the field of statistics as well as within hydrology and in all of the research projects, empirical measurements are used to estimate unknown parameters.

1.1.1 Paper [A] Conditional parametric models for storm sewer runoff

In Paper[A], the data originates from a waste water treatment plant in Denmark. The treatment plant is the outlet of a sewage system with a watershed of 10.89

km². The sewage system is built in the traditional manner with pipes and node points for pumping stations. Figure 1.1 shows a sketch of a sewage system.

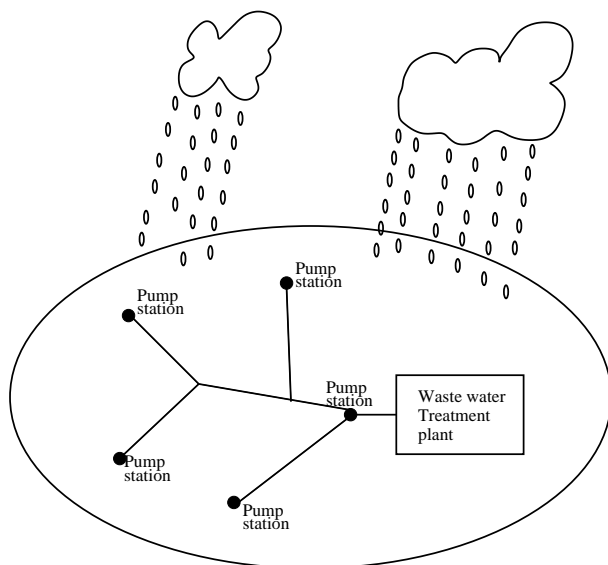


Figure 1.1: The sewage system

The input data is precipitation, measured at the waste water treatment plant. The output data is excess flow data from the last pumping station before the treatment plant.¹ The goal is to predict the flow in the last pumping station and use the predictions for on-line automatic control in the waste water treatment plant. Black box models have proven to provide good predictions in hydrological systems e.g., Carstensen et al. (1998) and thus such methods were tested. Linear FIR and linear ARX models were unsatisfactory and thus non-linear methods were used. The non-linear effects are mainly due to two factors; seasonality in the balance and saturation/threshold in the pipe system. Large parts of the measured precipitation do not enter the sewage system but evaporate or infiltrate into the ground. The infiltration rate depends on several factors and the wetness of the root zone plays an important role. Similarly, many factors affect the evaporation and especially the temperature plays a major role. Because of seasonal variations of temperature, plant growth and other physical factors, the variation of infiltration and evaporation varies seasonally and consequently the water balance does too. The other non-linear effect, the saturation/threshold is

¹The base flow in the sewage system, also known as dry weather flow, does not originate from rainfall. Consequently, the base flow is subtracted from the flow data and the resulting flow, the excess flow is used in the modelling approach.



Figure 1.2: Fnjorskadalur. (Photo Oddur Sigurdsson)

a consequence of limited capacity of the pumps in the sewage system. When a large amount of water enters the system the pump stations in the node points cannot serve all the water. Thus, water accumulates behind the pumping stations waiting to be served. During a very heavy rain storm the water enters the treatment plant with a delay, as compared to a normal rain storm. These two factors were taken into account in a conditional parametric model. Conditional parametric models are models where the parameters change as a function (conditioned) of some external variables. In this case the parameters changed as a function of seasonality and as a function of water quantity in the system. The method of conditional parametric modelling is a significant improvement compared to traditional linear modelling.

1.1.2 Paper [B]

Parameter estimation in a stochastic rainfall-runoff model

The subject in Paper[B] is a classic topic in hydrology, the rainfall-runoff relationship. The data originates from a 1132 km² mountainous watershed in Iceland. Figure 1.2 shows a part of the watershed. It shows the valley Fnjorskadalur and the river Fnjoska. The altitude range is about 1000 meters, stretching from 44 m to 1083 m. More than 50% of the watershed is above 800

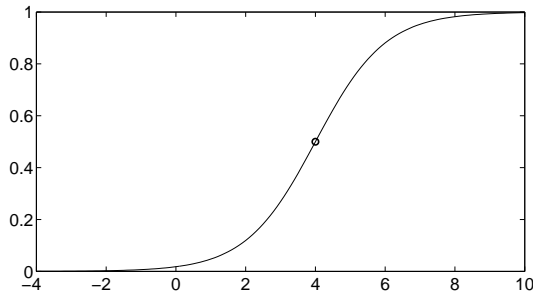


Figure 1.3: A sigmoid function with center 4 and scale parameter 1.

meters. The water level gauge is down in the valley, close to the river mouth. One meteorological observatory is in the watershed and it is located in the valley. Thus no meteorological observatory is located in the highlands nor close to the watershed in the highlands. The scarcity of meteorological observatories is well known in sparsely populated areas around the world, especially in mountainous areas. Because the watershed is large, and with a large altitude range, the weather condition in the watershed can be very different depending on location in the watershed. Furthermore, during winter, snow accumulates, and melts in spring, resulting in large spring floods in the river. Despite of limited data, a rainfall-runoff relationship was required. It was chosen to develop a stochastic conceptual model, and it is found necessary to use a stochastic model since too many effects are unknown and/or not measured.

The system is modelled in a continuous time by using stochastic differential equations. The model structure is kept as simple as possible and with as few parameters as possible in order to be able to use the data to estimate the parameter values. The stochastic differential equations describe a reservoir model with a snow routine. The watershed is not divided into elevation zones, but a smooth threshold function is used in the snow routine both for accumulation and melting, using positive degree day method. The smooth threshold function is the sigmoid function,

$$\phi(T) = \frac{1}{1 + \exp(b_0 - b_1 T)} \quad (1.1)$$

where T is temperature, b_0 and b_1 are constants. The constant b_0 is the center of the sigmoid function and b_1 controls the steepness. Figure 1.3 shows a sigmoid function with center $b_0 = 4$ and scale $b_1 = 1$.

In Figure 1.4 the modelling principle is illustrated. Precipitation enters the system and is divided into snow and rain, depending on the temperature. It can be rain only, snow only and partly snow and rain. The rain enters the first

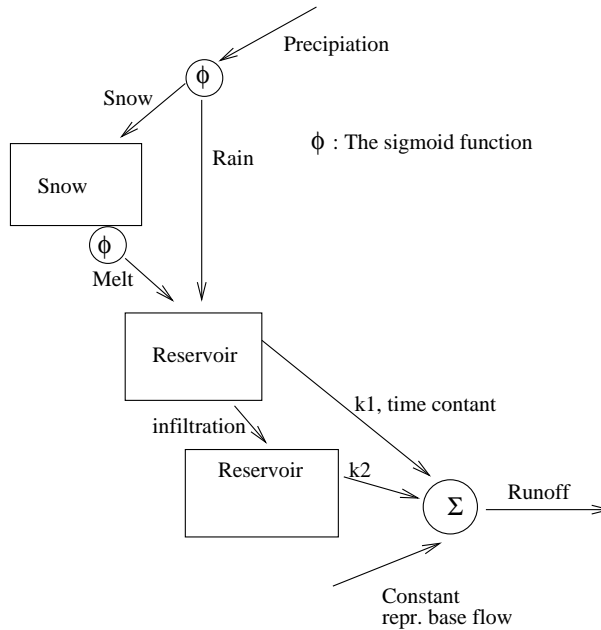


Figure 1.4: The modelling principles.

reservoir which delivers water partly directly into the river and partly into the second reservoir that finally delivers the water into the river. The snow, however, enters the snow container and stays in the snow container until it melts, and is then delivered into the first container. As mentioned earlier, the same smooth threshold function is used for precipitation division and snow melting. Thus, at a same time a precipitation can be divided into partly rain and partly snow while some ratio of the snow is melting. This modelling method computes precipitation division and melting on an average basis. This works well, particularly since no meteorological observatory is located in higher altitudes so that temperature lapse rate and precipitation lapse rate can be estimated and used as a basis for elevation division.

During the winter the snow container, because of its nature, swallows the snow and accumulates it until the temperature rises and the snow begins to melt. During the melting, the snow container delivers water into the system until the snow container is emptied. During summer, the snow container is inactive. Consequently, the snow routine causes the system to be both non-linear and stiff and, therefore, difficult to cope with numerically.

The parameters are estimated by using the program CTSM (Continuous Time

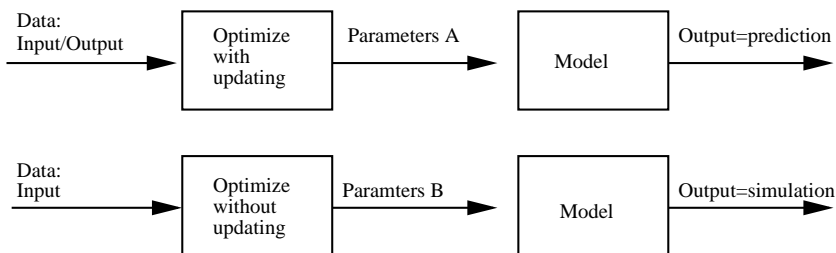


Figure 1.5: The input and output for different optimization principles.

Stochastic Modelling) (Kristensen et al. 2003). The estimation method is the maximum likelihood method and the principle of extended Kalman filter is used. Three different filtering methods or ODE solvers are implemented and it depends on the system's stiffness which one is the "optimal" to use. Furthermore, it is possible to choose to optimize the parameters with or without the traditional Kalman filter updating. Figure 1.5 illustrates the optimization principles. The optimized parameter values will not be the same we call them A and B. Optimization with Kalman filter updating results in a parameter Set A and those parameters are optimal for making model prediction. Contrarily, optimization without Kalman filter updating results in parameter Set B which is optimal for making model simulations if the true model exists.

1.1.3 Paper [C]

Assessment of serious water shortage in the Icelandic water resource system.

The topic in Paper[C] is a risk assessment of a water shortage in a hydropower plant. The data originates from the river Tungnaá in southern Iceland, measured at Mariufossar. The watershed is 3470 km², of which 555 km² is glacier. The data consist, of daily values of discharge over a period of 50 years. Figure 1.6 shows a hydropower plant and the corresponding reservoir. The water in the reservoir is led to the hydropower plant in pipes, located in the mountain, and if necessary bypass flow is led into the canyon which is on left side of the reservoir.



Figure 1.6: The hydropower plant Burfell and its reservoir. (Photo: Oddur Sigurdsson)

When a hydropower plant is designed, two major quantities are taken into consideration. One is the regulated flow, Q_{reg} , which is the flow of water delivered into the hydropower plant for electricity production. The other quantity is the size of the reservoir, V . For a given regulated flow Q_{reg} and for a given discharge series, a volume V exists, which is the smallest volume that can secure regulated flow Q_{reg} . The largest possible Q_{reg} which can be served without any water shortage is the mean value of the discharge. The relationship (Q_{reg}, V) is known as the regulation curve. Figure 1.7 shows the regulation curve for the discharge series at Mariufossar.

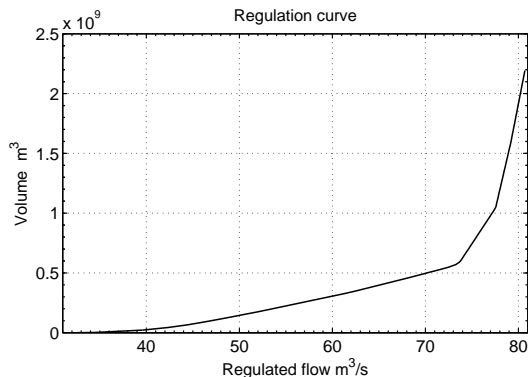


Figure 1.7: The regulation curve.

All available data are used to construct the regulation curve, and using a different discharge series will lead to another regulation curve. Thus, for a given point (Q_{reg}, V) on the regulation curve, the risk of a water shortage is zero using the data series which was used to construct the regulation curve. However, when the hydropower plant has been designed by choosing (Q_{reg}, V) , the future discharge series will not be exactly the same as the past discharge series and, thus, water shortage might occur. Consequently, a stochastic model must be developed in order to construct a simulated discharge series to be used for risk assessment. It is very important to have an estimation of the risk of water shortage in the lifetime of the hydropower plants, about 30-60 years.

A stochastic periodic model in the spirit of Yevjevich, (Yevjevich 1976) was developed and the available data were used to estimate the parameters in the stochastic model. The stochastic model is then used to simulate flow series in order to estimate the water shortage probabilities. The goal is to estimate probabilities of rare events and it turned out that it was necessary to simulate the daily flow for 50000 years in order to achieve a stable estimate of the risk of water shortage. Using the simulated data it was concluded that the water shortage probabilities can be described by the Weibull distribution. However, even though the distribution of the water shortage probabilities is known, simulations are required in order to estimate the parameters in the distribution.

1.1.4 Paper [D]

A grey box model describing the hydraulic in a creek.

In Paper[D] the subject is flow routing in a creek in a small watershed, in Northern Zealand in Denmark. The exact size of the watershed is not known. There are two measuring stations in the creek, see Figure 1.8. The available

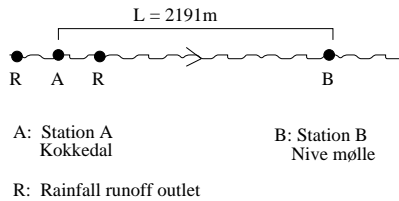


Figure 1.8: A sketch of the area in Usserod river.

data are precipitation and depth at two locations in the creek. The goal is to find a relationship between the depth at the upstream station and the depth at the downstream station and to predict the output depth at the downstream station. The Saint Venant equation of mass balance is used as a basis and the lateral inflow between the two measuring stations is modelled as a first order process with precipitation as input. The resulting model is a stochastic linear reservoir model described in continuous time by stochastic differential equations. The model is, however, different from the traditional reservoir model in that the lateral inflow of water between the two measuring stations is a state variable in the model and estimated by use of the Kalman filtering technique. This can be used in an environmental context so that it might be possible to estimate the concentration of chemical concentrations in the lateral inflow if the corresponding chemical concentrations are estimated both upstream and downstream. This can be very valuable in an environmental analysis. The program CTSM was used to estimate the parameters.

1.2 Comparison of the models

The hydrological subjects in this PhD project are on very different scales and with different aspects. However, all the projects are within the theory of hydrology. Thus the physical law, *conservation of mass* is the fundamental law. In hydrology this can be referred to as *the storage effect*, i.e., what comes in is either stored or comes out, see Figure 1.9.

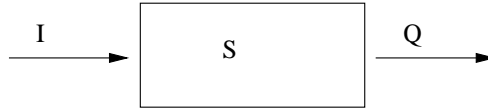


Figure 1.9: The storage effect

The storage equation is written

$$\frac{dS}{dt} = I(t) - Q(t) \quad (1.2)$$

where $I(t)$ is the input, $Q(t)$ is the output and $S(t)$ is the storage. The change in storage is the difference of input and output. All the projects/papers have to do with the storage but in different aspects.

The different storage interpretations can be seen graphically in Figure 1.10. In the following the different storage effects are summarized whereas an overview of the included papers are given in section 1.1

The subject of Paper[A] is a rainfall-runoff relationship in a sewage system. The input is precipitation and the output is excess flow. The storage is twofold. Firstly the storage is the time lag between input and output and, secondly, the storage is the long term storage. The model is an input-output model or a black box model, and since the input is precipitation and not effective precipitation the mass balance is not conserved in the model. This can be interpreted so that the storage container either swallows the rain or stores it on a long term basis. However, water comes out of the system eventually. Part of the water evaporates and some is permeated by plants. However, large part infiltrates into the root zone and becomes groundwater and can eventually be observed in creeks and rivers.

The subject in Paper[B] is also a rainfall-runoff relationship, the input is precipitation and the output is discharge. The model is not a mass balance model, since evaporation/transpiration are not taken into account, and the base flow is represented by a constant. However in this project there is no swallowing, a balance between input and output exists, only the "up-scaling" of the precipitation measurements is underestimated due to the amount of evaporation/transpiration and groundwater contribution. In this project the storage is time delay. It is a short-term time delay between rain and discharge during the summer and because of the snow storage it is a long-term time delay during winter time.

The topic in Paper[C] is in a different category. The topic is a risk assessment of a water shortage in a hydropower plant, i.e., the risk of emptying the reservoir. The input data is discharge series and the output is risk assessment, i.e.,

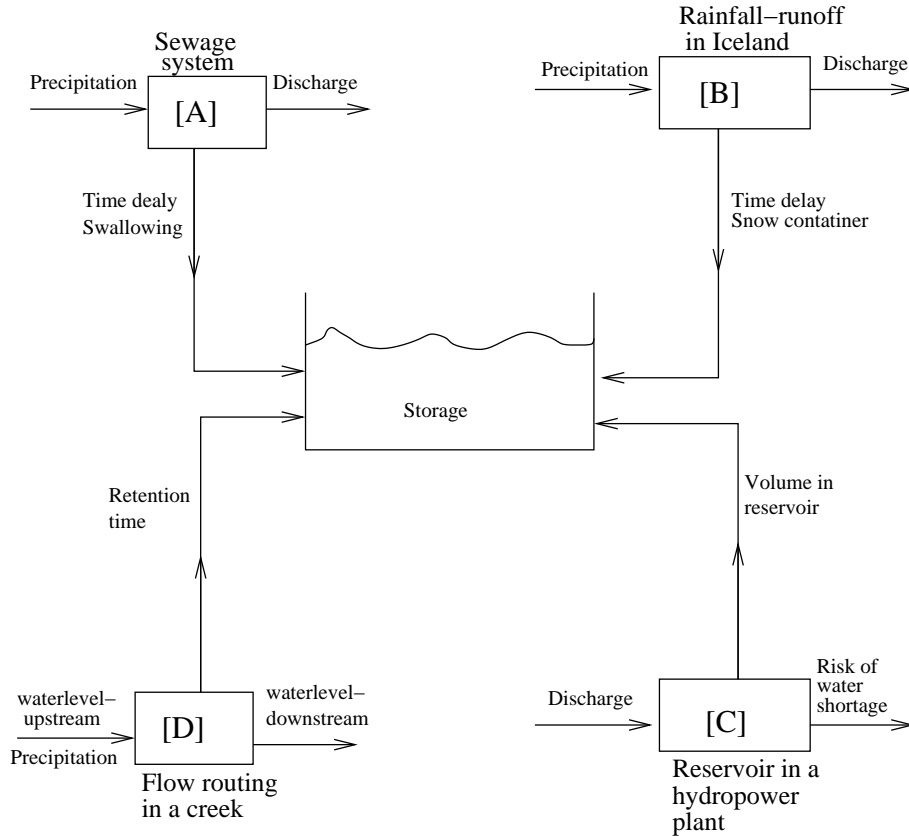


Figure 1.10: Overview of the storage effect.

probabilities versus volume of water shortage. This subject certainly involves storage, and in fact this might be the most obvious form of storage, a storage of water in a reservoir, measured in giga-liters, long-term storage of water from year to year.

Finally the subject in Paper [D] is a flow routing in a creek, the input is both a precipitation and the upstream depth. The water level at the downstream station is modelled as a function of the water level at the upstream station and precipitation. This is a small creek with short distance between the stations and no sub-creek merging in between. Consequently, the upstream water level has the largest impact on the downstream water level. Hence, the storage is mostly the retention time between the two measuring stations.

1.3 Why stochastic modelling?

In an ideal world, where all phenomena have been bought in the supermarket of physics, all occurrence can be described completely by physical equations. However, this is not the reality in our world, and particularly not in the field of hydrology. Consequently, use of stochastic models can be a useful option.

Models described by deterministic physical equation are often referred to as white box models. The stochastic models can be grouped into grey box models and black box models. The grey box models are described by physical equations and a noise factor. The noise factor is an extra term which is due to factors that are not described by the physical factors. The black box models are built up in such a way that statistical methods are used to find relation between input and output not necessarily based on physical processes.

The basic physical equation in hydrology is the equation of conservation of mass Eq. (2.1). The change of mass within a volume equals net outflow of the volume, i.e., the difference of the mass of inflow into the volume minus the mass of outflow out of the volume. In hydrology the control volume unit is a watershed. The total volume is found by integrating over the entire watershed, and the change of mass is found by the time derivative of the water inside the volume (watershed). The net outflow is found by inflow and outflow through the watershed's boundary. To carry out these calculations detailed information about precipitation, evaporation, transpiration, infiltration, surface runoff and groundwater runoff must be known. Information must be available in the whole watershed. In general such information does not exist and it is, therefore necessary to introduce stochastic terms in the models. Moreover, many of these processes are highly non-linear and cannot be described perfectly with mathematical equations, e.g., the infiltration (Viessman & Lewis 1996). Presently not enough is known about the processes to describe the perfectly. These model uncertainties reflect the inability to represent the physical process by use of deterministic equations and thus provide evidence for the stochastic modelling approach.

In addition it can be argued that both geophysical factors like the soil and many of the meteorological factors, indeed, have a stochastic behaviour. Moreover, use of a stochastic model can provide information about uncertainties in prediction (extrapolation) of the future.

Last but not the least, it is well known that the hydrological data are corrupted by errors due to measurement errors, both in the input data and output data. Additionally, errors exist due to the transformation from the measured values to the values requested e.g., transformation from water level measurements to

discharge.

1.4 Conclusion and discussion

The topic of this PhD project is stochastic modelling in hydrology and in all of the papers, parameters are estimated by using data. However, the statistical methods are of different types. One paper presents conditional parametric modelling, which is a black box type of model. Two papers present a parameter estimation in models described by stochastic differential equations, which are semi-physical models, or grey box models. Finally, one paper presents results which can be achieved by stochastic simulations only.

The topic of the rainfall-runoff relationship has been a study of interest for centuries. Numerous models of the rainfall-runoff relationship exist and it depends on the circumstances what kind of model is good to use, or possible to use. Sometimes detailed information about the watershed is available while in other cases the information is scarce. Models like SHE (Abbott et al. 1986*a*), (Abbott et al. 1986*b*), MIKE-SHE (Refsgaard & Storm 1995) and WATFLOOD (Singh & Woolhiser 2002) are physically based, distributed models. These types of models are often referred to as white box models. In a physically based model the hydrological processes of water movement are modelled either by finite difference representation of the partial differential equations of mass, momentum and energy conservation, and/or by empirical equations derived from independent experimental research. Spatial distribution of catchment parameters such as rainfall input and hydrological response is achieved in a grid network. All the physical processes are captured in the model, such as; interception, evapotranspiration, etc. However, as stated in (Refsgaard & Storm 1995), the application of a distributed, physically based model like MIKE SHE requires the provision of large amounts of parametric and input data. Moreover, the ideal situation where field measurements are available for all parameters rarely occurs. Hence, the problem of model calibration (parameter estimation) arises (Refsgaard et al. 1992) and also a decision of optimization criteria (Madsen 2000).

Contrarily, black box models have also been used in rainfall-runoff modelling. Black box models are completely data based, i.e., the model structure is determined by statistical methods and the data is used to estimate the parameters of the model.

In the 1970'ies linear black box models such as FIR and ARMAX models were quite popular, and in some cases they provide acceptable results. Nevertheless the rainfall-runoff process is believed to be highly non-linear, time-varying and

spatially distributed, e.g., (Singh 1964). With increased computer power non-linear models have become more popular. (Todini 1978) presented a threshold ARMAX model in a state space form. (Young 2002) model time variations by introducing the SDP approach, (State Dependent Parameter approach), which in the case of a non-linearity results in a two stage DBM approach. In recent years, various types of non-linear models have been developed such as neural networks e.g., (Shamseldin 1997) or (Hsu et al. 2002), Bayesian methods like (Campbell et al. 1999), fuzzy methods, e.g., (Chang et al. 2005) and non-parametric models e.g., (Iorgulescu & Beven 2004). In Paper [A] and Paper [B] several models are mentioned and (Singh & Woolhiser 2002) provides an overview of mathematical modelling of watershed hydrology.

In Paper [A] the method of conditional parametric models is introduced in hydrological modelling. A conditional parametric model is a semi parametric model, a mixture of a non-parametric, (Härdle 1990) and a parametric black box model. The name of the model originates from the fact that if the arguments of the conditional variables are fixed, then the model is an ordinary linear model, (Hastie & Tibshirani 1993), and (Anderson et al. 1994). In Paper [A] the basic modelling formulation are FIR and ARX models, except that the models parameters are non-parametrically described as a function of external variables. In the actual case, the parameters depend on the season and on the volume of water in the sewage system. The conditional variation is estimated by use of local polynomials as described in (Nielsen et al. 1997). The estimation is accomplished by using a software package LFLM (Locally weighted Fitting of Linear Models), which is an S-PLUS library package, see (Nielsen 1997). This approach turns out to provide improvements compared to linear modelling. By studying how the parameters vary as the conditional variables changes. This approach can also be used in a search for a more global modelling or structure identification. Hence, the approach is also valuable as a tool for an analysis, that might provide understanding of the system studied, usable in a grey box model interpretation.

In Paper [B] the modelling principle of white box modelling and black box modelling is combined in the grey box modelling approach. The principle is to develop a simple model, but still physically based in some sense, so that the parameters have at least a semi-physical or average-physical interpretation. However, the model is kept simple enough so that the available data can be used for parameter estimation. The model is formulated in a continuous-discrete time state space form. The system equations consist of stochastic differential equations. Hence the estimated parameters can be directly physically interpreted. The parameter estimation is a Maximum likelihood method, based on Kalman filter technique, for evaluating the likelihood function. This is implemented in a software package called CTSM (Continuous Time Stochastic Modelling), (Kristensen et al. 2003). One advantage of the stochastic state space approach

is that the model structure can be used for both prediction and simulation. The parameterization can be controlled in the software CTSM depending on, whether the model is to be used for prediction or simulation. In order to be able to estimate all the parameters, the model is kept more simple in structure than many of the existing conceptual models, such as the HBV model, (Bergström 1975) and (Bergström 1995) or the Tank model (Sugawara 1995). Some attempts have been made for parameter estimation in a state space models of similar type as the model in Paper [B]. (Lee & V.P.Singh 1999) applied an on-line estimation to the Tank model, but only for one storm at a time, calibrating the initial states manually. In (Georgakakos et al. 1988) the Sacramento model, org. in (Burnash et al. 1973), is modified and formulated in a state space form, but due to the model's complexity only some of the parameters are estimated, which, indeed, might result in locally optimal parameter values. The structure of the model presented in Paper[B] is simpler than in the two models mentioned above. In (Beven et al. 1995) it is stated that a number of studies have suggested that there is only enough information in a set of rainfall-runoff observations to calibrate 4 or 5 parameters, which is about the number of physical parameters estimated in Paper[B]. By using a smooth threshold function for separating the precipitation into snow and rain instead of elevation division keeps down the number of parameters. Physically this can be interpreted as some kind of averaging. The only data required for estimating the parameters of the model is two input series; precipitation and temperature and one output series, the discharge. In the light of limited data compared to the size and altitude range of the watershed. It is, in indeed, very satisfactory how well the model performs in the case study. The watershed is 1132 km², with an altitude range of about 1000 m, and 50% of the watershed is located above 800 m. The input series; temperature and precipitation are measured down in the valley, close to the river mouth. The number of physical parameter estimated is 8, additionally the initial states and the states variances are estimated. The calibration period is 6 years, while the validation period is 2 years (not used in calibration). This modelling approach provides a promising tool for further modelling in hydrology. Furthermore, (Kristensen et al. 2004a) showed that the stochastic state space model formulation gives significantly less biased parameter estimate than parameter estimates obtained by the optimization method based on deterministic model formulation. See also Section 4.7.

The topic in Paper [C] is a risk assessment of electrical power shortage in a hydropower plant. This is the same as risk assessment of a water shortage in the corresponding water resource system. A water shortage is met by flow augmentation from reservoirs. The management of these reservoirs are human interventions in the natural flow. One of the major questions in a simulation analysis of the Icelandic power system is the performance of the reservoirs as the electrical power system is hydropower based . During a heavy drought, the available water storage in the reservoir may not be sufficient to fulfill the demand

and, consequently, there will be a shortage of electrical power. It is therefore very important to have mathematical tools to estimate the risk of water shortage, when searching for management methods. The method of using all available flow series in order to design a reservoir, large enough to sustain a predefined flow output is well known in hydraulic engineering. The graphical version of the method can be seen in (Crawford & Linsley 1964) and this principle is still widely used. However, the method cannot predict the risk of water shortage. Stochastic methods in hydraulic design have been known for quite some time, e.g., (Plate 1992), but they are not yet extensively used in risk assessment. All the available data are used for design of the hydropower plant. Thus it is clear that the recurrence time of a drought in the reservoir is large. Consequently, the subject is to estimate small probabilities, probabilities which are in the tail of the corresponding distribution. A stochastic formulation of water shortage is a peak below threshold study, see (Medova & Kyriacou 2000). The case study is the river Tungnaá in southern Iceland. The data series consist of daily flow values over a period of 50 years. The mean value of the flow is $80.7 \text{ m}^3/\text{s}$. As an example, the results in the case study showed that the probability of a water shortage of 155 million m^3 is 0.5% and thus the recurrence time is 200 years. A water shortage of this magnitude means that the power station is out of order for about 3 weeks. If the economical lifetime of the hydropower station is 50 years, the probability that a large drought like that will occur is 25%. It is demonstrated that the only way to obtain a discharge series long enough for calculating a stable estimate of the drought risk is to produce a series by stochastic simulation.

The topic in Paper [D] is flow routing. In a broad sense the flow routing may be considered as an analysis to trace the flow through a hydrologic system, given the input. Numerous routing techniques exist, e.g., (Chow et al. 1988) and (Viessman & Lewis 1996). In Paper [D] a lumped stochastic model is developed to describe the downstream water level as a function of the upstream water level and precipitation. The Saint-Venant equation, e.g., (Chow et al. 1988), is used for deriving a stochastic linear reservoir model, represented as a state space model in continuous time by using stochastic differential equations. The parameters are estimated by using the program CTSM (Kristensen et al. 2003). The principle of linear reservoir model was proposed by (Nash 1957) and the concept was first introduced by (Zoch 1934, 1936, 1937) in an analysis of the rainfall-runoff relationship. The fact that the model in Paper [D] is stochastic allows for data to be used for parameter estimation including the parameters related to the system and observation errors. Furthermore, the model differs from the traditional reservoir model since the non-measured lateral inflow of water between the two measuring stations is a state variable in the model and estimated by use of the Kalman filtering technique. Using this in an environmental context means that it might be possible to estimate concentration of chemical concentrations in the lateral inflow if the corresponding chemical concentrations

are measured both upstream and downstream. This can be useful in an environmental analysis. It is found that the grey box modelling approach provides a strong modelling framework in flow routing. The possibility to combine the physical knowledge with data information valuable. It enables an estimation of non-measured variables and the stochastic approach makes it possible to provide uncertainty bounds on predictions and on parameter estimates.

In general it has been concluded that stochastic modelling in hydrology has the advantages of describing both non-linearities and non-stationaries. Furthermore, the grey box modelling approach provides a strong modelling approach which opens up to possibility for combine prior physical knowledge with data information. Hence, it bridges the modelling gap between the statistician and the physical expert.

CHAPTER 2

Introduction to hydrology

Water is the most vital substance on the Earth, the principal ingredient of all living things and a major force constantly shaping the surface of the earth. The first images of the surface of the Earth, as seen from the moon over two decades ago helped visualizing the Earth as a unit, an integrated set of systems; land masses, atmosphere, oceans, and the plant and animal kingdom.

This chapter provides a brief introduction to the concepts in hydrology used in this project. The chapter is mostly based on the books (Mays 1996), (Chow 1964) (Chow et al. 1988), (Burnash 1995), (Viessman & Lewis 1996), (McCuen 1989) and (Singh & Woolhiser 2002).

2.1 The history of water resources

The book (Mays 1996) gives an excellent overview of the history of water resources and human interaction with water up to 18th century. The following paragraphs are mostly based on this book.

Water is the key factor in the progress of civilization and the history of water resources cannot be studied without studying humanity. Humans have spent

most of their history as hunting and food gathering beings. It is only during the last 9000 to 10000 years that human beings have discovered how to raise crops and tame animals. From Iraq and Syria the agricultural evolution spread to the Nile and Indus valleys. During this agricultural evolution, permanent villages took the place of a wandering existence. About 6.000 to 7.000 years ago, farming villages of the Near and Middle East became cities. Farmers learned to raise more food than they needed, allowing others to spend time making things useful to their civilization. People began to invent and develop technologies, including how to transport and manage water for irrigation.

The first successful efforts to control the flow of water were made in Egypt and Mesopotamia. In ancient Egypt the construction of canals was a major endeavor of the Pharaohs. One of the first duties of provincial governors was the digging and repair of canals, which were used to flood large tracts of land while the Nile was flowing high. Problems of the uncertainty of the Nile flows were recognized. During very high flows the dikes were washed away and the villages were flooded, drowning thousands. During low flow the land did not receive water and no crops could grow. The building of canals continued in Egypt throughout the centuries.

The Sumerians in southern Mesopotamia built city walls and temples and dug canals that were the world's first engineering work. Flooding problems were more serious in Mesopotamia than in Egypt because the Tigris and Euphrates carried several times more silt per unit volume of water than the Nile. This resulted in rivers rising faster and changing their courses more often.

The Assyrians developed extensive public works. Sargon II invaded Armenia in -714, discovering the ganat (Arabic name). This is a tunnel used to bring water from an underground source in the hills down to the foothills. This method of irrigation spread over the Near East into North Africa over the centuries and is still used.

The Greeks were the first to show the connection between engineering and science, although they borrowed ideas from the Egyptians, the Babylonians and Phoenicians. Ktesibius (-285 - -247), invented several things e.g., the force pump, the hydraulic pipe, the water clock. Shortly after Ktesibius, Philon of Byzantium invented several things, one of which was the water wheel. One application of the water wheel was a bucket-chain water hoist, powered by an undershot water wheel. This water hoist may have been the first recorded case of using the energy of running water for practical use. Probably the greatest Hellenistic engineer, was Archimedes (-287 to -212). He founded the ideas of hydrostatics and buoyancy. The Hellenistic kings began to build public bath houses.

The early Romans devoted much of their time to useful public projects. They built roads, harbor works, aqueducts, baths, sewers etc. The Romans and Helenes needed extensive aqueduct systems for their fountains, baths and gardens. They also realized that water transported from springs was better for their health than river water. Knowledge of pipe making was in its infancy and the difficulty of making good large pipes was a hindrance. Most Roman piping was made of lead, and even the Romans recognized that water transported by lead pipes is a health hazard.

The fall of the Roman Empire in 476 extended over a 1000 year transition period called the Dark Ages. After the fall of the Roman Empire, water and sanitation in Europe declined, resulting in worse public health.

During the Renaissance, a gradual change occurred for purely philosophical concepts toward observational science. Leonardo da Vinci (1452-1519) made the first systematic studies of velocity distribution in streams. The French scientist Bernard Palissy (1510-1589) showed that rivers and springs originate from rainfall, thus refuting an age old theory that streams were supplied directly by the sea. The French naturalist Pierre Perrault (1608-1680) measured runoff, and found it to be only a fraction of rainfall. Blaise Pascal (1623-1662) clarified principles of the barometer, hydraulic press, and pressure transmissibility. Isaac Newton (1642-1727) explored various aspects of fluid resistance (inertia, viscosity and waves).

Hydraulic measurements and experiments flourished during the eighteenth century. New hydraulic principles were discovered, such as the Bernoulli (1700-1782) equation for forces present in a moving fluid and Chezy's (1718-1798) formula for the velocity in an open channel flow, also better instruments were developed. Leonard Euler (1707-1783) first explained the role of pressure in fluid flow and formulated the basic equation of motion.

Concepts of hydrology advanced during the nineteenth century. Dalton (1802) established a principle for evaporation, Darcy (1856) developed the law of porous media flow and Manning (1891) proposed an open channel flow formula. Hydraulics research continued in the nineteenth century, with Louis Marie Henry Navier (1785-1836) extending the equations of motion to include molecular forces. Jean-Claude Barre de Saint-Venant wrote in many fields on hydraulics. Others, such as Poiseuille, Weisbach, Froude, Stokes, Kirchoff, Kelvin, Reynolds and Boussinesq, advanced the knowledge of fluid flow and hydraulics during the nineteenth century.

At the beginning of the twentieth century quantitative hydrology was basically the application of empirical approaches to solve practical hydrological problems. Gradually, hydrologists did combine empirical methods with rational analysis

of observed data. One of the earliest attempts to develop a theory of infiltration was by Green and Ampt in 1911, who developed a physically based model for infiltration and in 1914 Hazen introduced frequency analysis of flood peaks. Sherman defined the unit hydrograph in 1932, as the unit impulse response function of a linear hydrologic system i.e., a function relating excess rainfall to direct runoff. In 1933 Horton developed a theory of infiltration to estimate rainfall excess and improved hydrograph separation techniques. In 1945 Horton developed a set of "laws" that are indicators of the geomorphologic characteristics of watersheds, now known as Horton's laws. In the years 1934 to 1944 Lowdermilk, Hursh and Brater, observed that subsurface water movement constituted one component of storm flow hydrographs in humid regions. Subsequently, Hoover and Hursh reported significant storm flow generation caused by a dynamic form of subsurface flow. The underground phase of the hydrologic cycle was investigated by Fair and Hatch in 1933, who derived a formula for computing the permeability of soil and in 1944 Jacob correlated groundwater levels and precipitation on the long Island, N.Y. The study of groundwater and infiltration led to the development of techniques for separation of base flow and interflow in a hydrograph. McCarthy and others developed the Muskingum method of flow routing in the 1934-1935 and the concept of linear reservoirs was first introduced by Zoch in the years 1934-1947, in an analysis of the rainfall and runoff relationship. In 1951 Kohler and Linsley developed the Antecedent index approach, which have been used in various models. The principle is that weighted summation of past daily precipitation amounts, is used as an index of soil moisture.

In 1960' the digital revolution broke out and since then numerous mathematical models have been developed. The models are of different types and developed for different purposes, although many of the models share structural similarities, because their underlying assumptions are the same. There exists models for simulation of watershed hydrology, for flood forecasting warning systems and for environmental managements. The type of models are different, there exists conceptual and models and detailed physically based distributed models. There exists numerous black box models and also grey box models, e.g., Paper [B]. The development of hydrological modelling will proceed on and on, mostly because of constantly improving modelling techniques. In Papers [A] and [B] several watershed models are mentioned and a fine overview of watershed models can be found in (Singh & Woolhiser 2002).

2.2 The hydrological cycle

Water on earth exists in a space called the hydrosphere which extends about 15 km up into the atmosphere and about 1 km down into the lithosphere, the crust of the earth. The cycle has no beginning or end. Figure 2.1 illustrates the hydrological cycle; the water evaporates from oceans and land surface to become part of the atmosphere, water vapor is transported and lifted in the atmosphere until it condenses and precipitates on the land or the ocean. Precipitated water may be intercepted by vegetation, become overland flow infiltrates into the ground, flows through the soil as subsurface flow and discharge into streams as surface runoff. Much of the intercepted water and surface runoff returns to the atmosphere through evaporation. The infiltrated water may percolate deeper to recharge groundwater, later emerging in springs or seeping into streams to form surface runoff, and finally flowing into the sea or evaporating into the atmosphere as the hydrologic cycle continues.

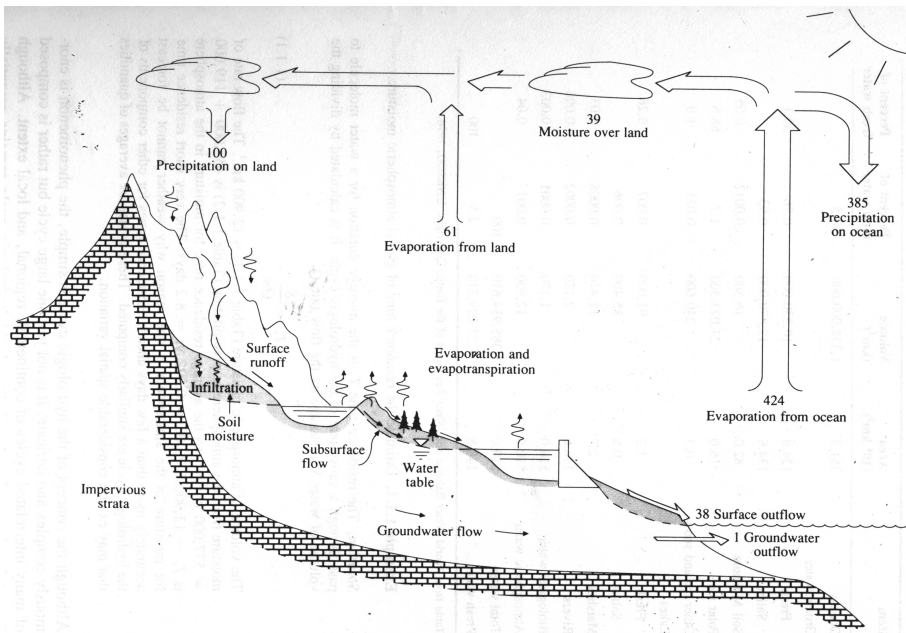


Figure 2.1: The hydrologic cycle, (Chow et al. 1988).

Although the concept of the hydrologic cycle is simple, the phenomenon is enormously complex and intricate. It is not just one large cycle but rather is composed of many interrelated cycles of continental, regional and local extents. The

hydrology of a region is determined by its weather patterns and by physical factors such as topography, geology and vegetation. Also as civilization progresses, human activities gradually encroach on the natural water environment, altering the dynamic equilibrium of the hydrologic cycle and initiating new processes and events.

2.3 The storage effect

By analogy, a hydrologic system is defined as a structure or volume in space, surrounded by a boundary, that accepts water and other inputs, operates on them and produces outputs. The basic physical law, is the law of conservation of mass, i.e., mass cannot be created or destroyed. Thus, for a fixed time independent region V , the net rate of flow of mass into the region is equal to the rate of increase of the mass within the surface. Figure 2.2 demonstrates this principle.

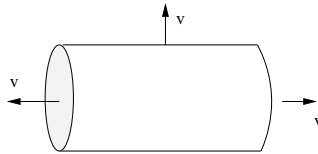


Figure 2.2: Flow in a control volume.

Mathematically this can be written:

$$\frac{d}{dt} \iiint_V \rho dV = - \iint_{\partial V} \rho \mathbf{v} dA. \quad (2.1)$$

where the d stands for the total derivative V stands for the volume, ρ is the density, \mathbf{v} is the velocity of the fluid, A is the area vector and ∂V is the surface of the control volume, i.e., the surface to be integrated over. The velocity of the flow is defined positive out of the surface. This is known as the integral equation of continuity for an unsteady, variable-density flow. If the flow has a constant density, the density ρ can be divided out of both terms of Eq. (2.1) leaving

$$\frac{d}{dt} \iiint_V dV = - \iint_{\partial V} \mathbf{v} dA. \quad (2.2)$$

The integral $\iiint_V dV$ is the volume of fluid stored in the control volume, denoted by S , thus the first term in Eq. (2.2) is the time rate of change of the storage

dS/dt . The second term is the net outflow and it can be split into inflow, $I(t)$, and outflow, $Q(t)$. The resulting equation is the storage equation

$$\frac{dS}{dt} = I(t) - Q(t) \quad (2.3)$$

which is the integral equation of continuity for an unsteady, constant density flow. When the flow is steady, $dS/dt = 0$. Figure 2.3 shows a schematic representation of the storage equation. Input enters the system, the system operates

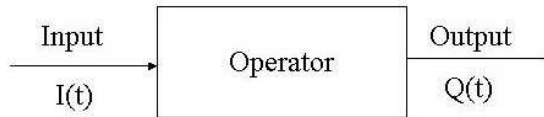


Figure 2.3: A system operation

on the input and delivers an output. In a systematic representation the operation function is often referred to as impulse response function, see Section 4.4.2.

The system considered in hydrologic analysis is the watershed. The watershed is defined as all the land area that sheds water to the outlet during a rainstorm i.e., all points enclosed within an area from which rain is falling at these points will contribute to the outlet. Big watersheds are made up of many smaller watersheds and thus it is necessary to define the watershed in term of a point. The shaded area of Figure 2.4 represents the watershed with outlet at point A whereas the watershed for point B is the small area enclosed within the dashed lines.

2.4 Surface water hydrology

Surface hydrology is the theory of movement of water along the surface or the Earth as a result of precipitation and snow melt. Runoff occurs when precipitation or snowmelt moves across the land surface. The land area over which rain falls is called the catchment while the land area that contributes surface runoff to any point of interest is called a watershed. The relationship between precipitation and runoff has been studied for decades. The hydrological subject in Papers [A], [B] and [D] is surface water hydrology.

A streamflow hydrograph is a graph (or table) showing the flow rate as a function of time at a given location in a stream. The spikes, caused by rain storms, are

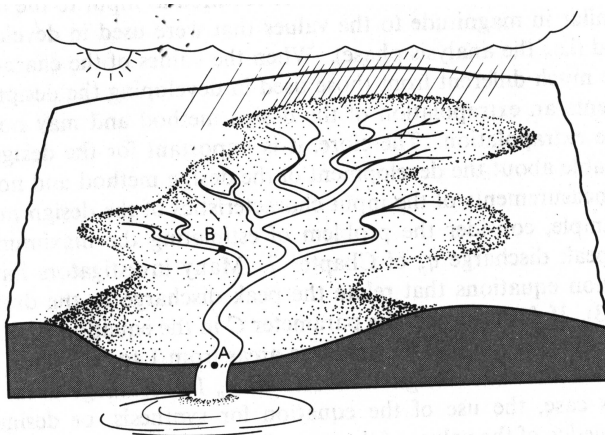


Figure 2.4: Delineation of a watershed boundary. (McCuen 1989).

called direct runoff or quickflow, while the slowly varying flow in rainless periods is called base flow.

Excess rainfall, or effective rainfall, is the rainfall which is neither retained on the land surface nor infiltrated into the soil. After flowing across the watershed, excess rainfall becomes direct runoff at the watershed outlet. The graph of excess rainfall vs. time is a key component in the study of rainfall runoff relationships. Figure 2.5 shows a hydrograph and some of the hydrographs components. The time base of a hydrograph is considered to be the time from

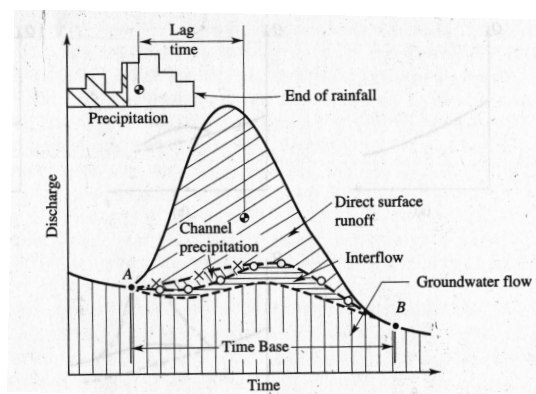


Figure 2.5: Storage flow relationship, (Viessman 1996).

which the concentration curve begins until the direct runoff component reaches

zero. Watershed lag time or basin lag time, is defined as the time from the center of mass of effective rainfall to the center of mass of direct runoff. If a uniform rain is applied on a tract, the portions nearest the outlet contribute to runoff at the outlet almost immediately. As rain continues, the depth of excess on the surface grows and discharge rates increase throughout. A time of concentration, t_c , is defined as the time required, with uniform rain, so that 100 percent of watershed (all portions of the drainage basin) is able to contribute to the direct runoff at the outlet. (Singh 1988) argues that the time of concentration t_c is 1.42 times the basin lag time. This fact is used in Paper [A].

2.4.1 The theory of unit hydrograph

The unit hydrograph is the unit pulse response function of a linear hydrologic system. First proposed by (Sherman 1932), the unit hydrograph of a watershed is defined as a direct runoff hydrograph resulting from 1 cm of excess rainfall generated uniformly of the drainage area at a constant rate for an effective duration.

A unit hydrograph is basically an impulse response function between excess rainfall to direct runoff. It fulfills the equation of continuity and thus the mass balance is conserved.

In practice the total rainfall and the total runoff (the discharge) are measured¹. Thus, for hydrograph calculations effective rainfall must be calculated and the runoff must be divided into baseflow and direct runoff. These calculations may be interpreted as data processing and not as an integral part of hydrograph theory. Several methods for base flow separation are used, such as, normal depletion curve, the straight line method, the fixed base method or the variable slope method, to mention some, for a description of these methods see e.g., (Chow et al. 1988). Likewise, for calculations of effective rainfall and total runoff many methods are used. These are methods for calculations of evaporation and infiltration. Methods for evaporation calculations might be the energy balance method, aerodynamic method, the combination method and e.g., Priestly-Taylor's method. Methods for infiltration calculations might be the Horton's equation, the Philip's equation, Green-Ampt's method (Chow et al. 1988), Huggin-Monke model and Holtan model (Viessman & Lewis 1996) and calculations of infiltration. For all the above reasons it is clear that the identification of a hydrograph very much depends on calculations earlier, it cannot be identified from the measured precipitation. The topics in Paper [A] and

¹In general only the water level is measured and not the discharge. The discharge is not measured in general, but the water level. The discharge is calculated from the water level data by use of rating curves, see Section 3.2

[B] are related to the theory of unit hydrograph as the flow is modelled as a function of precipitation. However, the input variable is total precipitation as it has proven advantage to develop models which do not rely on earlier calculations. Particularly, since a perfectly quantified general formula for separating the effective precipitation from the total precipitation does not exist (Viessman & Lewis 1996). The linear modelling approach is only an approximation of the relationship between effective rainfall and direct runoff.

2.4.2 Flow routing

In a broad sense the flow routing may be considered as an analysis to trace the flow through a hydrologic system. Thus given the input, i.e., a hydrograph at an upstream location, the flow routing is a procedure to determine the time and magnitude of the flow at a given point downstream. The flow routing techniques are divided into lumped flow routing and distributed flow routing. When distributed flow routing methods are used, the flow is calculated as a function of space and time through the system whereas if lumped routing methods are used the flow is calculated as function of time alone at a particular location. Routing by use of lumped methods is sometimes referred to as hydrologic routing, and routing by distributed methods is sometimes referred to as hydraulic routing. The topic in Paper [D] is flow routing.

2.4.2.1 Lumped flow routing

Lumped flow routing techniques are all founded on the equation of continuity represented in the operational form as Eq. (2.3) In general, the storage function may be written as an arbitrary function of $I(t)$, $Q(t)$ and their time derivatives

$$S = f \left(I, \frac{dI}{dt}, \frac{d^2I}{dt^2}, \dots, Q, \frac{dQ}{dt}, \frac{d^2Q}{dt^2} \right) \quad (2.4)$$

Sometimes it is possible to describe the storage as a function of only Q as, single-valued storage function $S = f(Q)$ as shown in Figure 2.6. For such reservoirs (control volumes), the peak outflow, occurs when the outflow hydrograph intersects the inflow hydrograph, because the maximum storage occurs when $dS/dt = I - Q = 0$, and the storage and outflow are related by $S = f(Q)$. This is shown in Figure 2.6, the point denoting the maximum storage, R and the point denoting the maximum outflow P , coincide. Hence, when the flow is steady, a single-value storage function always exists. When the reservoir is long and narrow like open channels or streams the storage-outflow relationship

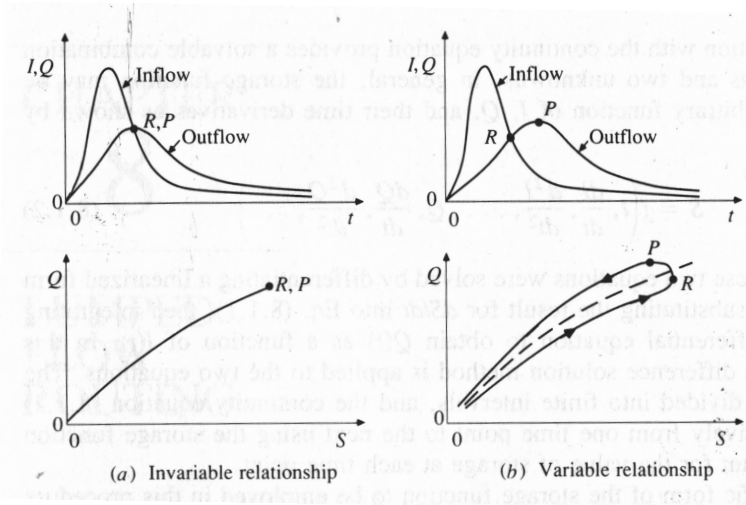


Figure 2.6: Storage flow relationship. (Chow et al. 1988).

is variable. The amount of storage due to backwater depends on the time rate of change of flow through the system. Figure 2.6 shows the relationship between the discharge and the system storage in a variable storage-outflow relationship. The relationship is no longer a single-valued function but exhibits a curve. Because of retarding effect due to backwater, the peak outflow usually occurs later than the time when the inflow and outflow hydrographs intersect, as indicated in Figure 2.6, i.e., the points R and P do not coincide.

The variable storage outflow relationship is one of the factors, which contributes to error in the $Q - h$ rating curve which is used to calculate the discharge from water level data, see Section 3.2.

Several lumped flow routing exists. The level pool method is an analytical method and can be used in situations where the storage function is single valued. The Muskingum method is a well known method in a variable discharge-storage relationship. Several other methods for variable discharge-storage relationship are commonly used, such as the SCS convex method, Muskingum-Chunge method and multiple storage method. For details see (Chow et al. 1988) and (Viessman & Lewis 1996).

Finally, the linear reservoir routing model will be described as the principles of the linear reservoir model is used in the flow routing model in Paper [D]. Furthermore, the modelling principle in the rainfall runoff model in Paper [B] is a reservoir model even though the model in Paper [B] is a non-linear threshold

model.

A linear reservoir is a reservoir where the storage is linearly related to its output by a storage constant k , which has the dimension time,

$$S = kQ \quad (2.5)$$

(Nash 1957) stated that a watershed may be represented by a series of n linear identical linear reservoirs, each having the same storage constant k . Figure 2.7 demonstrates this principle.

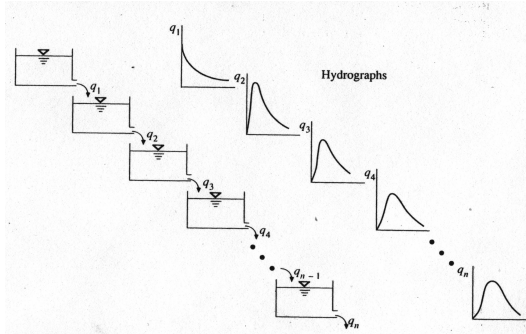


Figure 2.7: Linear reservoirs in series. Chow et al. (1988).

A transfer function (in continuous time) of the n linear reservoir model with storage constant k is

$$Q(t) = \underbrace{\frac{1}{(1+kD), \dots, (1+kD)}}_{n \text{ times}} I(t) \quad (2.6)$$

where D is the differential operator. A deterministic state space representation of this system is:

$$d \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{pmatrix} = \begin{pmatrix} -1/k & & & \\ 1/k & -1/k & & \\ & 1/k & -1/k & \\ & & 1/k & -1/k \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_n \end{pmatrix} dt + \begin{pmatrix} 1/k \\ 0 \\ \vdots \\ 0 \end{pmatrix} I dt \quad (2.7)$$

The impulse response function in this system, representing the outflow from the n -th reservoir is

$$h(t) = Q_n(t) = \frac{1}{k\Gamma(n)} \left(\frac{t}{k}\right)^{n-1} e^{-t/k} \quad (2.8)$$

which is the gamma distribution function.

2.4.2.2 Distributed flow routing

The flow of water through the soil and stream channels of a watershed is a distributed process because the flow rate, velocity, and depth vary in space throughout the watershed. This is described in a three dimensional space and time by the continuity equation Eq. (2.1)

For a description of flow routing in a channel the spatial variation in velocity across the channel can often be ignored. Hence, the flow process can be approximated as varying in one space dimension along the flow channel. The Saint-Venant equations, first developed by Barre de Saint-Venant in 1871, describe one-dimensional unsteady open channel flow. The Saint-Venant equation for mass balance is:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} - q = 0 \quad (2.9)$$

Where Q is the flow, x is the geometrical variable along the channel. Hence, $\partial Q/\partial x$ is the rate of change of the flow along the channel. The average cross-sectional is denoted A and q is the lateral inflow between the upstream location and downstream location.

The Saint-Venant equation for mass balance can be derived from the continuity equation, if following assumptions are true:

- The flow is one-dimensional, depth and velocity vary only in the longitudinal direction of the channel. This implies that the velocity is constant and the water surface is horizontal across any section perpendicular to the longitudinal axis.
- The flow is assumed to vary gradually along the channel so that hydrostatic pressure prevails and vertical accelerations can be neglected.
- The longitudinal axis of the channel is approximated as a straight line.
- The bottom slope of the channel is small and the channel bed is fixed, that is the scour effects and deposition are negligible.
- Resistance coefficients for steady uniform turbulent flow are applicable so that relations such as Manning's equation can be used to describe resistance effects
- the Fluid is incompressible and of constant density throughout the flow.

2.5 Subsurface water and groundwater

Surface water infiltrates into the soil and becomes soil moisture, subsurface flow (unsaturated flow) through the soil and groundwater flow (saturated flow) through soil or rock strata.

Soil properties control the rate at which water infiltrates into the soil, percolates through the subsurface and travels through the soil to surface water bodies. This rate affects the proportions of rainfall that appears as surface runoff and groundwater losses. Subsurface and groundwater outflow occur when subsurface water emerges to become surface flow in a stream or spring.

Infiltration is the process of water penetrating from the ground surface into the soil. When precipitation reaches the ground it hits the intercepting surfaces, such as trees, plants, grass and other structures. The water in excess of interception capacity then begins to fill surface depressions, and a film of water is also built up above the ground surface. Then some proportion of the water infiltrates downward through the surface of the earth and becomes soil moisture, recharges the aquifers that again support base flow during dry periods. The rate at which the infiltration occurs depends on the infiltration capacity of the soil, which i.e., to a large extent depends on the wetness of the soil. As a result of great spatial variation and the time variations in soil properties occurring as the soil moisture content changes, infiltration is a very complex process and mathematical equations can only describe it approximately (Chow et al. 1988), and, thus perfectly quantified general relation does not exist, (Viessman & Lewis 1996).

Hydrological data

Hydrological data are the basis for hydrological analysis. The essential data are precipitation, stream flow, evaporation and transpiration. Evaporation and transpiration data have not been available in the projects of this PhD program and therefore they will not be described here.

3.1 Precipitation

Precipitation data is the main factor in hydrological models. The precipitation data determine the total amount of water input into the model, and good precipitation data are very important for the quality of the model simulations (Sælthun & Killingtveit 1995).

Many types of rain gauges exist. The data used in this PhD project are both from Iceland and Denmark. Two countries use different rain gauges.

The meteorological institute in Iceland uses rain gauge with a Nipher wind shield. The gauges are located at 1.5 m height above ground, and even higher at some few locations because of the snow. Figure 3.1 shows a sketch of a gauge with a Nipher wind shield.

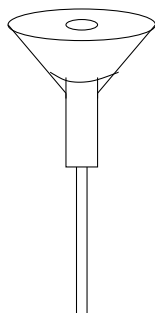


Figure 3.1: A rain gauge with a Nipher wind shield.

Rain drops into the gauge in a bottle. The bottle is emptied by manual labour twice within 24 hours.

In the projects where data originated from Denmark, the rain gauges used were tipping buckets. Figure 3.2 shows a sketch of a tipping bucket. The water drops

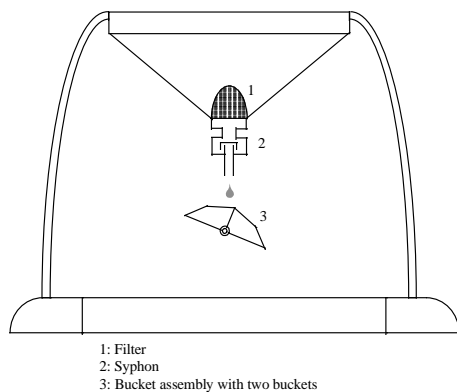


Figure 3.2: A tipping bucket rain gauge.

through a filter down to a syphon and then into a bucket with two boxes. The water enters only one box at a time. When this box has collected 0.2 mm of rain, it tips over, and the water starts to enter the second box. When the second box has collected 0.2 mm of rain, the bucket tips over again. Consequently, the precipitation data have values which are a multiplication of 0.2 mm.

3.1.1 The error

It is a well known fact that the amount of precipitation measured is an underestimate of the "ground true" precipitation. Experiments have been made in order to develop models to correct the underestimate. However such models depend on many factors, mainly the weather condition and the type of rain gauge used.

The main precipitation measurement error in Iceland is caused by aerodynamical effects near the rim of the gauges. No scientific experiments have been performed in Iceland in order to achieve a model to correct the underestimate. In 1987 a complete experimental field in Jokioinen in Finland was put up in order to develop models for operational correction of nordic precipitation data, see (Førland et al. 1996). Several models were developed. All of them included wind speed as a factor in the correction model. The Swedish, Danish and Norwegian countries included their national gauges in the experimental field. Unfortunately the accuracy of the models can be questioned since the average wind speed at Jokioinen is about 4 m/s whereas in the other Nordic countries the average wind speed is much higher. Hence, the models need to be extrapolated extensively. Furthermore, for practical purposes, the wind speed data must be available which often is not the case, at least not in Iceland.

The meteorologist Flosi Hrafn Sigurdsson states that measurements, not disturbed by wind, can be achieved by using a precipitation gauge, located in a hole in the ground with the gauges opening located at ground level. This type of gauge can only be used for measuring rain. During a summer period, May - September the ground level gauge in Reykjavik measured 21.7% more rain than the nipher gauge at 1.5 m height. Based on the ground level measurements and some other data, Flosi assumes that an average correction caused by wind for rain is 28% in Reykjavik and 32% in Hveravellir¹ The difference is mainly caused by stronger wind in Hveravellir. Furthermore, Flosi assumes that average wind correction for snow is 80% in Reykjavik and 100% in Hveravellir. More about Sigurdsson experiment and results can be seen in (Sigbjarnarson 1990).

3.2 Discharge

The discharge is a flow of water, having the unit $[m^3/s]$. Considering a cross sectional area across a river the discharge is the velocity, integrated over the

¹Hveravellir is located between Hofsjökull and Langjökull in middle of Iceland at a 646 m height.

cross sectional area. In general, the discharge is not measured on-line but the water level is measured.

The most commonly used methods for measuring water level are the floating principle and pressure measurements. Figure 3.3 shows a sketch of the floating

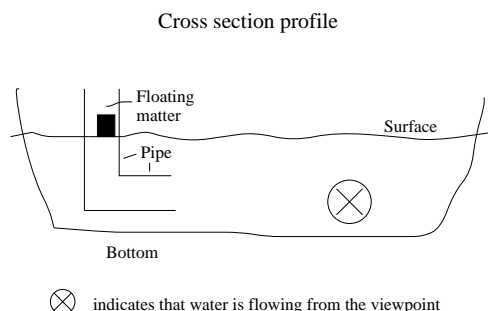


Figure 3.3: The floating principle

principle. An L shaped pipe (open in both ends) is installed in the river. One part of the pipe is in the river, parallel to the bottom, so that the stream is at the same level as the open area. The other part stands perpendicular to the river's surface. The water level inside the pipe is the same as outside the pipe. The water remains undisturbed in the pipe (no wind disturbance etc.) and the water level can be measured. This is done by using a floating device in the pipe.

By using the one to one relationship of pressure and depth, pressure measurements are also used to construct water level data. Figure 3.4 shows an outline

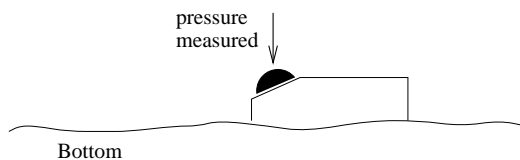


Figure 3.4: Sketch of a pressure transducer.

of a pressure transducer instrument. The transducer has a membrane and the pressure on the membrane is measured. (The measured pressure is corrected due to air pressure.) The pressure is sometimes measured by using bubbles. Then a tube is led into the river and small quantity of gas is put in the tube continuously. The pressure in the tube depends on the pressure at the tubes opening in the river.

The measured dept is used to calculate the discharge by using the fact that the flow equals the velocity intergrated over the cross sectional area. Conse-

quently, velocity measurements are required. Several methods and instruments for velocity measurements exist. The first mentioned method, and probably the oldest one, is to use a screw, as shown in Figure 3.5. The cross section is divided

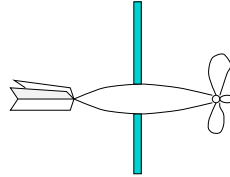


Figure 3.5: Outline of a screw velocity-measurement instrument

into sections as shown in Figure 3.6. The screw is used to measure the velocity

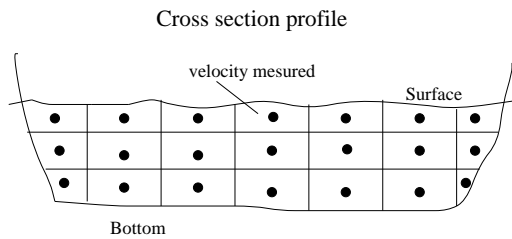


Figure 3.6: Principles of the velocity-area method.

in each section, and then an approximative velocity map can be drawn. This method is commonly used in Iceland

Another commonly used method is the use of magnetic flow meter. The operation of a magnetic flow meter is based upon Faraday's Law, which states that the voltage induced across any conductor as it moves at right angle through a magnetic field is proportional to the velocity of that conductor. Figure 3.7 shows a magnetic flow meter. This method is also quite commonly used in Iceland. However it is not convenient when the velocity is very large as the rocks in the bottom then move along the river.

The last mentioned method is the laser Doppler velocity meter, as shown in Figure 3.8. It sends a monochromatic laser beam toward the target and collects the reflected radiation. According to the Doppler effect the waves emitted from a source moving toward an observer are squeezed. Hence, the velocity of the object can be obtained by measuring the change in wavelength of the reflected laser light.

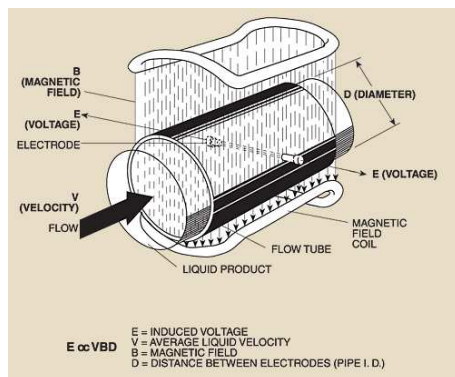


Figure 3.7: Magnetic flow meter

(Figure is from <http://www.omega.com/prodinfo/magmeter.html>).

3.2.1 The rating curve

Using velocity/flow measurements the relation between water level and discharge is found. This relation is known as the rating curve, or the $Q - h$ relationship (Q for discharge and h for depth/stage). The most commonly used formula is

$$Q = k(h - h_0)^N \quad (3.1)$$

where h_0 is the stage at which discharge is zero. Figure 3.9 shows water level and flow data.

3.2.2 The error

It is clear that the discharge data are far from being without a noise. The water level has to be measured and this process incorporates measurement errors as no instruments are perfect. Furthermore, the velocity has to be measured for corresponding measurement errors. Last but not the least, the $Q - h$ relation has to be found and this relationship is not perfectly described. Additionally the $Q - h$ relationship is dynamic since the cross section can change in time. Furthermore, in most real cases the $Q - h$ relationship is not unique. When Q varies with time it makes a loop, similar to the variable storage flow relationship $Q - S$, as shown in Figure 2.6. The hysteresis (i.e. loop-rating) in the $Q - h$ graph is created as for a fixed depth (h) the velocity is larger when the flow is increasing and smaller when flow is decreasing. Thus two points on each side of the top of a hydrograph will have different flow velocities and discharge (Q) even though the depth is the same. Usually the time variation of the flow is slow

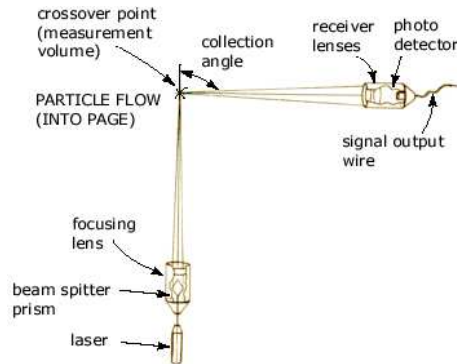


Figure 3.8: Laser Doppler velocity meter
 (Figure is from http://www.efunda.com/designstandards/sensors/laser_doppler/laser_doppler_effect_theory.cfm).

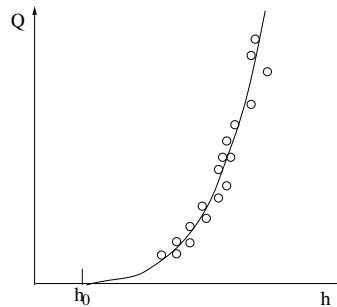


Figure 3.9: Water level data, flow data and a rating curve.

enough so this difference is negligible. But when that is not the case, enough data is usually not available to determine the complete loop.

3.3 River ice

In a colder climate icing in rivers disturbs water level measurements. The following section is based on; (Bengtson 1988), (Chow 1964), and (Rist 1962).

3.3.1 Construction of river ice

When the temperature of the surface water has dropped to the freezing point, net heat loss due to the atmosphere will cause ice production. Ice formation frequently starts with the development of many tiny disk formed crystals called *frazil* as shown in Figure 3.10, left. Frazil ice particles, frequently collected by



Figure 3.10: To left frazil ice and to right is ice pan.
(<http://www.clarkson.edu/~htshen>)

adhesion, form larger masses which move along with the current. As the ice content of the water increases, the water becomes oily or milky in appearance and the viscosity of the water increases.

Ice rapidly forms on the surface of most flowing rivers. At first this consists of agglomeration of broken surface crystals and frazil ice which unite to form round pans. Figure 3.10, right shows an ice pan. These pans grow by accretion and the open water between them becomes smaller. When the ice cover on the surface is complete, the river regime changes from open water flow to flow beneath the ice cover. However, the same flow is to be carried and as a consequence the water level is increased.

Ice can also form on underwater objects. This is called *anchor ice*, shown in Figure 3.11. The coating of anchor ice may be several inches thick and may then grow more rapidly on sharp corners. This anchor ice may eventually dam up the stream. Thus, it is possible to develop a staircase of a series of small ice dams with some still water trapped behind them. The increased viscosity due to the ice content, the damping of turbulent eddies, and the rise of the river bed due to the formation of anchor ice also cause an increase in the river stage.

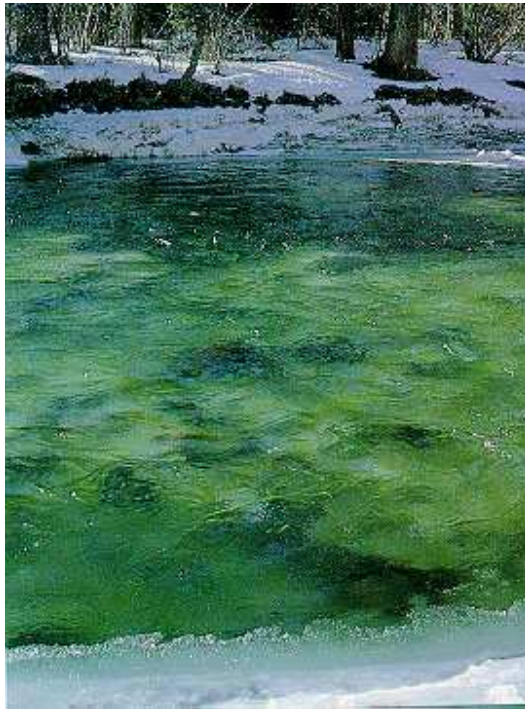


Figure 3.11: Anchor ice. (<http://www.clarkson.edu/htshen>)

3.3.2 Melting of river ice

The melting of ice in an icecovered river causes the river stage to rise in the spring. At one point the ice cover begins to break and to move with the stream. This initiates a chain reaction so that a complete ice cover can be removed in a matter of hours.

If the water rises fast the ice cover can be fragmented and forced to move while it still has almost its full strength. At particular locations depending on river morphology ice jams are formed. Very high water levels are caused by ice jams present during break-ups.

3.3.3 Ice reduction

The term ice reduction is used for the complete process of correcting the observed water levels so that these reach the levels that would have been observed if there were no ice present in the entire river system. The difference between the water level with ice present and the water level during summer conditions at the same discharge, is referred to as backwater from ice formation.

In countries with a predominantly continental climate, the river state can be categorized into following phases:

1. Freezing-over period
2. Ice cover period from early to late winter and
3. Breaking-up period in early spring
4. Ice free summer season

In some countries this cycle is so stable that the beginning and the end of each of its phases or periods may be predicted with an error of a few days only. In such circumstance a wintertime rating curve can be made and used during the winter.

In Iceland, this regularity is more or less absent. Freezing-over may start during a cold period in early winter, but before the rivers are frozen the temperature rises, the ice is broken up without any intervening ice cover period and an ice free period may then follow. Each winter this may be repeated several times.

The winter period in Iceland is rather long and the weather is changing. It is sometimes very hard to detect when an icing begins and if there is still ice in the river after a thawing period.

The Hydrological Service (HS) has no automatic procedures for detecting suspicious periods. All series are treated manually, and it is up to the operator to detect the period. At most gauging stations in Iceland, the Hydrological Service (HS) does not employ a local observer to follow and report on the building-up of ice jams in the river. Consequently, the usual situation is that no information from local sources on potential ice problems is available. In this case the expertise of the hydrologist working on ice correction is the only thing to rely on. The usual procedure is to use weather data from a nearby meteorological station showing both temperature and precipitation. The water level data is compared with the meteorological data and usually backwater effects are identified as an

abrupt increase in water level during periods where this would not be expected due to cold and dry weather. Sometimes this could be very difficult procedure especially if the backwater effects are not very remarkable and only last for short periods of time. Also, the effects of frazil ice and in particular anchor ice can be very difficult to detect, but ice problems due to these types of ice formation are quite common in Icelandic rivers. Discharge measurements during winter conditions are performed once every winter.

Finally, Figure 3.12 shows a graph of ice corrupted water level data in the River Fnjoska in northern Iceland in late winter 1996. Furthermore, the temperature and precipitation at Akureyri, also in northern Iceland, is shown. The temperature in January is cold and at the beginning of February the water level rises drastically and stays high until a long warm period in March sets in. Then the water level falls.

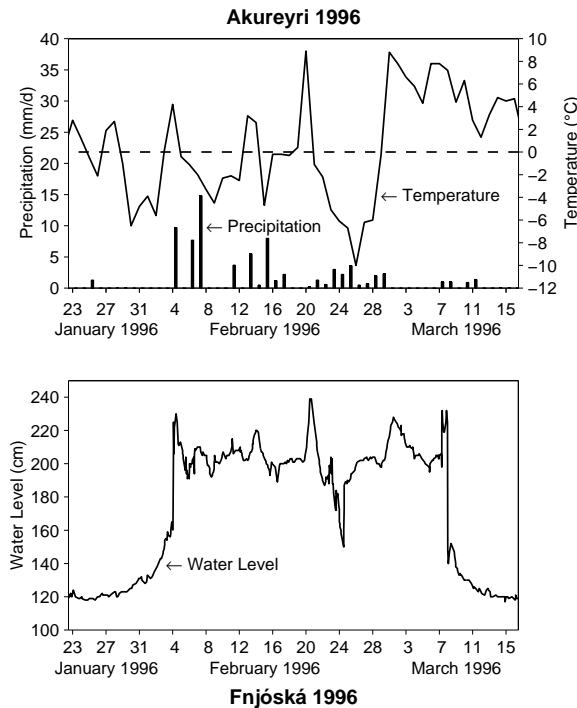


Figure 3.12: Ice disturbance.

The problem of river ice is, indeed, one of the reasons for applying stochastic models in hydrology.

The stochastic dynamic modelling

This PhD thesis consists of stochastic modelling of hydrological systems. Four projects have been studied and the results are illustrated in Papers [A], [B] [C] and [D].

Journal papers are often written in a compact form as only a limited number of pages are allowed. In this section the modelling approaches will be further described.

4.1 Model categorization

Modelling approach may be grouped into three categories; Black box models, white box models and grey box models.

Black box models are purely data based model. They approach the system in terms of input and output with the internals hidden in a black box.

White box models are the opposite of the black box ones. The internal of the system is fully known. In order the develop a white box the modeller must know

the system in details and occasionally the model becomes over-specified.

Grey box models are placed in between black- and white- box modelling approaches. A model is viewed as grey box if physical knowledge about the system is used along with the data. Thus the model is not completely described by physical equations but the equations and the parameters are physically interpretable.

In this project both grey box models and black box models have been used. The choice of model depends on the aim and on the available knowledge in each situation.

Models can be developed in continuous time and discrete time. As the time is continuous all the physical systems studied are defined in continuous time. However, the data are sampled in discrete time and occasionally it is more convenient to describe the model in discrete time. The dynamical data series studied are called the time series.

The systems, or the models, can be grouped into linear models and non-linear models, black box systems can be both linear and non-linear.

4.2 Non-linear models in general

Broadly speaking, the way a modeller often prefers to think about a non-linear system is a system where non-linear equations are used for describing it. In a discrete time a system is linear if the output, $y(t)$, is described as a linear function of past values of y , $y(t-1), \dots, y(1)$ and past values of input (if necessary) $u(t), \dots, u(1)$ and a white noise residuals. For model described in continuous time it is referred to as linear if the differential equations used to describe the system are linear. Furthermore the linearity can also be studied in the frequency domain, a linear system will respond to a single harmonic input with a single harmonic output.

A proper mathematical description of a non-linear system involves the Volterra series, who first published his ideas nearly a century ago. Later, Wiener, Varrett and Kalman took up the theme but from different standpoints (Tong 1990).

The aim of the modelling effort may be generally expressed as, finding a function h such that $\{\epsilon_t\}$ defined by

$$h(y_t, y_{t-1}, \dots) = \epsilon_t \quad (4.1)$$

where ϵ_t is a sequence of independent random variables. Suppose that the model is *causally invertible*, i.e., the equation above may be 'solved' such

$$y_t = h'(\epsilon_t, \epsilon_{t-1}, \dots). \quad (4.2)$$

Suppose that h' is sufficiently well-behaved so that it can be expanded in a Taylor series:

$$\begin{aligned} y_t = & \mu + \sum_{k=0}^{\infty} g_k \epsilon_{t-k} + \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} g_{kl} \epsilon_{t-k} \epsilon_{t-l} \\ & + \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} g_{klm} \epsilon_{t-k} \epsilon_{t-l} \epsilon_{t-m} + \dots \end{aligned} \quad (4.3)$$

The functions:

$$\mu = h'(0), \quad g_k = \left(\frac{\partial h'}{\partial \epsilon_{t-k}} \right), \quad g_{kl} = \left(\frac{\partial^2 h'}{\partial \epsilon_{t-k} \partial \epsilon_{t-l}} \right), \text{ etc.} \quad (4.4)$$

are called the **Volterra series** for the process $\{X\}$. The sequences $\{g_k\}, \{g_{kl}\}, \dots$ are called the **kernels** of the Volterra series.

For linear systems the following is true:

$$g_{kl} = g_{klm} = g_{klmn} = \dots = 0 \quad (4.5)$$

Hence, the system is completely characterized by either

$\{g_k\}$: Impulse response function

or

$\mathcal{H}(\omega)$: Frequency response function

In general there is no such thing as a transfer function for non-linear systems. However, an infinite sequence of generalized transfer functions may be defined as:

$$\begin{aligned} H_1(\omega_1) &= \sum_{k=0}^{\infty} g_k e^{-i\omega_1 k} \\ H_2(\omega_1, \omega_2) &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} g_{kl} e^{-i(\omega_1 k + \omega_2 l)} \\ H_3(\omega_1, \omega_2, \omega_3) &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} g_{klm} e^{-i(\omega_1 k + \omega_2 l + \omega_3 m)} \\ &\vdots \end{aligned}$$

Let u_t and y_t denote the input and the output of a non-linear system, respectively.

- For linear systems it is well known that:
 - L1 If the input is a single harmonic $u_t = A_0 e^{i\omega_0 t}$ then the output is a single harmonic of the same frequency but with the amplitude scaled by $|H(\omega_0)|$ and the phase shifted by $\arg H(\omega_0)$.
 - L2 The principle of superposition is valid, and the total output is the sum of the outputs corresponding to the individual frequency components of the input. (Hence, the system can be completely described if the response to all frequencies is known – that is what the transfer function supplies).
- For non-linear systems, however, neither of the properties (L1) or (L2) holds. More specifically:
 - NL1 For an input with frequency ω_0 , the output will, in general, also contain components at the frequencies $2\omega_0, 3\omega_0, \dots$ (frequency multiplication).
 - NL2 For two inputs with frequencies ω_0 and ω_1 , the output will contain components at frequencies $\omega_0, \omega_1, (\omega_0 + \omega_1)$ and all harmonics of the frequencies (inter-modulation distortion).

Further description can be found in, e.g., (Tong 1990) and (Madsen & Holst 2000).

4.3 Stochastic differential equations and parameter estimation

One of the most successful approach of grey box modelling is to use SDE, (Stochastic Differential Equations). The advantages of modelling in a continuous time is that the continuous time is a realization of the physical world and thus the models parameters are directly interpretable.

The modelling approach in papers [B] and [D] is by use of SDE's.

4.3.1 Continuous-discrete state space models

A general formulation of a continuous-discrete stochastic state space model is written as,

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta})d\boldsymbol{\omega}_t \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, t_k, \boldsymbol{\theta}) + \mathbf{e}_k \end{aligned} \quad (4.6)$$

where $t \in \mathcal{R}_+$ is time, $\mathbf{x}_t \in \mathcal{R}^n$ is a vector of state variables, $\mathbf{u}_t \in \mathcal{R}^m$ is a vector of input variables and $\boldsymbol{\theta} \in \mathcal{R}^p$ is a vector of parameters. The functions $\mathbf{f}(\cdot) \in \mathcal{R}^n$ and $\boldsymbol{\sigma}(\cdot) \in \mathcal{R}^{n \times q}$ are general non-linear function. The vector $\mathbf{y}_k \in \mathcal{R}^l$ is a vector of measurements, $\mathbf{x}_k = \mathbf{x}_{t=t_k}$ and $\mathbf{u}_k = \mathbf{u}_{t=t_k}$. The vector $\boldsymbol{\omega}_t$ is a q -dimensional standard Wiener process and $\mathbf{e}_k \in N(0, S(\mathbf{u}_k, t_k, \boldsymbol{\theta}))$ is an l -dimensional white noise process. The function $\mathbf{h}(\cdot) \in \mathcal{R}^l$ is a non-linear function.

Measurements can be used to estimate the parameter vector $\boldsymbol{\theta}$. This is performed by using the software CTSM which is based on the principle of maximum likelihood method (ML) for parameter estimation. In order for this to work out it is necessary to assume that the function $\boldsymbol{\sigma}(\cdot)$ is independent of the state variable \mathbf{x}_t . Thus, a general formulation of a continuous-discrete stochastic state space model of which parameters can be estimated by use of CTSM is:

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta})d\boldsymbol{\omega}_t \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, t_k, \boldsymbol{\theta}) + \mathbf{e}_k \end{aligned} \quad (4.7)$$

In Paper [B], the function $\mathbf{f}(\cdot)$ is non-linear, the function $\mathbf{h}(\cdot)$ is linear and independent of the input and the function $\boldsymbol{\sigma}$ is a constant, representing, a standard deviation of the increments. The equations in Paper [B] can thus be written as:

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\boldsymbol{\theta})d\boldsymbol{\omega}_t \\ \mathbf{y}_k &= \mathbf{C}(\boldsymbol{\theta})\mathbf{x}_t + \mathbf{e}_k \end{aligned} \quad (4.8)$$

Contrarily, in the Paper [D] the model is a linear time invariant model, and is written as

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{A}(\boldsymbol{\theta})\mathbf{x}_t dt + \mathbf{B}(\boldsymbol{\theta})\mathbf{u}_t dt + \boldsymbol{\sigma}(\boldsymbol{\theta})d\boldsymbol{\omega}_t \\ \mathbf{y}_k &= \mathbf{C}(\boldsymbol{\theta})\mathbf{x}_t + \mathbf{e}_k \end{aligned} \quad (4.9)$$

The measurements \mathbf{y}_k are used to estimate the parameter vector $\boldsymbol{\theta}$. It is a well known fact that the likelihood function for time series is a product of conditional densities

$$(\boldsymbol{\theta}; \mathcal{Y}_N) = \left(\prod_{k=1}^N p(\mathbf{y}_k | \mathcal{Y}_k, \boldsymbol{\theta}) \right) p(\mathbf{y}_0 | \boldsymbol{\theta}). \quad (4.10)$$

In order to obtain an exact evaluation of the likelihood function, the initial probability density $p(\mathbf{y}_0|\boldsymbol{\theta})$ must be known and all subsequent conditional densities must be determined. This can in theory be determined by Kolmogorov's forward equation (Jazwinski 1970). However, this is not very suitable in practice, and with the right assumptions this can be approximated with the Normal distribution

$$L(\boldsymbol{\theta}; \mathcal{Y}_N) \approx \prod_{k=1}^N \frac{1}{2\pi^{l/2}} \frac{\exp(\boldsymbol{\epsilon}_k^T \mathbf{R}_{k|k-1}^{-1} \boldsymbol{\epsilon}_k)}{\det(\mathbf{R}_{k|k-1})^{1/2}} \quad (4.11)$$

with mean

$$\boldsymbol{\epsilon}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1} = \mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{u}_{k|k-1}, t_{k-1}, \boldsymbol{\theta})$$

and variance

$$\mathbf{R}_{k|k-1} = V(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}).$$

Conditioning on \mathbf{y}_0 and taking the negative logarithm gives:

$$-\ln(L(\boldsymbol{\theta}; \mathcal{Y}_N)) \approx \frac{1}{2} \sum_{k=1}^N \left(\ln(\det(\mathbf{R}_{k|k-1})) + \boldsymbol{\epsilon}_k^T \mathbf{R}_{k|k-1}^{-1} \boldsymbol{\epsilon}_k \right) + \frac{1}{2} N l \ln(2\pi) \quad (4.12)$$

which involves calculating a sum rather than a product.

When the stochastic differential equation in Eq.(4.8) is LTI (Linear Time Invariant) or LTV (Linear Time Variant) the result in Eq.(4.11) is exact while in the case of a non-linear stochastic differential equation the result in Eq.(4.11) is approximative.

The calculations which lead to Eq. (4.11) are based on the Kalman filter equations. Thus, the continuous-discrete Kalman filter equations are written in Table 4.1, whereafter the assumptions and prerequisites are discussed. The continuous-discrete Kalman filter equations can be seen in (Jazwinski 1970).

The first equation is the **output prediction**, the second one is the one-step **prediction error/innovation**. The third one is the **covariance of the prediction error**. The fourth one is known as the **Kalman gain**. The fifth one is the **state updating**, $E[\mathbf{x}_t|\mathcal{Y}_t]$, and the sixth one is the **updating's variance**, $V[\mathbf{x}_t|\mathcal{Y}_t]$. The seventh one is the one-step **state prediction** and the eighth one is the **state prediction's variance**.

It is clear that for calculation of the output prediction $\hat{\mathbf{y}}_{k|k-1}$ and its variance $\mathbf{R}_{k|k-1}$ the variables $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ must be known. The change in time of $\mathbf{x}_{t|k-1}$ and $\mathbf{P}_{t|k-1}$ is described in continuous time. Thus the transition from one time to another involves integration.

Table 4.1: The Kalman filter equations

Linear SDE	Non – linear SDE
$\widehat{\mathbf{y}}_{k k-1} = \mathbf{C}\widehat{\mathbf{x}}_{k k-1} + \mathbf{D}\mathbf{u}_k$	$\widehat{\mathbf{y}}_{k k-1} = \mathbf{h}(\widehat{\mathbf{x}}_{k k-1}, \mathbf{u}_k, t_k, \boldsymbol{\theta})$
$\boldsymbol{\epsilon}_k = \mathbf{y}_k - \widehat{\mathbf{y}}_{k k-1}$	$\boldsymbol{\epsilon}_k = \mathbf{y}_k - \widehat{\mathbf{y}}_{k k-1}$
$\mathbf{R}_{k k-1} = \mathbf{C}\mathbf{P}_{k k-1}\mathbf{C}^T + \mathbf{S}_t$	$\mathbf{R}_{k k-1} = \mathbf{C}\mathbf{P}_{k k-1}\mathbf{C}^T + \mathbf{S}_t$
$\mathbf{K}_k = \mathbf{P}_{k k-1}\mathbf{C}^T\mathbf{R}_{k k-1}^{-1}$	$\mathbf{K}_k = \mathbf{P}_{k k-1}\mathbf{C}^T\mathbf{R}_{k k-1}^{-1}$
$\widehat{\mathbf{x}}_{k k} = \widehat{\mathbf{x}}_{k k-1} + \mathbf{K}\boldsymbol{\epsilon}_k$	$\widehat{\mathbf{x}}_{k k} = \widehat{\mathbf{x}}_{k k-1} + \mathbf{K}\boldsymbol{\epsilon}_k$
$\mathbf{P}_{k k} = \mathbf{P}_{k k-1} - \mathbf{K}_k\mathbf{R}_{k k-1}\mathbf{K}_k^T$	$\mathbf{P}_{k k} = \mathbf{P}_{k k-1} - \mathbf{K}_k\mathbf{R}_{k k-1}\mathbf{K}_k^T$
$\frac{d\widehat{\mathbf{x}}_{t k}}{dt} = \mathbf{A}\widehat{\mathbf{x}}_{t k} + \mathbf{B}\mathbf{u}_t$	$\frac{d\widehat{\mathbf{x}}_{t k}}{dt} = \mathbf{f}(\widehat{\mathbf{x}}_{t k}, \mathbf{u}_k, t_k, \boldsymbol{\theta})$
$t \in [t_k, t_{k+1}[$	$t \in [t_k, t_{k+1}[$
$\frac{d\widehat{\mathbf{P}}_{t k}}{dt} = \mathbf{A}\widehat{\mathbf{P}}_{t k} + \mathbf{P}_{t k}\mathbf{A}^T + \boldsymbol{\sigma}\boldsymbol{\sigma}^T$	$\frac{d\widehat{\mathbf{P}}_{t k}}{dt} = \mathbf{A}\widehat{\mathbf{P}}_{t k} + \mathbf{P}_{t k}\mathbf{A}^T + \boldsymbol{\sigma}\boldsymbol{\sigma}^T$
$t \in [t_k, t_{k+1}[$	$t \in [t_k, t_{k+1}[$
Shorthand notation	Shorthand notation
$\mathbf{A} = \mathbf{A}(\boldsymbol{\theta})$, and $\mathbf{B} = \mathbf{B}(\boldsymbol{\theta})$	$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}_t} \right _{\mathbf{x}=\widehat{\mathbf{x}}_{k k-1}, \mathbf{u}=\mathbf{u}_k, t=t_k, \boldsymbol{\theta}}$
$\mathbf{C} = \mathbf{C}(\boldsymbol{\theta})$, and $\mathbf{D} = \mathbf{D}(\boldsymbol{\theta})$	$\mathbf{C} = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right _{\mathbf{x}=\widehat{\mathbf{x}}_{k k-1}, \mathbf{u}=\mathbf{u}_k, t=t_k, \boldsymbol{\theta}}$
$\boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{\theta})$, and $\mathbf{S} = \mathbf{S}(\boldsymbol{\theta})$	$\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{u}_t, t, \boldsymbol{\theta})$, $\mathbf{S} = \mathbf{S}(\mathbf{u}_k, t_k, \boldsymbol{\theta})$

A **fundamental assumption** before integrating is to define the stochastic integral. The stochastic process must be defined as an **Ito process** and then the stochastic integral is an **Ito integral**, see (Øksendal 1995) for details.

In the case of a linear time invariant model like Eq. (4.9) the stochastic differential equation is

$$d\mathbf{x}_t = \mathbf{A}\mathbf{x}_t dt + \mathbf{B}\mathbf{u}_t dt + \boldsymbol{\sigma}d\boldsymbol{\omega}_t \quad t \in [t_k, t_{k+1}[\quad (4.13)$$

and then the solution is,

$$\mathbf{x}_{t_{k+1}} = e^{\mathbf{A}(t_{k+1}-t_k)}\mathbf{x}_{t_k} + \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\mathbf{B}\mathbf{u}_s ds + \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}d\boldsymbol{\omega}_s. \quad (4.14)$$

The prediction is the expectation value

$$\widehat{\mathbf{x}}_{k+1|k} = E\{\mathbf{x}_{t_{k+1}}|\mathbf{x}_{t_k}\}.$$

Because of the fact that the stochastic process is an Ito process it follows that

$$E\left\{\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}d\boldsymbol{\omega}_s\right\} = 0$$

which actually is not true for all stochastic integrals e.g., the Statonovich integral (Øksendal 1995). Thus, the Ito process assumption is fundamental. Given $\widehat{\mathbf{x}}_{k|k}$ the first two terms in Eq.(4.14) do not include a stochastic part. It follows that

$$\widehat{\mathbf{x}}_{k+1|k} = E\{\mathbf{x}_{t_{k+1}}|\mathbf{x}_{t_k}\} = e^{\mathbf{A}(t_{k+1}-t_k)}\mathbf{x}_{t_k} + \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\mathbf{B}\mathbf{u}_s ds \quad (4.15)$$

The state predictions covariance is

$$\begin{aligned} \widehat{\mathbf{P}}_{k+1|k} &= V\{\mathbf{x}_{t_{k+1}}\mathbf{x}_{t_{k+1}}^T|\mathbf{x}_{t_k}\} \\ &= V\left[e^{\mathbf{A}(t_{k+1}-t_k)}\mathbf{x}_{t_k}\right] + V\left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\mathbf{B}\mathbf{u}_s ds\right] \end{aligned} \quad (4.16)$$

$$+V\left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}d\boldsymbol{\omega}_s\right]. \quad (4.17)$$

The first term is

$$\begin{aligned} V\left[e^{\mathbf{A}(t_{k+1}-t_k)}\mathbf{x}_{t_k}\right] &= e^{\mathbf{A}(t_{k+1}-t_k)}V[\mathbf{x}_{k|k}]\left(e^{\mathbf{A}(t_{k+1}-t_k)}\right)^T \\ &= e^{\mathbf{A}(t_{k+1}-t_k)}\mathbf{P}_{k|k}\left(e^{\mathbf{A}(t_{k+1}-t_k)}\right)^T \end{aligned} \quad (4.18)$$

(using $V[Ax] = AV[x]A^T$ and $V[\mathbf{x}_{k|k}] = \mathbf{P}_{k|k}$). While the second term

$$V\left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\mathbf{B}\mathbf{u}_s ds\right] = 0, \quad (4.19)$$

because no stochastic part is involved. Finally, because of the fact that the noise term in Eq.(4.8) is independent of the state variables

$$\begin{aligned} V\left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}d\boldsymbol{\omega}_s\right] &= E\left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}d\boldsymbol{\omega}_s \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}d\boldsymbol{\omega}_s\right] \\ &= E\left[\int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-s)}\boldsymbol{\sigma}\boldsymbol{\sigma}^T\left(e^{\mathbf{A}(t_{k+1}-s)}\right)^T ds\right] \end{aligned} \quad (4.20)$$

since for an Ito integral the following is true

$$E\left[\left(\int_a^b \mathbf{g}(s)d\boldsymbol{\omega}_s\right)^2\right] = E\left[\int_a^b \mathbf{g}^2(s)ds\right]. \quad (4.21)$$

Again, the Ito process assumption is fundamental. Furthermore, because of the normal distribution assumption of the measurements noise term e_k , and the Wiener process assumption for the state space variables, the conditional distribution also becomes a normal distribution.

For the non-linear systems like Eq.(4.8) then

$$\mathbf{x}_{t_{k+1}} = \int_{t_k}^{t_{k+1}} \mathbf{f}(\mathbf{x}_s, \mathbf{u}_s, s, \boldsymbol{\theta}) ds + \int_{t_k}^{t_{k+1}} \boldsymbol{\sigma}(\boldsymbol{\theta}) d\boldsymbol{\omega}_s. \quad (4.22)$$

The state prediction $\hat{\mathbf{x}}_{k+1|k} = E\{\mathbf{x}_{t_{k+1}} | \mathbf{x}_{t_k}\}$ is the evaluation of the first integral as the expectation value of the latter integral is zero due to the Ito process assumptions. The evaluation of the integral can be calculated by linearizing the function $\mathbf{f}()$ and then use methods for linear models, or numerical methods can be applied, see more about this in Appendix E. For calculation of the covariance matrix $\hat{\mathbf{P}}_{k+1|k} = V\{\mathbf{x}_{t_{k+1}} \mathbf{x}_{t_{k+1}}^T | \mathbf{x}_{t_k}\}$ the non-linear function $\mathbf{f}()$ is approximated by its first derivative, as can be seen in the Kalman filter equations.

Finally, it is reasonable to believe that the conditional distribution can be well approximated by the normal distribution whether or not the function $\mathbf{f}()$ is linearized, or numerical methods for integration are applied.

4.4 The family of linear stochastic models

In this thesis it is argued that grey box modelling constitutes a very powerful modelling framework, allowing modelling for both non-linear and non-stationary systems. Furthermore, if attention is restricted to linear and time-invariant models, then a rich family of well-known linear stochastic models are obtained as a special case.

In this section the family of linear stochastic models are described, based on the general formulation of linear time invariant SDE's as in Eq. (4.9)

4.4.1 The transfer function form

The form of the linear, discrete time, transfer functions is also frequently called the Box-Jenkins transfer functions, since (Box & Jenkins 1976) are responsible for the great popularity of this class of models.

The relation between the discrete state space form and the transfer form is as follows: Consider the following well known formulation of a discrete time state space model with constant coefficients:

$$\begin{aligned} \mathbf{x}(t+1) &= \boldsymbol{\phi} \mathbf{x}(t) + \boldsymbol{\Gamma} \mathbf{u}(t) + \mathbf{v}(t) \\ \mathbf{y} &= \mathbf{C} \mathbf{x}(t) + \mathbf{e}(t) \end{aligned} \quad (4.23)$$

where $\{\mathbf{u}(t)\}$ is the input and $\{\mathbf{v}(t)\}$ and $\{\mathbf{e}(t)\}$ are mutual uncorrelated white noise processes with variance \mathbf{R}_1 and \mathbf{R}_2 , respectively.

This state space model might the solution of the linear time invariant SDE in Eq., (4.9) as shown in Eq. (4.14). Denote the constant sampling time as τ , where $\tau = t_{k+1} - t_k$, the notation in Eq.(4.23) implies $\tau = 1$. Then

$$\phi(\tau) = e^{\mathbf{A}\tau}. \quad (4.24)$$

By assuming that the input \mathbf{u}_s is constant in the time interval $[t, t + \tau[$ then

$$\Gamma(\tau) = \int_t^{t+\tau} e^{\mathbf{A}(t+\tau-s)} \mathbf{B} ds = \int_0^\tau e^{\mathbf{A}s} \mathbf{B} ds \quad (4.25)$$

where the last equation is true because of the time-invariant assumption. Finally,

$$\mathbf{v}(t, \tau) = \int_t^{t+\tau} e^{\mathbf{A}(t+\tau-s)} \boldsymbol{\sigma} d\omega_s \quad (4.26)$$

$\mathbf{v}(t, \tau)$ the variance \mathbf{R}_1 of the white noise process is $\mathbf{R}_1 = \int_t^{t+\tau} \phi(\tau) \boldsymbol{\sigma} \boldsymbol{\sigma}^T \phi(\tau)^T$

The z-transformation of the state space equations in Eq. (4.23)

$$\begin{aligned} z\mathbf{x}(z) &= \phi\mathbf{x}(z) + \Gamma\mathbf{u}(z) + \mathbf{v}(z) \\ \mathbf{y}(z) &= \mathbf{C}\mathbf{x}(z) + \mathbf{e}(z) \end{aligned} \quad (4.27)$$

Elimination of $\mathbf{x}(z)$ in the system of linear equations Eq. (4.27) yields

$$\mathbf{y}(z) = \mathbf{C}(z\mathbf{I} - \phi)^{-1}\Gamma\mathbf{u}(z) + \mathbf{C}(z\mathbf{I} - \phi)^{-1}\mathbf{v}(z) + \mathbf{e}(z) \quad (4.28)$$

The rational polynomials in $\mathbf{u}(z)$ are found ahead of z and $\mathbf{v}(z)$.

If $\{\mathbf{y}_t\}$ is a stationary process (the matrix A is stable) the noise processes in Eq. (4.28) can be combined into one stationary noise process only, (Goodwin & Payne 1977)

$$\mathbf{y}(z) = \mathbf{C}(z\mathbf{I} - \phi)^{-1}\Gamma\mathbf{u}(z) + [\mathbf{C}(z\mathbf{I} - \phi)^{-1}\mathbf{K} + \mathbf{I}]\boldsymbol{\epsilon}(z) \quad (4.29)$$

or alternatively in the transfer function form,

$$\mathbf{y}(z) = \mathbf{H}_1(z)\mathbf{u}(z) + \mathbf{H}_2(z)\boldsymbol{\epsilon}(z) \quad (4.30)$$

where $\{\boldsymbol{\epsilon}_t\}$ is white noise with variance R , and $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ are rational polynomials in z :

$$\mathbf{H}_1(z) = \mathbf{C}(z\mathbf{I} - \phi)^{-1}\Gamma \quad (4.31)$$

$$\mathbf{H}_2(z) = \mathbf{C}(z\mathbf{I} - \phi)^{-1}\mathbf{K} + \mathbf{I} \quad (4.32)$$

The function $\mathbf{H}_1(z)$ is referred to as the transfer function (Box & Jenkins 1976).

The matrix \mathbf{K} is the stationary Kalman gain. \mathbf{R} is determined from the values of \mathbf{R}_1 , \mathbf{R}_2 , ϕ and \mathbf{C} , since

$$\mathbf{K} = \phi \mathbf{P} \mathbf{C}^T (\mathbf{C} \mathbf{P} \mathbf{C}^T + \mathbf{R}_2)^{-1} \quad (4.33)$$

$$\mathbf{R} = \mathbf{C} \mathbf{P} \mathbf{C}^T + \mathbf{R}_2 \quad (4.34)$$

where \mathbf{P} is determined by the stationary Ricatti equation

$$\mathbf{P} = \phi \mathbf{P} \phi^T + \mathbf{R}_1 - \phi \mathbf{P} \mathbf{C} (\mathbf{C} \mathbf{P} \mathbf{C}^T + \mathbf{R}_2) \mathbf{C} \mathbf{P} \phi^T \quad (4.35)$$

The ARMAX class of models are obtained in cases where the denominators in (4.30) for \mathbf{H}_1 and \mathbf{H}_2 are equal. Hence, the models are written:

$$\mathbf{A}(z) \mathbf{y}(z) = \mathbf{B}(z) \mathbf{u}(z) + \mathbf{C}(z) \boldsymbol{\epsilon}(z) \quad (4.36)$$

where \mathbf{A} , \mathbf{B} , and \mathbf{C} are polynomials in z . In the time domain this is written as

$$\mathbf{A}(q^{-1}) \mathbf{y}(t) = \mathbf{B}(q^{-1}) \mathbf{u}(t) + \mathbf{C}(q^{-1}) \boldsymbol{\epsilon}(t) \quad (4.37)$$

where $q^{-1} \mathbf{y}(t) = \mathbf{y}(t-1)$. In the case $\mathbf{A}(z) = 1$, the model is a FIR model (Final Impulse Response model).

As shown above a transfer function can be found from the state space form by eliminating the state vector. In contrast, for a given transfer function model a whole continuum of state space models exists. The most frequently used solution is to choose a canonical state space model - see, e.g., (Goodwin & Payne 1977). Eventually, physical knowledge can be used to state a proper connection between desirable state variables, to be introduced for the state space form.

Note that compared to the discrete time state space model:

- The decomposition of the noise into system and measurement noise is lost.
- The state variable is lost, i.e., the possibility for physical interpretation is further reduced.

4.4.2 Impulse response function models

A non-parametric description of the linear system is obtained by polynomial division, i.e.,

$$\mathbf{y}(t) = \sum_{i=0}^{\infty} \mathbf{h}_i \mathbf{u}(t-i) + \mathbf{N}(t) \quad (4.38)$$

where N_i is a correlated noise sequence. The sequence $\{h_i\}$ is the **impulse response** (matrix) function.

In the frequency (or z -) domain:

$$\mathbf{y}(z) = \mathbf{H}(z)\mathbf{u}(z) + \mathbf{N}(z) \quad (4.39)$$

where $\mathbf{H}(z)$ is the transfer function, and for $z = e^{i\omega}$ the **frequency response function (gain and phase)** is obtained.

Comparing a state space model with a transfer function model the following is observed:

- The description of the noise process is lost.
- The non-parametric model hides the number of time constants, etc.

4.4.3 Periodic models

Periodic models can be linear in the stochastic terms. In Paper [C] the measurements are discharge and, as in all Icelandic rivers the discharge is periodic. The model is written as

$$Q(t) = P(t) + S(t)y(t) \quad (4.40)$$

$$y(t) = a y(t-1) + \epsilon(t) \quad \epsilon(t) \in N(0, \sigma^2) \quad (4.41)$$

In order to ensure non-negative flow in the stochastic simulations, the discharge series was transformed by using the logarithm. $Q(t)$ denotes the log transformed discharge, $P(t)$ denotes the periodic mean and $S(t)$ denotes periodic standard deviation. The estimations of $P(t)$ and $S(t)$ were based on the log transformed discharge data. By inserting Eq. (4.41) into Eq.(4.40) it becomes

$$Q(t) = P(t) + S(t)[a y(t-1) + \epsilon(t)] \quad (4.42)$$

and by induction

$$Q(t) = P(t) + S(t)\left[\sum_{i=1}^t a^{t-i}\epsilon(i)\right] \quad (4.43)$$

(the term $a^t S(t)y(0)$ can be neglected as $|a| < 1$). Thus, the Volterra series, defined in Eq.(4.3) and Eq.(4.4), does not include more than a single sum, i.e., $\mu = P(t)$ and $g_k = S(t)a^{t-k}$ and $g_{kl} = g_{klm} = g_{klmn} = \dots = 0$. This proves that the system is linear in the stochastic terms.

The periodic mean $P(t)$ is subtracted from the data $Q(t)$ and this difference is scaled by the periodic standard deviation $S(t)$. The resulting process, $y(t)$ in Eq. (4.40) is an $AR(1)$ process as denoted from Eq. (4.41) Hence, the transfer function is

$$y(z) = (z - 1)^{-1} \epsilon(z). \quad (4.44)$$

4.5 Non-linear models in discrete time

4.5.1 Parametric models

A discrete time system is said to be non-linear if its present output is not a linear combination of past input and output signal elements (Cadzow 1973).

An example of a non-linear discrete time series models are *ARCH* models, Autoregressive Conditional Heteroskedasticity models. An *ARCH(2)* is written as:

$$y_t = e_t \sqrt{\gamma + \phi_1 y_{t-1}^2 + \phi_2 y_{t-2}^2} \quad e_t \in N(0, \sigma) \quad (4.45)$$

ϕ_1 , ϕ_2 , and γ are parameters. The *ARCH* models are commonly used in finance modelling to model asset price volatility over time.

A wide class of non-linear models are so called **threshold** models. The presence of a threshold, r specifies an operating mode of the system, i.e., there are different models in different regimes, where the regimes are defined by the threshold value r .

Examples of threshold models are the *SETAR* models, Self-Exciting Threshold Auto-regressive models and the *TARSO* models (Open-loop Threshold AutoRegressive System). The *SETAR* models are extensions of the *AR* models whereas the *TARSO* models are extensions of the *ARX* models.

A *SETAR* model consists of k *AR* parts, one part for each different regime. A *SETAR* model is often referred to as a *SETAR(k, p)* model where k is the number of regimes and p is the order of the autoregressive parts. If the autoregressive parts have different orders it is referred to as *SETAR(k, p₁, ... p_k)*. The regimes are defined by values related to the output values and the shift from one regime to another depends on the past values of the output series y_t (hence the Self-Exciting part of the name). An example of a *SETAR(2, 1)* is:

$$y_t = \begin{cases} 0 & - & 0.9y_{t-1} & + & e_{at} & & y_{t-1} < 0, & e_{at}, \in N(0, \sigma_a) \\ 0.9 & + & 0.9y_{t-1} & + & e_{bt} & & y_{t-1} \geq 0, & e_{bt}, \in N(0, \sigma_b) \end{cases} \quad (4.46)$$

A *TARSO* model consists of k *ARX* parts, one part for each different regime and it is often referred to as a *TARSO*(k, p). However, the regimes are defined by values r related to the input. The switch from one regime to another depends on past values of the input series x_t .

In (Gudmundsson 1970) and (Gudmundsson 1975) both *SETAR* and *TARSO* models are tested for modelling river flow discharge in Icelandic rivers. The results can also be seen in (Tong 1990).

The threshold modelling principle is used in Paper [B]. However the model is defined in continuous time and by using a smooth threshold, i.e., shift between regimes is defined by a smooth function.

The class of non-linear discrete time models is enormous and extensive literature on the topic exists.

4.5.2 Non-parametric regression

The non-parametric regression analysis traces the dependence of a response variable without specifying the function that relates the predictor to the response. In the case of time series analysis, the predictor can be an input variable (external variable), past values of the output, and/or past values of the error. Denoting the regressor variable as \mathbf{x} as it can be a vector, i.e., more than single regression variable, the response is denoted as y . The idea is to find a curve which relates \mathbf{x} and y . This is often referred to as smoothing.

Smoothing of a data set $\{\mathbf{x}_t, y_t\}$ involves the approximation of the mean response curve m in the regression relationship

$$y_t = m(\mathbf{x}_t) + e_t \quad t = 1, \dots, N \quad (4.47)$$

If repeated observations at a fixed point x are available the estimation of $m(\mathbf{x})$ can be done by using the average of the corresponding y -values. However, in the majority of cases repeated responses at a given point cannot be obtained and only a single response variable y and a single predictor variable x exists.

In the trivial case when $m(\mathbf{x}_t)$ is a constant, an estimation of m reduces to the point estimation of location, since average over the response variables y yields an estimate of m . However, in practical studies it is unlikely that the regression curve is constant. Rather the assumed curve is modelled as a smooth continuous function of a particular structure which is 'nearly constant' in small neighborhoods around \mathbf{x} . This local average should be constructed in such a way that it is defined only from observations in a small neighborhood around

\mathbf{x} , since y -observations from points far away will, in general, have very different mean values. This local averaging procedure can be viewed as the basic idea of smoothing. More formally this procedure can be defined as

$$\hat{m}(\mathbf{x}) = \frac{1}{N} \sum_{t=1}^N w_t(\mathbf{x}) y_t. \quad (4.48)$$

The estimator $\hat{m}(\mathbf{x})$ is called **smoother** (Härdle 1990). It is a weighted average of the response y_t in a neighborhood around \mathbf{x} . The amount of averaging is controlled by the weight sequence $\{w(\mathbf{x})_{t=1}^N\}$ which is tuned with a smoothing parameter. This smoothing parameter regulates the size of the neighborhood around \mathbf{x} .

If the weights $w_t(\mathbf{x})$ are positive and if the sum of the weights is one for all \mathbf{x} then $\hat{m}(\mathbf{x})$ is a least squares estimate at point x since

$$\arg \min_{\theta} \frac{1}{N} \sum_{t=1}^N w_t(\mathbf{x}) (y_t - \theta)^2 \quad (4.49)$$

Thus, the basic idea of local averaging is equivalent to the procedure of finding local least squares estimate.

A local average over too large a neighborhood would cast away the good with the bad. In this situation an extremely “over-smooth” curve would be produced, resulting in a biased estimate. On the other hand, defining the smoothing parameter so that it corresponds to a very small neighborhood would not sift the chaff from the wheat. Only a small number of observations would contribute to the estimate, which makes the non-parametric regression curve rough and wiggly. Finding the choice of the smoothing parameter that balances the trade-off between over-smoothing and under-smoothing is called the smoothing parameter selection problem.

A simple approach to a definition of the weight sequence $w_t(\mathbf{x}), t = 1, \dots, N$ is to describe the shape of the weight function by a density function with a scale parameter that adjusts the size and the form of the weights near \mathbf{x} . It is common to refer to the shape function as a **kernel** (Härdle 1990). A kernel is a continuous, bounded and symmetric real function k which integrates to one. A kernel has a shape parameter and scale parameter, the scale parameter is also referred to as the **bandwidth** (Härdle 1990). Commonly used kernel functions are a Gauss bell, a tricube function and an Epanecnikov kernel (Härdle 1990). The choice of weight function does not have a large impact (Silverman 1986). Thus, in kernel smoothing the choice of the bandwidth is the smoothing parameter selection problem.

An extension of the kernel estimation is local-polynomial regression. The fitted values are produced by locally weighted regression rather than by locally weighted averaging. A local linear regression can be formulated as

$$\arg \min_{\theta} \frac{1}{N} \sum_{t=1}^N w_t(\mathbf{x})(y_t - (\theta_0 + \theta_1(\mathbf{X}_t - \mathbf{x})))^2 \quad (4.50)$$

where \mathbf{X}_t denote the regressor at time t and \mathbf{x} is a grid point and thus $w_t(\mathbf{x})$ defines a neighborhood of points around \mathbf{x} and the local linear estimate of $m(\mathbf{x})$ is

$$\hat{m}(\mathbf{x}) = \hat{\theta}_0. \quad (4.51)$$

Local polynomial regression tends to be less biased than kernel regression, particularly on the boundary. Figure 4.1 shows a kernel estimate and a local-line regression. Note that the kernel is more biased in the boundaries. Using local lines instead of local constant allows a larger bandwidth without at bias problem.

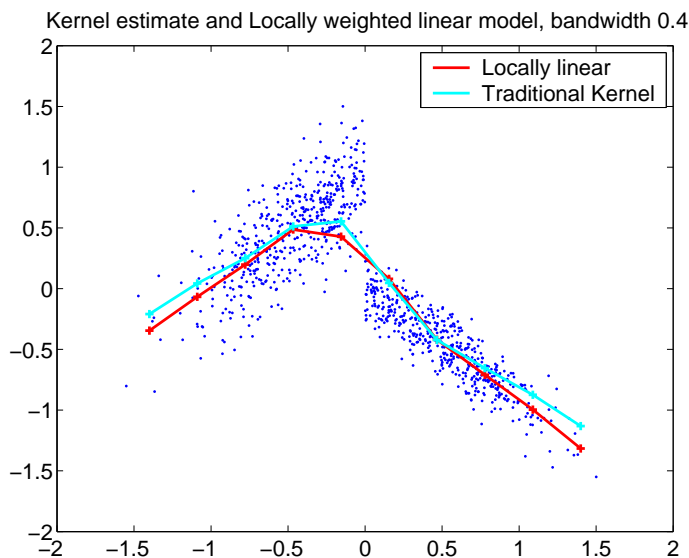


Figure 4.1: A comparison of the kernel estimated and locally linear models.

Several other smoothing techniques exist, e.g., orthogonal polynomials, spline smoothing and others. For more see (Härdle 1990) or (Burden & Faires 1989).

4.5.3 Conditional parametric models

A generalization of linear models are varying-coefficient models (Hastie & Tibshirani 1993). A varying-coefficient model is formulated as a linear model where the coefficients are assumed to change smoothly as an unknown function of other variables. When all the coefficients depend on the same variable the model is denoted as a conditional parametric model. The general formulation is

$$y_t = \mathbf{z}_t^T \theta(\mathbf{x}_t) + e_t; \quad t = 1, \dots, N, \quad e_t \in N(0, \sigma^2) \quad (4.52)$$

The variables \mathbf{z} and \mathbf{x} are predictors, $\mathbf{z}_t \in \mathcal{R}^k$ is the traditional predictor variable and $\mathbf{x}_t \in \mathcal{R}^r$ is a predictor variable which affects the variation of the coefficients, referred to as the explanatory variable. When the functional relationship $\theta(\mathbf{x}_t)$ is unknown (i.e., cannot not be parameterized) the relationship, be modelled by the principle of local estimation. This can be accomplished by using kernels and local polynomials

In a linear regression model the parameters $\theta_1, \dots, \theta_k$ are constants. In a conditional parameter model each of the $\theta_j \quad j = 1, \dots, k$ is modelled as a smooth function, estimated locally. Using a linear function this results in:

$$\theta_j(\mathbf{x}) = \theta_{j0} + \boldsymbol{\theta}_{j1}^T \mathbf{x} \quad (4.53)$$

Hence,

$$y_t = z_{1t}\theta_{10} + z_{1t}\boldsymbol{\theta}_{11}^T \mathbf{x} + \dots + z_{kt}\theta_{k0} + z_{kt}\boldsymbol{\theta}_{k1}^T \mathbf{x} \quad (4.54)$$

where $\boldsymbol{\theta}$ is estimated locally with respect to \mathbf{x} . If $z_j = 1$ for all j this becomes a local polynomial regression, in line with the method introduced by (Cleveland & Develin 1988). If $\boldsymbol{\theta}(\cdot)$ is also a local constant, the method of estimation reduced to determining the scalar $\hat{\theta}_j(\mathbf{x})$ so that $\sum_{t=1}^n w_t(\mathbf{x})(y_t - \hat{\theta}_j(\mathbf{x}))^2$ is minimized, i.e., the method is reduced to traditional kernel estimation, see (Härdle 1990) or (Hastie & Loader 1993).

In practice a new design matrix is defined as:

$$\mathbf{s}_t^T = [(z_{1t}, z_{1t}x_{1t}, \dots, z_{1t}x_{rt}), \dots, (z_{kt}, z_{kt}x_{1t}, \dots, z_{kt}x_{rt})] \quad (4.55)$$

and a new column vector as:

$$\boldsymbol{\theta}_{j\mathbf{x}} = \begin{pmatrix} \theta_{j0} \\ \theta_{j1} \\ \vdots \\ \theta_{jr} \end{pmatrix} \quad (4.56)$$

and

$$\boldsymbol{\theta}_{\mathbf{x}} = [\boldsymbol{\theta}_{1\mathbf{x}}^T, \dots, \boldsymbol{\theta}_{j\mathbf{x}}^T, \dots, \boldsymbol{\theta}_{k\mathbf{x}}^T]^T. \quad (4.57)$$

The vector \mathbf{y}_t can then be written as

$$\mathbf{y}_t = \mathbf{s}_t^T \boldsymbol{\theta}_{\mathbf{x}} + e_t \quad t = 1, \dots, N, \quad (4.58)$$

The parameter vector $\boldsymbol{\theta}_{\mathbf{x}}$ is fitted locally to \mathbf{x} . This is accomplished by using the traditional weighted least squares, where the weight on observation t is related to the distance from \mathbf{x} to \mathbf{x}_t , so that

$$w_t(\mathbf{x}) = W(\|\mathbf{x}_t - \mathbf{x}\|/d(\mathbf{x})), \quad (4.59)$$

where $\|\mathbf{x}_t - \mathbf{x}\|$ is the Euclidean distance between \mathbf{x}_t and \mathbf{x} and $d(\mathbf{x})$ is the bandwidth. Hence, it is clear that the fitted values \hat{y}_t are a linear combination of the measurements y_1, \dots, y_t .

When the local estimate in Eq. (4.58) $\hat{\boldsymbol{\theta}}_{\mathbf{x}}$ is obtained, the elements of $\hat{\boldsymbol{\theta}}(\mathbf{x})$

$$\hat{\boldsymbol{\theta}}_j(\mathbf{x}_t) = [1, x_{1t}, x_{2t}] \hat{\boldsymbol{\theta}}_{j\mathbf{x}} \quad (j = 1, \dots, k). \quad (4.60)$$

In case of an ARX model as in Paper [A] the vector \mathbf{z} consists of lagged values of the output y and lagged values of the input u . An ARX(2,6) with time delay 2 as in Paper [A]

$$\begin{aligned} y_t &= a_1(\mathbf{x}_{t-m})y_{t-1} + a_2(\mathbf{x}_{t-m})y_{t-2} \\ &\quad + b_2(\mathbf{x}_{t-m})u_{t-2} + \dots + b_7(\mathbf{x}_{t-m})u_t \quad e_t \in N(0, \sigma) \end{aligned} \quad (4.61)$$

where m is the time delay in the explanatory variable \mathbf{x} if any. In time series notation this is written as

$$A_{\mathbf{x}_{t-m}}(q^{-1})y_t = B_{\mathbf{x}_{t-m}}(q^{-1})u_t + e_t \quad e_t \in N(0, \sigma) \quad (4.62)$$

where

$$A_{\mathbf{x}_{t-m}}(q^{-1}) = 1 - a_1(\mathbf{x}_{t-m})q^{-1} - a_2(\mathbf{x}_{t-m})q^{-2} \quad (4.63)$$

$$B_{\mathbf{x}_{t-m}}(q^{-1}) = b_2(\mathbf{x}_{t-m})q^{-2} + \dots + b_7(\mathbf{x}_{t-m})q^{-7} \quad (4.64)$$

Thus, for each, fixed value of the explanatory variable \mathbf{x}_{t-m} the transfer function form in the z domain is

$$y(z) = (A_{\mathbf{x}_{t-m}}(z))^{-1} B_{\mathbf{x}_{t-m}}(z)u(z) + (A_{\mathbf{x}_{t-m}}(z))^{-1} e_t. \quad (4.65)$$

4.6 Overview of the statistical methods

The statistical methods used in this PhD project have been described. This section provides an overview of the methods and the model relations are shown.

Two models are grey box models, formulated by SDE's: one model is linear (Paper [D]) and the other is a non-linear threshold model (Paper [B]). Two models are black box models; of which one model is periodic model but linear in the stochastic terms (Paper [C]). The other model is a conditional black box model (Paper [A]).

Figure 4.2 shows an overview of the models and demonstrates how the models are related. A linear time invariant stationary differential equation can be written in a discrete state space form, which can be rewritten in the well known transfer function form. The periodic model can also be rewritten in a transfer function form. The non-linear differential equation can in each time step t be written as a discrete state space equation, either by using a 1st order Taylor approximation and derive the discrete state space equation from the linearized equation, or a numerical solution to the non-linear equation can be found and thereby relate $\mathbf{x}(t+1)$ and $\mathbf{x}(t)$ in a discrete state space equation. This can then be simplified further and written as a transfer function model. Finally, the conditional parametric model can be written on the transfer function form for each fixed value of the conditional variable.

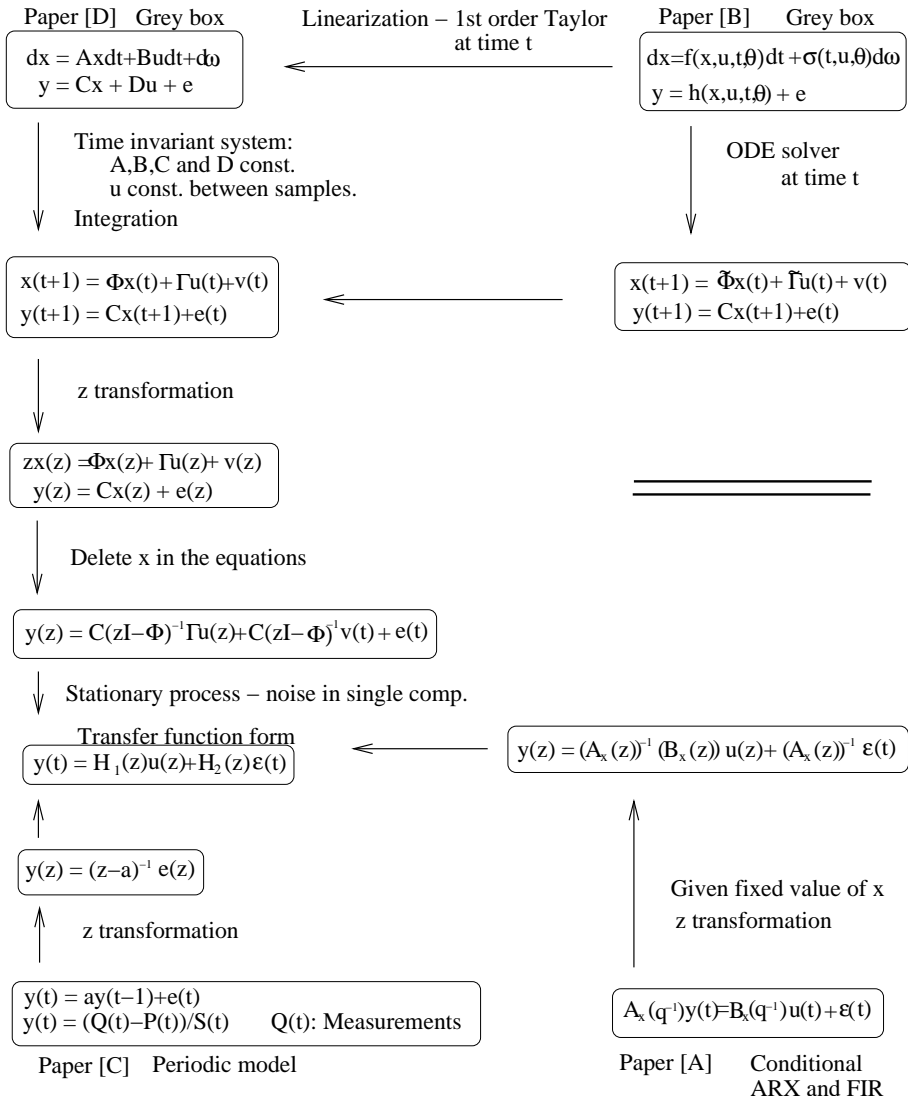


Figure 4.2: An overview of the statistical methods and their relations

4.7 Motivation for using grey box modelling

The advantage of the grey box modelling approach in continuous time is that it combines the benefits of both the white box and black box modelling principles. Hence, the grey box approach delivers a strong modelling framework, the possibility of combining prior physical knowledge with data information.

White box models often tend to be over specified and in hydrology sometimes the physical processes are so complicated that true physical equations have not yet been proposed (Viessman & Lewis 1996). However, in most cases it will be almost impossible to establish an exact model of all the subprocesses needed for describing the true system. Furthermore, the ideal situation of knowing all the parameters of the model is often absent, especially in hydrology (Refsgaard et al. 1992). This calls for some form of calibration or estimation from data. It is quite common that physical models are formulated as:

$$\begin{aligned} dx_t &= f(x_t, u_t, t, \theta) dt \\ y_k &= h(x_k, u_k, t_k, \theta) + e_k \end{aligned} \quad (4.66)$$

Compared to Eq.(4.7) there is no noise in the system equation, i.e., the dynamical part in Eq. (4.66), indicates that the physical dynamics is perfectly described in the model used.

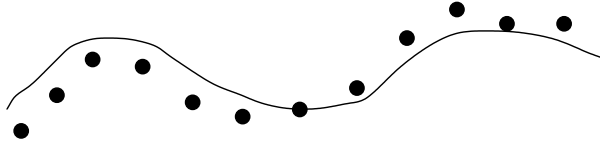


Figure 4.3: The line demonstrates a dynamic process modelled with differential equations without system noise, whereas the dots denote typical observations related to such a system.

A physical modelling typically results in deviations between the output predicted by the model and observations as indicated in Figure 4.3. Hence, the model error is serial correlated. This autocorrelation in the model errors calls for a dynamic model which includes system noise, because if no system noise is present for the true system then the prediction errors must be independent. In conclusion a situation as sketched in Figure 4.3 calls for using stochastic differential equations (SDEs) as an alternative to ordinary differential equations (ODEs).

Furthermore, in (Kristensen et al. 2004a) it is shown that in cases where true system contains noise in the system equation (the system is not perfectly described by the system equation) a calibration method which does not take this

into account will lead to biased estimates, whereas parameter estimation methods which account for the system noise, as the method implemented in CTSM, will provide the true values of the parameters.

A modelling framework based on SDE's also provides useful techniques for finding the most adequate proposal for an extension of a given model. Consider a given model based on a dynamic description provided by a SDE. The SDE contains, as mentioned previously, a drift term and a diffusion term. If the SDE provides a perfect description of the system then the diffusion part vanishes, and hence the diffusion terms describe the part of the dynamics which is not adequately described by the drift part of the model. The influence of the diffusion part is described by the incremental covariances (typically of a Wiener process). This implies that large elements in the diagonal of the covariance matrix indicate that the relevant dynamic part of the model calls for some improvement. In (Kristensen et al. 2004b) a systematic approach for modelling extension is proposed. The idea is to locate the dynamic equations with large incremental covariances, and then to extend the state space with extra equations to enable a random walk variation of some of the parameters in the original equation.

However, occasionally, sufficient knowledge about the system so that grey box model can be developed does not exist. Some black box modelling approaches can be used as a first step in the modelling development. In Paper [A] conditional parameter modelling approach is used for flow prediction in a sewage system. The model performs well and significantly better than the traditional linear black box models. The modelling approach can be used for a further analysis of the system which. This might provide an understanding which can then be used for formulating a grey box model.

The direct arguments for including the noise part of the system equation are the following:

- The ODE or drift part of the model contains only an approximation of the true system.
- The measurements of the input to the system are encumbered with measurement noise.
- Unrecognized inputs. Some variables, not directly considered in the model, might affect the time evolution of the states.

Above it has been argued that the SDE based modelling approach contains many advantages compared to the traditionally ODE approach with observation noise.

For statistical modelling discrete time stochastic state space models are often

considered. Compared to this modelling class the continuous time based modelling offered by the SDE approach offers several advantages:

- Non-linearity and non-stationary systems are easily modelled.
- The model in continuous time contains less parameters than equivalent models in discrete time.
- Prior physical knowledge can be directly incorporated into the model.
- The models parameters are directly physically interpretable.
- Non-equidistant data can be used directly for estimation – hence missing data can be dealt with by using the SDE based approach directly.
- The framework enables a direct collaboration between the physical expert and the statistician.

The fact that the direct formulation of the dynamics in continuous time creates a background for a direct collaboration between the expert and the statistician, is most likely one of the most important aspects of the grey box modelling approach. Traditionally, the statistical models (ARMA, Box-Jenkins, GLM, etc.) are rather easy to deal with for the statistician, but the physical expert most often are not able to interpret the parameters and the results. On the other hand, for white box models, the physical expert is able to formulate a model describing the dynamics, but the statistician, in general, is not able to estimate the all the model parameters. Thus, the grey box approach based on stochastic differential equations bridge the modelling gap between the statistician and the physical expert.

APPENDIX A

Conditional parametric models for storm sewer runoff

Accepted in *Water Resources Research*.

WATER RESOURCES RESEARCH, VOL. ???, XXXX, DOI:10.1029/

Conditional parametric models for storm sewer runoff

H. Jonsdottir,^{1,2} H. Aa Nielsen,¹ H. Madsen,¹ J. Eliasson,² O. P. Palsson²
and M. K. Nielsen³

Abstract. The method of conditional parametric modeling is introduced for flow prediction in a sewage system. It is a well known fact that in hydrological modeling the response (runoff) to input (precipitation) varies depending on soil moisture and several other factors. Consequently, nonlinear input-output models are needed. The model formulation described in this paper is similar to the traditional linear models like FIR (Final Impulse Response) and ARX (Auto Regressive eXogenous) except that the parameters vary as a function of some external variables. The parameter variation is modeled by local lines, using kernels for local linear regression. As such, the method might be referred to as a nearest neighbor method. The results achieved in this study were compared to results from the conventional linear methods, FIR and ARX. The increase in the coefficient of determination is substantial. Furthermore, the new approach conserves the mass balance better. Hence, this new approach looks promising for various hydrological models and analysis.

1. Introduction

Hydrology is one of the oldest fields of interest in science and has been studied on both small and large scales for about 6000 years. The goal of the present work is to achieve good predictions of flow in a sewage system. Black box models have been providing good prediction results, often much better than conceptual or physical models, depending on how well the actual system is known. *Carstensen et al.* [1998] showed that data driven models are more reliable for on-line applications in sewers than stationary deterministic models.

Black box models have been used in hydrology for decades; *Sherman* [1932] presented the first black box model by introducing the theory of unit hydrograph. The unit hydrograph is an impulse response function and as such is estimated directly as a FIR model, i.e. the flow is modeled as lagged values of precipitation. The unit hydrograph describes the relation between effective precipitation and quick flow. Hence, for the flow data, a base flow separation must be performed and the effective precipitation must be calculated from the precipitation data. Quite often physical equations are used for effective precipitation calculations, e.g. Horton's infiltration formula, *Horton* [1935] or Philip's equation, *Philip* [1969]. Effective rain identification can also be incorporated in the hydrograph modeling process itself, e.g. *Hsu et al.* [2002].

For the purpose of flow predictions, ARX and ARMAX (Auto Regressive Moving Average eXogenous) models are in most cases more successful than FIR models. This means that the flow is modeled not only as a function of precipitation, but also by using past flow values and in that case all the available information is applied. *Todini* [1978] used an

ARMAX model for on-line flow predictions and *Novotny and Zheng* [1990] used an ARMAX model for deriving watershed response function and their paper provides an overview of how ARMAX models, transfer functions, Green's functions and the Muskingum routing method are related.

Both the FIR models and the ARMAX models are linear time-invariant models. These models are simple and easy to use and in many cases provide acceptable results, particularly when the volume of the flood is large compared to the infiltrated volume. Nevertheless, the rainfall runoff process is believed to be highly nonlinear, time-varying and spatially distributed, e.g. *Singh* [1964], *Chiu and Huang* [1970], or *Pilgrim* [1976]. With increased computer power, nonlinear models have become increasingly popular. *Capkun et al.* [2001] handle the nonlinearity by using an ARX model and by modeling the variance as a function of past rainfall. Bayesian methods have also been applied; *Campbell et al.* [1999] used such a procedure for parameter estimation in their nonlinear flood event model. *Iorgulescu and Beven* [2004] used nonparametric techniques for the identification of rainfall-runoff relationship using direct mapping from the input space to the output space with good results. During the last decade neural networks have been popular as in *Hsu et al.* [1995], *Shamseldin* [1997] and more recently the SOLO-ANN model by *Hsu et al.* [2002]. *Karlson and Yakowitz* [1987] used a nonparametric regression method, which they refer to as the nearest neighbor method. They compare FIR, ARMAX and nearest neighbor models. Their results favor the nearest neighbor and the ARMAX models; however, they do not distinguish between the ARMAX and the nearest neighbor models. *Porporato and Ridolfi* [1996] used a nearest neighbor model and found that the local linear model with small neighborhoods gave the best results. In *Porporato and Ridolfi* [1997] strong nonlinear deterministic components were detected in the discharge series. They used noise reduction techniques specifically proposed for the field of chaos theory to preserve the delicate nonlinear interactions, and then used nonlinear prediction (NLP) with good results. In *Porporato and Ridolfi* [2001] these methodologies are followed up for multivariate systems. *Previdi and Lovera* [2004] tackle the nonlinearity by using time-varying ARX models, which they refer to as Non-Linear Parameter Varying Models (NLPV). The parameter variation is defined as an output of a non-linear function and the optimization is performed by using Neural Networks. In *Young et al.* [2001]

¹Department of Informatics and Mathematical Modelling, Bldg. 321 DTU, DK-2800 Lyngby, Denmark.

²Faculty of Engineering, University of Iceland, Hjardarhaga 2-6, 107 Reykjavik, Iceland.

³Waste Water Control aps - WWC, Kollemosevej 47, DK-2830, Virum, Denmark

the time variable parameters are considered to be state dependent and the method is thus referred to as the SDP approach. For non-linear phenomena this approach results in a two stage approach, called the DBM approach (Data Based Mechanistic approach) Young [2002]. In recent years fuzzy methods have been tested for flood forecasting, e.g. Chang *et al.* [2005] and Nayak *et al.* [2005].

In the present paper conditional parametric models are used to develop models for flow predictions in a sewage system. A conditional parametric model is a linear regression model where the parameters vary as a smooth function of some explanatory variable. Thus the method presented here are in a line with the SDP and the NLPV methodologies. The name conditional parametric model originates from the fact that if the argument of the functions is fixed then the model is an ordinary linear model, Hastie and Tibshirani [1993] and Anderson *et al.* [1994]. In the models presented here, the parameters vary locally as polynomials of external variables, as described in Nielsen *et al.* [1997]. In contrast to linear methods like FIR and ARX, this methodology allows fixed input to provide different output depending on external circumstances.

This paper is organized as follows: In Section 2 the models are described, followed by Section 3 with a description of the parameter estimation method. Section 4 contains results and in Section 5 there are discussions about on-line prediction and control in sewage systems. Finally, in Section 6 conclusions are drawn.

2. The Models

In the present paper the excess outflow is modeled as a function of total precipitation (the base flow in the sewage system does not originate in rainfall). To avoid the calculation of infiltration it was decided to use the total precipitation as measured on-line. This is very convenient, particularly since the infiltration rate depends on several physical factors and no perfectly quantified general formula exists, Viessman and Lewis [1996]. Some of the more recently developed models identify the effective precipitation along with the hydrograph e.g. Nalbantis *et al.* [1995]. The goal is to predict flow in the sewage system as a function of measured precipitation; consequently division of the precipitation into effective rain and infiltration/evaporation is not important. For the purpose of flow prediction, conditional parametric models are applied, Nielsen *et al.* [1997]. These models are an extension of the well known linear regression model where the parameters vary as functions of some external variable. In this research two types of models were tested: conditional parametric FIR models and conditional parametric ARX models. The models are formulated as:

$$\text{FIR: } y_t = \sum_{i=0}^{q_1} h_i(\mathbf{x}_{t-m}) z_{t-i} + e_t \quad e_t \in N(0, \sigma_{\text{FIR}}^2) \quad (1)$$

$$\text{ARX: } y_t = \sum_{i=1}^p a_i(\mathbf{x}_{t-m}) y_{t-i} + \sum_{i=0}^{q_2} b_i(\mathbf{x}_{t-m}) z_{t-i} + e_t \quad e_t \in N(0, \sigma_{\text{ARX}}^2) \quad (2)$$

where y_t is the output (flow), z_t is the input (precipitation), \mathbf{x}_t is the explanatory variable and m is the time delay if any. Here, the explanatory variable is season and/or threshold, see Section 4. The order of the FIR model in Eq. (1) is denoted q_1 and the order of the ARX model in Eq. (2) is denoted (p, q_2) .

In the FIR model the function \mathbf{h} , represented by the coefficients $h_i(\mathbf{x}_{t-m})$ $i = 1, \dots, q_1$ is known as the impulse response function. It demonstrates how the system responds to the input. In the ARX model $A_{\mathbf{x}_{t-m}}(q^{-1})$ is defined as

the p -th order polynomial operator

$$A_{\mathbf{x}_{t-m}}(q^{-1}) = a_1(\mathbf{x}_{t-m})q^{-1} + \dots + a_p(\mathbf{x}_{t-m})q^{-p} \quad (3)$$

where q^{-1} is the backward shift operator. Similarly $B_{\mathbf{x}_{t-m}}(q^{-1})$ is defined as:

$$B_{\mathbf{x}_{t-m}}(q^{-1}) = b_0(\mathbf{x}_{t-m}) + b_1(\mathbf{x}_{t-m})q^{-1} + \dots + b_{q_2}(\mathbf{x}_{t-m})q^{-q_2} \quad (4)$$

as the q_2 -th order polynomial operator. Then for a fixed \mathbf{x}_{t-m} the impulse response function can be derived from the transfer function $A_{\mathbf{x}_{t-m}}(q^{-1})/B_{\mathbf{x}_{t-m}}(q^{-1})$ as described for example by Ljung [1987]. In this case the impulse response function includes coefficients h_0, h_1, \dots up to infinity. However, in practice, only the first n coefficients are used.

Ashan and O'Connor [1994] define the gain factor G of a unit hydrograph as

$$G = \frac{1}{A} \sum_0^n h_i \quad (5)$$

where A is the area of the watershed, n is the order of the model, the coefficients h_i are the coefficients in a unit hydrograph, where the input is effective rain and the output is excess flow. In an ideal situation the gain factor is one, but Høybye and Rosbjerg [1999] state that such a linear relationship does not exist. Furthermore, Ashan and O'Connor [1994] state that the overall model efficiency is in general very sensitive to the magnitude of the gain factor. In this paper the input is total precipitation; however, the value in Eq. (5), will be referred to as the gain factor. The gain factor will not be one. However, the gain provides valuable information about the system, it provides the fraction of the total precipitation that becomes excess rainfall and thus also the fraction that infiltrates into the ground. Furthermore, the change in the gain factor as the external variable \mathbf{x} changes provides a valuable information for understanding the system.

3. The Estimation Method

The models used are locally linear regression. In order to describe the ARX and FIR models together the notation is changed to the notation of a linear regression

$$y_t = \mathbf{z}_t^T \theta(\mathbf{x}_t) + e_t; \quad t = 1, \dots, N, \quad e_t \in N(0, \sigma^2) \quad (6)$$

where the output or the response, y_t is a stochastic variable; $\mathbf{z}_t \in \mathcal{R}^k$ is the input; $\mathbf{x}_t \in \mathcal{R}^r$ is an explanatory variable. The parameter vector $\theta(\cdot) \in \mathcal{R}^k$ is a vector of smooth functions of \mathbf{x}_t , and $t = 1, \dots, N$ are observation numbers. In the case of a FIR model, the variable \mathbf{z}_t is the lagged values of the precipitation and $\theta(\mathbf{x}_t)$ are the coefficients in a hydrograph. In the case of an ARX model, the \mathbf{z}_t consist of the lagged values of precipitation and the lagged values of flow, where $\theta(\mathbf{x}_t)$ consists of the corresponding parameters. If \mathbf{x}_t is constant across all the observations, the model reduces to a traditional linear regression model, hence the name. The estimation of $\theta(\cdot)$ is accomplished by estimating the functions at a number of distinct values of \mathbf{x} . Given a point \mathbf{x} , each θ_j , $j = 1, \dots, k$ is approximated by a local linear function

$$\theta_j(\mathbf{x}_t) = \theta_{j0} + \theta_{j1}^T \mathbf{x}_t \quad j = 1, \dots, k \quad (7)$$

The coefficients θ_{j0} and θ_{j1}^T are estimated by using weighted least squares (by using kernels).

If \mathbf{x}_t is 2-dimensional $\theta_j(\mathbf{x}_t)$ can be written as

$$\theta_j(\mathbf{x}_t) = \theta_{j0} + \theta_{j1}x_{1t} + \theta_{j2}x_{2t} \quad j = 1, \dots, k \quad (8)$$

hence

$$y_t = z_{1t}\theta_{10} + z_{1t}\theta_{11}x_{1t} + z_{1t}\theta_{12}x_{2t} + \dots + z_{kt}\theta_{k0} + z_{kt}\theta_{k1}x_{1t} + z_{kt}\theta_{k2}x_{2t} + e_t \quad (9)$$

Then a row in a new design matrix can be defined as

$$\mathbf{u}_t^T = [z_{1t}, z_{1t}x_{1t}, z_{1t}x_{2t}, \dots, z_{jt}, z_{jt}x_{1t}, z_{jt}x_{2t}, \dots, z_{kt}, z_{kt}x_{1t}, z_{kt}x_{2t}] \quad (10)$$

and by defining the column vector

$$\theta_{j\mathbf{x}} = [\theta_{j0}, \theta_{j1}, \theta_{j2}] \quad (11)$$

and

$$\theta_{\mathbf{x}} = [\theta_{1\mathbf{x}}^T, \dots, \theta_{j\mathbf{x}}^T, \dots, \theta_{k\mathbf{x}}^T]^T \quad (12)$$

the flow vector \mathbf{y}_t can be written as

$$y_t = \mathbf{u}_t^T \theta_{\mathbf{x}} + e_t \quad t = 1, \dots, N, \quad (13)$$

The parameter vector $\theta_{\mathbf{x}}$ is fitted locally to \mathbf{x} . This is accomplished by using the traditional weighted least squares, where the weight on observation t is related to the distance from \mathbf{x} to \mathbf{x}_t , so that

$$w_t(\mathbf{x}) = W(\|\mathbf{x}_t - \mathbf{x}\|/d(\mathbf{x})), \quad (14)$$

where $\|\mathbf{x}_t - \mathbf{x}\|$ is the Euclidean distance between \mathbf{x}_t and \mathbf{x} . The function $W: \mathcal{R} \rightarrow \mathcal{R}$ is a nowhere increasing function. In this paper the tricube function

$$W(v) = \begin{cases} (1 - v^3)^3, & v \in [0; 1] \\ 0, & v \in [1; \infty) \end{cases} \quad (15)$$

is used. The scalar $d(\mathbf{x}) > 0$ is called the bandwidth. If $d(\mathbf{x})$ is constant for all values of \mathbf{x} , it is denoted a fixed bandwidth. On the other hand, if $d(\mathbf{x})$ is chosen so that a certain fraction of the observations is within the bandwidth, it is denoted as nearest neighbor bandwidth. The advantage of the tri-cube weighting function is that it is a smooth function like the Gauss bell, but unlike the Gauss bell the tri-cube function is zero outside the bandwidth, which makes the computational effort smaller. The choice of weighting function or kernel does not have a large impact, see *Silverman* [1986].

In general, if \mathbf{x} has a dimension of two or larger, scaling of the individual elements of \mathbf{x} before applying the method should be considered, e.g. *Cleveland and Develin* [1988]. A rotation of the coordinate system, in which \mathbf{x} is measured, could also be relevant. When the local estimate in Eq. (13) $\hat{\theta}_{\mathbf{x}}$ is obtained, the elements of $\hat{\theta}(\mathbf{x})$ in Eq. (6) are calculated as

$$\hat{\theta}_j(\mathbf{x}_t) = [1, x_{1t}, x_{2t}] \hat{\theta}_{j\mathbf{x}} \quad (j = 1, \dots, k). \quad (16)$$

When $\mathbf{z}_j = 1$ for all j this method is almost identical to the method introduced by *Cleveland and Develin* [1988]. Furthermore, if $\theta(\cdot)$ is a local constant, then the method of estimation reduces to determining the scalar $\theta_j(\mathbf{x})$ so that $\sum_{t=1}^n w_t(\mathbf{x})(y_t - \theta(\mathbf{x}))^2$ is minimized, i.e. the method reduces to traditional kernel estimation, see also *Härdle* [1990] or *Hastie and Loader* [1993].

Furthermore, it is worth mentioning that, as for traditional linear regression, the fitted values \hat{y}_i , $i = 1, \dots, N$ are linear combinations of the observations, see *Nielsen et al.* [1997].

As noted earlier, the method of conditional parametric modeling has certain similarities to the SDP method, e.g. *Young et al.* [2001], as well as the NLPV as used in *Previdi and Lovera* [2004]. All these models are time varying AR-MAX type of models. In the NLPV approach, the nonlinear optimization is by use of neural network, which is completely black box oriented. The nonlinear SDP methodology results in the two stage DBM approach *Young* [2002]. In the first stage an appropriate model structure is identified by considering a class of linear transfer function models whose parameters are allowed to vary over time. In the second stage any identified (significant) parameter variation is modeled using a parametric approach, and the parameters of the resulting parametric non-linear model are estimated using non-linear least squares or maximum likelihood estimation. Hence, the resulting model is a non-linear model with fixed parameters. *Young* [2005] provides a fine overview and comparison of the SDP and NLPV modeling approaches.

The suggested method is typically a one-stage approach. The resulting model is a conditional parametric model, where the total parametrization is a combined parametric and non-parametric model. Consequently, in every neighborhood, there exists an approximately linear parametric model. Furthermore, by studying the values of the parameter $\theta(\mathbf{x})$ as the external variable \mathbf{x} changes, a complete parameterized model might be developed if that is desired. A complete parameterized model has both advantages and disadvantages: It often provides a better "physical" understanding of the system. However, the parameters are under all circumstances estimated by use of existing data and if the external circumstances change, this involves extrapolation. Local estimates (estimates in a neighborhood, kernel estimation) will adapt to new circumstances quickly. In this paper the non-linearity is described directly without any use of recursive/adaptive estimation. In the case of time-varying models, adaptive estimation, as described in *Nielsen et al.* [2000], can be superimposed on the method. Hence, the approach makes it possible to track time variation in a non-linear model, this extension is however, not the focus of the present paper.

4. Results

4.1. Description of the Data and the Circumstances

The data originate from the company Waste Water Control aps. in Denmark and consist of 68 rain events which occurred in the period 1st January 2003 to 4th May 2004. The 68 rain events cover many types of rain, of a varying intensity and length. The data consist of pairs of measured precipitation [mm/(6 min)] and excess flow [m³/(6 min)]. The sampling time is, as indicated, 6 minutes. The excess flow is calculated from the total flow by subtracting the system's base flow. The base flow, or the dry weather flow, in the sewage system does not originate from rain and is defined as a constant plus a daily variation, see *Carstensen et al.* [1998]. The sewage system is built up in the traditional manner as a net, with pumping stations located at some of the node points. During heavy rain events the volume of water entering the pipes can exceed the pumping station's capacity causing some a kind of saturation/threshold in the system.

The area of the watershed is 10.89 km². The impermeable area is dominated by urban area and the soil is mostly clay. The data contained 3 heavy rain events where the threshold/saturation phenomena can be seen. A water balance study was performed which showed a yearly variation in the water balance. This seasonality is mostly due to the soil

moisture content in the root zone and variation in ground-water level; the soil is much drier during the summer than during the winter.

4.2. Model Construction

An analysis of the data using linear models with non-varying coefficients showed that the time delay from input to output is 2 lags, i.e. 12 minutes. It has been found that at most 19 lags are needed in a FIR model, as in Eq. (1). In the ARX models, as in Eq. (2), the "best" linear model, using AIC criteria, is $ARX(2,6)$ with a time delay of 2, i.e. the output y_t is a function of $y_{t-1}, y_{t-2}, z_{t-2}, \dots, z_{t-7}$. For the sake of convenience these model degrees were used in the whole study, i.e. the same number of lags was used for in the conditional parametric models.

The numerator in the ARX models is quite high compared to what is often seen in hydrology. However, most rainfall runoff studies are on a daily basis. In this project the sampling time is 6 minutes, consequently the numerator needs to be higher. The linear model order was chosen by use of AIC/BIC criteria (the AIC and BIC indicated the same model order) and use of some other criteria might have led to lower orders. However, *Porporato and Ridolfi* [1996] indicate that the degree of the numerator should not be less than the basin concentration time, and in order to capture the entire subsequent runoff the numerator should be even greater. The basin concentration time in the sewage system, using the 6 minute sampling time is about 6 lags (the basin lag is estimated to be 4 lags and, referring to *Singh* [1988], the time of concentration is 1.42 times the basin lag time, which is close to 6 lags). The order of the numerator in the ARX model is 6. Evidently the FIR model has a larger model degree than the ARX model.

As mentioned earlier the non-linear effects are mostly due to seasonal variations and the saturation/threshold effects in the pipes. The seasonal variation is modeled as the first term in a Fourier series, i.e. a sinus wave

$$x_t^s = C \sin(\omega t + \phi) \quad (17)$$

where x_t^s is the explanatory variable due to season. The water balance study showed the largest response to precipitation in February and the smallest in August. Consequently the parameters ω and ϕ are chosen such that x_t^s peaks in mid February. The parameter C is set to 100 which is a necessary scaling in the 2-dimensional model presented later in this section. In practise the seasonal variation is not as regular as a sinus wave. However, since only 16 months of data are available it is not possible to estimate a seasonality function without restrictions as in Eq. (17). The seasonality in the parameters can most likely be modeled globally as in a PARMA model e.g. *Rasmussen et al.* [1996].

The saturation/threshold effect is modeled either as a function of the rain-intensity or as a function of the flow, depending on the model type. In a FIR model the conditional variable representing the saturation/threshold is precipitation intensity, x_t^p and set as

$$x_t^p = (u_{t-2} + u_{t-3} + u_{t-4} + u_{t-5} + u_{t-6})/5 \quad (18)$$

i.e. the average rain intensity in lags 2 to 6. This choice is based on the facts that the time delay is 2 lags, and the time of concentration is 6 lags.

In the ARX models the saturation/threshold is modeled as a function of the flow itself instead of the rain intensity. This is in fact more physically correct because the threshold occurs because there is more water in the pipes than the pumps in the node points can serve, even though all this water is caused by heavy rain. Hence, the explanatory variable is defined as

$$x_t^f = y_{t-1} \quad (19)$$

There were only 3 heavy rain events during this period and since 19 coefficients need to be identified in the FIR model, the 3 events with heavy rainfall were not quite enough to identify the 19 coefficients within an acceptable confidence level, meaning that several combinations of solutions might be possible. However, some solutions were found and those were used for prediction. As a consequence of this sparse data it was not possible to identify a FIR model where the coefficients varied both with the season and the threshold. On the other hand in the ARX models the constants are rather well identified and it was possible to identify coefficients depending on two variables, season and flow.

The local estimation requires bandwidth decisions. The bandwidth determines the smoothness of the estimate. If the bandwidth is small the variance is large and the bias is small. If the bandwidth is large the variance is small but the bias increases. An "optimal" bandwidth is a bandwidth which is a compromise of these two factors. In the traditional kernel estimation, as in *Härdle* [1990], the estimates are local constant; here the estimates are local lines. This allows a larger bandwidth without the cost of a bias problem. The bandwidths are different, depending on the model types. In each case the bandwidths were found by manual optimization, the bandwidth needs to be small enough to detect differences in the conditional variable. However, the larger it is, the less variation in the estimates. For example in the ARX model where the conditional variable is the season, it was possible to use a large bandwidth. The seasonal variable is almost evenly distributed, and the optimal bandwidth included 2400 data points, which is about 65% of the data. On the other hand in the FIR model with rain intensity as a conditional variable, the bandwidth included only 55 data points, which is about 1.5% of the data. This is because there are few events with heavy rain and thus, by using a larger bandwidth the few data-points no longer have an effect.

The calculations are performed by using a program named LFLM (locally weighted fitting of linear models) which is an S-PLUS/R library package. For a description see *Nielsen* [1997].

4.3. Modeling Results

Model validation demands some measure of the model's quality. This measure is not a single number which can be used for each and every model and in each and every situation. In this project the main goal was to achieve accurate predictions. Thus, the optimization (model calibration) is a least squares method and as such the model's performance is validated with respect to that. In hydrology several other factors might be of higher importance, like the overall water balance, the timing of the peak flow or other things.

The coefficient of determination R^2 , often referred to as the Nash efficiency, is a widely used model criterion in hydrology and it is a fine measure of the model's efficiency with respect to the least squares minimization. The residuals used for model validation are one step prediction errors, using the calibration data series. A cross validation would have been adequate. However, the model's parameters were estimated locally and will thus adapt to the data in use; hence, cross validation does not have the same meaning as when the parameters are estimated globally. Table 1 shows the R^2 for the conditional parametric models. As the mass balance is an important dimension in hydrology, the mean value of the error was also calculated. The mean value of the error demonstrates the mass balance on average. If the mass balance is well conserved, the mean value of the error will

be zero. Table 2 shows the mean value of the error. The tables both show the overall performance and also the seasonal performance. For each season the seasonal calculations are based two months in each season for better distinction between the seasons. For the seasonal calculations the 3 most heavy rain events were excluded. These were events, where the saturation/threshold effect exists, those were grouped together and the calculations were performed for them separately in order to measure how the models perform in that situation. As a reference R^2 was also calculated for the corresponding well known linear models FIR and ARX. It is well known that the coefficient of determination does not penalize over-parameterizations. However, both AIC and BIC studies led to this model order as does a physical study like the basin concentration time as discussed in Section 4.2. Specifically the conditional parameter models are compared to a linear models of same orders. Thus the FIR models are comparable, and the ARX models are as well. The estimation was based on all the events, and the tables show total results for all the 68 rain events.

Table 1. The coefficient of determination R^2 for various models and different conditions. The unit is $m^3/6 min$ as the sampling time is 6 minutes. The models are: Linear FIR model, Seasonal FIR model and Threshold/Saturation FIR model. Linear ARX model, seasonal ARX model, Threshold/Saturation ARX model and both seasonal and threshold/saturation ARX model. The R^2 calculations are 1-step prediction, performed using overall data. For winter, spring, summer and fall, only 2 months were used for better seasonal distinction. Finally, the R^2 is calculated for the three heaviest rain events, Those events are excluded in the seasonal calculations.

	Cond.	All	W 8 ev.	Sp. 13 ev.	Su. 13 ev.	F 12 ev.	Heavy rain 3 ev.
FIR	Lin.	0.79	0.64	0.73	0.68	0.90	0.83
	Seas.	0.84	0.70	0.78	0.85	0.92	0.89
	Thr.	0.82	0.63	0.75	0.79	0.93	0.93
ARX	Lin.	0.94	0.91	0.94	0.91	0.96	0.93
	Seas.	0.95	0.92	0.94	0.93	0.96	0.94
	Thr.	0.96	0.93	0.95	0.94	0.97	0.97
	S×Thr.	0.97	0.95	0.96	0.97	0.98	0.99

In a comparison of the three FIR models, the seasonal FIR model has the best performance both with respect to the R^2 and to the mass balance. The seasonal FIR is clearly an improvement of the traditional FIR model with constant parameters. Even in a situation with heavy rain, the seasonal FIR outperforms the traditional FIR. The threshold FIR is the best during heavy rain events, as a result of its design, however the bias is quite large.

All the conditional parametric ARX models outperform the traditional ARX, both the 1-dimensional models and the 2-dimensional model, which is the best both with respect to the Nash-efficiency and the bias.

In Table 2 it can be seen how the traditional linear models underestimate the runoff during the winter and overestimate it during the summer, especially the FIR models. The FIR models have a larger bias, and even the seasonal FIR is not quite acceptable in all seasons, especially in the

spring season. Thus, the FIR models are not acceptable for predictions.

Table 2. The mean value of the error, $m^3/6 min$. (For nomenclature see Table 1).

	Cond.	All	W 8 ev.	Sp. 13 ev.	Su. 13 ev.	F 12 ev.	Heavy rain 3 ev.
FIR	Lin.	1.13	11.18	19.54	-22.80	-2.4	-37.22
	Seas.	0.23	0.76	15.79	-1.14	-6.8	-10.23
	Thre.	3.40	14.13	20.47	-17.12	-2.3	-17.50
ARX	Lin.	0.05	1.67	2.79	-3.84	-0.48	-7.58
	Seas.	-0.20	0.01	2.72	-0.27	-1.45	-2.58
	Thre.	0.12	1.77	3.3	2.85	-1.29	-4.74
	S×Thr.	-0.01	0.02	2.37	0.19	1.48	-1.76

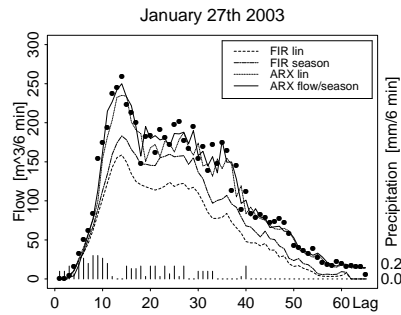


Figure 1. Event in winter. The data is shown as points and the precipitation as bars, with the scale on the right axis. The time lag is 6 minutes and the figure shows a conditional ARX model, where the parameters depend on season and flow, and a conditional FIR model where the parameters depend on season. For comparison the conventional time-invariant linear models FIR and ARX are also shown.

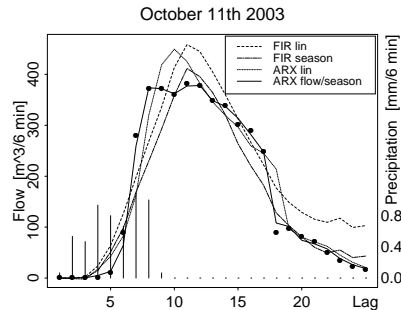


Figure 2. Event in autumn. The data is shown as points and the precipitation as bars, with the scale on the right axis. The time lag is 6 minutes and the figure shows a conditional ARX model, where the parameters depend on season and flow, and a conditional FIR model where the parameters depend on season. For comparison the conventional time-invariant linear models FIR and ARX are also shown.

As expected, the ARX models outperform the FIR models. However, it must be stressed that the FIR models and the ARX models need different inputs for prediction. Using a FIR model, 1-step prediction demands past precipitation which is also required when using an ARX model, but past values of the flow are additionally required. For k -step prediction both the FIR and the ARX models need past and present values of the precipitation and also $(k-2)$ prediction of the precipitation. Additionally the ARX models demand past, present and $(k-1)$ step prediction of the flow. It must be mentioned though that using predicted values of precipitation as input will never be quite as reliable as using measured values since the predicted values have much larger variance; this is also true for the predicted values of the flow. Moreover, the parameter estimates are performed with the assumption that the input is measured not predicted.

For visual comparison two events were chosen. These are a 'typical' event in winter time and an event with heavy rain, showing the saturation/threshold. Note that even though single events are shown in the figures, the estimate is based on all the events. A single conditional ARX model and a single conditional FIR model will be drawn along with the traditional linear models. The figures show the best conditional FIR model, the 1-dimensional seasonal FIR and the best conditional ARX model, the 2-dimensional ARX along with the linear FIR and ARX, for comparison.

Figure 1 shows an event in the winter time; the duration of the event is about 6.5 hours. Note that the linear FIR model underestimates the runoff as demonstrated in Table 2, and a seasonal FIR model is clearly an improvement on the traditional linear FIR. The linear ARX model is better, but not as good as the conditional parametric ARX. Figure 2 shows the same for an event with heavy rain and thus the threshold/saturation effect; the duration of this event is about 2.5 hours. In this case both the linear FIR and the linear ARX overestimate the flow peak, as does a seasonal FIR, while the conditional ARX nicely captures the flat and long peak.

It might be argued that in real applications a confidence interval for the predictions would be required. This is indeed true; confidence intervals for the predicted output are valuable. However, since the model is non-linear, it is believed that prediction intervals should be estimated by methods like quintile regression as in *Nielsen et al.* [2006]. This is not covered in this paper.

Finally, conditional parametric models with local estimates can also be used to study the circumstances of the watershed and thus provide a useful information for developing a non-linear global parametric model if wanted.

For example, the seasonality can be studied. For this purpose a seasonal ARX model is used. In this study the conditional variable representing the seasonality is a sinus wave and the parameters are estimated as local lines, depending on the values of the sinus. Table 3 shows estimated coefficients in the ARX for fixed values of the season. Due to symmetry, it is not possible to distinguish between spring and autumn. A comparison of the autoregressive parameters a_1 and a_2 , shows that a_2 is larger than a_1 during the winter while a_2 is close to zero during the summer and a_1 is the dominating autoregressive parameter. The negative value of the parameter b_5 in August is physically incorrect, and this is probably due to sparse data, since there is only one summer season and August is close to and on the boundary of the seasonal variation parameter.

Table 3. Local parameter estimates in a seasonal ARX model. S_1 =February, S_2 =April/December, S_3 =June/October, S_4 =August.

Seas.	a_1	a_2	b_2	b_3	b_4	b_5	b_6	b_7
S_1	0.34	0.51	33.45	42.75	68.56	21.98	19.13	5.48
S_2	0.62	0.20	32.66	50.67	48.33	23.57	21.12	5.23
S_3	0.82	-0.03	23.80	52.73	22.09	21.88	11.95	18.04
S_4	0.73	0.08	20.36	70.23	21.59	-13.30	14.13	19.44

Using the estimated parameters, the impulse response function can be calculated, as shown in Figure 3, which also shows the sum of the coefficients and the corresponding gain factor, calculated by Eq. (5). Note that the impulse response function has the longest tail during the winter and shortest tail during the summer, and it also reaches larger values during the winter than the summer. Consequently, during the winter about 10% of the total water reaches the sewage system, while during the summer about 6% enters the sewage system. A similar analysis has been performed for the flow dependence of the impulse response function and it mostly shows that when the flow is large the impulse response function is flatter, it peaks later, the values are smaller, and the tail is longer.

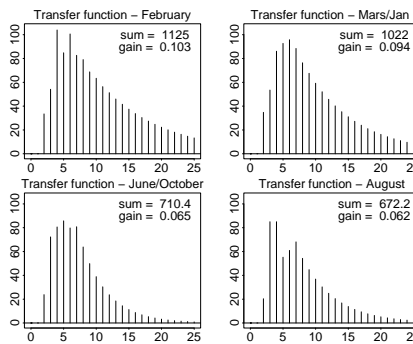


Figure 3. Impulse response function estimates for four different seasons calculated using a seasonal ARX model.

5. Discussion

The FIR models provide 2-step prediction i.e. information 12 minutes ahead, since the time delay between precipitation and flow is 2 lags. The ARX models provide 1-step prediction, since flow at time $t-1$ is used for prediction. For real time on-line prediction and automatic control it might be necessary to achieve information with longer a time horizon, say 30 minutes i.e. 5-step prediction. For both the FIR and the ARX models a 5-step prediction requires 3-step prediction of precipitation, i.e. on-line weather forecast. However, since it is only a question of a couple of minutes, it might be possible to use on-line precipitation measurements a bit further from the treatment plant, i.e. a weather station capturing the frontal rain a little bit earlier. Evidently this depends on the wind and frontier movement direction although in many cases the wind during rain is from the (south) west, which is the dominating wind direction. The wind direction might also be a conditional variable in the model if enough data are available. On the other hand, for

the ARX model the situation is a little bit more complicated because the flow y_{t-1} is a conditional variable. The most practical thing would be to provide on-line flow data from a couple of node points in the sewage system net, node points which are distributed geographically in the sewage system. The flow in the node points is naturally delayed compared to the flow in the waste water treatment plant, and obviously the delay is different depending on the geographical localization. However, if data from the node points are available in general, it would be most convenient to use the flow in the node points as an input in a model for on-line prediction and control, at the waste water treatment plant, and thereby remove much of the unaccountable rain distribution.

6. Summary and Conclusion

Conditional parametric models have been developed and tested for rainfall-runoff modeling in a sewage system. The models are FIR and ARX models with the coefficients varying as a function of external variables. The input of the models is the total precipitation as measured on-line, and the output is the excess flow prediction. The base flow is separated by using simple equations since the base flow in the sewage system does not originate in rainfall.

Both the conditional parametric FIR and the conditional parametric ARX provide results which are significantly superior to results from conventional linear models. As expected the ARX models provide the best 1-step predictions.

In this study the conditional variables are used to capture seasonal fluctuations and threshold/saturation due to the limited capacity of the system pumps and pipes.

Use of this modeling approach has a good potential for developing good prediction models. Furthermore, the method can also be used for sensitivity analysis while constructing a physical model of the system flow. The method of conditional modeling is a useful contribution to the tools of nonlinear modeling techniques used in hydrology.

Acknowledgments. The authors wish to thank the company Waste Water Control aps in Denmark for delivering the data and information about the system. Furthermore, the authors wish to thank the Hydrological Service at the National Energy Authority of Iceland for lending their facilities.

References

- Anderson, T. W., K. T. Fang, and I. Olkin (Eds.) (1994), *Multivariate Analysis and its Applications*, chap. Coplots, Nonparametric Regression, and Conditionally Parametric Fits, pp. 21–36, Institute of Mathematical Statistics, Hayward, Institute of Mathematical Statistics, Hayward.
- Ashan, M., and K. M. O'Connor (1994), A simple non-linear rainfall-runoff model with a variable gain factor, *Journal of Hydrology*, 155, 151–183.
- Campbell, E. P., D. R. Fox, and B. C. Bates (1999), A bayesian approach to parameter estimation and pooling in nonlinear flood event models, *Water Resources Research*, 35(1), 211–220, doi:10.1029/1998WR900043.
- Capkun, G., A. Davison, and A. Musy (2001), A robust rainfall-runoff transfer model, *Water Resources Research*, 37(12), 3207–3216, doi:10.1029/2001WR000295.
- Carstensen, J., M. K. Nielsen, and H. Strandbæk (1998), Prediction of hydraulic load for urban storm control of municipal wwtp, *Water Science and Technology*, 37(12), 363–370.
- Chang, L.-C., F.-J. Chang, and Y.-H. Tsai (2005), Fuzzy exemplar-based inference system for flood forecasting, *Water Resources Research*, 41(2), W02005, doi: 10.1029/2003WR003514.
- Chiu, C. L., and J. T. Huang (1970), Nonlinear time-varying model of rainfall runoff relation, *Water Resources Research*, 6(1), 1277–1286.
- Cleveland, W. S., and S. J. Develin (1988), Locally weighted regression: An approach to regression analysis by local fitting, *Journal of The American Statistical Association*, 83, 596–610.
- Härdle, W. (1990), *Applied Nonparametric Regression*, Syndicate of the University of Cambridge, United Kingdom.
- Hastie, T., and C. Loader (1993), Local Regression: Automatic Kernel Carpentry, *Statistical Science*, 8(2), 120–129.
- Hastie, T., and R. Tibshirani (1993), Varying-coefficient models, *Journal of the Royal Statistical Society*, 55, 757–796.
- Horton, R. E. (1935), Surface Runoff Phenomena: Part I, Analysis of the Hydrograph, *Tech. rep.*, Horton Hydrol. Lab. Pub. 101. Ann Arbor, MI: Edwards Bros.
- Høybye, J., and D. Rosbjerg (1999), Effect of input and parameter uncertainties in rainfall-runoff simulations, *Journal of hydrologic engineering*, 4, 214–224, doi:10.1061/(ASCE)1084-0699(1999)4:3(214).
- Hsu, K., H. V. Gupta, and S. Sorooshian (1995), Artificial neural network modeling of the rainfall-runoff process, *Water Resources Research*, 31(10), 2517–2530, doi: 10.1029/95WR01955.
- Hsu, K., H. V. Gupta, Z. Gao, and S. S. B. Imam (2002), Self-organizing linear output map (SOLO): An artificial neural network suitable for hydrologic modeling and analysis, *Water Resources Research*, 38(12), 1302, doi:10.1029/2001WR000795.
- Iorgulescu, I., and K. Beven (2004), Nonparametric direct mapping of rainfall-runoff relationships: An alternative approach to data analysis and modelling, *Water Resources Research*, 40(8), W08403, doi:10.1029/2004WR003094.
- Karlson, M., and S. Yakowitz (1987), Nearest-Neighbor Methods for Nonparametric Rainfall-Ruoff Forecast, *Water Resources Research*, 27(7), 1300–1308.
- Ljung, L. (1987), *System identification: Theory for the user*, Prentice-Hall, NJ.
- Nalbantis, I., C. Obled, and J. Rodriguez (1995), Unit Hydrograph and effective precipitation identification, *Journal of Hydrology*, 168, 127–157.
- Nayak, P., K. P. Sudheer, D. Rangan, and K. S. Ramasastri (2005), Short-term flood forecasting with neurofuzzy model, *Water Resources Research*, 41(4), W04004, doi: 10.1029/2004WR003562.
- Nielsen, H. A. (1997), LFLM version 1.0 - An SPLUS/R library for locally weighted fitting of linear models, *Tech. Rep. 22*, Department of Mathematical Modelling, Technical University of Denmark, available from <http://www.imm.dtu.dk/han/software.html>.
- Nielsen, H. A., T. S. Nielsen, and H. Madsen (1997), Conditional parametric ARX-models, *11th IFAC Symposium on System Identification*, 2, 475–480.
- Nielsen, H. A., T. S. Nielsen, A. K. Joensen, H. Madsen, and J. Holst (2000), Technical note - tracking time-varying-coefficient function, *International Journal of Adaptive Control and Signal Processing*, 14, 813–827, doi:10.1002/1099-1115(200012).
- Nielsen, H. A., H. Madsen, and T. S. Nielsen (2006), Using quantile regression to extend an existing wind forecasting system with probabilistic forecast, *Wind Energy*, 9, 95–108, doi: 10.1002/we.180.
- Novotny, V., and S. Zheng (1990), Rainfall-runoff transfer function by arma modeling, *Journal of Hydraulic Engineering*, 115, 1386–1400.
- Philip, J. (1969), *Theory of infiltration*, Division of Plant Industry Commonwealth Scientific and Industrial Research Organization, Canberra, Australia.
- Pilgrim, D. H. (1976), Travel times and nonlinearity of flood runoff from tracer measurements on a small watershed, *Water Resources Research*, 31, 2517–2530.
- Porporato, A., and L. Ridolfi (1996), Clues to the existence of deterministic chaos in river flow, *International Journal of Mod. Physics B*, 10(15), 1821–1862.
- Porporato, A., and L. Ridolfi (1997), Nonlinear analysis of river flow time sequences, *Water Resources Research*, 33(6), 1353–1367.
- Porporato, A., and L. Ridolfi (2001), Multivariate nonlinear prediction of river flows, *Journal of Hydrology*, 248, 109–122.
- Prevdi, F., and M. Lovera (2004), Identification of non-linear parametrically varying models using separable least squares., *International Journal of Control*, 77(16), 1382–1392, doi: 10.1080/0020717041233318863.

- Rasmussen, P. F., J. D. Salas, L. Fagherazzi, J.-C. Rassam, and B. Bobe (1996), Estimation and validation of contemporaneous PARMA models for streamflow simulation, *Water Resources Research*, 32(10), 3151–3160, doi:10.1029/96WR01528.
- Shamseldin, A. Y. (1997), Application of a neural network technique to rainfall-runoff modelling, *Journal of Hydrology*, 199, 272–294.
- Sherman, L. K. (1932), Streamflow from rainfall by the unit-graph method, *Eng. News Rec.*, 108, 501–505.
- Silverman, B. (1986), *Density estimation for statistics and data analysis*, Chapman and Hall, New York.
- Singh, V. P. (1964), Nonlinear instantaneous unit hydrograph theory, *Journal of Hydraulic Div. Am. Soc. Civ. Eng.*, 90(HY2), 313–347.
- Singh, V. P. (1988), *Hydrologic Systems, Vol 1, Rainfall-runoff Modelling*, Prentice Hall.
- Todini, E. (1978), Using a Desk-Top Computer for an On-Line Flood Warning System, *IBM Journal of Research and Development*, 22, 464–471.
- Viessman, W., and G. L. Lewis (1996), *Introduction to hydrology*, HarperCollins College Publishers, New York.
- Young, P. C. (2002), Data-based mechanistic and top-down modelling, *Proceedings of the International Environmental Modelling and Software Society Conference (iEMS 2002)*, Lugano, 24–27 Jun, 2002, pp. 363–374.
- Young, P. C. (2005), Comments on identification of non-linear parametrically varying models using separable least squares by F. Previdi and M. Lovera: Black box or open box, *International Journal of Control*, 78(2), 122–127, doi:10.1080/002071705000073772.
- Young, P. C., P. McKenna, and J. Bruun (2001), Identification of non-linear stochastic systems by state dependent parameter estimation, *International Journal of Control*, 74(18), 1837–1857, doi:10.1080/00207170110089824.

Harpa Jonsdottir, Faculty of Engineering,
University of Iceland, Hjardarhaga 2-6,
107 Reykjavik, Iceland. (halloharpa@gmail.com)

APPENDIX B

Parameter estimation in a stochastic rainfall-runoff model

Published in *Journal of Hydrology* 2006, Vol 326, p. 379-393.



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Hydrology 326 (2006) 379–393

Journal
of
Hydrology

www.elsevier.com/locate/jhydrol

Parameter estimation in stochastic rainfall-runoff models

Harpa Jonsdottir ^{a,*}, Henrik Madsen ^a, Olafur Petur Palsson ^b

^a Department of Informatics and Mathematical Modelling, Bldg. 321 DTU, DK-2800 Lyngby, Denmark

^b Department of Mechanical and Industrial Engineering, University of Iceland, Hjarðarhaga 2-6, 107 Reykjavík, Iceland

Received 14 November 2003; revised 31 October 2005; accepted 9 November 2005

Abstract

A parameter estimation method for stochastic rainfall-runoff models is presented. The model considered in the paper is a conceptual stochastic model, formulated in continuous-discrete state space form. The model is small and a fully automatic optimization is, therefore, possible for estimating all the parameters, including the noise terms. The parameter estimation method is a maximum likelihood method (ML) where the likelihood function is evaluated using a Kalman filter technique. The ML method estimates the parameters in a prediction error settings, i.e. the sum of squared prediction error is minimized. For a comparison the parameters are also estimated by an output error method, where the sum of squared simulation error is minimized. The former methodology is optimal for short-term prediction whereas the latter is optimal for simulations. Hence, depending on the purpose it is possible to select whether the parameter values are optimal for simulation or prediction. The data originates from Iceland and the model is designed for Icelandic conditions, including a snow routine for mountainous areas. The model demands only two input data series, precipitation and temperature and one output data series, the discharge. In spite of being based on relatively limited input information, the model performs well and the parameter estimation method is promising for future model development.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Conceptual stochastic model; Rainfall-runoff model; Parameter estimation; Maximum likelihood; Extended Kalman filter; Prediction and simulation

1. Introduction

All hydrological models are approximations of reality, and hence the output of a system can never be predicted exactly and the problem is how to achieve an acceptable and operational model.

The numerous hydrological models which already exist vary in their model construction, partly because the models serve somewhat different purposes. There are models for design of drainage systems, models for flood forecasting, models for water quality, etc. Singh and Woolhiser (2002) give a comprehensive overview of mathematical modelling of watershed hydrology, however, a brief overview will be given here. The HBV model, see Bergström (1975, 1995), is a standard model in the Scandinavian countries

* Corresponding author. Tel.: +354 5696051; fax: +354 5688896;

E-mail address: hj@os.is (H. Jonsdottir).

Nomenclature			
a	Low pass filtering constant [1/day]	N	Snow cover [m]
b	Constant in $\psi(\cdot)$, controlling the smoothness	P	Measured precipitation [mm]
b	Center of the threshold function $\phi(\cdot)$	pdd	Positive degree day constant for melting [m/([°C]day)]
b_1	Sharpness of the threshold function $\phi(\cdot)$	S_1	Upper surface water reservoir [m]
c	Precipitation correction factor	S_2	Lower surface water reservoir [m]
f	Filtration from upper reservoir to lower reservoir [1/day]	s_{11}	White noise process for the observations
K	Constant representing the base flow [m/day]	$T(t)$	Measured temperature [°C]
k	Constant in $\psi(\cdot)$, controlling the smoothness	$T_s(t)$	Low pass filtered temperature [°C]
k_1	Routing constant, from upper surface reservoir [1/day]	Y	Measured discharge [m/day]
k_2	Routing constant, from lower surface reservoir [1/day]	$d\omega_I$	One dimensional Wiener process
M	Constant in $\psi(\cdot)$, controlling the upper limit value	$\phi(x)$	Smooth threshold function (sigmoid function) $1/(1 + \exp(b_0 + b_1x))$
		$\psi(x)$	Indicator function for the snow $\psi(x) = M \exp(-b \exp(-kx))$
		σ_{ii}	Incremental covariance of the Wiener process

and has also been used around the globe. The HBV model has been classified as a semi-distributed conceptual model and is based on the theory of linear reservoirs. The model has a number of free parameters, which are found by calibration. The model presented here is in the spirit of the HBV model, but additionally, the approach suggested in this paper includes a procedure for fully automatic parameter estimation. The NAM/MIKE11/MIKE21 models, see e.g. Nielsen and Hansen (1973); Gottlieb (1980); Havnø et al. (1995), are used for flood forecasting in Denmark and other European countries. The NAM model is based on similar principles as the HBV model and a further development of the model led to the MIKE11 software package which is a one dimensional modelling system for simulation of flow, sediment-transport and water quality. MIKE21 is a two dimensional version. The TOPMODEL, Beven and Kirkby (1979), has been used in Great Britain. The model is a set of conceptual tools that can be used to reproduce the hydrological behaviour of the catchment area in a distributed or semi-distributed way. The parameters are physically interpretable and the watershed is classified by using the so-called topographic index. The SHE model is a physically

based, distributed watershed modelling system, developed jointly by the Danish Hydraulic Institute, the British Institute of Hydrology and SOGREAH in France. The SHE model is widely used, see Abbott et al. (1986); Bathurst (1986); Singh and Woolhiser (2002); Jain et al. (1992); Refsgaard et al. (1992). The MIKE SHE model is a further development of the SHE modelling concept Refsgaard and Storm (1995) and it has been used in many European countries. The ARNO model, Todini (1996), is a semi-distributed conceptual model, and it is well known in Italy. Like the HBV and NAM models the Tank model, Sugawara (1995), is a model based on linear reservoirs and it has been used in Japan. The Xinjiang model, Zhao and Liu (1995); Zhao (2002), is a distributed, basin model for use in humid and semi-humid regions where the evaporation plays a major role. The model has been widely used in China since 1980. In Canada, the WATFLOOD model Singh and Woolhiser (2002), is being used. The WATFLOOD model is a distributed hydrological model based on the GRU (Group Response Unit) concept, i.e. all similarly vegetated areas within a sub-watershed are grouped as one response unit. The NWS River Forecast system, based on the Sacramento

Model, Burnash (1995), is a standard model in the United States for flood forecasting and in Australia the RORB model, Layrenson and Mein (1995), is commonly employed for flood forecasting and drainage design.

Increased computer power and data storage capabilities have opened the possibility for working with more detailed distributed models. Hence, many of the recently developed models are physically based distributed models, and they are occasionally used together with GIS (Geographic Information Systems). These models both utilize a large amount of information, but can also provide various information, however, if the input data (or information) are not available the model is of little use.

Black box models have also been used for flood forecasting, starting with linear transfer function models in the beginning of 1970s and since then various kinds of linear and nonlinear models. In recent years neural network models have been popular. Sajikumar and Thandaveswara (1999) used an artificial neural network as a nonlinear rainfall-runoff model for the river Lee in the UK and for the river Thuthapuzha in India. Shamseldin (1997) used neural networks for rainfall-runoff modelling which was tested on six different catchment areas.

The main advantage of black box models in hydrology is that they are not as data demanding as the physical models; this refers to all kinds of physical information about the watershed as well as long record of flow and precipitation.

Some of the conceptual models are not very data demanding and it is important to work with those kind of models as well, i.e. models with few input data, few parameters and limited prior information. The parameters in a lumped conceptual model can be interpreted as some kind of an average over a large area, but in general the most likely parameter values cannot be given, and the final parameter estimation must, therefore, be performed by calibration against observed data. Refsgaard et al. (1992) stated that in principle the parameters in a physically based model can be estimated by field measurements, but such an ideal situation requires comprehensive field data, which cover all the parameters. This situation rarely occurs and the problem of calibration will arise. Because of the large number of parameters in a physically based model the parameter estimation can

not be done by free optimization for all parameters, however, an over all parameter estimation is possible for simpler models.

The state space formulation and the Kalman filter has been used in hydrology for years, representing both black box models and grey box models, i.e. conceptual physical models where parameter values are estimated using data. Szollosi-Nagy (1976) used a state space formulation for on-line parameter estimation in linear hydrography using a FIR model (Finite Impulse Response model). Todini (1978) presented a threshold ARMAX model, formulated in a state space form and the parameters estimated off line, i.e. in a batch form. Refsgaard et al. (1983) reformulated the NAM model in a state space form where two of the model parameters were time varying i.e. on-line estimated. Haltiner and Salas (1988) used ARMAX models, both with off-line (batch) parameter estimation and on-line parameter estimation method in the SRM model, see also Martinec (1960); Martinec and Rango (1986). All the above-mentioned models are formulated in a discrete time. In Georgakakos (1986a,b) rather large physical models are presented using a state space formulation. However, the model parameters are constants and not estimated. In Georgakakos et al. (1988) the Sacramento model (orig. in Burnash et al. (1973)), is modified and formulated in state space form and some of the parameters are estimated. Rajaram and Georgakakos (1989) represent a model for acid decomposition in a lake watershed system formulated in a continuous-discrete state space form, and they estimated the parameters. Lee and V.P. Singh (1999) applied an on-line estimation to the Tank model (see e.g. Sugawara (1995)), for single storm at a time, calibrating the initial states manually. Lee and V.P. Singh (1999) also gave a short overview of application of the Kalman filter to hydrological problems upto 1999. In Ashan and O'Connor (1994) a general discussion about the use of Kalman filter in hydrology is found.

In the following a stochastic lumped, conceptual rainfall-runoff model is developed. The model is formulated as a continuous-discrete time stochastic state space model. The dynamics are described by stochastic differential equations and the observations are described by equations relating the discrete time observations to the state variables at time points where

observations are available. The main advantage of this model formulation is that the stochastic part permits a description of both the model and the measurement uncertainty, and hence more rigorous statistical methods can be used for parameter optimization. Furthermore, the stochastic modelling approach allows a much simpler model structure than a deterministic modelling approach since some of the variations observed in data are described by the stochastic part of the model. The model presented is a watershed model designed for discharge forecasting. It is a simple lumped reservoir model with two input variables, precipitation and temperature and one output variable, the discharge. Because of the simplicity of the model and few parameters it is possible to estimate all the parameters including threshold parameter in the snow routine and the system noise, which often has been difficult to identify in hydrology. The method suggested for parameter estimation (in batch form) is a maximum likelihood method, where the one step ahead prediction errors required for evaluating the likelihood function are evaluated using the Kalman filter technique. Moreover, the state space formulation allows the model to be used for simulation as well, however, good simulation results require different parameter values, Kristensen et al. (2004). Parameter values, which are suitable for simulation can be achieved by fixing the system noise to a small value and then estimating the remaining parameters. Conversely good prediction results are obtained by using parameter values where all the parameters have been optimized, including the parameters describing the system noise.

The paper is organized as follows. In Section 2 the availability of data is discussed. The model is described in Section 3 and the method for parameter estimation is described in Section 4. Section 5 includes discussion of some estimation principles. In Section 6 results are demonstrated and in Section 7 conclusions are drawn.

2. The data

The goal is to develop a model, which can be used in mountainous areas with snow accumulation. Such areas are often thinly populated and the meteorological observatories are often rather spread. However,

there is a need for flood forecasting for various reasons, such as warning related to the spring floods or for operational planning of hydropower plants. The data used in this project originates from Iceland which has in general only mountainous catchment areas and it certainly is thinly populated with only 2.8 inhabitants per km², and only a few meteorological observatories exist.

Precipitation is the main input for hydrological models as the precipitation and the evaporation control the water balance. In Iceland, the evaporation plays only a minor role, but the precipitation is important. However, there is a shortage of good precipitation data, which indeed has an effect on the prediction performance of the model. The poor quality of precipitation data arises both from a rain gauges bias towards too small values, and a limited number of rain gauge measurement stations. Due to the influence of wind the amount of precipitation measured is an underestimate of the 'ground true' precipitation. Unfortunately, no experiments have been made in Iceland in order to develop models to adjust for this bias. Experiments, like for instance the Nordic project in Jokioinen in Finland during the years 1987–1993 (Førland et al. (1996)), had the purpose of developing models to describe the underestimate of the different rain gauges depending on weather condition. This experiment is of little use here since the wind speed in Iceland in general is much higher than in Finland. Furthermore, most of the meteorological stations in Iceland are located along the coastal line in the inhabited areas and most rivers, especially the larger ones, stretch far into the country and have thus watershed in high mountainous areas. Occasionally, there are no meteorological stations in the whole watershed and if any they are typically located near the coast. Precipitation lapse rate is also difficult to track since in practice the lapse rate depends highly on wind speed and direction in the mountainous areas.

The discharge data are calculated from water level data using the $Q-h$ formula. The errors of the discharge data are caused both by uncertainty of the water level data and uncertainty of the parameters in $Q-h$ formula. The errors of the water level measurements more or less only occur during the winter because of the icing, which causes the water level to rise even though the flow of water is not increasing. This has to be corrected manually.

In Iceland few discharge measurements with very large discharge exist. This is due to the fact that even though annual spring floods occur, it is difficult to predict peaks several days ahead, and the number of rivers that can be measured simultaneously is limited.

The discharge used in this project originates from the river Fnjóská in Northern Iceland, see Fig. 1. The river is a direct runoff river with no glacier in the watershed, whereas many larger rivers in Iceland have a glacier factor. The watershed is about 1132 km², the altitude range is between 44 and 1084 m, and 54% of the catchment area is above 800 m. The catchment area is dominated by grit and rocks; a very small part of the region in the valley is cope and grassland. A meteorological observatory is located in the watershed, at Lerkihlid, about 20 km from the outlet of the watershed, and it is situated 150 m above sea level. No meteorological observatory is located in the highlands

which could have given information about the weather condition in the catchment area there. The data used are diurnal averages of the discharge, diurnal averages of the temperature and the total precipitation for the past 24 hours. Fig. 2 shows the discharge, the temperature and the precipitation for the whole period of 8 years, starting 1st of September 1976 and ending 31st of August 1984.

3. The stochastic model

The stochastic model proposed is a simple smooth threshold model with a snow routine, and the basic idea is similar to the idea behind the HBV and NAM models, see Bergström and Fossman (1973); Bergström (1975); Nielsen and Hansen (1973), and

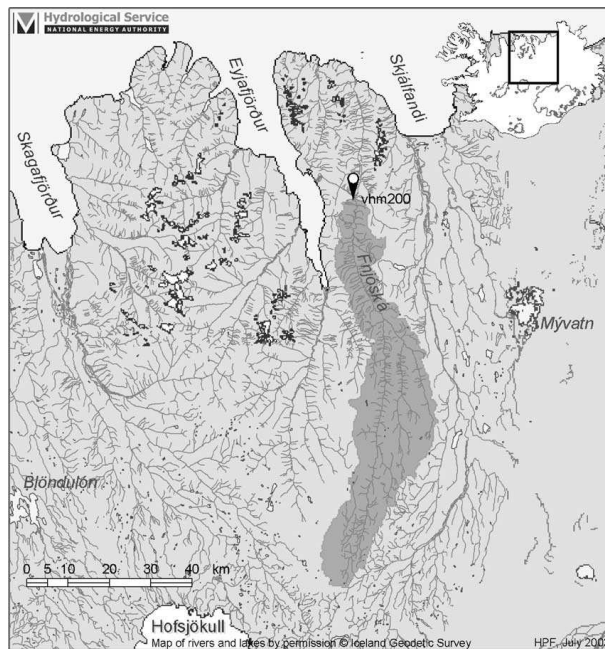


Fig. 1. The watershed of the river Fnjóská is about 1132 km² and located in Northern Iceland.

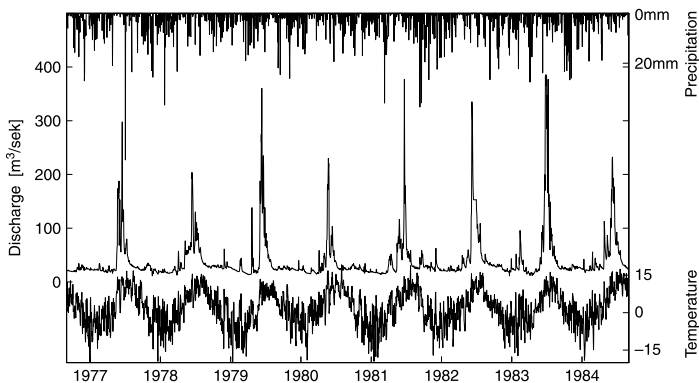


Fig. 2. The data series for discharge, temperature and precipitation starting 1st of September 1976 and ending 31st of August 1984.

Gottlieb (1980). A diagram of the model structure is shown in Fig. 3.

The water is stored in reservoirs and the outflow of the reservoirs are routed to the stream with different time constants. The main distinction between the non-stochastic HBV and NAM models and the stochastic model suggested here is that the water flow is modelled as a function of only precipitation and temperature, and there are no factors for evaporation and infiltration into the ground. On the other hand no manual calibration is required since the stochastic model allows for statistical methods for parameter estimation. The total precipitation is divided into snow and rain using a smooth threshold function $\phi(T(t))$, where $T(t)$ is the air temperature. The threshold function is formulated as the sigmoid function

$$\phi(T(t)) = \frac{1}{1 + \exp(b_0 - b_1 T(t))} \tag{1}$$

The same smooth threshold is used for the melting process, where the melting $M(t)$ is formulated using the positive degree day method

$$M(t) = \text{pdd } T(t) \phi(T(t)) \tag{2}$$

where pdd is the positive degree day constant, which typically is calibrated. No attempt is made to model

the actual physical process of melting, i.e. the fact that in the beginning of the melting process the water first stays in the snow pack and is not released until the snow pack is wet enough. However, in order to take this into account the temperature $T(t)$ is low pass filtered

$$dT_s(t) = [-aT_s(t) + aT(t)]dt + dw(t) \tag{3}$$

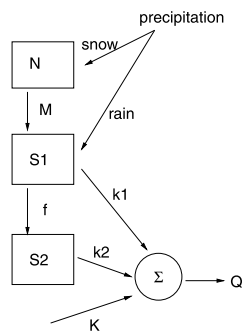


Fig. 3. The model structure. The precipitation is divided into snow and rain, N is snow a container, S_1 and S_2 are upper and lower surface reservoirs, M is melting, f is filtration between the reservoirs and k_1 and k_2 are the routing constants. K is a constant representing the base flow and Q is the discharge.

The consequence is that a single warm day does not give as much impact as a warm day followed by another warm day.

The suggested stochastic state space model is:

$$dT_s(t) = [-aT_s(t) + aT(t)]dt + \sigma_1 dw_1(t) \quad (4)$$

$$dN(t) = [-pddT_s(t)\phi(T_s(t))\psi(N(t)) + (1 - \phi(T_s(t)))cP(t)]dt + \sigma_2 dw_2(t) \quad (5)$$

$$dS_1(t) = [pddT_s(t)\phi(T_s(t))\psi(N(t)) - (f + k_1)S_1(t) + (\phi(T_s(t))cP(t))]dt + \sigma_3 dw_3(t) \quad (6)$$

$$dS_2(t) = [fS_1(t) - k_2S_2(t)]dt + \sigma_4 dw_4(t) \quad (7)$$

$$Y(t) = k_1S_1(t) + k_2S_2(t) + K + e_1(t) \quad (8)$$

where N is the amount of snow in the snow container in meters and the function ψ is a smooth indicator function, controlling whether there is snow to melt or not,

$$\psi(N) = \text{Mexp}(-b\exp(-kN)) \quad (9)$$

S_1 and S_2 are water content reservoirs, f is filtration from the upper surface reservoir to the lower surface reservoir, k_1 and k_2 are the routing constants, and c is a precipitation correction factor. $\sigma_1, \dots, \sigma_4$ are constants representing the variances of the system noise and the noise terms $dw_1(t), \dots, dw_4(t)$ are assumed to be independent standard Wiener processes and all are assumed independent of measurement noise $e_1(t)$. The base flow is assumed to be constant. An extension of the model with a ground water reservoir would improve the physical reality of the model and it might be a task for future research.

4. Parameter estimation

In this section a maximum likelihood method for estimation of the parameters of the continuous–discrete time stochastic state space models is outlined. The procedure is implemented in a program called CTSM (Continuous Time Stochastic Modelling), and for a further description of the mathematics and numerics behind the program, see Kristensen et al. (2003), and Kristensen et al. (2004).

The hydrological model described by Eq. (4)–(8) is a continuous–discrete time stochastic state space model. The stochastic differential equations describe the dynamics of the system in continuous time as stated by Eq. (4)–(7), and the algebraic equation Eq. (8) describes how the measurements are obtained as a function of the state variables at discrete time instants. Using a slightly different and more compact notation, the mathematical formulation of the continuous–discrete time stochastic state space model is

$$dx_t = f(x_t, u_t, t, \theta)dt + \sigma(u_t, t, \theta)d\omega_t \quad (10)$$

$$y_k = h(x_k, u_k, t_k, \theta) + e_k \quad (11)$$

where $t \in \mathbb{R}_+$ is time, $x_t \in \mathbb{R}^4$ is a vector of the state variables (since $x_t = [T_s(t), N(t), S_1(t), S_2(t)]^T$), $u_t \in \mathbb{R}^2$ is a vector of the input variables (since $u_t = [T(t), P(t)]$), $\theta \in \mathbb{R}^p$ is a vector of the unknown parameters. The vector $y_k \in \mathbb{R}$ is a vector of measurements (i.e. the discharge). The notation $x_k = x_{t=t_k}$ and $u_k = u_{t=t_k}$ is used. Furthermore, the functions $f(\cdot) \in \mathbb{R}^4$, $\sigma(\cdot) \in \mathbb{R}^{4 \times 4}$ and $h(\cdot) \in \mathbb{R}$ are nonlinear functions, ω_t is a 4-dimensional standard Wiener process and $e_k \in N(0, S(u_k, t_k, \theta))$ is a Gaussian white noise process. With this model formulation the parameters are constants and estimated *off-line* or in a batch form. An *on-line* estimation of (some) parameters is possible by extending the state vector with the relevant parameters. As mentioned, Eq. (10) is known as the system equation and Eq. (11) is known as the measurement equation.

The measurements y_k are in discrete time. It is well known that the likelihood function for time series models is a product of conditional densities (see e.g. Restrepo and Bras (1985)). By introducing the notation

$$\mathcal{Y}_k = (y_k, y_{k-1}, \dots, y_1, y_0) \quad (12)$$

where $(y_k, y_{k-1}, \dots, y_1, y_0)$ is the time series of all measurements up to and including the measurement at time t_k . The likelihood function can be written as

$$L(\theta; \mathcal{Y}_N) = \left(\prod_{k=1}^N p(y_k | \mathcal{Y}_k, \theta) \right) p(y_0 | \theta) \quad (13)$$

In order to obtain an exact evaluation of the likelihood function, the initial probability density $p(y_0 | \theta)$ must be known and all subsequent conditional

densities can be determined by successively solving Kolmogorov’s forward equation and applying Bayes’s rule, Jazwinski (1970). This approach is not feasible in practice. However, since the diffusion term in the system equation, Eq. (10), in the continuous–discrete state space model is a Wiener process, which is independent of the state variables, and the error term in the measurement equation, Eq. (11), it follows that for LTI (Linear Time Invariant) and LTV (Linear Time Variant) models the conditional densities are Gaussian, Jazwinski (1970). In the NL (Non Linear) case it is reasonable to assume that under suitable regularity conditions, the conditional densities can be well approximated by the Gaussian distribution, Kristensen et al. (2004). This assumption can be tested after the estimation e.g. by considering the sequence of residuals. Thus, assuming that the conditional densities $p(y_k|U_k, \theta)$ are Gaussian the likelihood function becomes

$$L(\theta|U_N) = \left(\prod_{k=1}^N \frac{\exp\left(-\frac{1}{2} \varepsilon_k^T \mathbf{R}_{k|k-1}^{-1} \varepsilon_k\right)}{\sqrt{\det(\mathbf{R}_{k|k-1})} (\sqrt{2\pi})^l} \right) p(y_0|\theta) \tag{14}$$

where $\varepsilon_k = y_k - \hat{y}_{k|k-1} = y_k - E\{y_k|U_{k-1}, \theta\}$ is the one step prediction error and $\mathbf{R}_{k|k-1} = V\{y_k|U_{k-1}, \theta\}$ is the associate conditional covariance. For given parameters and initial states, ε_k and $\mathbf{R}_{k|k-1}$ can be computed by means of a Kalman filter in the linear case or an extended Kalman filter in the nonlinear case. The continuous-discrete Kalman filter equations are (Kristensen et al. (2004) or Jazwinski (1970)):

$$\hat{y}_{k|k-1} = \mathbf{h}(\hat{x}_{k|k-1}, \mathbf{u}_k, t_k, \theta) \quad \text{(Output prediction)} \tag{15}$$

$$\mathbf{R}_{k|k-1} = \mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{S} \quad \text{(Output variance prediction)} \tag{16}$$

$$\varepsilon_k = y_k - \hat{y}_{k|k-1} \quad \text{(Innovation)} \tag{17}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}^T\mathbf{R}_{k|k-1}^{-1} \quad \text{(Kalman gain)} \tag{18}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \mathbf{K}\varepsilon_k \quad \text{(Updating)} \tag{19}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{R}_{k|k-1}\mathbf{K}_k^T \quad \text{(Updating)} \tag{20}$$

$$\frac{d\hat{x}_{t|k}}{dt} = \mathbf{f}(\hat{x}_{t|k}, \mathbf{u}_k, t_k, \theta) \quad t \in [t_k, t_{k+1}] \tag{21}$$

(State prediction)

$$\frac{d\hat{\mathbf{P}}_{t|k}}{dt} = \mathbf{A}\hat{\mathbf{P}}_{t|k} + \mathbf{P}_{t|k}\mathbf{A}^T + \sigma\sigma^T \quad t \in [t_k, t_{k+1}] \tag{22}$$

(State var. pred.)

where

$$\mathbf{A} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}_t} \right|_{\mathbf{x}=\hat{x}_{t|k-1}, \mathbf{u}=\mathbf{u}_k, t=t_k, \theta}, \tag{23}$$

$$\mathbf{C} = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}_t} \right|_{\mathbf{x}=\hat{x}_{t|k-1}, \mathbf{u}=\mathbf{u}_k, t=t_k, \theta}$$

and

$$\sigma = \sigma(\mathbf{u}_k, t_k, \theta), \quad \mathbf{S} = \mathbf{S}(\mathbf{u}_k, t_k, \theta) \tag{24}$$

Given information upto and including time t the prediction $\hat{x}_{k+1|k} = E\{x_{t_{k+1}}|x_{t_k}\}$ and $\mathbf{P}_{k+1|k} = E\{x_{t_{k+1}}x_{t_{k+1}}^T|x_{t_k}\}$ are needed for the Kalman filter equations. Eq. (22) is a linear differential equation, which can be solved analytically. This analytical solution is used to calculate the ‘initial’ problem $\mathbf{P}_{k+1|k}$. On the other hand equation Eq. (21) is nonlinear with a nontrivial solution. The software CTSM offers three options for handling this:

Linearization by first order Taylor, the linear equation is solved analytically, iteratively in a subsampled interval.

Numerical solution of the ODE equation Eq. (21) by using a Predictor/Corrector scheme, also occasionally referred to as Gears method or Adams method (see e.g. Dahlquist and Björck (1988)).

Numerical solution of the ODE equation Eq.(21) by using BDF (Backward Difference Formula) (see e.g. Dahlquist and Björck (1988)).

For a detailed description of all the methods see Kristensen et al. (2003). The BDF formula demands a Newton-like method for solving a nonlinear zero-point equation and is thus the most time consuming algorithm. However, for stiff systems the BDF formula is the most reliable method (Dahlquist and Björck (1988)) and consequently this option has been used in the following.

The hydrological model described by Eq. (4)–(8) is a model with four states, the low pass filtered temperature T_S , the snow container N , and upper

and lower surface containers S_1 and S_2 . It is found important to include the snow container as a state variable and not to treat the water from snow melt as an input as some times is done, e.g. Refsgaard et al. (1983). Treating the snow-melt as an input requires manual control of the snow balance. However, by including the snow container into the state vector leads to numerical complications. The system is singular during summer, fall, and most of the winter or more accurately while snow is not melting. The system is non-stiff during spring floods, i.e. when melting is significant and extremely stiff during the transition points in between.

On of the strengths of using the maximum likelihood method for parameter estimation it that it follows from the central limit theorem that the estimator $\hat{\theta}$, is asymptotically Gaussian with mean θ and covariance

$$\hat{\Sigma}_{\hat{\theta}} = H^{-1} \quad (25)$$

where the information matrix H is given by

$$h_{ij} = -E \left\{ \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln(L(\theta|Y)) \right\} \quad i, j = 1, \dots, p. \quad (26)$$

An approximation of H can be obtained by evaluating $h_{ij} = \partial^2 / (\partial \theta_i \partial \theta_j) \ln(L(\theta|Y))$ in the point $\theta = \hat{\theta}$. The asymptotic Gaussianity of the estimator also allows marginal t -test to be performed like a test for the hypothesis:

$$H_0 : \theta_j = 0 \quad H_1 : \theta_j \neq 0 \quad (27)$$

The one step prediction of the output $\hat{y}_{k|k-1}$, the state update $\hat{x}_{k|k}$, and the state prediction $\hat{x}_{k|k-1}$, corresponding to each time instant t_k are generated by the (extended) Kalman filter. A simulation $\hat{x}_{i|0}$ and $\hat{y}_{i|0}$ can be obtained using the (extended) Kalman filter equations without the updating.

5. Some comments on parameter estimation in hydrological models

Model calibration has been a topic in hydrology since the computer evolution in 1960 and since then parameter optimization has been practiced. A solution to a rainfall-runoff prediction problem is to optimize the parameters such that the model performs the 'best'

fit to data. On the other hand what is best the fit to data? This is a selective question with a selective answer. Best fit can be such that the sum of squared simulation error is minimized, or the sum of squared prediction error is minimized, or models, which conserve the water balance best, or those who have the best timing of flood peaks. In the recent years multi objective calibration and Pareto optimality have been applied in rainfall-runoff modelling, see e.g. Madsen (2000). However, two estimation methods have frequently been used in hydrology. Those are, the Output Error method (OE), and the Prediction Error method (PE). The OE method minimizes the sum of squared simulation error and is used in white box modelling but also in other contexts. This method is always off line. The PE method minimizes the sum of squared one step prediction error, this method offers both off-line and on-line estimation. In order to allow for a comparison between the methods the off-line method is considered in the following. Young (1981) gives an overview and comparison of parameter estimation for continuous time models, which includes PE and OE principles. The maximum likelihood method as presented here is a PE method, whereas the OE method, can in statistical implications include Maximum Likelihood terms for the case where there is no system noise, Young (1981). Using the Kalman filter notation, the sum of squares of the error terms for the OE method is written as $\sum (y_k - \hat{y}_{k|0})^2$. This corresponds to a state space representation without system noise and all the errors incorporated in the measurement noise, which means prediction without updating, i.e. a simulation.

Comparing the computational time for the two modelling approaches, the state space formulation and the Kalman filter in general involve more calculations since a state filtering through the whole data series is needed for each evaluation of the objective function. It is thus questionable whether this time-consuming estimation method is worth the time. Kristensen (2002) performed a simulation study for continuous discrete models by comparing the PE method as implemented in the program CTSM and the OE method as implemented by Bohlin and Graebe (1995). The calculations for the OE method were performed by using the MoCaVa software (Bohlin (2001)), which runs under Matlab. Some of the results and discussions are also demonstrated in Kristensen et al.

(2004). The results show that the PE estimation method gives significantly less biased estimate of the parameters than the OE method. For simulations with no system noise the methods were similar, but the more noise the greater is the difference between the two methods resulting in larger bias for the OE method. Moreover, the PE estimation method provides uncertainty information in terms of standard deviations of the estimates and other statistical tools for model evaluation. Therefore, for the purpose of short-term prediction such as in flood warning systems it is truly recommended to use the PE method even though the method is more computational demanding. Once the parameters are estimated the output prediction is not time consuming. Only if the model focus on good long-term prediction capabilities, the OE method is to be preferred Kristensen et al. (2004). It must, however, be kept in mind that the input, i.e. precipitation and temperature are always needed as input and long-term prediction for precipitation and temperature variables are not particularly precise and for that reason long-term prediction might not be so reliable. Last, but none the least, in a state space formulation it is easy to handle missing values in observations automatically, and this prevents the user from having to resort other models (e.g. black box models) to fill in gaps in the data.

It is worth mentioning that Rajaram and Georgakakos (1989) presented a parameter estimation of stochastic hydrologic models formulated in a continuous-discrete state space form, with the parameters estimated in a batch form. Their method mainly differs from the one presented here in two ways. Firstly, the filtering, or the state prediction is calculated by a fourth order predictor–corrector scheme, while here a BDF method is used. Secondly, and probably the most important difference, in the methodology presented here the system error, $\sigma(u_r, t, \theta)d\omega_t$ is estimated. Conversely in the methodology presented by Rajaram and Georgakakos (1989), the estimation of the state error $\sigma(u_r, t, \theta)d\omega_t$ demands a human input. In Rajaram and Georgakakos (1989) the state error $\sigma(u_r, t, \theta)d\omega_t$ is decomposed into three error terms; error term from input, error term associated with estimation of uncertain constants (such as topographic or rating curve constants) and error term in model structure. Only the last term is estimated, the two first must be set as a degree of

believe by the trained hydrologists if they are not exactly known, Rajaram and Georgakakos (1989).

6. Results

The stochastic model in Eq. (4)–(8) is used to investigate how the parameter estimation method performs for the hydrological problem described in Section 2. The parameters are estimated by using the first 6 years of the data while the last two years are used for validation. For a comparison between the PE and OE method the optimization was performed using both methods. First in a PE setting by estimating *all* the parameters, including the system noise, and then in an OE settings by fixing the system noise term parameters to a small value. The former parameter values are optimal for prediction and the latter for simulations. The estimated parameter values are shown in Table 1.

The units for the snow container N and the upper and lower reservoirs S_1 and S_2 are given in meters. The total volume is calculated by multiplying with the watershed area. Fig. 4 illustrates the results from the PE method and Fig. 5 illustrates the results from the OE method.

Note from Fig. 5, that the OE formulation produces the same prediction and simulation and hence the coefficient of determination (Nash and Sutcliffe, 1970), is the same, $R^2=0.69$, in both cases. Conversely, the PE method produces very different results for prediction and simulation with coefficients of determination as $R^2_{\text{prediction}}=0.93$ and $R^2_{\text{simulation}}=0.43$, respectively. Furthermore, it is interesting to compare some of the estimated parameter values yielded by the two different estimation techniques. The precipitation correction factor c , the threshold parameter b_0 (for snow/rain) and the positive degree-day constant pdd are much larger for the OE estimation than for PE estimation; the difference being almost factor 2. The routing constants k_1 and k_2 are, however, smaller in the OE estimation, whereas the filtration f is similar. Note also that the PE method estimates some memory in the temperature, i.e. $a=1.475$ while the OE estimates no memory in the temperature i.e. $a=4.939$. Finally, the total noise is incorporated in the measurement noise in the OE settings, resulting in larger prediction error

Table 1
Estimation results and comparison of the PE and OE method

Par.	PE method		OE method		Unit
	Estimate	Std. dev.	Estimate	Std. dev.	
N_0	0.000		0.000		<i>m</i>
S_0	0.00099	0.00020	0.00010	3×10^{-7}	<i>m</i>
S_1	0.00037	0.00022	0.00016	5×10^{-7}	<i>m</i>
$T_{s,0}$	3.000		3.000		°C
b_0	4.511	0.012	8.131	0.073	°C
b_1	1.000		1.000		
M	1.000		1.000		
B	100		100		
K	200		200		
c	1.518	0.00075	2.788	0.054	
pdd	0.00342	9×10^{-7}	0.00585	6×10^{-6}	m/°C day
F	0.031	0.00065	0.049	0.002	1/day
k_1	0.674	0.05704	0.216	0.010	1/day
k_2	0.097	0.00424	0.049	0.070	1/day
A	1.475	0.01411	4.939	0.023	
σ_{T_s}	10^{-8}		10^{-8}		°C
σ_N	0.0074	0.00012	1×10^{-6}		<i>m</i>
σ_{S_1}	0.0011	0.00042	1×10^{-6}		<i>m</i>
σ_{S_2}	0.0008	0.00002	1×10^{-6}		<i>m</i>
K	0.00198	0.00005	0.00175	0.00003	m/day
s_1	2.2×10^{-10}	7.6×10^{-12}	1.4×10^{-6}	5.0×10^{-8}	m/day
R_{pred}^2	0.93		0.69		
R_{sim}^2	0.43		0.69		

and hence, the confidence band around the prediction is larger than in the PE settings.

In the following the results from the PE estimation will be discussed. The routing constants k_1 and k_2 , and

the filtration f are measured in the unit 1/day, i.e. 24 h. For hourly values the constants can be multiplied by 1/24. The routing constants are rather small but bearing in mind that the size of the watershed is

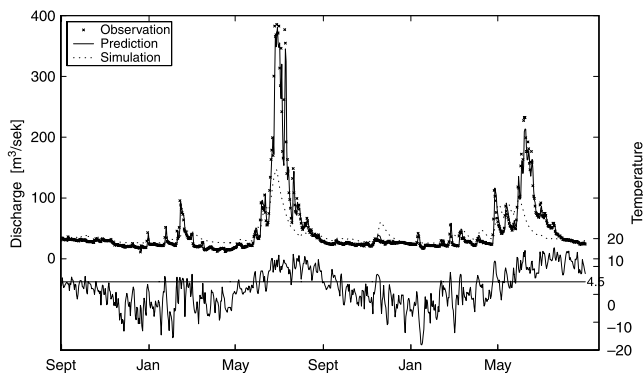


Fig. 4. The results from PE estimation. The validation period from 1st of September 1982 to 31st of August 1984. The figure shows the river discharge the one step prediction and the simulation. The temperature shown is the low pass filtered air temperature.

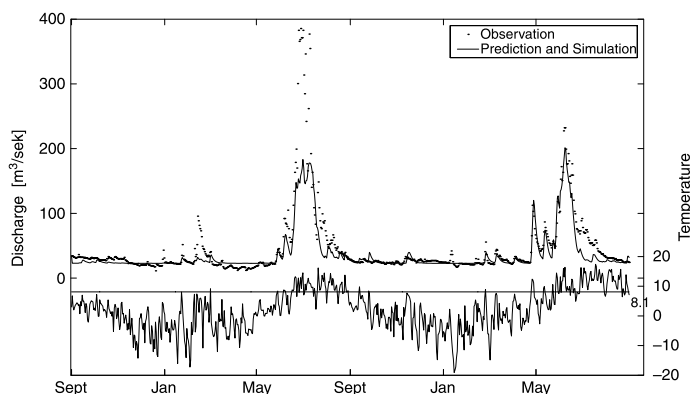


Fig. 5. The results from OE estimation. The figure shows river discharge, the one step prediction and the simulation. The validation data are from 1st of September 1982 to 31st of August 1984.

1132 km², these small constants are realistic. The upper routing constant is 0.674/day and thus the corresponding time constants is about 1 and a half day, whereas, the lower routing constant is 0.097/day, and hence the time constant is about 10 and a half day. The filtration constant is 0.031/day and the corresponding time constant about 32 days. Consequently most of the spring flood is delivered through the river via the first reservoir. The threshold function Eq. (1) for dividing precipitation into snow and rain is the same as the threshold function for melting snow. The parameter b_1 controls the steepness and has been set to one, and the parameter b_0 controls the center and is estimated to 4.50 °C. Thus, the threshold function is about zero when the temperature is 0 °C and then no snow is melting and all precipitation is solid. When the temperature is about 9 °C snow is melting everywhere and all precipitation is rain. In between some precipitation is snow and some as rain, and a proportion of the snow is melting (if there is snow in the snow-container). Recall that the watershed is 1132 km² with an altitude ranging from 44 to 1084 m and the meteorological observatory is located about 20 km from the watershed outlet at an altitude about 150 m. The center of mass of the watershed altitude is about 830 m and, if it is assumed that the temperature in altitude 830 m is zero when the temperature is 4.5 °C at the observatory, it leads to

a temperature lapse rate of $4.5/6.8 = 0.66$ °C/100 m, which is physically realistic. Hence, this smooth threshold function has the effect that it is not necessary to divide the area into elevation zones.

The precipitation correction constant c is estimated as 1.5. This correction is both correcting the underestimate of the rain gauge and the average increase in precipitation due to altitude. The factor c controls the input-output balance of the model. A water balance model with ground water container and evapotranspiration would have had a much larger correction constant. However, it should be mentioned that it is not possible to identify (estimate) both the correction constant and evapotranspiration given only measurement of the precipitation and discharge. Finally, Fig. 6 shows the state estimates of the contents of the snow-container, N , and the upper and lower surface containers S_1 and S_2 as predicted by the model using the PE parameters.

The unit is meter and the total volume is calculated by multiplying with the watershed area, 1132 km². The estimated noise terms σ_N , σ_{S_1} and σ_{S_2} have a order of magnitude 10^{-3} and thus the noise terms more or less only have an effect when the states are around zero, with the consequence that the states might become slightly negative. This has not lead to problems in this case. The problem might be solved by transforming the model using the logarithm.

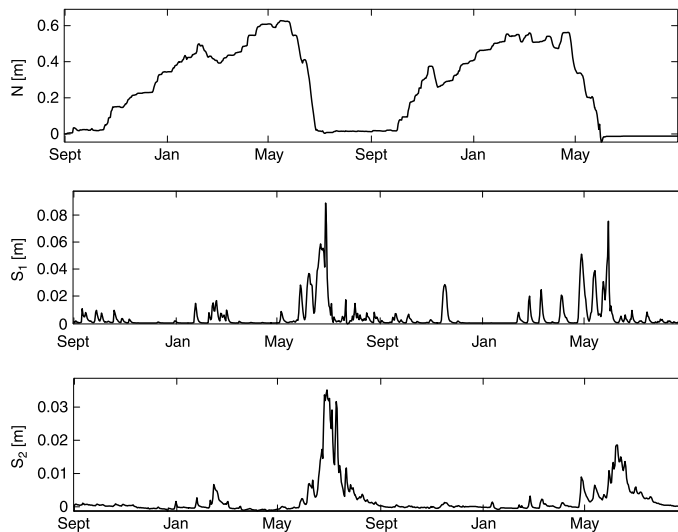


Fig. 6. State estimates of the contents of the snow container and the upper and lower surface containers.

As a flood forecasting model it is concluded that this simple model is satisfactory. Using the simple model makes it possible to estimate all the parameters, thus allowing the data to be used for an automatic calibration. The optimization, using 6 years of data, takes several hours on a PC computer but as mentioned earlier, once the parameters are estimated the update of the Kalman filter and the prediction are not computational demanding. CTSM can be run on a parallel computer using several CPUs and then the computer time will be much lower.

For the purpose of flood forecasting the most interesting development of the model would be to include some of the parameters in the state vector and thus allow for time varying parameters. Particularly since the parameters have been estimated, these estimates could act as good initial states for the time varying parameters. This could particularly be done for the base flow constant.

Finally, it is interesting to point out some revisions, which might improve the performance of simulation using an OE estimation. The large threshold temperature $b_0 = 8.1$ °C indicates that it might be necessary to

divide the area into two elevation zones, still using smooth threshold functions but with different centers. The former spring flood is much higher and narrower than the latter and such a narrow flood is difficult to produce. It might be necessary to have three surface containers and thus three time constants for the flow. It would also be interesting to let the pdd constant vary in time. Rango and Martinec (1995) state that the positive degree day factor should gradually increase during the melting season and this could certainly be introduced in the PE settings as well. A time varying pdd might though have larger differences in cases where the melting season is longer such as for glacier rivers.

7. Conclusions

All precipitation runoff models are approximations of the reality and hence they cannot be expected to provide a perfect fit to data. The process is highly non-stationary and the dynamics related to the snow is extremely non-linear. Furthermore, the deviations between the model prediction and the

data (the residuals) are almost always serially correlated. This calls for a stochastic model with both system noise and measurement noise.

In this paper a simple conceptual stochastic rainfall-runoff model is suggested. A method for estimation of the parameters of the model is outlined. The estimation method is a generic maximum likelihood method for parameter estimation in systems described by continuous-discrete time state space models, where the system equation consists of stochastic differential equations. Hence, the dynamics are described in continuous time, which allows for a direct use of prior physical knowledge, and the estimated parameters can be physically interpreted directly.

A further advantage of the stochastic state space approach is that the same model structure can be used for both prediction and simulation. It is advocated that the only difference lies in a different parameterization of the system error leading to different parameter values.

The presented model is simple and demands only two input variables, namely precipitation and temperature, and a single output, the discharge. The results for simulation are reasonable but not fully satisfying and it is concluded that a slightly more complicated model is needed even though it is questionable whether it is possible to obtain a better performance due to the poor precipitation data as in this study. However, the results obtained for prediction (flood forecasting) are satisfying.

Acknowledgements

The authors wish to thank the National Energy Authority in Iceland for delivery of discharge series, and The Icelandic Meteorological Office for meteorological data series. Furthermore, we wish to thank PhD Niels Rode Kristensen for technical support with the program CTSM.

References

Abbott, M., Bathurst, J., Cunge, J., O'Connell, P., Rasmussen, J., 1986. An introduction to the the European Hydrological System—

- System Hydrologique Europeen, 'SHE' 2: structure of a physically-based, distributed modelling system. *Journal of Hydrology* 87, 61–77.
- Ashan, M., O'Connor, K.M., 1994. A reappraisal of the kalman filtering technique, as applied in river flow forecasting. *Journal of Hydrology* 161, 197–226.
- Bathurst, J.C., 1986. Physically-based distributed modelling of an upland catchment using the Systeme Hydrologique Europeen. *Journal of Hydrology* 87, 79–102.
- Bergström, S., 1975. The development of a snow routine for the HBV-2 model. *Nordic Hydrology* 6, 73–92.
- Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water resources Publication, Colorado.
- Bergström, S., Fossman, A., 1973. Development of a conceptual deterministic rainfall-runoff model. *Nordic Hydrology* 4, 147–170.
- Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin* 24, 43–69.
- Bohlin, T., 2001. A Grey-box Process Identification Tool: Theory and Practice. Technical Report IR-S3-REG-0103, Department of Signals, Sensors and Systems, Royal Institute of Technology, Stockholm, Sweden.
- Bohlin, T., Graebe, S.F., 1995. Issues in nonlinear stochastic grey-box identification. *International Journal of Adaptive Control and Signal Processing* 9, 465–490.
- Burnash, R., 1995. NWS river forecast system—catchment modeling. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water resources Publication, Colorado.
- Burnash, R.J.C., Ferral, R., McGuire, R., 1973. A Generalized Streamflow Simulation System—Conceptual modeling for Digital Computers. Technical Report, National Weather Service, NOAA and State of California Department of Water Resources, Joint Federal State River Forecast Center Sacramento, California.
- Dahlquist, G., Björck, Å., 1988. *Numerical Methods*. McGraw-Hill, New Jersey.
- Førland, E.J., Allerup, P., Dahlström, B., Elomaa, E., Jónsson, T., Madsen, H., Perälä, J., Rissanene, P., Vedin, H., Vejen, F., 1996. Manual for Operational Correction of Nordic Precipitation Data. Technical Report No. 24/96 ISSN 0805-9918, The Norwegian.
- Georgakakos, K.P., 1986a. A generalized stochastic hydrometeorological model for flood and flash-flood forecasting I. *Water Resources Research* 22, 2083–2095.
- Georgakakos, K.P., 1986b. A generalized stochastic hydrometeorological model for flood and flash-flood forecasting II. *Water Resources Research* 22, 2096–2106.
- Georgakakos, K.P., Rajaram, H., Li, S.G., 1988. On improved Operational Hydrologic Forecasting. Technical Report IHHR No. 325, Institute of Hydraulic Research, University of Iowa.
- Gottlieb, L., 1980. Development and applications of runoff model for snowcovered and glacierized basins. *Nordic Hydrology* 11, 255–272.

- Haltiner, J., Salas, J.D., 1988. Short term forecasting of snowmelt runoff using armax models. *Water Resources Bulletin* 24, 1083–1089.
- Havnø, K., Madsen, M.N., Dørgé, J., 1995. MIKE11 a generalized river modelling package. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water resources Publication, Colorado.
- Jain, S., Storm, B., Bathurst, J., Refsgaard, J., Singh, R., 1992. Application of the SHE to catchments in India Part 2. Field experiments and simulation studies with the SHE on the Kolar subcatchment of the Narmand river. *Journal of Hydrology* 140, 25–47.
- Jazwinski, A.H., 1970. *Stochastic Processes and Filtering Theory*. Academic Press, San Diego.
- Kristensen, N.R., 2002. Fed-Batch Process Modelling for State Estimation and Optima Control. PhD Thesis, Department of Chemical Engineering, The Technical University of Denmark. ISBN 87-90142-83-7.
- Kristensen, N.R., Melgaard, H., Madsen, H., 2003. CTSM Version 2.3—A Program for Parameter Estimation in Stochastic Differential Equations. Available from www.imm.dtu.dk/ctsm.
- Kristensen, N.R., Madsen, H., Jørgensen, S.B., 2004. Parameter estimation in stochastic grey-box models. *Automatica* 40, 225–237.
- Layrenson, E., Mein, R., 1995. RORB a hydrograph synthesis by runoff routing. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water resources Publication, Colorado.
- Lee, Y.H., Singh, V.P., 1999. Tank model using Kalman filter. *Journal of Hydrologic Engineering* 4, 344–349.
- Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology* 235, 276–288.
- Martinec, J., 1960. The degree day factor for snowmelt-runoff forecasting. In: *General Assembly of Helsinki Commission on Surface Runoff*, number 51 in IAHS, pp. 468–477.
- Martinec, J., Rango, A., 1986. Parameter values for Snowmelt runoff modeling. *Journal of Hydrology* 84, 197–219.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part i—a discussion of principles. *Journal of Hydrology* 10, 282–290.
- Nielsen, S.A., Hansen, E., 1973. Numerical simulation of the rainfall-runoff process on a daily basis. *Nordic Hydrology* 4, 171–190.
- Rajaram, H., Georgakakos, K.P., 1989. Recursive parameter estimation of hydrologic models. *Water Resources Research* 25, 281–294.
- Rango, A., Martinec, J., 1995. Revisiting the degree-day method for snowmelt computations. *Water Resources Bulletin* 31, 657–669.
- Refsgaard, J., Storm, B., 1995. MIKE SHE. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water resources Publication, Colorado.
- Refsgaard, J.C., Rosbjerg, D., Markussen, L.M., 1983. Application of the Kalman filter to real time operation and to uncertainty analyses in hydrological modelling. In: Plate, E., Buras, N. (Eds.), *Scientific Procedures Applied to the Planning, Design and Management of Water Resources Systems Number 147* in IAHS, pp. 273–282.
- Refsgaard, J., Seth, S., Bathurst, J., Erlich, M., Storm, B., Jørgensen, G., Chandra, S., 1992. Application of the SHE to catchments in India Part 1. General results. *Journal of Hydrology* 140, 1–23.
- Restrepo, P.J., Bras, R.L., 1985. A view of maximum-likelihood estimation with large conceptual hydrologic models. *Applied Mathematics and Computation* 17, 375–404.
- Sajikumar, N., Thandaveswara, B., 1999. A non-linear rainfall-runoff model using an artificial neural network. *Journal of Hydrology* 216, 35–55.
- Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modelling. *Journal of Hydrology* 199, 272–294.
- Singh, V.P., Woolhiser, D.A., 2002. Mathematical modeling of watershed hydrology. *Journal of Hydrologic Engineering* 7, 270–292.
- Sugawara, M., 1995. Tank model. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water resources Publication, Colorado.
- Szollasi-Nagy, S., 1976. An adaptive identification and prediction algorithm for the real time forecasting of hydrological time series. *Hydrological Sciences Bulletin* 21, 163–176.
- Todini, E., 1978. Using a desk-top computer for an on-line flood warning system. *IBM Journal of Research and Development* 22, 464–471.
- Todini, E., 1996. The ARNO rainfall-runoff model. *Journal of Hydrology* 175, 339–382.
- Young, P., 1981. Parameter estimation for continuous-time models—a survey. *Automatica* 17, 23–39.
- Zhao, R.-J., 2002. The Zinanjang model applied in China. *Journal of Hydrology* 135, 371–381.
- Zhao, R.J., Liu, X., 1995. The Zinanjang model. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water Resources Publication, Colorado.

APPENDIX C

**Assessment of serious water
shortage
in the Icelandic water
resource system**

Published in *Physics and Chemistry of the Earth* 2005, Vol 30, p.420-425 .



Available online at www.sciencedirect.com



Physics and Chemistry of the Earth 30 (2005) 420–425

**PHYSICS
and CHEMISTRY
of the EARTH**

www.elsevier.com/locate/pce

Assessment of serious water shortage in the Icelandic water resource system

H. Jonsdottir^{a,*}, J. Eliasson^b, H. Madsen^a

^a Department of Informatics and Mathematical Modelling, Bldg. 321 DTU, DK-2800 Lyngby, Denmark

^b Department of Civil and Environmental Engineering, University of Iceland, Hjarðarhaga 2-6, 107 Reykjavík, Iceland

Abstract

Water resources are economically important and environmentally extremely vulnerable. The electrical power system in Iceland is hydropower based and due to the country's isolation, power import is not an option as elsewhere in Europe. In the hydropower system, a water shortage is met by flow augmentation from reservoirs. The management of these reservoirs are a human intervention in a natural flow and therefore necessarily limited by environmental regulations. During a heavy drought, the available water storage in the reservoir may not be sufficient to cater for the demand and consequently there will be a shortage of electrical power. This is politically acceptable as long as it only touches heavy industries but not power deliveries to the common market. Empty or near empty reservoirs cause power shortage that will be felt by homeowners and businesses, until spring thaw sets in and inflow to the reservoirs begins. If such a power shortage event occurs, it will cause heavy social problems and a political decision making will follow. It is commonly agreed, that management methods leading to such a disastrous event as a general power shortage in the whole country, are not acceptable. It is therefore very important to have mathematical tools to estimate the risk of water shortage in the system when searching for the best management method. In view of the fact that the subject is to estimate the risk of events that have to be very rare, i.e. with large recurrence time, stochastic simulation is used to produce synthetically run-off records with adequate length, in order to estimate very rare droughts. The method chosen is to make the run-off series stationary in the mean and the variance and simulating the resulting stationary process. When this method is chosen, future trends in the run-off from climate change and glacier reduction can easily be incorporated in the model. The probabilities of extreme droughts are calculated and their frequencies are compared to theoretical distributions.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Droughts; Stochastic simulation; Hydropower plant

1. Introduction

Computer simulations have been used to analyze the capacity of the Icelandic power system since about 1970. The simulation system has steadily been upgraded and extended to meet the various requirements for specified information on risks and capacity figures. However, simulations with stochastic flow models have not been much used so far, except for a few attempts in the years

1970–1990. One of the main questions is the risk of emptying the main reservoir and the magnitude of the following drought. Such a drought will inevitably cause a major power shortage. If this power shortage is long enough (more than a few days) it will cause serious social and economic problems such as degradation of food stocks in cold storage, operational failure of large district heating systems and immense difficulty in communication and telecommunication.

Stochastic methods have been known in hydraulic design for quite some time, e.g., Plate (1992) but they are still not extensively used in risk assessment. One of the major questions in the simulation analysis of

* Corresponding author. Tel.: +354 5696051; fax: +354 5688896.
E-mail address: hj@os.is (H. Jonsdottir).

the Icelandic power system is the performance of the reservoirs and the magnitude and the energy shortage if they run dry (Johannsson and Eliasson, 2002, 2003).

The method of using all available flow series in order to design a reservoir large enough to sustain a predefined flow output is well known in hydraulic engineering. The non-computerized graphical version can be seen e.g., in Linsley and Franzini (1964). This method is still largely used by engineers in reservoir capacity planning, but the method cannot predict the risk of water shortage. It is however, evident that the longer the inflow series is, the more reliable is the resulting volume capacity, but there is no way of presenting this knowledge in an explicit form. However simulation with stochastic flow models can provide that. To demonstrate this principle we have selected a reservoir in the river Tungnaá in Iceland, originally proposed in 1960 but not yet built.

2. Regulated flow and volume of reservoir

The major decision of a hydropower construction is how much power is to be produced. The power produced is linearly dependent on the flow. To supply the power net a flow is needed that is very constant compared to the natural flow; this constant outflow is known as the regulated flow. These decisions are made on a basis of discharge time series. Figs. 1 and 2 show the discharge in the river Tungnaá.

A diagram of a simple hydropower plant with one reservoir is shown in Fig. 3. Such a model is used by the National Power Company of Iceland in order to calculate water values in their system simulation studies (Johannsson and Eliasson, 2002).

It is clear that the maximum regulated flow is the average flow Q_{mean} of the whole series, shown in Fig. 1, and then no flow is bypassed at any point in time.

The reservoir is high up in the mountains. Its purpose is flow augmentation for a series of power stations, at lower elevations, downstream in the river basin. The

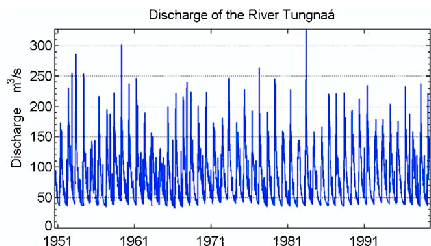


Fig. 1. Natural flow in the river Tungnaá September 1st 1951–August 31st 2001 (Nat. Energy Authority of Iceland).

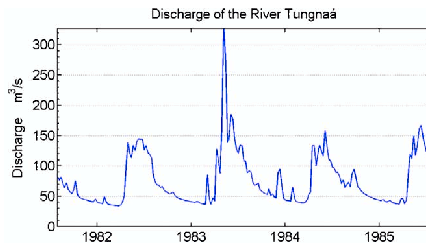


Fig. 2. Natural flow in the river Tungnaá September 1st 1981–August 31st 1985 (Nat. Energy Authority of Iceland).

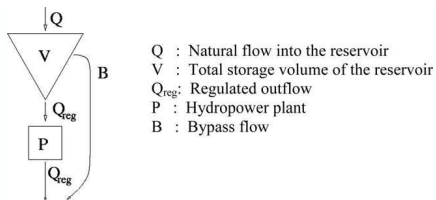


Fig. 3. Schematic drawing of reservoir (V), and power station (P).

water level in the reservoir does not affect the power capacity of any of these stations. In the following analysis it will be assumed for the sake of simplicity, that Q_{reg} is constant in time which implies that P is constant and the outflow from the power station is constant and equal to Q_{reg} . In practice Q_{reg} will be somewhat larger in wintertime than in summertime; this will increase the storage volume requirement somewhat, so the V values discussed in this article can be regarded as minimum values.

In general the water balance is calculated as the total inflow minus the total regulated flow at any given time step i.e.

$$Balance(i) = \int_0^i Q(\tau) d\tau - i \cdot Q_{reg} \tag{1}$$

Fig. 4 shows the water balance for the data in Fig. 1 with regulated flow as the average flow, i.e. $Q_{reg} = Q_{mean}$. The water balance means the balance between total inflow and total outflow. The only evaporation and rainfall to be considered is on the reservoirs surface itself and that water amount is negligible.

As mentioned the largest possible flow which can be regulated is the average flow Q_{mean} . Define V_{max} as the smallest reservoir which can serve the maximum regulated flow, Q_{mean} . Considering the time series of the water balance, $Balance(i)$, defined by Eq. (1) and shown in Fig. 4, it is clear that the volume $V_{max} =$

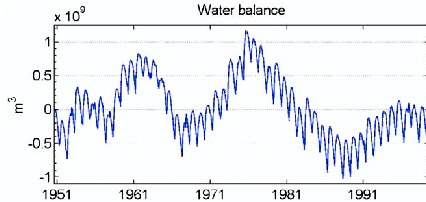


Fig. 4. The water balance.

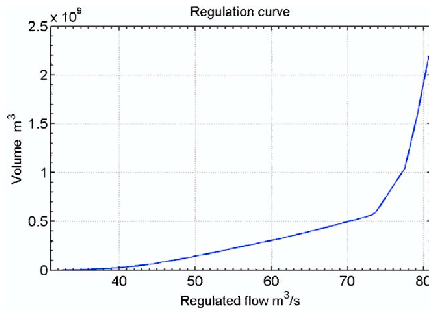


Fig. 5. Regulation curve for the Tungnaá river.

$\max[\text{Balance}(i)] - \min[\text{Balance}(i)]$. This volume serves the purpose of securing zero bypass and thus maintains the average flow as regulated flow, i.e. $Q_{\text{mean}} = Q_{\text{reg}}$. In general for a given regulated flow Q_{reg} there is a corresponding volume V which is the smallest reservoir volume that can secure Q_{reg} without any water shortage occurring according to the flow series in Fig. 1. In other words, for a given V , Q_{reg} can be calculated as the maximum regulated flow such that no water shortage occurs. Fig. 5 shows this curve for the Tungnaá data in Fig. 1, note the point $(Q_{\text{mean}}, V_{\text{max}}) = (80.7 \text{ m}^3/\text{s}, 2192.8 \times 10^6 \text{ m}^3)$.

The curve (Q_{reg}, V) is completely based on the time series for Q and from a deterministic point of view the risk of water shortage is zero, when using a point (Q_{reg}, V) from the curve. In order to estimate the risk of water shortage a stochastic model is required to perform simulation studies.

3. Stochastic model

The model chosen is a stochastic periodic model as suggested by Yevjevich (1976). The length of the period is denoted as T and number of periods is denoted as n . Let Q denote the matrix of discharge data

$$Q = \begin{pmatrix} Q(1,1) & \cdots & Q(1,T) \\ \vdots & \ddots & \vdots \\ Q(n,1) & \cdots & Q(n,T) \end{pmatrix}$$

Define $P(t)$ as a periodic mean, and $S(t)$ as a periodic standard deviation. $P(t)$ is estimated as

$$P(t) = \frac{1}{n} \sum_{j=1}^n Q(j,t) \quad t = 1, \dots, T \tag{2}$$

and similarly $S(t)$ is estimated as

$$S(t) = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (Q(j,t) - P(t))^2} \tag{3}$$

The standardized residuals are calculated as

$$Y(j,t) = \frac{Q(j,t) - P(t)}{S(t)} \quad j = 1, \dots, n \text{ and } t = 1, \dots, T. \tag{4}$$

The matrix of standardized residuals is reorganized as a row vector by $Y_{\text{vect}} = Y(1,1), \dots, Y(n,1), Y(2,1), \dots, Y(2,n), \dots, Y(T,n)$ and fitted to a seasonal AR model $\phi(B)\Phi(B^T)Y(i) = \epsilon(i)$, $i = 1, \dots, T \cdot n$. The operator ϕ is a polynomial of degree p in the backward shift operator B , i.e. $\phi(B)Y(i) = (1 - a_1B - \dots - a_pB^p)Y(i) = Y(i) - a_1Y(i-1) - \dots - a_pY(i-p)$. Similarly the operator Φ is a polynomial of degree p in the seasonal backward shift operator B^T , thus representing the seasonal component of the AR model, if needed. Then the stochastic periodic model (Yevjevich, 1976) is written as

$$\begin{aligned} \tilde{Q}(t) &= P(t) + S(t)\tilde{Y}(t) \quad t = 1, \dots, T \\ \phi(B)\Phi(B^T)\tilde{Y}(t) &= \epsilon(t) \quad \epsilon(t) \sim N(0, \sigma^2) \end{aligned} \tag{5}$$

where $\tilde{Y}(t)$ is simulated by using the seasonal AR model as in Eq. (5) and $\tilde{Q}(t)$ is the periodic discharge, simulated by using the $\tilde{Y}(t)$.

In this project a sampling time of one week is chosen. The daily discharge data are low pass filtered with a seven day average in order to decrease the variance of the data, but yet the weekly sampling time is small enough for decision making. Thus the length of the period is 52 time steps, and the data available span 49 seasonal periods.

To ensure that physical laws are conserved, such as nonnegative flow it was chosen to transform the data by using the logarithm base (\log_e) of the data. For the transformed data the appropriate model was found i.e.

$$\begin{aligned} \tilde{Q}_{\log}(t) &= P_{\log}(t) + S_{\log}(t)\tilde{Y}_{\log}(t) \quad t = 1, \dots, T \\ \phi(B)\Phi(B^T)\tilde{Y}_{\log}(t) &= \epsilon(t) \quad \epsilon(t) \sim N(0, \sigma^2) \end{aligned} \tag{6}$$

where $P_{\log}(t)$ and $S_{\log}(t)$ are calculated from the log-transformed data. Afterwards the residuals $Y_{\log}(t) = (Q_{\log}(t) - P_{\log}(t))/S_{\log}(t)$ are modelled. This transformation implies that the residuals in the original model $\epsilon(t)$

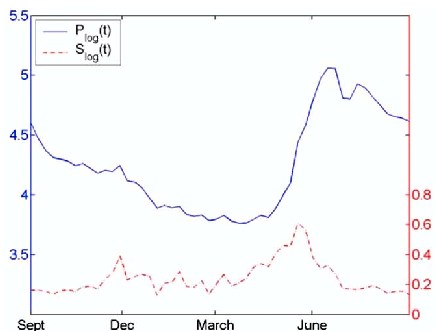


Fig. 6. The periodic average $P_{\log}(t)$ and the periodic standard deviation $S_{\log}(t)$ of the log-transformed data.

as defined in Eq. (5) are lognormal distributed, which indeed has a physical meaning. The calculations are performed in the program Splus and the result is as follows:

$$(1 - 0.62B)Y_{\log}(t) = \varepsilon(t) \quad \varepsilon(t) \sim N(0, 0.06) \quad (7)$$

i.e. an AR(1) model without a seasonal component. Fig. 6 shows the curves $P_{\log}(t)$ and $S_{\log}(t)$. Note the similarity between the derivative of the flow (i.e. $\nabla P_{\log}(t)$) and the standard deviation (i.e. $S_{\log}(t)$).

4. The simulation study

The simulation was performed using the theoretical model in Eq. (6) where the process $Y_{\log}(t)$ is an AR(1) model without a seasonal component as estimated in Eq. (7), the resulting simulated series is denoted $Q_{\text{sim}}(t)$. Two droughts are defined as independent if they either occur in two different years or if they occur in the same year and the reservoir is refilled between the two droughts, two or more dependent droughts are grouped together in single independent droughts. The water shortage is calculated as the total shortage of water within a single independent drought. Let X denote the total water shortage within a drought, then the required probability is the probability of a water shortage larger than x , i.e. $P(X \geq x)$ where X denotes the random variable of water shortage. Simulations were performed for several pairs (Q_{reg}, V) on the regulation curve shown in Fig. 5. The goal is to estimate probabilities of events that are very rare and it was found necessary to simulate for 50,000 years in order to achieve a stable estimate of the water shortage probabilities. Note that the simulations of 50,000 years does not imply prediction 50,000 years into the future but a stochastic generation for 50,000 years given that the weather condition will be like the past 50 years which were used for parameter estimation

in the stochastic model. However, since the deviation series $Y(t)$ is stationary in mean and variance, climate change predictions for future trends in runoff series such as glacier melt can be taken into account either as deterministic or stochastic variables depending on the climate change model output.

None of the simulation results included events with two independent droughts within the same year thus the probabilities of water shortage is estimated as

$$P(X \geq x) = \frac{\text{Number of years with water shortage larger or equal than } x}{\text{Number of years in simulation}} \quad (8)$$

Thus there are estimated n probability values p_1, \dots, p_n , where n is the total number of droughts which occurred in the simulation and

$$p_j = P(X \geq x_j) \quad (9)$$

where x_j is the j th largest water shortage, thus $p_1 = 1/50,000, p_2 = 2/50,000, \dots, p_n = n/50,000$.

Note that $P(X \geq x) \approx P(X > x) = 1 - F(x)$, where $F(x)$ is the probability distribution function. The presented results are from simulations where the reservoir is $1315.7 \times 10^6 \text{ m}^3$ and regulated flow is $78.36 \text{ m}^3/\text{s}$, which is a point on the regulation curve illustrated in Fig. 5. The simulations were repeated 100 times in order to obtain information about the variations. Consequently for each probability p_j there correspond 100 different x values $x(j, 1), x(j, 2), \dots, x(j, 100)$, which yielded the estimate $p_j = P(X \geq x(j, i)), i = 1, \dots, 100$. The conditional distribution of the random variable $\{x(j, i)|p(j)\}$ is assumed to be a normal distribution $N(\mu(j), \sigma^2(j))$. The mean, $\mu(j)$, and the variance $\sigma^2(j)$ are estimated (using 100 observations) with the maximum likelihood method and the estimated mean $\hat{\mu}(j)$ will be denoted as x_{mean} and the pairs $(x_{\text{mean}}(i), p(i))$ will be referred to as the simulation result. The generalized extreme value distribution (GEV), is fitted to the accumulated probabilities $(1 - p)$. The GEV distribution can be parameterized as (Reiss and Thomas, 1997)

$$P(X \geq x) = \text{GEV}_{\alpha, \beta, \xi}(x) = \exp \left[- \left(1 + \xi \frac{x - \alpha}{\beta} \right)^{\frac{1}{\xi}} \right] \quad (10)$$

The parameters are estimated using the least square method and the result is shown in Table 1.

Fig. 7 shows the (x_{mean}, p) , the 2.5% quantile, $(x_{0.025}, p)$, the 97.5% quantile, $(x_{0.975}, p)$ and the fitted distribution $(x_{\text{mean}}, p_{fit})$, i.e. $1 - \text{GEV}(x)$.

The difference of the simulation result and the fitted probabilities can hardly be detected, but from probabilities around 0.1% (i.e. recurrence time 1000 years) the fitted distribution converges to zero faster than the simulated probabilities. This can better be detected in a

Table 1
Estimated parameters in the GEV distribution using the least square method

	α	β	ξ
Estimation	-897.95	376.86	-0.2744

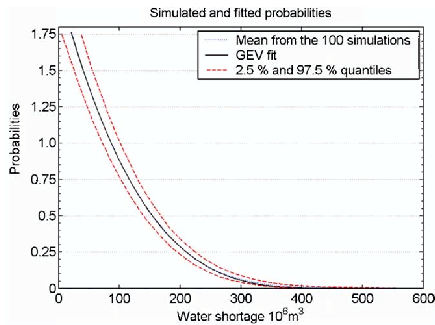


Fig. 7. The estimated mean value of the repeated simulations, approximated 2.5% and approximated 97.5% quantiles and the fitted GEV distribution.

quantile diagram. There is an interest in the tail of the distribution and consequently it was chosen to divide the data into as many intervals as reasonable. Hence, 160 intervals with equal probabilities are generated, with the expected number of observations in each interval as 5.48. Fig. 8 shows a quantile diagram for the GEV(x) distribution, with $(\alpha, \beta, \xi) = (-897.95, 376.86, -0.2744)$ and estimated probabilities from histogram with the 160 intervals. The chi-square test statistics for test of distribution using the same intervals is $z = 12.0228$, and the

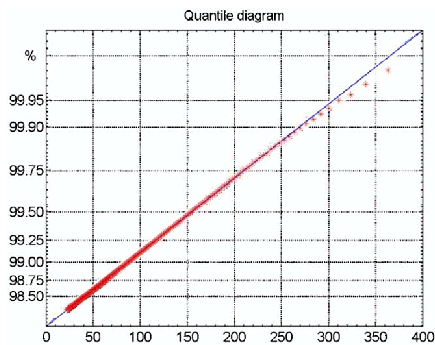


Fig. 8. Quantile diagram for the data compared to the GEV distribution with the parameter estimate according to Table 1 $(\alpha, \beta, \xi) = (-897.95, 376.86, -0.2744)$.

90% quantile is $\chi(156)_{0.9}^2 = 179$ and it follows that a hypotheses that the water shortage is GEV distributed with the estimated parameters as shown in Table 1, $(\alpha, \beta, \xi) = (-897.95, 376.86, -0.2744)$ is accepted. However, it must be kept in mind that the simulations are necessary in order to estimate the parameters.

The shape parameter ξ in the generalized extreme value distribution has been estimated as negative, thus the extreme value distribution is identified as the Weibull distribution (Reiss and Thomas, 1997).

The domain the of distribution is the interval $]-\infty, -\beta/\xi + \alpha] =]-\infty, 475.23]$. Note that about 98% of the probability mass is below zero, i.e. with negative x values. The mean and the standard deviation of the distribution are (Reiss and Thomas, 1997):

$$E(x) = \mu = \alpha - \frac{\beta}{\xi} + \frac{\beta}{\xi} \Gamma(1 - \xi) = -762.96 \quad (11)$$

$$V(x) = \sigma^2 = \left(\frac{-\beta}{\xi}\right)^2 \Gamma(1 - 2\xi) + \Gamma^2(1 - \xi) = 1791.39^2 \quad (12)$$

The Weibull distribution can be re-parameterized with domain $]-475.27, \infty[$. Then the distribution function is reflected about the y -axis and the random variables multiplied by -1 . Hence, the distribution function becomes

$$1 - \exp\left(-\left(1 + \xi \frac{-x - \alpha}{\beta}\right)^{-1/\xi}\right) \text{ with } x \geq -(-\beta/\xi + \alpha) \quad (13)$$

setting $k = -1/\xi > 0$, $b = k\beta = -\beta/\xi$ and $a = -(\alpha + k\beta) = -\alpha + \beta/\xi$ this becomes

$$1 - \exp\left(-\left(\frac{x - a}{b}\right)^k\right) \text{ with } x \geq a \quad (14)$$

which is a more commonly used parameterization in hydrology. Using this interpretation the random variable is interpreted as the result of the water balance equation

$$X(i) = \int_0^i Q_{\text{sim}}(\tau) d\tau - i \cdot Q_{\text{reg}} \quad (15)$$

and water shortage will occur if $X(i)$ is negative. The stochastic formulation of a water shortage, using this interpretation, is a peak below threshold study, with threshold zero, see e.g. Medova and Kyriacou (2000). On the other hand for practical purposes it is more convenient to work with water shortages as positive variables with decreasing probabilities. (Note that the dynamic variable $X(t)$ in Eq. (15) is an unstable time series i.e. with pole equal to one, since $(1 - B)X(t) = u(t)$ with $u(t) = Q_{\text{sim}}(t) - Q_{\text{reg}}$).

Using the quantile diagram in Fig. 8, the probability of water shortage of 155 million m^3 is 0.5% and thus the recurrence time for a water shortage of 155 million m^3 is

200 years or larger. A water shortage of 155 million m³ means that the power station is out of operation for about three weeks. There is a 15–30% probability that a large drought like that will occur in the economical lifetime of the project, which is 30–60 years for hydropower stations in Iceland.

5. Conclusions

The risk of water shortage in a hydropower plant has been estimated through stochastic modeling and simulations. In general the available data are used for design of a hydropower plant. Thus the recurrence time of drought is large and therefore a very long time series is needed in order to estimate the drought risk. The stochastic simulations produce a time series long enough for achieving an estimate of the probability distribution function of a water shortage. Furthermore the repeated simulations provide an estimate of uncertainty of the probability function estimation. The simulation study in this project was performed for a simple system with one hydropower plant and one reservoir, but using the tools already developed for power system studies in Iceland, it is straightforward to extend the model for more complicated systems.

As mentioned the recurrence time for water shortage of 155 million m³ is estimated to be 200 years or larger, which means the power station is out of operation for about three weeks. For a water power station with a life time of 50 years, there is a probability of 25% that water shortage of this magnitude will occur in the economical

lifetime. On top of that, there is a great probability that water shortage will occur in other reservoirs as well due to spatial correlation in Icelandic run-off data. A power failure of this magnitude will most likely be considered socially and politically unacceptable with disastrous consequences for power system management practices.

Acknowledgment

The authors wish to thank The National Energy Authority in Iceland for providing the discharge data used in this study.

References

- Johannsson, S., Eliasson, E.B., 2002. Simulation model of the hydrothermal power system in Iceland. Technical Report, Annad veldi ehf. Available from: <www.veldi.is>.
- Johannsson, S., Eliasson, E.B., 2003. Simulation model of transmission constrained hydrothermal power system. Technical Report, Annad veldi ehf. Available from: <www.veldi.is>.
- Linsley, R.K., Franzini, J., 1964. *Water Resources Engineering*. McGraw-Hill, New York.
- Medova, E., Kyriacou, M., 2000. Extreme values and the measurement of operational risk. *Oper. Risk* 1 (8), 12–15.
- Plate, E.J., 1992. Stochastic design in hydraulics. In: *Proceedings of the International Symposium on Stochastic Hydraulics*.
- Reiss, R.-D., Thomas, M., 1997. *Statistical Analysis of Extreme Values*. Birkhäuser Verlag, Switzerland.
- Yevjevich, V., 1976. Structure of natural hydrologic time series. In: Shen, H.W. (Ed.), *Stochastic Approaches to Water Resources*, Colorado.

APPENDIX D

A grey box model describing the hydraulics in a creek

Published in *Environmetrics* 2001 p. 347-356.

A grey-box model describing the hydraulics in a creek

Harpa Jónsdóttir*, Judith L. Jacobsen and Henrik Madsen

Department of Mathematical Modelling, Bldg. 321 DTU, DK-2800 Lyngby, Denmark

SUMMARY

The Saint-Venant equation of mass balance is used here to derive a stochastic lumped model describing the dynamics of the flow in a river. The flow dynamics are described by the evolution of the cross-sectional area of the flow at two locations in the river. The unknown parameters of the model are estimated by combining the physical equations with a set of data. This method is known as grey-box modelling. The data consist of water level measurements, taken every minute at two locations in a river, over a period of nine days. The data are sub-sampled to a sampling period of 15 minutes before further processing, and a maximum likelihood method is used to estimate the parameters of the model.

Three different models were applied to the data set. All three are linear reservoir models with an estimate of the dynamic lateral inflow as a function of precipitation. The first model is a single reservoir model, which proved to be too simple to adequately describe the effect of precipitation. The second model is also a single reservoir model, but the data from the downstream station were translated forward in time, corresponding to a time delay in the system (a retention time). This model responds in a physically reasonable manner to precipitation, capturing very well the flow peaks caused by rain events. The third model is based on two reservoirs, and like the second model, it responds reasonably to precipitation. Its description of the dynamics seems quite good, though it does not capture the flow peaks quite as well as the second model. However, it is shown, that this model statistically provides the best description of the system. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS: grey-box modelling; maximum likelihood estimation; linear reservoir; stochastic hydraulic model.

1. INTRODUCTION

The water quality of a river is mainly affected by the chemical composition of the basin. Chemicals are carried into the river either in dissolved or in particulate form and, in most cases, the chemical profile of the inflow differs from that of the river itself, thus affecting the flora of the river. An adequate modelling of the impact of varying chemical composition on the ecology frequently requires a hydraulic model, which permits estimates of the influx into the system.

In the present work, three hydraulic models are applied and compared. We have used the grey-box approach to establish the models, which are simple lumped models along with stochastic terms, resulting in stochastic ordinary differential equations. The lateral inflow is described as a function of precipitation, and by using Kalman filtering it is possible to perform an estimation of the inflow.

*Correspondence to: H. Jónsdóttir, Department of Mathematical Modelling, Building. 321 DTU, DK-2800 Lyngby, Denmark.

The data were obtained from a slowly flowing creek, and the ultimate purpose of the data collection was to model the water quality by considering the oxygen level. In order to set up such a model, the hydraulics must be reasonably described. This is facilitated by the hydraulic models proposed in this paper. The only measured variable related to water volume is the depth, and therefore our model is not a true water flow model as we do not know the water flow nor the so-called $Q-h$ relations. Instead, the models describe the evolution of the cross-sectional area, because it is more related to flow than the depth.

Three linear reservoir models are tested. The first one is a single reservoir model. The second one is also a single reservoir model but with a time delay (the retention time), and the third model contains two reservoirs. Whitehead and Young (1975) developed a grey-box model for stream flow forecasting, using a deterministic single reservoir model and a black-box rainfall runoff model for the remaining variations. Young and Beck (1974) and Beck and Young (1975) used a linear reservoir model with one reservoir and a transportation delay in a BOD-DO model. Furthermore, Jacobson *et al.* (1997) modelled the water level using a stochastic lumped model with three reservoirs, formulated by stochastic differential equations.

2. THE DATA

The data originate from a creek in North Zealand, approximately 25 km north of Copenhagen. Water level measurements were obtained from two measuring stations which are about 2 km apart. The upstream and downstream stations are referred to as U and D, respectively. There are two rainfall runoffs near the upstream station, one located about 10 m further downstream and the other 2 m upstream. The precipitation measuring station is located approximately 2.5 km upstream from station U. Figure 1 shows a sketch of the area and the measuring stations. The data span a period of nine days in August 1996, with a sampling period of one minute. The goal is to model the overall dynamics, especially the increased water flow during rain events. Data have been low-pass filtered and sub-sampled by averaging over 15-minute periods. The choice of this sampling period was made by comparing the data with the phenomena of interest. The data indicate a retention time of approximately one hour between the two stations. To minimize the risk of information loss, a sampling period of 15 minutes was chosen. The equations below are based on mass balance wherefore it was decided to use the cross sectional areas rather than the water level in the modelling work. Figure 2 shows the cross-sectional areas and the rain intensity.

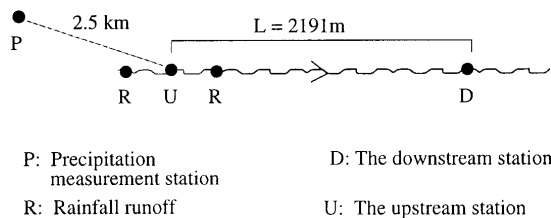


Figure 1. Overview of the area

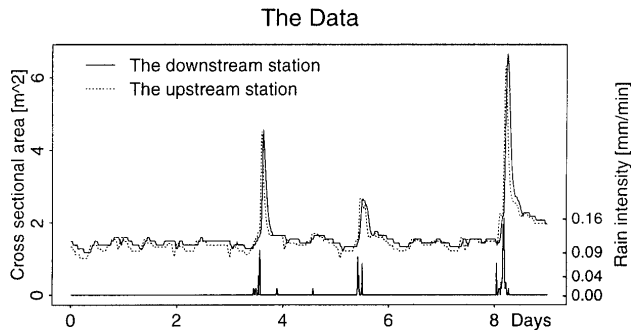


Figure 2. The cross-sectional areas and the rain intensity

3. THE MATHEMATICAL MODEL

The conservation of mass is used to derive a stochastic lumped model describing variations in time of the cross-sectional areas. Our goal is to develop a hydraulic model which can be used as a basis for a water quality model, for example, a model for the oxygen concentration. The only data available related to the hydraulics are water level measurements, so in order to develop a physical model, some simplifications must be made. The flow in a river can be described by the Saint-Venant equation

$$\frac{\partial A(x, t)}{\partial t} + \frac{\partial Q(x, t)}{\partial x} = q(x, t) \quad (1)$$

where A is the cross-sectional area (m^2), Q is the flow of water (m^3/min), q is the lateral inflow per unit length ($(m^3/\text{min})/m = (m^2/\text{min})$), x is the location along the river (m), and t is the time (min).

Measurements of A are available at two locations in the creek. The flow can be expressed as $Q(x, t) = v(x, t) A(x, t)$, where $v(x, t)$ is the average cross-sectional velocity. A constant retention time is assumed. This assumption may be questionable, since an increased water level probably leads to an increase in the average velocity, resulting in a shorter retention time. However, since a period of nine days with relatively small rain events is considered, this simplification is probably reasonable. The velocity $v(x, t)$ is set to be a constant average transportation velocity $v = L/s$, where L is the distance between the two stations and s is the retention time. Equation (1) can thus be rewritten as

$$\frac{\partial}{\partial t} A(x, t) + \frac{L}{s} \frac{\partial}{\partial x} A(x, t) = q(x, t) \quad (2)$$

The next step is to discretize the mass balance equation by $\frac{\partial}{\partial x} A(x, t) \approx \frac{\Delta A(x, t)}{\Delta x}$ and setting x equal to the location of the downstream station and let Δx be the distance L between the two stations. This leads to

$$\frac{\partial}{\partial t} A(x_D, t) + \frac{A(x_D, t) - A(x_U, t)}{s} = q(x_D, t) \quad (3)$$

where x_D is the location of the downstream station and x_U is the location of the upstream station. Using the rule of total differential¹ in Equation (3) and, since this is an Euler description, $\frac{\partial}{\partial t}$ can be replaced with $\frac{d}{dt}$, resulting in

$$\frac{d}{dt}A(x_D, t) = A(x_U, t)/s - A(x_D, t)/s + q(x_D, t) \quad (4)$$

This equation can also be obtained by simplifying a linear reservoir model, as in Jacobson *et al.* (1997), where a similar model was obtained using that approach. We shall therefore refer to Equation (4) as the single reservoir model. Let the lateral inflow be a function of precipitation, described by a first-order process,

$$\frac{d}{dt}q(t) = aq(t) + bP(t) + k \quad (5)$$

where $P(t)$ is the precipitation. In some situations it might be necessary to introduce a time delay in Equation (5) between the rain event and the inflow by using $P(t - \tau)$. This is not the case here, since the precipitation has an almost immediate impact on the lateral inflow through the rainfall runoffs. Finally, a noise term is added to the equations and thereby a stochastic model is obtained. The resulting single reservoir model is expressed in matrix notation as

$$\begin{bmatrix} dq(t) \\ dA_D(t) \end{bmatrix} = \begin{bmatrix} a & 0 \\ 1 & -1/s \end{bmatrix} \begin{bmatrix} q(t) \\ A_D(t) \end{bmatrix} dt + \begin{bmatrix} b & 0 & k \\ 0 & 1/s & 0 \end{bmatrix} \begin{bmatrix} P(t) \\ A_U(t) \\ 1 \end{bmatrix} dt + \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix} \quad (6)$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} q(t) \\ A_D(t) \end{bmatrix} + e(t) \quad (7)$$

where the last equation is introduced to describe the fact that only the cross-sectional area is measured, and the measurement $y(t)$ is encountered with the measurement noise $e(t)$. In an abbreviated notation, the equations are written as the following continuous discrete-time state space model,

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)dt + \mathbf{B}\mathbf{U}(t)dt + d\mathbf{w}(t) \quad (8)$$

$$\mathbf{Y}(t) = \mathbf{C}\mathbf{X}(t) + e(t) \quad (9)$$

Equation (8) is known as the system equation which describes the dynamics of the physical system. The vector \mathbf{X} is the state vector and \mathbf{U} is a vector of inputs. The matrix \mathbf{A} characterizes the dynamic behaviour of the system and the matrix \mathbf{B} specifies how the input signals enter the system. The noise term $d\mathbf{w}(t)$ is the model error, and it is assumed to be a diagonal Wiener process, i.e., the increments are mutually independent and Gaussian distributed. In our formulation, the cross-sectional area at the downstream station and the lateral inflow are state variables, whereas the precipitation and the cross-sectional area at the upstream station are inputs. Equation (9) is known as the measurement equation. The vector $\mathbf{Y}(t)$ contains the measured variables and the matrix \mathbf{C} specifies which linear combination

¹ $\frac{d}{dt}A(x_D, t) = \frac{\partial}{\partial x}A(x_D, t)\frac{dx}{dt} + \frac{\partial}{\partial t}A(x_D, t)\frac{dt}{dt}$

of the states actually are measured. The vector \mathbf{e} contains the measurement errors which are assumed to be mutually independent Gaussian distributed and independent of the model error $d\mathbf{w}$.

Inspection of Equations (6)–(7) reveals that the parameter s has the unit minutes and is interpreted as the retention time. The parameter b is independent of time, with the units m^2/mm . 1 mm of rain will thus result immediately in $b\text{m}^2$ of lateral inflow, while the extra inflow due to rain declines as described by a/min .

4. PARAMETER ESTIMATION AND MODEL VALIDATION

In this section an outline of the estimation procedure is presented, and some validation methods are described.

The unknown parameters are estimated using a maximum likelihood method. A test of significance of the estimated parameters can therefore be performed by using a t -test. It is well known that the likelihood function for time series data becomes a product of conditional densities. Based on the fact that the model in Equations (6)–(7) is linear, and the assumptions about the Gaussian nature of the error components in Equations (8) and (9), it is easily shown that the conditional densities are Gaussian. The Gaussian distribution is completely characterized by the conditional mean and the conditional variance, which can be calculated recursively by the Kalman filter (see, for example, Harvey (1994)).

Because the data are given in discrete time, the stochastic differential equations have to be integrated through the sampling interval in order to evaluate the likelihood function. Given the constant sampling period τ , the solution to Equations (8) is

$$\mathbf{X}(t + \tau) = e^{A\tau} \left(\mathbf{X}(t) + \int_t^{t+\tau} \mathbf{B}U(s)e^{-(s-t)A} ds + \int_t^{t+\tau} e^{-(s-t)A} d\mathbf{w}(s) \right) \quad (10)$$

(see Kloeden and Platen (1992) for a solution of stochastic differential equations in the form of Equation (8)). Since the matrix \mathbf{B} is a constant, the expression can be simplified by moving \mathbf{B} outside the integral.

The input $U(s)$ is measured only at discrete time points, t and $t + \tau$. For all t , the input U inside the interval $[t, t + \tau]$ is obtained by linear interpolation. Equation (10) can then be written as

$$\mathbf{X}(t + \tau) = \boldsymbol{\Phi}(\tau)\mathbf{X}(t) + \mathbf{B}\boldsymbol{\Psi}(\tau)U(t) + \mathbf{B}\boldsymbol{\Psi}(\tau)(U(t + \tau) - U(t)) + \mathbf{v}(t, \tau) \quad (11)$$

where

$$\begin{aligned} \boldsymbol{\Phi}(\tau) &= e^{A\tau}; \quad \boldsymbol{\Psi}(\tau) = \int_0^\tau e^{-(r-\tau)A} dr; \quad \boldsymbol{\Psi}(\tau) = \int_0^\tau e^{-(r-\tau)A} \frac{r}{\tau} dr \\ \mathbf{v}(t, \tau) &= \int_t^{t+\tau} e^{(t+\tau-s)A} d\mathbf{w}(s). \end{aligned} \quad (12)$$

Equation (11) is the system equation of the discrete-discrete state space model originated from the continuous-discrete time space model. The estimation procedure has been implemented in a program called CTLSM (continuous time linear stochastic modelling). See Melgaard and Madsen

(1993) for a description of the program, and Jacobsen and Madsen (1996) for a further description of the method.

The main purpose of validating a model must be to verify whether or not the model describes the physical phenomena, and in a statistical approach, whether the model describes the data. We concentrate on the residuals, which are the one-step prediction errors. The assumptions on the error terms previously mentioned should lead to white noise residuals. Hence, all the well-known tests for white noise residuals from time series analysis can be used (see, e.g., Box and Jenkins (1976)).

5. RESULTS AND DISCUSSION

In this section the results corresponding to the suggested model Equations (6)–(7) are first discussed. These results point to some extensions of the model which are then introduced. The results of the parameter estimation are shown in Table I. The results in the first column correspond to the single reservoir model given by Equations (6)–(7). The parameter b is estimated to be less than zero, which means that the lateral inflow abates and actually becomes negative because of rain. Physical intuition tells us that this cannot be true, and the conclusion is that Equations (6)–(7) are too simple to describe the system. The model must therefore be modified and this is done in two ways. The first modification is as follows: instead of using the input (the cross-sectional area at the upstream station and the precipitation) at time t , the input at time $t - s$ is considered where s is the retention time. The same method is used by Young and Beck (1974) in a DO-BOD model. In the second modification, an unobserved state between the upstream station and the downstream station is introduced, i.e., the reservoir model is extended to contain two reservoirs. Jacobsen *et al.* (1997) presented a grey-box model with several reservoirs using the same data set. The main difference between the models here and the models presented in Jacobsen *et al.* (1997) is that here we have the non-observed lateral inflow as a state variable. Table I shows the parameter estimates for all the models. It is readily seen that all the parameters are significantly different from zero.

Each model will now be discussed. As mentioned before, using the single reservoir model Equations (6)–(7), the ‘precipitation response’ parameter b is estimated to be negative. The explanation for this is as follows. From Equations (6)–(7) it is seen that the lateral inflow $q(t)$ is a function of precipitation, but it also depends on the value of $A_D(t) - A_U(t)$. Figure 3 shows this difference. It is seen that $A_D(t) - A_U(t)$ is negative at the beginning of the rain events. The reason for this is that almost all the lateral inflow enters the system through the rainfall runoffs located

Table I. Maximum likelihood estimates of the unknown parameters, the numbers in parentheses are the standard deviation of the parameter estimates. All the parameters are significantly different from zero

Parameter	Single reservoir (Equations (6)–(7))	Single reservoir with delay (Equations (13)–(14))	Two reservoirs (Equations (16)–(17))	Units
a	– 0.0293 (0.0042)	– 0.0411 (0.0032)	– 0.0115 (0.0026)	(1/min)
b	– 0.0031 (0.0011)	0.0271 (0.0033)	0.0044 (0.0011)	(m ² /mm)
s	66.603 (3.635)	22.769 (2.0992)	61.33 (1.72)	(min)
k	3.5×10^{-5} (1.1×10^{-5})	1.04×10^{-4} (1.78×10^{-5})	2.3×10^{-5} (7.3×10^{-6})	(m ² /mm)

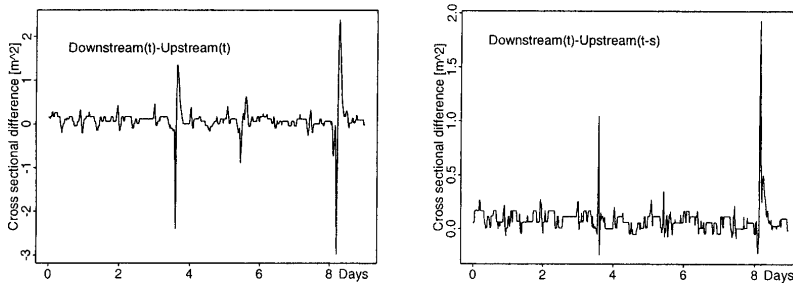


Figure 3. The cross-section difference. To the left is the actual difference $A_D(t) - A_U(t)$ and to the right is the delayed difference, $A_D(t) - A_U(t - s)$, where the retention time s is set to one hour

near the upstream station, and it takes some time for this extra amount of water to reach the downstream station. Hence, the model (6)–(7) obtains the best fit to the data by a negative response to rain.

Figure 4 shows the measured and simulated values of the cross-sectional area and the estimate of the non-observed lateral inflow. Considering the single reservoir model defined by Equations (6)–(7), it is seen that the simulation results are quite accurate under stationary conditions, but the peaks due to rain are not adequately described. The estimated lateral inflow $q(t)$ is far from being realistic. The variations are too sharp and the peaks are too narrow compared with the peaks of the cross-sectional area.

The single reservoir model with the time delay is considered next. In the previous model the retention time was estimated to be 63 minutes. The data intervals are 15 minutes, and consequently the delay is approximated by 4 time steps. Figure 3 shows the cross-sectional difference $A_D(t) - A_U(t - 60)$. This difference does not reach as large negative values as the undelayed difference $A_D(t) - A_U(t)$. The model becomes:

$$\begin{bmatrix} dq(t) \\ dA_D(t) \end{bmatrix} = \begin{bmatrix} a & 0 \\ 1 & -1/s \end{bmatrix} \begin{bmatrix} q(t) \\ A_D(t) \end{bmatrix} dt + \begin{bmatrix} b & 0 & k \\ 0 & 1/s & 0 \end{bmatrix} \begin{bmatrix} P(t - 60) \\ A_U(t - 60) \\ 1 \end{bmatrix} dt + \begin{bmatrix} dw_1(t) \\ dw_2(t) \end{bmatrix} \quad (13)$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} q(t) \\ A_D(t) \end{bmatrix} + e(t) \quad (14)$$

The parameter s can no longer be interpreted as a retention time. The ‘precipitation response’ parameter b is now estimated to be positive, i.e., the model responds to rain by a positive lateral inflow. Figure 4 shows the measured and simulated values of the cross-sectional area, and the estimated lateral inflow $q(t)$. The simulation captures the peaks very well, but still the lateral inflow does not behave as intuitively expected. As in the previous model, the variations are too sharp. Note that the estimated values of $q(t)$ during rain periods are much higher than the in previous model. This is because $q(t)$ has a different interpretation. Here $q(t)$ depends on the value of $A_D(t) - A_U(t - s)$ and this value actually represents the integrated lateral inflow during the retention time. This can be seen by solving

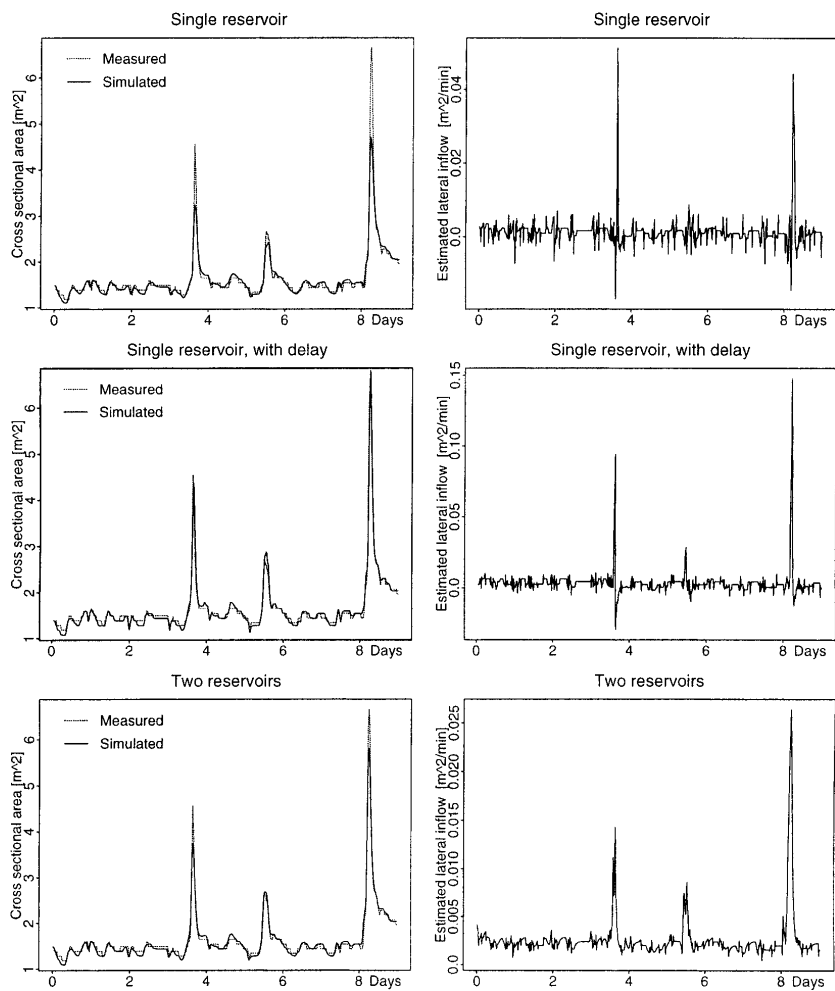


Figure 4. To the left are measurements and simulations using the three models and to the right are the estimated lateral inflows $q(t)$

the Saint-Venant equation of mass balance along with all our simplifications (i.e., Equation (2)), which gives the relation

$$A_D(t) = A_U(t - s) + \int_{t-s}^t q(\tau) d\tau \tag{15}$$

Finally the single reservoir model Equations (6)–(7) is extended to a two reservoir model. This is done by introducing a virtual station between the two measurement stations, denoted $A_F(t)$. The equations then become:

$$\begin{bmatrix} dq(t) \\ dA_F(t) \\ dA_D(t) \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ 1 & -2/s & 0 \\ 0 & 2/s & -2/s \end{bmatrix} \begin{bmatrix} q(t) \\ A_F(t) \\ A_D(t) \end{bmatrix} dt + \begin{bmatrix} b & 0 & k \\ 0 & 2/s & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P(t) \\ A_U(t) \\ 1 \end{bmatrix} dt + \begin{bmatrix} dw_1(t) \\ dw_2(t) \\ dw_3(t) \end{bmatrix} \tag{16}$$

$$y(t) = [0 \quad 0 \quad 1] \begin{bmatrix} q(t) \\ A_F(t) \\ A_D(t) \end{bmatrix} + e(t) \tag{17}$$

We have chosen to let the lateral inflow $q(t)$ affect only the first reservoir, since it is known that almost all the transient inflow between the two stations enters the river stretch close to the upstream station. Notice that this is another way of dealing with the time delay. Figure 4 shows the measured cross-sectional area along with the simulation and the estimated lateral inflow. The simulation is quite good but it does not capture the peaks as well as the single reservoir model with time delay. However, the lateral inflow $q(t)$ seems realistic, since it varies around some value close to zero and then increases when it rains. The duration of the extended inflow corresponds to the duration of the increased cross-sectional area. The parameter b is estimated to be positive, which indicates that the system responds to rain by a positive inflow.

For a comparison of the models, several model validation statistics were carried out, and these are shown in Table II. The correlations in time of the single-step prediction errors are studied in the frequency domain but are not shown in the table. A white noise error is equally distributed in the frequency domain and by using a Kolmogorov–Smirnov test in the frequency domain, it was concluded that none of the model errors are white noise processes, although the model with two reservoirs is quite close. White noise test statistics have also been performed by using the Portmanteau lack-of-fit test for white noise, see e.g., Box and Jenkins (1976), as shown in the bottom row of Table II.

Table II. The following model validation statistics are shown: the mean value of the one step prediction error and its standard deviation, the mean value of the simulation error and the Portmanteau lack-of-fit test statistic

	Single reservoir Equations (6)–(7)	Single reservoir with delay Equations (13)–(14)	Two reservoirs Equations (16)–(17)	Units
Mean of prediction error	0.0012	0.0011	0.0006	(m ²)
Variance of prediction error	0.0012	0.0014	0.0010	(m ²)
Mean of simulation error	0.0187	0.0065	0.0066	(m ²)
Portmanteau test statistic	73.33	203	67.45	

The single reservoir model with a time delay is far from having a white noise prediction error, whereas the simple single reservoir model is much closer. The mean of the prediction error is smallest for the two reservoir model, and even though the simulation seemed to be best in the single reservoir model with time delay, the mean value of the simulation error of the two reservoir model is almost the same. From a statistical point of view, the two reservoir model is surely the best. This is also the case from a physical point of view, since this model is the only one capable of estimating the lateral inflow $q(t)$ reasonably well. Even though there are some unexplained dynamics in the two reservoir model, it is expected that for most practical purposes it is acceptable.

6. CONCLUSION

In this paper, three lumped parameter models describing the variation in the cross-sectional area in a creek are established. These are a single reservoir model, a single reservoir model with a time delay and a two reservoir model. All the models are based on the mass balance equation, and the main simplification is that the retention time is assumed to be constant. All the models provide a dynamic estimate of the lateral inflow, which was not measured, this can be very useful in an environmental context. A comparison of the models shows that the two reservoir model gives a better description than the single reservoir models. The prediction error is close to white noise, which was not the case for the other two models, and the mean value of the prediction error is lowest. The two reservoir model is the only model which provides a physically consistent estimate of the lateral inflow. It is concluded that for most practical purposes the two reservoir model describes the data reasonably well.

In this study it has been found advantageous to use the grey-box approach. Contrary to black-box models, the parameters of a grey-box model have a physical meaning. In contrast to traditional physical modelling, or white-box modelling, we have been able to estimate the coefficients of the differential equations. Furthermore, the stochastic approach makes it possible to provide uncertainty bounds on predictions.

ACKNOWLEDGEMENTS

We thank PH-consult and the Hørsholm City Council for providing data.

REFERENCES

- Beck MB, Young PC. 1975. A dynamic model for DO-BOD relationships in a non-tidal stream, *Water Research* **9**: 769–776.
- Box GEP, Jenkins GM. 1976. *Time Series Analysis, Forecasting and Control*; Holden-Day: San Francisco.
- Harvey AC. 1994. *Forecasting, Structural Time Series Models and the Kalman Filter*; Cambridge University Press: Cambridge.
- Jacobsen JL, Madsen H. 1996. Grey Box Modelling of oxygen levels in a small stream. *Environmetrics* **7**: 109–121.
- Jacobsen JL, Madsen H, Harremoes P. 1997. A stochastic model for two-station hydraulics exhibiting transient impact. *Water, Science and Technology* **36**(5): 19–26.
- Kloeden PE, Platen E. 1992. *Numerical Solutions of Stochastic Differential Equations*, 2nd edn. Applications of Mathematics, Stochastic Modelling and Applied Probability, Springer-Verlag: Heidelberg.
- Melgaard H, Madsen H. 1993. CTLSM version 2.6 – a program for parameter estimation in stochastic differential equations, Technical Report No. 1/1993, IMM, Building 321, DTU, DK-2800 Lyngby.
- Whitehead PG, Young PC. 1975. A dynamic-stochastic model for Water quality in part of the Bedford–Ouse river system. In *Computer Simulation of Water Resources Systems*, Vansteenkiste GC (ed). North Holland: Amsterdam; 417–438.
- Young PC, Beck MB. 1974. The modelling and control of water quality in a river system. *Automatica* **10**: 455–468.

APPENDIX E

The program CTSM

E.1 Introduction

The program CTSM is a software package for parameter estimation for models formulated as Stochastic Differential Equations.

The original version of the program was developed by Professor PhD Henrik Madsen in 1985 with the name CTLSM, Continuous Time Linear Stochastic Modelling. The program was written in Fortran using optimization routine VA13CD from the HARWELL Subroutine Library. The optimization routine is a quasi Newton method, using finite difference approximation to the gradient. The Hessian is updated by the BFGS updating formula. In 1991 there was a numerical revision created by Henrik Melgaard. In 1993 the first version with a k-step optimization and k-step predictions were developed and in 1994 the first version with non-linear routine where build. The routine was based a linear approximation of the non-linear function and sub-sampling methods using the extended Kalman filter.

In 2000 Niels Rode Kristensen developed the first general non-linear program, CTSM (Continuous Time Stochastic Modelling), still using the optimization routine VA13CD from the HARWELL Subroutine Library. The first graphical version, programmed in java was developed in 2001. In April 2003 a version with different filtering routes where developed, this is the first version with an ODE solver for filtering in non-linear models. The user can choose between three different filtering routines; The function $f()$ is linearized in sub-sampled intervals, an ODE solver with Adams method for non-stiff systems and an ODE solver with BDF for stiff systems (Backward Differentiation formula), see Section E.2 for an outline of the filtering methods and Appendix F for an introduction to stiff systems. The latest version came out in December 2003, this version is the first graphical version which provides a smoothing and k -step prediction along with one step prediction and pure simulation.

In the following sections some of the numerical methods used in the program are outlined, a further description is available in the manual (Kristensen et al. 2003).

The task is to find parameters such that the logarithm of the likelihood function, Eq. (4.12) is minimized. A single value of the likelihood function involves calculation of one step prediction ϵ_t and its variance $R_{t|t-1}$, in all the data points $t = 1, \dots, N$. The optimization procedure might thus be classified into two numerical tasks:

- The filtering, i.e., calculation of all the one step prediction. This involves
 - Computation of the exponential $e^{\mathbf{A}\tau_s}$. Test for singularity and in case of singularity, use of singularity routines.
 - Numerical integration - ODE solvers
- The optimization
 - Calculation of a gradient, which is performed by finite difference approximations.
 - Penalty calculations, since the optimization is a constrained.

E.2 Filtering methods

Calculation of the predictions is referred to as the filtering. In this context it refers to prediction of the state variable $\hat{\mathbf{x}}_{k+1|k}$. Different filtering routines can be applied:

- **Linear models:** The stochastic differential equation is solved analytically as shown in Eq.(4.14). The numerical task is to compute the exponential $e^{\mathbf{A}\tau}$. This might involve eigenvalue problems in case of a singularity. The program CTSM contains a singularity test routines and special singularity routines.
- **Non-linear models:** In this case, three different methods are implemented:
 1. Sub-sampling approximations; the time interval $[t_k, t_{k+1}[$ is sub-sampled into $[t_k, \dots, t_j, \dots, t_{k+1}[$ and the equations are linearized at each sub-sampling instant.

$$\frac{d\hat{\mathbf{x}}_{t|j}}{dt} = \mathbf{f}(\hat{\mathbf{x}}_{j|j-1}, \mathbf{u}_j, t_j, \boldsymbol{\theta}) + \mathbf{A}(\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_j) + \mathbf{B}(\mathbf{u}_t - \mathbf{u}_j) \quad [t_{k_j}, t_{k_{j+1}}[\quad (\text{E.1})$$

$$\frac{d\hat{\mathbf{P}}_{t|j}}{dt} = \mathbf{A}\hat{\mathbf{P}}_{t|j} + \mathbf{P}_{t|j}\mathbf{A}^T + \boldsymbol{\sigma}\boldsymbol{\sigma}^T \quad t \in [t_{k_j}, t_{k_{j+1}}[\quad (\text{E.2})$$

using same shorthand notation as in Table 4.1. An analytic solution to the linear differential equation is found in each sub-sampled interval as for linear models.

2. Numerical ODE solution using Adams method with predictor-corrector scheme, for non stiff systems. The principle is as follows (shown for one dimension in order to keep focus on the method).

The task is to solve initial value problem:

$$\frac{dx}{dt} = f(t, x) \quad x(t_k) = c \quad (\text{E.3})$$

The well known Euler method is to divide into subintervals and approximate the derivative $\frac{dx}{dt}$ with difference quotient $(\tilde{x}_{k+1} - \tilde{x}_k)/h$, where \tilde{x}_k denotes a numerical approximation to x_k and h denotes the length of the interval. This leads to the difference equation

$$\frac{\tilde{x}_{k+1} - \tilde{x}_k}{h} = f(t_k, \tilde{x}_k) \quad \text{or} \quad \tilde{x}_{k+1} = \tilde{x}_k + hf(t_k, \tilde{x}_k). \quad (\text{E.4})$$

The weakness of the Euler's method is that the step needs to be small in order to obtain acceptable accuracy (Burden & Faires 1989). The Euler's method is called a one step method because the approximation t point t_k only involves information from one previous point. Methods using approximations k previous values are called k step methods or k th order methods. Adams predictor-corrector scheme is a multi step method. An example of a predictor-corrector scheme can be a fourth order Adams-Bashford, for predicting \hat{x}_4

$$\hat{x}_4^{(0)} = \hat{x}_3 + \frac{h}{24}[55f(t_3, \hat{x}_3) - 59f(t_2, \hat{x}_2) + 37f(t_1, \hat{x}_1) - 9f(t_0, \hat{x}_0)] \quad (\text{E.5})$$

then the predicted value $\hat{x}_4^{(0)}$ is used in a three order Adams-Moulton formula

$$\hat{x}_4^{(1)} = \hat{x}_3 + \frac{h}{24}[9f(t_4, \hat{x}_4^{(0)}) - 19f(t_3, \hat{x}_3) + 5f(t_2, \hat{x}_2) - f(t_1, \hat{x}_1)] \quad (\text{E.6})$$

Equations like Eq.(E.6) are known as implicit formulas since \hat{x}_4 occurs on both sides. Summarizing; the first equation is used to predict the value x_4 and the predicted value is then used in the latter, implicit formula equation for improving (correcting) the approximation obtained by the explicit formula Eq.(E.5), a detailed description of the method can be seen in (Dahlquist & Björck 1988).

3. Numerical ODE solution using the Backward Difference Formula, There exists several BDF formulas, in (Dahlquist & Björck 1988) it is formulated as:

$$hD = -\ln(1 - \nabla) = \nabla + \frac{1}{2}\nabla^2 + \frac{1}{3}\nabla^3 + \dots + \frac{1}{m}\nabla^m \quad (\text{E.7})$$

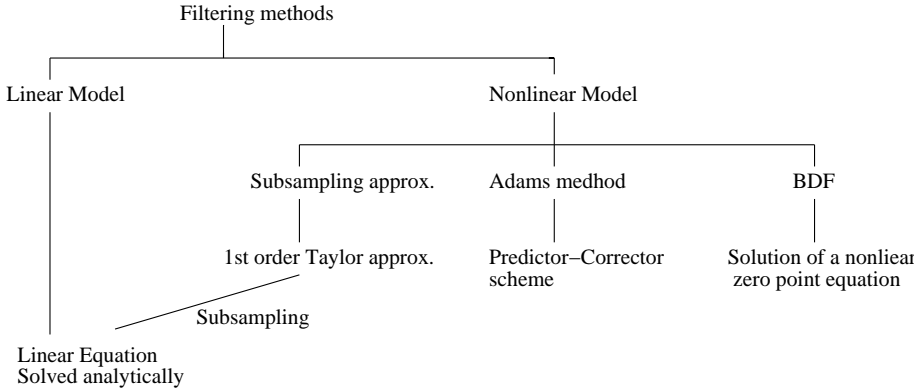


Figure E.1: Overview of the filtering methods in the program CTSM.

or

$$hf(x_{k+1}) = \nabla x_{k+1} + \frac{1}{2} \nabla^2 x_{k+1} + \dots + \frac{1}{m} \nabla^m x_{k+1} \quad m \leq 6 \quad (E.8)$$

i.e.

$$0 = -hf(x_{k+1}) + \nabla x_{k+1} + \frac{1}{2} \nabla^2 x_{k+1} + \dots + \frac{1}{m} \nabla^m x_{k+1} \quad m \leq 6 \quad (E.9)$$

where ∇ is the backward difference operator and D is the differentiation operator. As m increases the local truncation error decreases, but the stability properties become worse (Dahlquist & Björck 1988). Hence, the numerical task is to solve a non-linear zero point problem, often implemented by using "Newton-like" formulas.

Figure E.1 shows an overview of the filtering methods.

E.3 Optimization routine

The optimization method used in CTSM is a quasi-Newton method based on the BFGS updating formula and a soft line search algorithm to solve the non-linear optimization problem Eq.(4.12). In analogy with ordinary Newton-Rapson methods for optimization, quasi-Newton methods seek a minimum of the non-linear objective function, the likelihood function $-\ln(L(\boldsymbol{\theta}; \mathcal{Y}_N))$. In this section this will be denoted by $\mathcal{F}(\boldsymbol{\theta})$ i.e., define the short term notation

$\mathcal{F}(\boldsymbol{\theta}) = -\ln(L(\boldsymbol{\theta}; \mathcal{Y}_N))$ The minimum is found where the gradient is zero, i.e.,

$$\nabla \mathcal{F}(\boldsymbol{\theta}) = \mathbf{0}. \quad (\text{E.10})$$

As well as the Newton-Rapson, the quasi-Newton is based on the Taylor expansion of first order of the gradient $\nabla \mathcal{F}(\boldsymbol{\theta})$.

$$\nabla \mathcal{F}(\boldsymbol{\theta}^i + \boldsymbol{\delta}) = \nabla \mathcal{F}(\boldsymbol{\theta}^i) + \left. \frac{\partial \nabla \mathcal{F}(\boldsymbol{\theta})}{\partial(\boldsymbol{\theta})} \right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^i} \boldsymbol{\delta} + o(\boldsymbol{\theta}). \quad (\text{E.11})$$

the partial derivative in Eq. (E.11) is the Hessian. The gradient is approximated by finite difference approximation and the Hessian is updated with the BFGS updating formula, see (Kristensen et al. 2003) for mathematical formulas.

The optimization routine is a routine which finds minima within the limited area. i.e., not on the boundary. Thus the constrains must be defined such that the optimum parameter values is in between i.e.,

$$\theta_j^{\min} < \theta_j < \theta_j^{\max}. \quad (\text{E.12})$$

The traditional way of attacking this task is to to defining a new objective function $\widehat{\mathcal{F}}(\boldsymbol{\theta})$ by adding a penalty function $P(\lambda, \boldsymbol{\theta}, \theta_j^{\min}, \theta_j^{\max})$ to the objective function $\mathcal{F}(\boldsymbol{\theta})$

$$\widehat{\mathcal{F}}(\boldsymbol{\theta}) = \mathcal{F}(\boldsymbol{\theta}) + P(\lambda, \boldsymbol{\theta}, \theta_j^{\min}, \theta_j^{\max}). \quad (\text{E.13})$$

A proper choices of Lagrange multiplier λ , and the limiting values θ_j^{\min} and θ_j^{\max} the penalty function has no influence of the estimation. However, the penalty function will force the finite difference derivative to increase when θ_j is close to one of the limits.

APPENDIX F

Stiff systems

F.1 Introduction

When a system consists of more than one first order differential equations the possibility of stiffens arises. The phenomenon of stiffness is difficult to define in precise mathematical turns in a satisfactory manner (Lambert 1991). Stiffness has to do with numerical stabilities and step-lengths. A frequently used statement is:

[S1]: *Stiffness occurs when stability requirements, rather than those of accuracy constrain the step-length.*

Another statement is:

[S2]: *Stiffness occurs when some components of the solution decay much more rapidly than others.*

Broadly speaking this means that there are different time scales in the system. A frequently used definition is

[S3]: *A linear constant coefficient system is stiff if all of its eigenvalues have negative real part and the stiffness ratio is large.*

The stiffness ratio is defined as $\lambda_{max}/\lambda_{min}$ if the eigenvalues are Real (as in this project). Frequently the ratio $\lambda_{max}/\lambda_{min}$ is called the matrix condition number (Montgomery & Runger 2002). Furthermore, (Montgomery & Runger 2002) state that if the condition number is less than 100 the system is non-stiff whereas it starts to show stiffness characteristics when the condition number exceeds 100.

In (Lambert 1991) there is shown that none of these state ments quite cover the phenomena of stiffness. In (Lambert 1991) the following definition is used

[S4]: *If a numerical method with a finite region of absolute stability, applied to a system with any initial conditions, is forced to use in a certain interval of integration a step-length which is excessively small in relation to the smoothness of the exact solution in that interval, then the system is said to be stiff in that interval.*

In non-linear systems, the Jacobian $\partial f/\partial x$ can be calculated and the characteristics of the Jacobian can be studied. For more about stiff systems see

(Lambert 1991).

Consider system

$$\begin{aligned}x_1' &= -2x_1 + 1x_2 & x_1(0) &= 1 \\x_2' &= 998x_1 - 999x_2 & x_2(0) &= 1\end{aligned}\tag{F.1}$$

A solution to this equation is:

$$x_1 = e^{-t} + 1e^{-1000t}\tag{F.2}$$

$$x_2 = e^{-t} - 998e^{-1000t}\tag{F.3}$$

The e^{-1000} term is completely negligible in determining the values of x_1 and x_2 as soon as one is away from the origin. However, a general forward equation would demand a step size $h \ll 1/1000$ for the method to be stable.

In general for a set of linear differential equations:

$$\frac{d\mathbf{x}}{dt} = -\mathbf{C}\mathbf{x} \quad \mathbf{x}(t_0) = \mathbf{c}\tag{F.4}$$

where \mathbf{C} is a positive definite matrix. A first order Euler yields

$$\tilde{\mathbf{x}}_{k+1} = (\mathbf{I} - \mathbf{C}h)\tilde{\mathbf{x}}_k = (\mathbf{I} - \mathbf{C}h)^{n+1}\tilde{\mathbf{x}}_0\tag{F.5}$$

A matrix \mathbf{A}^n tends to zero as $n \rightarrow \infty$ only if the largest eigenvalue of \mathbf{A} has magnitude less than unity, thus $\tilde{\mathbf{x}}_n$ is bounded as $n \rightarrow \infty$ only if the largest eigenvalue of $(\mathbf{I} - \mathbf{C}h)$ is less than 1, or

$$h < \frac{2}{\lambda_{\max}}\tag{F.6}$$

where λ_{\max} is largest eigenvalue of the matrix \mathbf{C} . Implicit differences is

$$\tilde{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_k - h\mathbf{C}\tilde{\mathbf{x}}_{k+1}\tag{F.7}$$

or

$$\tilde{\mathbf{x}}_{k+1} = (\mathbf{I} + \mathbf{C}h)^{-1}\tilde{\mathbf{x}}_k\tag{F.8}$$

if the eigenvalues of \mathbf{C} are λ the the eigenvalues of $(\mathbf{I} + \mathbf{C}h)^{-1}$ are $(1 + \lambda h)^{-1}$ which has magnitude less than h for all h , thus the method is stable for all step sizes h . This explains why implicit methods are desirable option when the system is stiff. Note that the penalty for the stability of the implicit methods is that the inverse of a matrix must be found at each step.

F.2 The stiffness of the non-linear system in Paper [B]

The non-linear system in Paper [B] is a stiff system. Furthermore, the system swings between being singular to being stiff to being singular again which is a real computational challenge, which succeeded to accomplish.

The system is stiff in that sense that the method controls the convergence. In order to achieve a solution at all, and to achieve a stable solution it was necessary to use an implicit method for ODE solver as described in Section E.2

Furthermore, it is also stiff with respect to the stiffness ratio as in [S3].

The linear approximation (a part of the Jacobian) to be considered is the partial derivative with respect to the state variables in the state space model:

$$\mathbf{A} = \left. \frac{\partial f}{\partial x_t} \right|_{\mathbf{x}=\hat{\mathbf{x}}_j|_{j-1}, \mathbf{u}=\mathbf{u}_j, t=t_j, \theta} \quad (\text{F.9})$$

which is

$$\begin{pmatrix} -a & 0 & 0 & 0 \\ -pdd\phi(T_s)\Psi(N) & -pddT_s\phi(T_s)\Psi'(N) & 0 & 0 \\ -pddT_s\phi'(T_s)\Psi(N) - cP\phi'(T_s) & & & \\ 0 & -pddT_s\phi(T_s)\Psi'(N) & -(f+k_1) & 0 \\ 0 & 0 & f & -k_2 \end{pmatrix} \quad (\text{F.10})$$

The matrix is a lower diagonal matrix and thus the eigenvalues are the values on the diagonal. The values $-a$, $-(f+k_1)$ and $-k_2$ are constants which values varies from 0.05 to 5, see Table 1 in Paper [B]. The largest is 100 times the larger than the smallest and according to (Montgomery & Runger 2002) then the system begins to behave stiff.

Figure F.1 shows the threshold function for the snow container $\Phi(N)$ and Figure F.2 shows its derivative $\Phi'(N)$

When there is enough snow to melt the function $-pddT_s\phi(T_s)\Psi'(N)$ is zero and the system is singular. In the transforming period then there is little snow left to melt and N converges to zero the derivative alters from being 0 to about 80

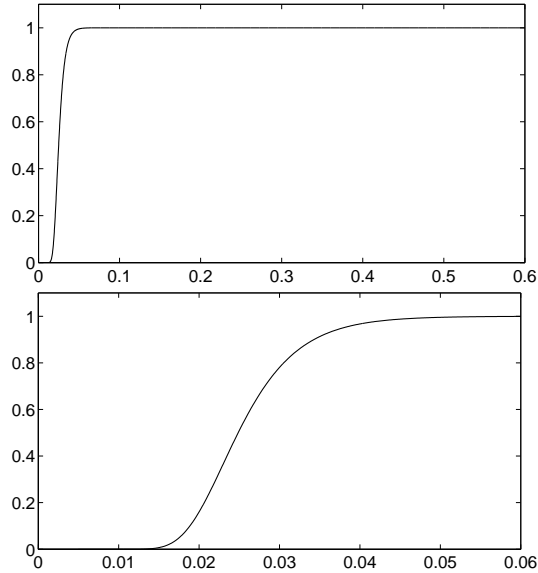


Figure F.1: The threshold function for snow
 $\Psi(N) = 100 \exp(-100 \exp(-200N))$.

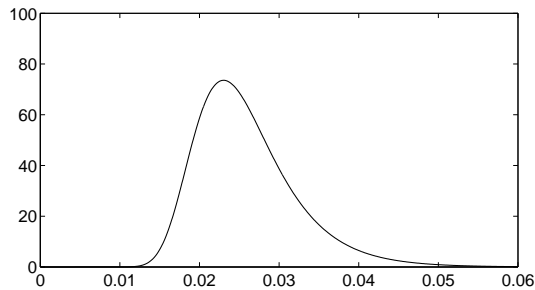


Figure F.2: The derivative of the threshold function for snow.

or even larger during the optimization procedure. Hence, the eigenvalue alters from being 0 to being about 4, depending on the value of the temperature T and then the derivative becomes zero. During the optimization the numerical routine uses a singularity route or a non-singularity routine depending on the situation. In the shifting phase some eigenvalues can be very small, resulting in a large stiffness ration.

It was necessary to have the threshold function for snow, $\Psi(N)$ steep to secure nonnegative values in the snow container. However, the steeper the threshold

value, the larger is its derivative, resulting in the stiffer system.

Bibliography

- Abbott, M., Bathurst, J., Cunge, J., O'Connell, P. & Rasmussen, J. (1986*a*), 'An introduction the the European Hydrological System - System Hydrologique Europeen, "SHE" 1: History and phylosophy of a physically-based distributed modelling system', *Journal of Hydrology* **87**, 45–59.
- Abbott, M., Bathurst, J., Cunge, J., O'Connell, P. & Rasmussen, J. (1986*b*), 'An introduction the the European Hydrological System - System Hydrologique Europeen, "SHE" 2: Structure of a physically-based, distributed modelling system.', *Journal of Hydrology* **87**, 61–77.
- Anderson, T. W., Fang, K. T. & Olkin, I., eds (1994), *Multivariate Analysis and its Applications*, Institute of Mathematical Statistics, Hayward, Institute of Mathematical Statistics, Hayward, chapter Coplots, Nonparametric Regression, and Conditionally Parametric Fits, pp. 21–36.
- Bengtson, L. (1988), River ice forecasting, Technical Report 21, Nordic Hydrological Programme.
- Bergström, S. (1975), 'The development of a snow routine for the HBV-2 model', *Nordic Hydrology* **6**, 73–92.
- Bergström, S. (1995), The HBV model, in V. P. Singh, ed., 'Computer Models of Watershed Hydrology', Water resources Publication, Colorado.
- Beven, K., Lamb, R., Quinn, P., Romanowich, R. & Freer, J. (1995), TOP-MODEL, in V. P. Singh, ed., 'Computer Models of Watershed Hydrology', Water resources Publication, Colorado.
- Box, G. E. P. & Jenkins, G. M. (1976), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco.

- Burden, R. L. & Faires, J. D. (1989), *Numerical Analysis*, PWS-KENT Publishing Company, Boston.
- Burnash, R. (1995), NWS river forecast system - catchment modeling, *in* V. P. Singh, ed., 'Computer Models of Watershed Hydrology', Water resources Publication, Colorado.
- Burnash, R. J. C., Ferral, R. & McGuire, R. (1973), A generalized streamflow simulation system - Conceptual modeling for digital computers., Technical report, National Weather Service, NOAA and State of California Department of Water Resources, Joint Federal State River Forecast Center Sacramento, California.
- Cadzow, J. A. (1973), *Discrete time systems*, Prentice Hall, New Jersey.
- Campbell, E. P., Fox, D. R. & Bates, B. C. (1999), 'A bayesian approach to parameter estimation and pooling in nonlinear flood event models', *Water Resources Research* **35**(1), 211-220.
- Carstensen, J., Nielsen, M. K. & Strandbæk, H. (1998), 'Prediction of hydraulic load for urban storm control of municipal wwtp', *Water Science and Technology* **37**(12), 363-370.
- Chang, L.-C., Chang, F.-J. & Tsai, Y.-H. (2005), 'Fuzzy exemplar-based inference system for flood forecasting', *Water Resources Research* **41**(2), W02005 1-12.
- Chow, V. T. (1964), *Handbook of applied hydrology*, McGraw-Hill Book Company, USA.
- Chow, V. T., Maidment, D. R. & Mays, L. W. (1988), *Applied hydrology*, McGraw-Hill Book Company, New Jersey.
- Cleveland, W. S. & Develin, S. J. (1988), 'Locally weighted regression: An approach to regression analysis by local fitting', *Journal of The American Statistical Association* **83**, 596-610.
- Crawford, N. H. & Linsley, R. K. (1964), Digital simulation in hydrology: Stanford watershed model mark iv (swm iv), Technical Report 39, Dept. of Civ. Eng., Stanford University, California.
- Dahlquist, G. & Björck, Å. (1988), *Numerical Methods*, McGraw-Hill Book Company, New Jersey.
- Førland, E. J., Allerup, P., Dahlström, B., Elomaa, E., Jónsson, T., Madsen, H., Perälä, J., Rissanene, P., Vedin, H. & Vejen, F. (1996), Manual for operational correction of nordic precipitation data, Technical Report No. 24/96 ISSN 0805-9918, The Norwegian.

- Georgakakos, K. P., Rajaram, H. & Li, S. G. (1988), On improved operational hydrologic forecasting, Technical Report IIHR No. 325, Institute of hydraulic Research, University of Iowa.
- Goodwin, G. & Payne, R. (1977), *Dynamic system identification*, Academic Press.
- Gudmundsson, G. (1970), 'Short term variation of a glacier-fed river', *Tellus* **22**, 341–353.
- Gudmundsson, G. (1975), 'Seasonal variations and stationarity', *Nordic Hydrology* **6**, 137–144.
- Härdle, W. (1990), *Applied Nonparametric Regression*, Syndicate of the University of Cambridge, United Kingdom.
- Hastie, T. & Loader, C. (1993), 'Local Regression: Automatic Kernel Carpentry', *Statistical Science* **8**(2), 120–129.
- Hastie, T. & Tibshirani, R. (1993), 'Varying-coefficient models', *Journal of the Royal Statistical Society* **55**, 757–796.
- Hsu, K., Gupta, H. V., Gao, Z., Sorooshian, S. & Imam, B. (2002), 'Self-organizing linear output map (SOLO): An artificial neural network suitable for hydrologic modeling and analysis', *Water Resources Research* **38**(12), 1302.
- Iorgulescu, I. & Beven, K. (2004), 'Nonparametric direct mapping of rainfall-runoff relationships: An alternative approach to data analysis and modelling', *Water Resources Research* **40**(8), W08403 1–11.
- Jazwinski, A. H. (1970), *Stochastic processes and filtering theory*, Academic Press Inc., San Diego.
- Kristensen, N. R., Madsen, H. & Jørgensen, S. B. (2004a), 'Parameter estimation in stochastic grey-box models', *Automatica* **40**, 225–237.
- Kristensen, N. R., Madsen, H. & Jørgensen, S. B. (2004b), 'A method for systematic improvement of stochastic grey-box models.', *Computer and Chemical Engineering* **28**, 1431–1449.
- Kristensen, N. R., Melgaard, H. & Madsen, H. (2003), *CTSM Version 2.3 - a program for parameter estimation in stochastic differential equations*. Available from <http://www.imm.dtu.dk/ctsm>.
- Lambert, J. D. (1991), *Numerical methods for ordinary differential systems - The initial value problem*, John Wiley and Sons Ltd., England.

- Lee, Y. H. & V.P.Singh (1999), 'Tank Model Using Kalman Filter', *Journal of Hydrologic Engineering* **4**, 344–349.
- Madsen, H. (2000), 'Automatic calibration of a conceptual rainfall-runoff model using multiple objectives', *Journal of Hydrology* **235**, 276–288.
- Madsen, H. & Holst, J. (2000), Modelling Non-Linear and Non-Stationary Time Series, December 2000, Technical Report No. 1/2000, IMM.
- Mays, L. (1996), *Water Resources Handbook*, McGraw-Hill, USA.
- McCuen, R. H. (1989), *Hydrologic Analysis and Design*, Prentice-Hall, Inc., New Jersey.
- Medova, E. & Kyriacou, M. (2000), 'Extreme values and the measurement of operational risk', *Operational Risk* **1**, 12–15.
- Montgomery, D. C. & Runger, G. C. (2002), *Applied Statistics and Probability for Engineers*, Wiley.
- Nash, J. E. (1957), 'The form of the instantaneous unit hydrograph', *IASH publication no. 45* **3-4**, 114–121.
- Nielsen, H. A. (1997), Lflm version 1.0 - an spls/r library for locally weighted fitting of linear models, Technical Report 22, Department of Mathematical Modelling, Technical University of Denmark. Available from <http://www.imm.dtu.dk/han/software.html>.
- Nielsen, H. A., Nielsen, T. S. & Madsen, H. (1997), 'Conditional parametric arx-models', *11th IFAC Symposium on System Identification* **2**, 475–480.
- Øksendal, B. (1995), *Stochastic Differential Equations*, Springer Verlag, Berlin.
- Plate, E. J. (1992), Stochastic design in hydraulics, in 'Proc. Int. Symposium on Stochastic Hydraulics'.
- Refsgaard, J., Seth, S., Bathurst, J., Erlich, M., Storm, B., Jørgensen, G. & S.Chandra (1992), 'Application of the SHE to catchments in India Part 1. General results', *Journal of Hydrology* **140**, 1–23.
- Refsgaard, J. & Storm, B. (1995), MIKE SHE, in V. P. Singh, ed., 'Computer Models of Watershed Hydrology', Water resources Publication, Colorado.
- Rist, S. (1962), 'Winter ice of thorsa river system', *Jokull* **12**, 1–30.
- Sælthun, N. R. & Killingtveit, Å. (1995), *Hydropower development, Vol VII*, Norwegian Institute of Technology, Trondheim.
- Shamseldin, A. Y. (1997), 'Application of a neural network technique to rainfall-runoff modelling', *Journal of Hydrology* **199**, 272–294.

- Sherman, L. K. (1932), 'Streamflow from rainfall by the unit-graph method', *Eng. News Rec.* **108**, 501–505.
- Sigbjarnarson, G. (1990), *Vatnid og landid*, Orkustofnun, Reykjavik.
- Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York.
- Singh, V. P. (1964), 'Nonlinear instantaneous unit hydrograph theory', *Journal of Hydraulic Div. Am. Soc. Civ. Eng* **90**(HY2), 313–347.
- Singh, V. P. (1988), *Hydrologic Systems, Vol 1, Rainfall-runoff Modelling*, Prentice Hall.
- Singh, V. P. & Woolhiser, D. A. (2002), 'Mathematical modeling of watershed hydrology', *Journal of Hydrologic Engineering* **7**, 270–292.
- Sugawara, M. (1995), Tank model, in V. P. Singh, ed., 'Computer Models of Watershed Hydrology', Water resources Publication, Colorado.
- Todini, E. (1978), 'Using a Desk-Top Computer for an On-Line Flood Warning System', *IBM Journal of Research and Development* **22**, 464–471.
- Tong, H. (1990), *Non-linear Time Series - A Dynamic System Approach*, Oxford University Press, Britain.
- Viessman, W. & Lewis, G. L. (1996), *Introduction to hydrology*, HarperCollins College Publishers, New York.
- Yevjevich, V. (1976), Structure of natural hydrologic time series, in H. W. Shen, ed., 'Stochastic approaches to water resources', Hsieh Wen Shen, Colorado.
- Young, P. C. (2002), 'Advances in real-time flood forecasting', *The Royal Society* **360**, 1433–1450.
- Zoch, R. T. (1934, 1936, 1937), 'On the relation between rainfall and stream-flow', *Monthly Weater Review* **62,64,65**, 315–322, 105–121, 135–147.

Index

- Bernoulli, 21
- Adams-Bashford, 120
- Adams-Moulton, 120
- agricultural evolution, 20
- Ampt, 22
- anchor ice, 40
- Archimedes, 20
- ARMAX, 55
- ARX, 2, 62
- backwater, 42
- bandwidth, 59, 62
- base flow, 2, 10, 22, 26, 32
- basin lag time, 27
- BDF formula, 120
- black box model, 12, 22, 45, 65
- Boussinesq, 21
- Brater, 22
- calibration, 65
- catchment, 25
- Chezy, 21
- Chow, 29, 30
- civilization, 19
- conceptual models, 4, 22
- condition number, 124
- conditional parametric model, 61
- conservation of mass, 9, 12, 24
- continuity equation, 24, 31
- CTSM, 5, 9, 117
- Dalton, 21
- Darcy, 21
- data
 - observed, 22
- direct runoff, 26, 28
- discharge, 10, 23, 35, 42
- distributed models, 22
- distributed routing, 28, 31
- effective rainfall, 26–28
- Egypt, 20
- eigenvalues, 125
- energy, 20
- estimation, 49, 65–67
- Euler, 21
- evaporation, 10, 12, 21, 23, 27
- excess rainfall, 22, 26
- explanatory variable, 61
- Fair, 22
- finite differenc, 119
- FIR, 2, 55
- flow
 - excess, 10
 - overland, 23
 - routing, 22
 - subsurface, 22, 23
- flow of water, 20
- flow routing, 9, 11, 28
 - distributed, 28
 - lumped, 28

- fluid
 - resistance, 21
- fluid flow, 21
- force pump, 20
- forward equation, 125
- frazil, 40
- frequency multiplication, 48
- Froude, 21

- gamma distribution, 30
- Greeks, 20
- Green, 22
- grey box model, 48
- grey box model, 12, 45, 53, 65, 66
- groundwater, 10, 12, 22, 23, 32

- Hatch, 22
- Hazen, 22
- Hoover, 22
- Horton, 22
 - laws, 22
- Hursh, 22
- hydraulic routing, 28
- hydrograph, 22, 25, 27
- hydrology, 12
- hydrologic routing, 28
- hydrologic system, 24
- hydrologic systems, 2
- hydrological data, 33
- hydrology, 1, 3, 19, 21, 25
- hydropower plant, 7

- ice, 40
- ice corrupted, 43
- ice reduction, 42
- icing, 39
- Implicit differences, 125
- impulse response function, 25, 30, 47, 56
- inertia, 21
- infiltration, 2, 12, 22, 23, 26, 32
- inter-modulation distortion, 48
- interflow, 22
- irrigation, 20

- Ito integral, 51
- Ito process, 51, 52

- Jacobian, 124

- Kalman filter, 6, 50
- Kalman gain, 50, 55
- Kelvin, 21
- kernel, 47, 59
- Kirchoff, 21
- Kohler, 22

- lag time, 27
- Leonardo da Vinci, 21
- level pool method, 29
- linear model, 119
- linear reservoir, 9, 22, 29, 30
- linear system, 47
- Linsley, 22
- Lowdermilk, 22
- lumped flow routing, 29
- lumped routing, 28

- Manning, 21
- mass balance, 10, 27
- maximum likelihood, 118
- McCarthy, 22
- McCuen, 26
- Mesopotamia, 20
- model
 - black box, 65
 - black box, 2, 10, 12, 22, 45
 - conceptual, 4, 22
 - distributed, 22
 - grey box, 12, 45, 48, 53, 65, 66
 - input-output, 10
 - white box, 12, 65
 - white box , 45
- Muskingum, 22
- Muskingum method, 29

- Navier, 21
- Newton, 21
- non-linear, 46, 49
- non-linear model, 119

- non-linear systems, 48
- Numerical integration, 119
- ODE solver, 119
- open channel, 21
- optimization, 6, 119
- overland flow, 23
- Palissy, 21
- parameter estimation, 48
- Pascal, 21
- percolate, 23
- permeability, 22
- Perrault, 21
- pipe, 21
- Poiseulle, 21
- precipitation, 2, 5, 9, 10, 12, 22, 23, 25, 33, 35
- pressure, 21
- quickflow, 26
- rain gauge, 33, 34
 - Nipher, 33
 - tipping bucket, 34
- rainfall
 - effective, 26, 28
 - excess, 26
- rainfall runoff, 3, 4, 10, 26
- rating curve, 38
- regulation curve, 7
- reservoir, 5, 22
 - linear, 30
- response
 - function, 22
- retain, 26
- Reynolds, 21
- risk assessment, 7
- Romans, 21
- routing, 22, 28
 - distributed, 31
 - hydraulic, 28
 - hydrologic, 28
 - linear reservoir, 29
 - lumped, 29
- runoff, 22, 26
 - direct, 28
 - surface, 23
- Saint Venant, 9, 21
- Saint-Venant equations, 31
- sanitation, 21
- SDE, 48, 65
- SETAR, 57
- sewage system, 10
- sewer, 21
- Sherman, 22
- singularity, 119
- smoother, 59
- smoothing, 58
- state space model, 49, 54
- state space models, 55
- stationary process, 54
- stiff system, 118, 123
- stochastic differential equation, 50
- stochastic differential equations, 48
- Stokes, 21
- storage, 9–11
 - constant, 30
- storatge, 11
- streamflow, 25
- subsurface flow, 22, 23
- subsurface water, 32
- superposition, 48
- surface runoff, 12, 23
- threshold, 57
- threshold function, 4
- time base, 26
- time of concentration, 27
- tipping bucket, 34
- transfer function, 30, 47, 53, 55, 56
- transpiration, 10, 12
- unit hydrograph, 22
- varying-coefficient models, 61
- vegetation, 23

-
- viscosity, 21
 - water, 20, 21
 - flow of, 20
 - resources, 19
 - transported, 21
 - wheel, 20
 - water clock, 20
 - water resources, 19
 - water shortage, 7, 8
 - watershed, 1, 3, 4, 9, 12, 25–27
 - waves, 21
 - Weisbach, 21
 - white box model, 12, 45, 65
 - white noise, 46, 49
 - Wiener process, 49, 52
 - Yevjevich, 8
 - Zoch, 22