UPPSALA
UNIVERSITET

UPTEC W 17 025
Examensarbete 30 hp
Augusti 2017

Modeling hydrometeorological extremes in Alpine catchments

Theo Voulgaridis

ABSTRACT

Modeling hydrometeorological extremes in Alpine catchments

Theo Voulgaridis

Uncertainties with a modeling framework consisting of a weather generator, two precipitation disaggregation models and the hydrological HBV model was assessed with respect to hydrometeorological extremes in Tyrol, Austria. Extreme precipitation events are expected to increase in intensity and frequency in the Alps during a warmer climate. The Alpine regions may be particularly vulnerable to such changes in climate where many floods in Europe occurred during recent years and caused major damage and loss of life.

Weather generators typically provide time series at daily resolution. Different disaggregation methods have therefore been proposed and successfully tested to increase temporal resolution in precipitation. This is essential since flood peaks may be maintained for as little as minutes. Here, the non-parametric method of fragments was tested and compared with the multiplicative microcanonical cascade model with uniform splitting on the reproduction of precipitation extremes. It is also demonstrated that the method of fragments model can be transformed to disaggregate temperature with slight changes in the model structure. Preliminary test results show that the simulation of discharge peaks can be improved by disaggregating temperature in comparison with using daily averages as input in the HBV model.

Test results show that precipitation extremes were simulated within confidence bounds for Kelchsauer and Gurglbach when using historical observations as input. These two catchments had longer records of data available in comparison with Ruetz where the majority of simulated precipitation extremes were found outside confidence ranges. This indicates that the model is data driven. Synthetic data series were constructed with the weather generator from historical data and disaggregated with the two disaggregation models. The differences between the models were bigger for Ruetz where less observed data was available. The method of fragments simulates extremes with closest resemblance to extremes. This is also true for the reproduction of wet spells and simulated variance.

To account for parameter uncertainty in the HBV model, it is highly motivated to simulate discharge with different but suitable parameter sets to account for equifinality. However, the large amount of data produced when disaggregating the weather generated time series transcended the data capacity of the HBV model and made it crash. Other uncertainties related to the framework are the use of theoretical probability distributions in the weather generator and the dependence of high resolution data for the disaggregation model. Despite these uncertainties, the framework is closer to a physical understanding of the causes of floods than the uncertain frequency analysis method. The framework is also applicable to land-use and climate change studies.

Keywords: method of fragments, HBV, rainfall disaggregation, weather generation

Department of Earth Sciences, Program of Air, Water and Landscape Sciences, Uppsala University, Villavägen 16 SE 752 36, Uppsala, Sweden. ISSN 1401-5765

REFERAT

Modellering av hydrometeorologiska extremvärden i alpina avrinningsområden

Theo Voulgaridis

Osäkerheter med ett modelleringsramverk bestående av en vädergenerator, två disaggregeringsmodeller för nederbörd och den hydrologiska HBV-modellen utvärderades med avseende på hydrometeorologiska extremvärden i Tirol, Österrike. Extrem nederbörd förväntas öka i intensitet och frekvens i Alperna under ett varmare klimat. Alperna kan vara särskilt känsligt för sådana förändringar där många av de senaste översvämningarna i Europa har inträffat och resulterat i stora materiella skador och dödsfall.

Vädergeneratorer skapar typiskt nederbördsdata med daglig upplösning. Flödestoppar vid översvämningar utvecklas ofta på tidskalor som är betydligt mindre varav nedskalning av nederbördsdata är nödvändigt. Den här uppsatsen jämför fragmentsmetoden med den multiplikativa mikrokanoniska modellen med likformig delning med avseende på simulering av extrem nederbörd. Eftersom dagliga medeltemperaturer hade en icke försumbar påverkan på flödestoppar i HBV-modellen, transformerades fragmentsmetoden med några förändringar för att också disaggregera temperatur. Preliminära resultat visade att disaggregerade temperaturserier förbättrade modelleringen av flödestoppar i HBV-modellen.

Resultaten från studien visar att fragmentsmetoden reproducerar extrem nederbörd inom beräknade konfidensintervall för de två områdena med mest tillgängliga historiska data, Kelchsauer och Gurglbach. Resultaten var sämre för Ruetz där mindre data fanns till förfogande, vilket indikerar att fragmentsmetoden är databegränsad. Med vädergeneratorn skapades väderserier från historisk daglig nederbördsdata. Dessa disaggregerades med hjälp av fragmentsmodellen och den mikrokanoniska modellen. Skillnaderna mellan modellerna var störst för avrinningsområdet Ruetz. fragmentsmodellen simulerade extremvärden, varians och antal dagar med nederbörd med störst likhet till historiska observationer.

Då flera kombinationer av parametrar i HBV-modellen kan ge samma fel med avseende på en kriteriefunktion, är det motiverat att simulera avrinning med flera olika men lämpliga parameteruppsättningar. Den stora datamängd som disaggregerade 100-årsserier på timlig upplösning gav upphov till gjorde att HBV-modellens datakapacitet överskreds. Vidare så har flera osäkerheter uppmärksammats med modelleringsramverket. Bland annat använder vädergeneratorn teoretiska sannolikhetsfördelningar för att modellera nederbörd dessutom är effektiviteten av disaggregeringsmodellerna begränsade av högupplösta tidsserier. Modelleringsförfarandet är dock närmare en fysisk förståelse för orsakerna bakom översvämningar i jämförelse med frekvensanalys som har kritiserats för att vara en osäker metod. Dessutom möjliggör kombinationen av modellerna studier av förändringar i klimat och markanvändning.

Nyckelord: vädergenerator, hydrologisk modellering, HBV-modellen

Institutionen för geovetenskaper, luft-, vatten och landskapslära, Uppsala Universitet, Villavägen 16 SE 752 36, Uppsala, Sverige. ISSN 1401-5765

PREFACE

The seed for this thesis was planted when Kaycee, Kenechukwu Okoli, at the Department of Earth Sciences, had a lecture on uncertainties related to the estimation of flood flows. The criticism was not new, references for this thesis goes back to the 1950s. The central question was, how can one use distribution fitting to estimate design floods, when for the most part, the part of the distribution we are interested in (the tail), is far away from where the observations provide information on the distribution profile. Kaycee ended the lecture by calling for more research on the physical understanding of flood flows. This was a research question that I wanted to engage in. If we are to liberate hydrology from the uncertainties related to frequency analysis, what should we replace it with?

This is what I have tried to figure out during this 30 credits master thesis course that will be the icing on the cake, the end of a five-year long quest for the pursue of a master of science degree in environmental and water engineering. I hope you will enjoy reading my thesis.

This thesis was made possible thanks to the academic supervisor Giuliano Di Baldassarre, my supervisor Korbinian Breinl, Kenechukwu Okoli , Eduardo Reynolds and Jan Seibert (Department of Earth Sciences, Program of Air, Water and Landscape Sciences; Hydrology) whom all assisted and gave me good advice and assistance to improve my thesis. Kenechukwu also assisted with an Matlab algorithm to fit theorethical probability distributions to historic discharge observations with the method of maximum likelihood (Okoli 2017, personal communication).

A special thanks to Korbinian Breinl, my supervisor, who provided raw weather and discharge data (at hourly resolution) and weather generated realizations (at daily resolution) extrapolated from historic observations as well as the disaggregated precipitation time series (at hourly resolution) simulated with the microcanonical model with uniform splitting (Breinl 2017, personal communication). He also gave me good assistance and his email reply was never more than a few minutes away. Thanks for the cooperation and thanks for giving me valuable insights on the constituents of a researcher's everyday life.

Theo Voulgaridis
Uppsala August 2017

UPTEC W 17 025, ISSN 1401-5765
Published digitally at the Department of Earth Sciences, Geotryckeriet, Uppsala University, Uppsala 2017

POPULÄRVETENSKAPLIG SAMMANFATTNING

Modellering av hydrometeorologiska extremvärden i alpina avrinningsområden

Theo Voulgaridis

Varje år utsätts ett stort antal människor i världen för översvämningar med dödlighet och stora ekonomiska skador som följd. Även ekosystem är utsatta, där risker för permanenta skador i habitat eller artmiljöer och mark kan bli den logiska påföljden. Många gånger orsakas översvämningar som en följd av komplexa och invecklade samband mellan exempelvis nederbördsintensitet, grundvattennivåer och rådande markförhållanden. Det är därför riskfyllt att diskutera enskilda faktorer utan ett helhetsperspektiv vid behandling av översvämningsproblematik.

En ytterligare komplicerande faktor är att klimatförändringar på många håll i världen förutspås leda till skiften i nederbördsmönster, atmosfäriska förhållanden och hydrologiska förlopp. Vattenånga kan exempelvis hålla mer vatten under varmare tillstånd, något som förenklat förutspås leda till både mer och intensivare nederbördsmönster. I Europa kan de Alpina områdena vara särskilt sårbara. Där har många av de största samtida översvämningarna skett och åsamkat både dödsfall och stora materiella skador. Den mest ödesdigra översvämningen skedde 2002 i Österrike, Italien och Tyskland och resulterade i skador motsvarande 27 miljarder US $. Många anledningar bidrar till att göra de alpina områdena sårbara, särskilda atmosfäriska cirkulationsmönster sägs vara en huvudorsak. Områden som är belägna på en högre höjd är också exponerade för större volymer och mer frekvent nederbörd. Andra bidragande faktorer är exempelvis smältande snömängder, mindre vegetationsmängder och markdjup vilket minskar fömågan att bromsa, det vill säga, uppehålla vatten. Extrem nederbörd förväntas öka under alla årstider för de flesta Alpina områdena under ett varmare klimat. Ett tvågradig temperaturökning av klimatet förutspås leda till att en extrem nederbördsmängd som tidigare förväntades ske ungefär en gång var 100e år (i medeltal), kan komma att återkomma vart 20e år.

Trots att bakgrunden till en viss översvämning kan vara komplex, är det relativt vanligt att bara använda flödesdata som grund för en viss extremflödesprognos, genom en så kallad frekvensanalys. Då extrema föruteelser sällan inträffar, finns det ofta få representativa observationer av sådana händelser. Problemet med frekvensanalys är därför att prognosen baseras på relativt lite information om det som eftersöks, nämligen extremvärden. Metoden ger inte heller någon direkt information om vad som kan ha orsakat extremflödet eller översvämningen. Kritik av frekvensanalys har därför länge varit en uppmärksammad fråga inom den hydrologiska forskningsgemenskapen. Målet med metoden är att bestämma ett så kallat dimensionerande flöde, ett hypotetiskt maxflöde som exempelvis kan användas för att dimensionera en damm. En annan utmaning med metoden är att det ofta inte finns flödesdata att tillgå och att det inte går att ta hänsyn till förändrade nederbördsmönster om man vill undersöka vad en sådan förändring skulle betyda för hydrologin i området.

Istället efterfrågas därför metoder som är mer realistiska och närmre avbildningar av naturliga och fysiska fenomen. Sådana modeller finns där förenklade matematiska ekvationer syftar till att avbilda fysikaliska processer. Sådana modeller kallas konceptuella. I sådana modeller går det att mata in väderdata så som temperatur och nederbörd. Beroende av värdet på olika parametrar i ekvationerna, parametrar som motsvarar olika hydrologiska egenskaper, kommer olika indata resultera i olika utflöden från modellen. Fördelen med ett sådant tillvägagångssätt är att förändringar i markanvändning och klimat kan simuleras. Det kan göras genom att koppla ihop modeller som genererar väderdata, så kallade vädergeneratorer med den konceptuella modellen. Vädergeneratorer har påvisats kunna återskapa nederbördsmängder- och mönster väl. En utmaning med vädergeneratorerna är att det ofta är eftersträvansvärt att skapa konstgjort väder med daglig upplösning trots att intensiva flödestoppar kan genereras under tidsintervall som är mycket kortare. Det är därför nödvändigt att dissaggregera, omfördela dagliga regnmängder till mindre tidsskalor för att bättre kunna simulera extremflöden. Flera publicerade verk och lyckade försök har gjorts med modeller som ökar upplösningen i nederbördsdata från dagliga värden till timvärden. Genom att koppla ihop vädergeneratorer, disaggregeringsmodeller och hydrologiska modeller är det möjligt att ta sig an många av de utmaningar som har presenterats ovan.

Det här arbetet har fokuserat på att undersöka osäkerheter med ett sådant modelleringsramverk, där HBV-modellen använts som hydrologisk modell. Osäkerheter mellan två olika disaggregeringsmodeller (fragmentsmodellen och en mikrokanonisk) har också undersökts. Resultaten av de ovan nämnda försöken ställs sedan i förhållande till konventionell frekvensanalys som också testades under arbetets gång. Mätdata som använts i studien kommer från de tre Alpina avrinningsområdena Ruetz, Kelchsauer och Gurgblach belägna i Tirol, Österrike.

Det första som kunde bekräftas var att frekvensanalys gav väldigt osäkra prognoser för de undersökta områdena. Ett så kallat dimensionerande flöde för en händelse som i medeltal förväntas återkomma med 50-årsintervall, varierade mellan 100-300 m^3s^{-1}. Med ett sådant resultat är det både svårt och osäkert att komma med rådgivning om flödesdimensionering. Fragmentsmodellen återskapade extremnederbörd på timskala (inom de osäkerhetsmarginaler som är vetenskapligt accepterade att använda) för Kelchsauer och Gurglbach. För Ruetz, där det fanns mindre nederbördsdata att tillgå, var resultaten sämre och majoriteten av de simulerade datapunkterna hamnade strax utanför marginalerna. HBV-modellen förbereddes genom att kalibrera, det vill säga, försöka sätta värden på de olika parametrarna i ekvationerna som beskriver avrinningsområdet. I valideringen, förfarandet där det undersöks hur väl modellen är kalibrerad för att återskapa observerade mätdata kunde det konstateras att simuleringar av extremvärden ofta var skiljt från de observerade. Däremot visade ekvationen som mäter modelleffektivitet att modellen reproducerade värden med accepterade skillnader. En osäkerhetsanalys för HBV-modellen tydde också på att parametrarna för studieplatserna var väldigt odefinierade, många olika parameteruppsättningar gav samma värde på modelleffektivitet.

Eftersom vädergeneratorn skapar både nederbörd och temperatur på daglig basis, jämfördes resultat mellan simuleringar från observerade temperaturer med flödessimuleringar där temperaturmedelvärden istället användes. Detta ansågs ha en icke försumbar påverkan på maxflödena så disaggregering av temperatur var också nödvändig. Temperaturmodellen konstruerades utifrån fragmentsmodellen. Det var möjligt eftersom temperatur och nederbörd är relaterade till varandra.Resultat från temperaturmodellen indikerade att disaggregerad temperaturdata simulerade flöde-stoppar med större likhet till observerade flödestoppar i jämförelse med användandet av temperaturmedelvärden.

100 år av nederbördsdata med daglig upplösning användes sedan för att jämföra disaggregeringsmodellerna. fragmentsmodellen producerade resultat med större likhet till historiska observationer för alla avrinningsområden i jämförelse med den mikrokanoniska modellen. Modellerna var mer olika för Ruetz, där mindre data fanns att tillgå. Vidare disaggregerades väderdata från vädergeneratorn för att testa hela modelleringsramverket. Eftersom disaggregeringen leder till stora datamängder är det ofrånkomligt att dela upp datamängder i mindre delar om simuleringar med flera parameterkombinationer ska göras. Det beror på att HBV-modellens nuvarande struktur inte kan hantera så stora mängder data.

Flera körningar är rekommenderat eftersom osäkerheten för de olika parametrarna som kalibrerats fram till HBV-modellen var stora. Frekvensanalys i den här studien bekräftas vara en väldigt osäker metod. Det ger inte heller någon information om vad som orsakat extremflödena. Vidare löser det inte problem med små datamängder och det går inte att simulera förändringar i klimat och markanvändning. fragmentsmodellen återskapade tidigare extremväder för Kelchsauer och Gurglbach. Resultaten visades vara beroende av hur mycket observerade högupplösta data som fanns tillgängligt. Vidare simulerade fragmentsmodellen resultat som var närmare de observerade än den mikrokanoniska modellen med vädergenererade data. fragmentsmodellen påvisades också kunna transformeras till en nedskalningsmodell för temperatur.

Som helhet finns det många osäkerheter kopplade till modelleringsramverket. Det är också mer komplicerat än frekvensanalys. Trots osäkerheterna, är ramverket närmre en fysisk förståelse för orsakerna bakom översvämningar än frekvenanalys. Modelleringsförfarandet möjliggör också studier av förändringar i klimat och markanvändning och kan vara en lösning när datamängder är små och på för stora tidsskalor.

Table of Contents

Abstract I

Referat II

Preface III

Populärvetenskaplig sammanfattning IV

1 Introduction 1
- 1.1 Objectives . . . 4

2 Background and Theory 4
- 2.1 Frequency Analysis . . . 4
- 2.2 HBV Light . . . 5
  - 2.2.1 HBV Light calibration . . . 6
  - 2.2.2 HBV model routines . . . 7
- 2.3 Weather generation introduction . . . 8
  - 2.3.1 Weather generator theory . . . 9
  - 2.3.2 Weather generator in this study . . . 11
- 2.4 Precipitation Disaggregation Introduction . . . 11
  - 2.4.1 Random multiplicative cascade models . . . 12
  - 2.4.2 Microcanonical cascade model with uniform splitting . . . 13
- 2.5 Method of fragments Introduction . . . 16
  - 2.5.1 Method of fragments algorithm . . . 16

3 Material and Methods 17
- 3.1 Study area and Data . . . 17
  - 3.1.1 Hydrology in the Alps . . . 18
- 3.2 Modeling Framework . . . 18
- 3.3 Frequency Analysis Methodology . . . 19
- 3.4 HBV Light Methodology . . . 19
- 3.5 Weather Generation Methodology . . . 21
- 3.6 Method of fragments Methodology . . . 21
- 3.7 Microcanonical Model Methodology . . . 22
- 3.8 Temperature Disaggregation Methodology . . . 22

4 Results 24
- 4.1 Frequency Analysis . . . 24
- 4.2 HBV calibration and Validation . . . 26
  - 4.2.1 HBV sensitivity study . . . 30
- 4.3 Precipitation Disaggregation . . . 31
  - 4.3.1 Method of fragments . . . 31
  - 4.3.2 Disaggregation models with weather generated data as input . . . 34
- 4.4 Temperature Disaggregation . . . 37
  - 4.4.1 Temperature Disaggregation With Observed Data . . . 37

4.4.2 Temperature disaggregation influence in the HBV model 38
4.4.3 Temperature disaggregation with weather generated data 39

5 Discussion **39**
5.0.1 Frequency analysis 39
5.0.2 HBV 40
5.0.3 Precipitation disaggregation 41
5.0.4 Temperature disaggregation 42
5.0.5 Modeling Framework 43

6 Conclusions **45**

7 References **46**
7.0.1 Personal communication 49

1 INTRODUCTION

Understanding potential changes in precipitation patterns due to anthropogenic emissions of greenhouse gases is central to constrain and improve projections of expected changes to the global hydrological cycle. Many climate models predict that extreme rainfall events will increase in frequency in an anthropogenic warmed climate (Allan & Soden, 2008; Gobiet et al., 2014; Westra et al., 2013). These changes are controlled by interactions of many thermodynamic processes. For example, warmer conditions increase the moisture capacity of air, an important process that is expected to lead to more intense precipitation in a warmer climate (Gobiet et al., 2014). It has also been suggested that changes in rainfall extremes might be underestimated by models (Allan & Soden, 2008). Mountainous areas, as for example the Alpine regions in Europe, where many of the floods in Europe during recent years occurred and caused major damage and loss of life, may be particularly vulnerable to such changes (Baldassare & Ranzi, 2003; Gobiet et al., 2014). The costliest flood affecting Europe and the Alpine countries Germany, Austria and Italy in 2002, resulted in material damages of approximately 27 billion US$ (inflation-adjusted) (Kundzewicz et al., 2012).

Extreme precipitation is expected to increase in intensity in all seasons for most of the Alpine regions. In relation to current conditions, this corresponds to a reduction of return periods for floods. In fact, a two degree increase in temperature with a 10 % increase in precipitation can transform a 100-year winter flood to a flood event with a return period of 20 years (Gobiet et al., 2014). Furthermore, intensification of precipitation up to + 30% during fall has been suggested for the northern Alps, equivalent to more than a halving of return periods which may lead to more repeated and severe flooding. Winter and spring floods are also projected to increase in magnitude and frequency in a warmer climate (Gobiet et al., 2014).

Many factors contribute towards making the Alpine regions vulnerable (Weingartner et al., 2003; Gobiet et al., 2014). Orographic mechanisms that extract ambient moisture are one major reason causing heavy precipitation events in the Alps (Gobiet et al., 2014). Moreover, higher areas have a greater total volume and more frequent rain than areas on lower altitudes. Lower regions of mountains are especially prone to convective (uplift) short term heavy precipitation. The soil depth typically declines with increasing altitude leading to quicker hydrological responses. The vegetation growth in mountain areas is limited to short periods resulting in reduced protection against erosion, minor interception of precipitation and low evaporation rates. Furthermore, the gravitational influences from higher slopes in the mountain regions are a determining factor for increased runoff generation and reduced retention times. Precipitation intensities in the European Alps have been reported to amount up to 100 $mm\ h^{-1}$ under present climate conditions (Weingartner et al., 2003). It's further reported that precipitation intensities of more than 70 $mm\ h^{-1}$ and 240 $mm\ d^{-1}$ has been reported to be quite common in the Alpine areas (Kobold & Brilly, 2006). The immense societal consequences that any shift in precipitation intensity or frequency could result in, highlights the importance of this research area (Westra et al., 2013).

However, the physical processes that give rise to floods are complex and controlled by a range of intricate interactions between variables such as snowmelt, precipitation regime, catchment characteristics and the current state of the catchment (Merz & Blöschl, 2003). Despite this complexity, flood modeling is often limited to fitting a theoretical probability distribution to a sample of observed discharge peaks (Merz & Blöschl, 2003; Moran, 1958; Chow et al., 1988). These distributions are supposed to represent the cumulative effect of all physical processes governing runoff generation (Merz & Blöschl, 2003; Chow et al., 1988). Since the true distribution of the sample is unknown, it is of common practice to guess the shape by constructing probability plots and histograms of the occurrences as guidance (Kottegoda & Rosso, 2008). A consequence of this simplification is that this methodology does not provide any information about the physical causes of the floods and their relation to flood probabilities. In addition, data series may provide sufficient knowledge about daily occurrences but not the extremes which are to be modeled. Since the sample does not provide information on the shape of the tail and information on the physical basis of the floods, extrapolation tends to perform poorly beyond the conditions of the sample (Merz & Blöschl, 2003; Moran, 1958; Chow et al., 1988). In other words, the projection of low frequent extremes may be highly uncertain.

In extent, fitting a suitable distribution for a given dataset might be difficult and this way of modeling may be limited to reproduce only some of the desired properties in the observed dataset. To circumvent such challenges, models that make fewer assumptions about distributional and parametric properties need to be introduced. Such models may be resampling models (Haberlandt et al., 2011) or conceptual models, which are a compromise between black-box and more physically based models. One such conceptual model is the precipitation-runoff HBV (Hydrologiska Byråns Vattenbalansavdelning) model which has been widely used in research practices such as flood assessment (Seibert, 2012; Ding et al., 2016; Breinl, 2016; Seibert, 2012; Bergström 1992). The HBV model was originally introduced by Bergström (1976) and simulates discharge with precipitation, temperature and estimates of monthly long-term potential evaporation rates as input. An advantage of precipitation-runoff models like the HBV is that no assumptions on the distributional properties of the dataset need to be made (Seibert, 2012). Other advantages, compared to distributional fitting, are the models applicability to studies of changes in land use and climate (Seibert, 1999a; Lindström et al., 1997).

However, modeling of low frequent events such as extreme floods, presumes a rigorous calibration and validation process so that internal procedures are simulated correctly and parameter values constrained. Most of the parameters, usually around 10-15, used in conceptual models needs to be determined during calibration (Seibert, 2000). Many different parameter sets may produce equally good results, something referred to as equifinality. This motivates the use of different parameter groups to account for model uncertainty (Beven & Freer, 2001).

Manual calibration is time consuming and potentially subjective. Therefore, automatic multi- or single-criteria calibrations are often used as an alternative. The conceptual HBV precipitation-runoff model allows for different automatic calibration procedures that can be easily applied and evaluated (Seibert, 2012). However, even when using conceptual models like HBV, one must often, as with the distributional fitting to a sample, extrapolate beyond the probabilities that can be justified from the given observations. A difference between the two procedures is increased confidence in the validation procedure on the conceptual model, since this is performed on an independent time interval (Seibert, 1999a). Hydrological modeling, even when using conceptual models like HBV, is closer to understanding the physical processes of catchment hydrology (Lindström et al., 1997; Seibert, 1999a, 1999b, 2012) compared to distributional fitting (Merz & Blöschl 2003; Moran, 1958; Chow et al., 1988).

Long series of high-resolution precipitation data is a necessity to model and understand the impact of specific floods (Müller & Haberlandt, 2015; Westra et al., 2012; Kobold & Brilly, 2006). In fact, data at hourly resolution is often essential for flood design purposes since flood peaks may be maintained for as little as hours or minutes (Pui et al., 2009; Kobold & Brilly, 2006). High resolution data is particularly motivated since many of the expected changes in precipitation patterns are predicted at sub-daily time intervals (Westra et al., 2013). In contrast, available precipitation data series are often short, the resolution inadequate and spatial coverage lower in relation to what is needed (Müller & Haberlandt, 2015; Westra et al., 2012). To overcome the issue of short time series, precipitation can be synthetically generated using weather generators, providing long time series where extremes are better captured (Haberlandt et al., 2011; Breinl, 2016). Extending historical records or to generate new ones have been helpful for hydrological design purposes (Pui et al., 2009; Molnar & Burlando, 2005). Coupling weather generators with hydrological modeling enables uncertainty and impact assessment on changing land use and climate conditions (Bergström et al., 2001; Booij, 2005; Breinl, 2016).

Weather generators do however generally provide time series at daily resolution (Pui et al., 2009; Breinl, 2016) which, as mentioned above, may be insufficient in flood assessment (Pui et al., 2009; Kobold & Brilly, 2006). However, a main reason behind generating daily weather data is the absence of continous long term precipitation data (Breinl et al., 2015; Müller & Haberlandt, 2015; Westra et al., 2012).

Precipitation disaggregation has therefore been proposed and successfully tested to increase resolution in precipitation data (Olsson, 1998; Güntner et al., 2001; Müller & Haberlandt, 2015; Molnar & Burlando, 2005; Pui, et al., 2009, 2012). Disaggregation of precipitation data has also been used in flood analysis (Pui, et al., 2009). However, many of the mentioned studies (Olsson, 1998; Güntner et al., 2001; Müller & Haberlandt, 2015; Molnar & Burlando, 2005; Pui et al., 2009, 2012) have focused on the precipitation reproducing properties and not the resulting runoff. Therefore, it is of high interest to assess a modeling framework consisting of a weather generator, precipitation disaggregation models coupled with the hydrological HBV model as an alternative to estimate design floods with conventional frequency analysis. The HBV model is computationally inexpensive and thus allow for many runs and proper uncertainty analysis (Ding et al., 2016; Breinl, 2016; Seibert, 2012; Bergström, 1992).

1.1 OBJECTIVES

The objectives with this thesis were to investigate uncertainties with a modeling framework to estimate design floods in relation to uncertainties with frequency analysis. The framework consisted of a weather generator to extend historical records of precipitation and temperature data, a disaggregation procedure to increase resolution in the synthetically generated weather data which then could be used as input in the hydrological HBV model to estimate design floods in Alpine catchments. The framework was examined to provide an alternative to conventional frequency analysis methodology where fitting different probability distributions to annual peak discharges with maximum likelihood or the method of moments is the common practice. Another objective was to compare the output of the disaggregation models and assess their uncertainties.

The specific objectives were:

- Use conventional frequency analysis to estimate design floods with return periods of 20 and 50 years for the three study catchments and assess uncertainties with the method
- Compute the method of fragments model for precipitation and temperature disaggregation and assess simulation results and uncertainties
- Couple weather generated temperature and precipitation data with the method of fragments and the multiplicative microcanonical cascade model with uniform splitting and evaluate their simulation results and uncertainties
- Evaluate uncertainties with the modeling framework for estimating design floods

2 BACKGROUND AND THEORY

2.1 FREQUENCY ANALYSIS

Hydrologic systems are occasionally exposed to extreme events such as storms and floods. The frequency of occurrence of a specific flood or storm is inversely related to the magnitude of the event. An extreme event is, in comparison with more moderate events, per definition something that occurs less frequent. The aim of frequency analysis is to correlate certain events in hydrologic data with theoretical probability distributions. All probability distributions are functions of random variables showing their probability of occurrence (Chow et al., 1988). Fitting a theoretical probability distribution to a sample of observed discharge peaks is a common engineering procedure (Haberlandt et al., 2011; Merz & Blöschl, 2003). When doing so, one assumes that the hydrologic data is independent, identically distributed and is produced by a rainfall storm system which is considered stochastic (partly deterministic and partly random), space-independent (the system is regarded as a single point or equally the variable does not change in relation to its position) and time-independent (the hydrological events do not influence each other) (Chow et al., 1988).

In order to not compromise these criteria, one needs to be careful when choosing data. In practice that often means choosing the annual maximum peak flow at wanted resolution with the assumption that the discharge is independent from year to year (Chow et al., 1988). If the highest annual discharge is drawn from a historical series at hourly resolution that means it is the largest of 8760 values. These values may therefore be located in the extreme tail of their parent probability distribution (Chow et al., 1988).

Not surprisingly, the probability distribution of extremes may differ from the distribution from which the extreme values were drawn. The distribution of extreme values has been seen to converge to one out of three forms of asymptotic extreme value distribution types referred to as Type I, II and III. Common asymptotic distributions used in hydrological applications of extremes are The Extreme Value Type I (EV1) or Gumbel distribution, the Log-Pearson Type III distribution which is the standard distribution for annual maximum flows in the United States, Pearson Type III distribution and the General Extreme Value (GEV) distribution which all of the other asymptotic distributions derive from. Another common distribution used in flood design practices is the Lognormal distribution (Chow et al., 1988).

The central limit theorem states that if a sequence of random variables X_i are distributed independently and identically with mean μ and variance σ^2, the sum of n random variables $Y = \sum X_i$ tends to follow the normal distribution with mean $n\mu$ and variance $n\sigma^2$ as n becomes large regardless of the original distribution function of X (Chow et al., 1988).

There are two commonly applied methods to fit a distribution to a sample of hydrological data. The first method is called *the method of moments* and the other is called *the method of maximum likelihood*. When fitting a distribution to a sample, all the original information of the sample is compacted to the probability function and its associated parameters. Fitting a distribution to a sample of observations with the method of moments can be done by calculating a frequency factor, K_t, to estimate a flood with return period T. The degree of fit for a sample to a proposed distribution can be evaluated using a Chi-square test (χ^2) for example. A confidence level is then chosen for the null hypothesis to determine if there is a significant relation between the fitted distribution and the data. A more theoretically appropriate method for fitting a distribution to a sample, is however the method of maximum likelihood. This is true in the sense that the method produces the most efficient parameters, i.e. those which approximate the sample parameters with least average difference (Chow et al., 1988).

2.2 HBV LIGHT

The HBV model is a widely used conceptual runoff model (Bergström, 1976, 2001; Seibert, 2012; Breinl, 2016; Kobold & Brilly, 2006; Ding et al., 2016). The HBV model is fast and straightforward in comparison to more complex, fully distributed physical hydrological models which come at a higher computational cost (Arnold et al., 1998). A more complex model does not necessarily mean better modeling results (Das et al., 2008) something that motivates the use of a conceptual model like the HBV. The HBV model simulates discharge with rainfall, temperature and estimates of monthly long-term potential evaporation rates as input and has been used in many flood simulation studies (Ding et al., 2016; Breinl, 2016; Seibert, 2012; Bergström, 1992). It is a semi-distributed model, i.e., catchments can be separated into subcatchments as well as into different elevation and vegetation zones. The model consists of different routines, which are a snow routine, a soil routine and a response routine where runoff is computed from a function of water storage. The response routine is calculated at a specific location (lumped) in contrast to the snow and soil routine where calculations are carried out for each different elevation zone. (Seibert, 2012).

In a deterministic lumped model like the HBV, a given input with the same parameter values always result in the same output, an output which is spatially averaged in a single point in space without dimensions (Chow et al., 1988).

A straightforward an updated user-friendly version of the HBV model, *HBV light*, has been developed at the University of Zürich and was used in this study (Seibert, 2012).

2.2.1 HBV Light calibration

Monte-Carlo simulations and the evolution based Genetic Algorithm and Powell optimization (GAP) can be used in the HBV light version for automatic calibration (Seibert, 2012). The evolution-based generic algorithm mimics evolution by giving parameter sets with satisfactory simulation more chances to generate new sets than those with poorer results. Multi-criteria calibration using many objective functions to access the efficiency of the model has been seen to constrain parameter values (Seibert, 2000). However, the values of different objective functions, which all judge the results of the parameter sets by dissimilar criteria, are not directly comparable and therefore hard to combine. One way to allow for comparison is to assign weights to the different objective functions in the model and then combine these reconstructed functions into a so called fuzzy measure (Seibert, 1999).

Assigning weights can be done manually in the HBV light model before running the automatic calibration. Finding weights that balance the objective functions so that the resulting parameter ranges are constrained, may be complicated. The LindstromMeasure objective function, which is available in the HBV light model version, is an empirically derived model performance measure, which combines different objective functions into a fuzzy value. The objective functions used in the LindstromMeasure are already weighted and could therefore be used to constrain ranges of parameter values as well as to assess model performance. The LindstromMeasure contains the coefficient of efficiency, R_{eff} (1), originally developed by Nash & Sutcliffe (1970) and the volume error which is weighted to 0.1 in equation (2). R_{eff} evaluates if the model is a better measure to estimate runoff in comparison with a benchmark series, usually the mean (Seibert, 2001).

$$R_{eff} = 1 - \frac{\sum(Q_{obs} - Q_{sim})^2}{\sum(Q_{obs} - \overline{Q}_{obs})^2} \tag{1}$$

$$LindstroemMeasure = R_{eff} - 0.1\frac{|\sum(Q_{obs} - Q_{sim})|}{\sum(Q_{obs})} \tag{2}$$

2.2.2 HBV model routines

The HBV model can be used on any temporal resolution, as long as the time interval, t, remains constant. The form of precipitation is conditioned to temperature. If precipitation falls when the temperature is below the temperature threshold value, TT (°C), precpitation is simulated to be snow and vice versa. When the temperature is below TT precipitation is simulated as snow and multiplied with a snow correction factor, $SFCF$. This is done to compensate for systematic errors in snowfall measurements and the "missing" evaporation from snowpack in the model. The melting of snow is calculated with a degree day method (3). The snowpack can store a certain amount of meltwater. If this portion, CWH, is not exceeded, meltwater is retained within the snowpack. $CFMAX$ ($mm\,°C^{-1}$ d^{-1}) (3) is the degree day factor. Some of the meltwater refreezes within the snowpack, a relationship governed by a refreezing factor, CFR, in equation 4 (Seibert, 1999).

$$melt = CFMAX(T(t) - TT) \tag{3}$$

$$refreezing = CFR \cdot CFMAX(TT - T(t)) \tag{4}$$

Melting water from snowmelt and precipitation (P) are then divided into either the soil or groundwater box conditional to the relationship of the content of water in the soil box, soil moisture SM (mm) and the maximum water content FC (mm) (5). Actual evaporation from the soil box depends on the ratio $\frac{SM}{FC}$. If the ratio is bigger than LP, actual evaporation equals potential evaporation while a linear reduction is used when the ratio is below LP (6) (Seibert, 1999).

$$\frac{recharge}{P(t)} = \left(\frac{SM(t)}{FC}\right)^{BETA} \tag{5}$$

$$E_{act} = E_{pot}\left(\frac{SM(t)}{FC \cdot LP}, 1\right) \tag{6}$$

The upper groundwater box, SUZ (mm), is first filled when precipitation or meltwater is added to the groundwater box. The parameter $PERC$ ($mm\,d^{-1}$) describes the maximum proportion of percolating water from the upper to the lower groundwater box SLZ (mm). Evaporation and precipitation in lakes are conceptualized by adding or subtracting water from the lower groundwater box. Runoff from the groundwater boxes depends on K_0, K_1 and K_2 (d^{-1}) representing the outflow equations. It is optional whether one want to use two or three of these linear equations. Runoff is calculated as the sum of the outflow equations depending on if SUZ is above the threshold value, UZL (mm), or not (7). A final transformation of the runoff is then done in the triangular weighting function (9) defined by the parameter $MAXBAS$ (h) to simulate discharge.

$$Q_{GW}(t) = K_2 \cdot SLZ + K_1 \cdot SUZ + K_0 max(SUZ - SLZ, 0) \tag{7}$$

$$Q_{sim}(t) = \sum_{i=1}^{MAXBAS} c(i) Q_{GW}(t - i + 1) \tag{8}$$

Where $c(i)$ equals (9).

$$\int_{i-1}^{i} \frac{2}{MAXBAS} - \left| u - \frac{MAXBAS}{2} \right| \frac{4}{MAXBAS^2} du \tag{9}$$

By using deviations of the temperature $T(t)$ from its long term-mean value, a correction factor, C_{ET} the long-term mean potential evaporation, $E_{pot,M}$, for a certain day of the year can be corrected to the real value, E_{pot}, of day t ($°C^{-1}$) (10).

$$E_{pot}(t) = (1 + C_{ET}\,(T(t) - T_M))\, E_{pot}, M 0 \leq E_{pot}(t) \leq 2E_{pot}, M \tag{10}$$

Where $0 \leq E_{pot}(t) \leq 2E_{pot}, M$.

To account for the differences in temperature and precipitation in relation to altitude, P_{CALT} (11) and T_{CALT} (12) are used.

$$P(h) = P_0 \left(1 + \frac{PCALT(h - h_0)}{10000} \right) \tag{11}$$

$$P(h) = T_0 - \frac{TCALT(h - h_0)}{100} \tag{12}$$

Further details on the model structure and its functions can be found in Bergström (1992) and elsewhere.

2.3 WEATHER GENERATION INTRODUCTION

As mentioned above, rainfall runoff modeling for flood risk assessment require, like many other hydrology fields or water resources practices, continuous high-resolution precipitation data. In contrast, available precipitation time series are often short (Müller & Haberlandt, 2015; Westra et al., 2012). Continuous time series are of particular interest in modeling floods since it enables to account for both the antecedent conditions prior to the event as well as for the precipitation event influencing the flood. Both these factors could have a potential significant impact on the resulting magnitude of the flood (Westra et al., 2012). Long and continuous time series are also motivated to capture the intermittency in precipitation and sequences of storm pulses (Pui et al., 2009; Molnar & Burlando, 2005).

Moreover, simulating hydrological processes with observed data as the only input results in estimations based on one individual realization. If another series with similar attributes as the observed were used, it may change the simulation results (Richardson, 1981). Different types of weather generators have therefore been proposed to extend historical records in several studies (Haberlandt et al., 2011; Richardson, 1981; Breinl et al., 2015). Weather generators are also used to assess anthropogenic effects of proposed modifications in the hydrological system, as meteorological data are used as input for models that simulate hydrological processes (Richardson, 1981).

Many different stochastic rainfall generation models have been proposed during recent years for short time step rainfall. Due to difficulties with modeling the high intermittence of high temporal resolution data (which require many parameters) the classical alternating renewal models are more suitable for daily or longer time step precipitation. Such models are based on a series of dry-wet spell events. Moreover, the model structure and the parameter estimation procedure is straightforward and based on point observations. Other weather generators are autoregressive-moving-average models (ARMA). ARMA models operate under the premise that current values depend directly on one or more previous time steps (depending on the order of the autoregressive component) and a moving average element. These models are often applied to the simulation of continuous time series, like temperature (Haberlandt et al., 2011).

2.3.1 Weather generator theory

Precipitation can be described as a stochastic process, i.e., a process Y_t with sequences of random variables X_i (that usually are correlated). A precipitation time series represents one such realization of the process. Dimensionality, persistence and stationarity are important characteristics in stochastic processes. A univariate or one-dimensional system considers one variable as a function of time while multivariate systems include more than one variable that may depend on both time and space. Many hydrological processes are continuous and the variables controlling the process may inherit information from earlier time steps (dependent variables). Such processes are categorized as processes with Markov-character, which means that all required information for Y_t can be acquired exclusively from t previous time steps (Haberlandt, et al., 2011).

If the stochastic process instead is governed by a probability distribution independent of time, the process is called stationary or time invariant. Stochastic processes can be separated into three different types: the normal type, the point process and the alternating process. The normal process is characterized by constant or steady varying behavior, a process without sudden change (e.g. a river flow). Point processes are instead characterized by being events of short duration occurring randomly in time. The alternating routine comprises two exclusive normal processes. An intermittent routine is a special case for which one of the alternating processes is zero (Haberlandt et al., 2011).

Precipitation is a complex version of an intermittent process (wet-dry periods) in space and time. The occurrence and intensity of rainfall are in many weather generators seen as dissimilar processes. Precipitation intermittency and variability increase with decreasing time steps. At daily resolution, a precipitation time series displays temporal intermittency. Finding a suitable model which can simulate both occurrence and precipitation amount is therefore needed. A common method for doing this is to use two separate models, one for occurrence and one model for precipitation depths. Markov chain models are suitable for modeling the occurrence of precipitation (Haberlandt et al., 2011; Richardson, 1981; Breinl et al., 2015).

Discrete observations in the time series represent altered states in the Markov chain and a sequence of states is described using transition probabilities. Typically, transition probabilities depend on a few previous time steps of the chain and are conditional. Similarly, the probability governing the sequent time step depends on q foregoing time steps. The order of q determines the order of the Markov chain. Precipitation occurrence is generally modeled using a two-state Markov chain of first order. A day with precipitation is represented by a number, as for example 1, corresponding to a wet state and a day without precipitation by 0 corresponding to a dry state. Four transition probabilities are used in Markov chain models (Haberlandt et al., 2011).

(i) p_{10} wet state followed by a dry state

(ii) p_{00} a dry state followed by a dry state

(iii) p_{11} a wet state followed by a wet state

(iv) p_{01} a dry state followed by a wet state

p_{01} and p_{00} sums up to 1 which also goes for p_{11} and p_{10}. These probabilities can be derived from the observed time series. A precipitation time series can then be simulated using a uniform $[0, 1]$ random number generator which compares the generated value to that of the transition probability. A number corresponding to a value less then the transition probability will result in a state which is wet on the following day or dry respectively if the value is higher than the transition probability. The precipitation depths for different wet states can be estimated using a suitable probability distribution assuming that precipitation amounts are serially independent (Haberlandt et al., 2011).

Assessment on the output of hydrological models when using synthetic rainfall data from weather generators as input, has only been done in a few studies to the author's knowledge (e.g. Booij, 2005; Breinl, 2016; Haberlandt et al., 2011).

2.3.2 Weather generator in this study

A weather generator must develop an underlying structure to account for certain relationships between meteorological processes. For example, temperature is more probable to be below normal on a rainy day and the maximum and minimum temperature is likely to be small on a cloudy day. Here temperature is conditioned to precipitation since this is the primary variable (Richardson, 1981).

In this study the weather generator of Breinl et al (2015b) was used one-to-one with permission for the generation of daily weather at the study sites. The simulations from the weather generator were then used for disaggregation. As many other weather generators, this model uses two-state Markov chains of second order for the generation of daily precipitation occurrences. Precipitation amounts are simulated using a two-step process which first resamples observations using parametric functions and then reshuffles the parametric values (Breinl et al., 2015b).

Here, a gamma distribution was used (different models for each month) which is the most common distribution for this purpose. The gamma distribution was also suitable since it has been applied in the geographically neighboring state with similar environment (Breinl et al., 2013). This methodology allows for creation of unobserved extremes which is of great interest to make projections on future scenarios (Breinl et al., 2015b). Many other traditional weather generators can only reproduce observed precipitation events. Temperature data are simulated conditioned to rainfall by using ARMA models (Breinl et al, 2013, 2015b). For more information on the weather generator, see (Breinl et al, 2013, 2015a, 2015b; Richardson, 1981).

2.4 PRECIPITATION DISAGGREGATION INTRODUCTION

As mentioned in the introduction, data at hourly resolution is often indispensable for flood design modeling (Pui et al., 2009). In fact, flash floods, floods associated with isolated and localized very intense precipitation in small and medium sized catchments, may be especially hard to simulate since peak discharges are maintained only for hours or minutes (Kobold & Brilly, 2006).

Disaggregation of precipitation data is thus required and provides an attractive alternative to achieve sufficient temporal resolution (Olsson, 1998). It is important that these models preserve central statistical properties such as the characteristics of extreme events, intermittency and scaling distributions seen in the observed rainfall (Molnar & Burlando, 2005). The floods seen in Europe and elsewhere during recent years emphasize the importance of evaluating short-term processes of runoff generation adequately (Güntner et al., 2001). Therefore, different precipitation disaggregation methods such as the random multiplicative cascades (microcanonical and canonical) and the nonparametric method of fragments have been tested for obtaining fine resolution rainfall data for design flood analysis (Pui et al., 2009).

These models have been seen to reproduce conventional statistics (such as mean and variance) and empirical wet spells satisfactory. If the output of the precipitation disaggregation models is to be used for flood design purposes, the generation of realistic dry and wet spells are especially important to account for antecedent soil conditions. In general, the method of fragments model turned out to outperform the aforementioned models and was the only method to satisfactorily mimic extreme weather behavior at hourly scale (Pui et al., 2009). However, the method of fragments effectivity to reproduce observed precipitation can also serve as its limit. The model's underlying logic is based on resampling observed precipitation events with the closest resemblance to the day to be disaggregated. Therefore, it is expected to produce statistics with close affinity to that of the observed time series. This may also limit the model's ability and flexibility to produce unobserved events. The same problem faces the cascade models as well, since their parameters are estimated based on historical observations, assuming that parameters are stationary even under changed climate conditions. Therefore, it could be highly useful to test different disaggregation models on their capability to simulate events beyond historic observations when using synthetic precipitation with unobserved measures as input (Pui et al., 2009).

In contrast, cascade modeling for disaggregating precipitation data has been used successfully to capture fundamental features of precipitation processes over a range of scales. The distribution of dry and rainy periods as well as the scaling structures of observed data are very well reproduced with a microcanonical disaggregation approach while the number of dry intervals as well as the durability of dry and wet periods tend to be slightly underestimated. This indicates that the precipitation producing mechanism inherent a certain cascade type behavior. It also suggests that precipitation models which are scaling-based could provide an important tool in hydrological practices (Olsson, 1998). This has been strengthened by Müller & Haberlandt (2015) who state that uniform splitting, a microcanonical cascade model version, is able to reproduce rainfall characteristics well. Their model was also able to mimic statistics from observed time series while observed extreme values were reproduced reasonably well, thereby slightly overestimating the precipitation amounts for higher non exceedance probabilities (Müller & Haberlandt, 2015).

2.4.1 Random multiplicative cascade models

Mass is distributed gradually in a multiplicative manner into consecutive levels in discrete multiplicative cascade models. If $V > 0$, i.e., the box is wet, mass is distributed from one cascade level to the nearest higher, which corresponds to an increase in resolution depending on the branching number, b. A branching number b equal to two refers to a doubling of resolution and mass is distributed between two boxes which each represent a time interval T and an associated volume V (Olsson, 1998). The transition of mass from higher to consecutive cascade levels occurs via a so-called cascade generator, W, which distributes the mass in a multiplicative manner. Depending on how mass is distributed, i.e., how the cascade generator is programmed, determines the version of cascade model. Cascade generators that preserve mass on average during the disaggregation are noted canonical while a generator that preserves mass exactly are called microcanonical (Molnar & Burlando, 2005). Precipitation disaggregation in multiplicative cascade models is reminiscent of observed precipitation scaling patterns. These models have been successfully applied to precipitation modeling (Molnar & Burlando, 2005; Müller & Haberlandt, 2015; Pui et al., 2009; Olsson, 1998). In Figure (1) the branching number b is equal to two in all the consecutive levels to achieve a resulting temporal resolution of 0.75 h (Müller & Haberlandt, 2015).

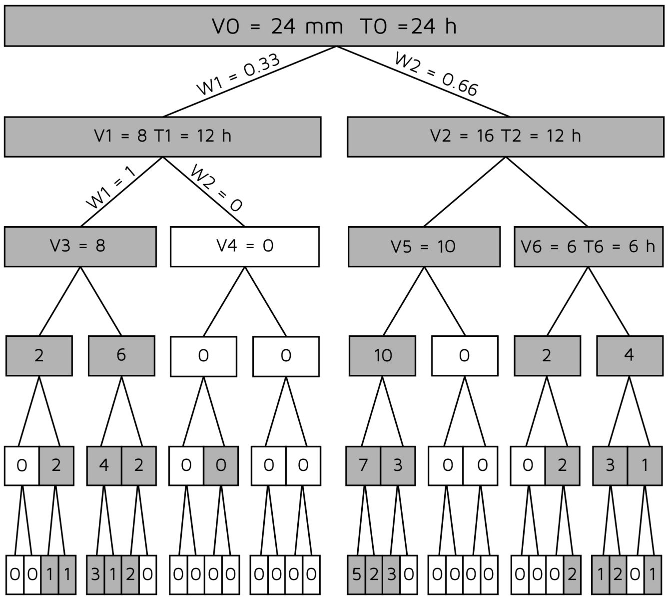


Figure 1. A principle sketch of a microcanonical cascade model. Here, a 24 mm rain is disaggregated from daily, $24h \rightarrow 12h \rightarrow 6h \rightarrow 3h \rightarrow 1.5h \rightarrow$ to $0.75h$. Boxes of zero are considered dry and nonzero boxes are wet.

It should be noted that even if comparisons between different disaggregation models without doubt offer useful insights, it is impossible to draw any general conclusion on best model structure for precipitation disaggregation. This has to do with precipitation properties and generating mechanisms being significantly different at different climate regions (Pui et al., 2012).

2.4.2 Microcanonical cascade model with uniform splitting

In this study, a microcanonical cascade model with uniform splitting was used (Müller & Haberlandt, 2015).

For every increase in resolution, the precipitation volume V from a coarser level i is distributed in a multiplicative manner via the cascade generator W into a successive level with a finer resolution. The mass is preserved exactly during disaggregation. To achieve a temporal resolution of one hour, this method uses branching number b equal to three in the first disaggregation step and then b equal to two in the consecutive levels. This leads to a temporal resolution of one hour ($24h \rightarrow 8h \rightarrow 4h \rightarrow 2h \rightarrow 1h$). For b equal to three, there are seven principal possibilities of how precipitation can be distributed when disaggregated with the condition that the boxes (time steps) T_i and T_{i+1} are short enough to have a nonzero probability of zero precipitation. Each split is related to a probability, P. There are seven different possibilities of disaggregation when b is equal to three (Olsson, 1998; Müller & Haberlandt, 2015; Güntner et al., 2001). If mass is distributed so that the whole volume of precipitation is allocated to the first time interval, T_i, i.e., if the weighted multiplicators from the cascade generator have the values $W_1 = 1$, $W_2 = 0$ and

$W_3 = 0$, corresponds to one wet interval of the particular day with probability P(1/0/0). For each split, the sum of W_1, W_2 and W_3 is equal to one. The probability P(1/0/0) does not affect the position of the wet interval, only the number of wet boxes to avoid over-parametrization. The position of a wet box is assigned randomly (Müller & Haberlandt, 2015).

The other possibilities of precipitation disaggregation using b equal to three are the probabilities for two or three wet intervals during a day. The probability for two wet intervals P(0.5/0.5/0) can occur in three different ways. The probability of three wet intervals can occur in one way, i.e., P(0.33/0.33/0.33) (Müller & Haberlandt, 2015).

$$W_1, W_2, W_3 = \begin{cases} 1, 0 \text{ and } 0 \text{ with probability} & P(1/0/0) \\ 0, 1 \text{ and } 0 \text{ with probability} & P(0/1/0) \\ 0, 0 \text{ and } 1 \text{ with probability} & P(0/0/1) \\ \frac{1}{2}, \frac{1}{2} \text{ and } 0 \text{ with probability} & P(\frac{1}{2}/\frac{1}{2}/0) \\ 0, \frac{1}{2} \text{ and } \frac{1}{2} \text{ with probability} & P(0/\frac{1}{2}/\frac{1}{2}) \\ \frac{1}{2}, 0 \text{ and } \frac{1}{2} \text{ with probability} & P(\frac{1}{2}/0/\frac{1}{2}) \\ \frac{1}{3}, \frac{1}{3} \text{ and } \frac{1}{3} \text{ with probability} & P(\frac{1}{3}/\frac{1}{3}/\frac{1}{3}) \end{cases}$$

The precipitation volume from a coarser level is distributed uniformly amongst all boxes that are defined as wet. These parameters can be estimated through calculating the number of occurrences for each combination from a high-resolution historical time series. For higher precipitation depths, the probability for two or three wet 8-h intervals increases. Therefore, a volume threshold is identified to account for these differences and to separate precipitation into different volume classes conditional to total daily precipitation depths. The threshold was chosen to $q_{0.998}$ in this study in line with Müller & Haberlandt (2015). A microcanonical model with uniform splitting has branching number b equal to three in the first step followed by b equal to two in the consecutive steps Figure (2) (Müller & Haberlandt, 2015).

For level two to five where the branching number b is equal to two, there are three principal possibilities of how rainfall can be distributed when disaggregated. The three different possibilities of precipitation disaggregation are shown below. If mass is distributed so that the whole volume of precipitation is allocated to the first time interval, T_i, i.e., if the weighted multiplicators from the cascade generator have the values $W_1 = 1$ and $W_2 = 0$ corresponds to P(1/0). The weighted multiplicators are not independent of each other. For each split, the sum of W_1 and W_2 is equal to one. With probability P(0/1) splitting is achieved vice versa since no precipitation is assigned to the first time step W_1. The third possibility of disaggregation is a redistribution of mass from the coarser level to both finer time steps so that $W_1 = x$, $0 < x < 1$, $W_2 = 1 - W_1$. This corresponds to the probability P(x/(1-x)) where x is a random variable in all disaggregation steps. The density function $f(x)$ can be estimated if a probability for each value of x is assigned and if these values are associated with a certain theoretical probability distribution $W_{x/x}$ (Olsson, 1998; Müller & Haberlandt, 2015; Güntner et al., 2001).

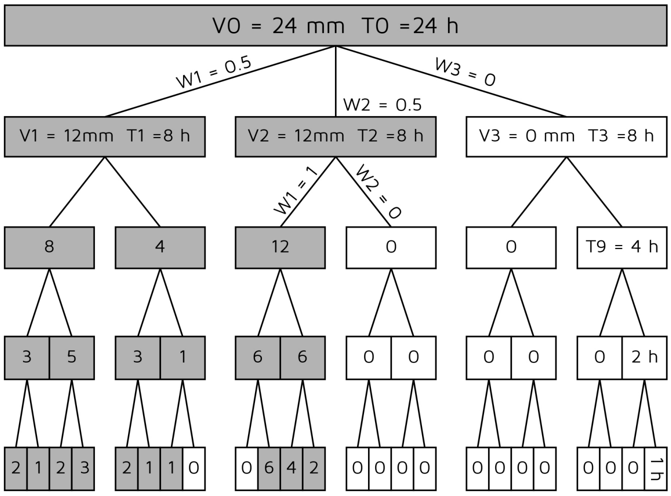


Figure 2. A schematic visualization of a microcanonical cascade model with uniform splitting. In the first level the branching number is b = 3. Here, a 24 mm rain is disaggregated from daily to hourly values $(24 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1)$. Boxes of zero are considered dry and nonzero boxes are wet.

$$W_1, W_2 = \begin{cases} 1 \text{ and } 0 \text{ with probability} & P(1/0) \\ 0 \text{ and } 1 \text{ with probability} & P(0/1) \\ x \text{ and } 1-x \text{ with probability} & P(x/(1-x)); \quad 0 < x < 1 \end{cases}$$

Parameter dependencies have been found on the position in the precipitation sequence as well as for the precipitation volume of each time interval. This dependence could be seen due to the growth of P(x/(1-x)) for increasing volumes as well as for boxes inside a precipitation sequence in comparison to boxes in the edge of sequences where P(x/(1-x)) was lower. To account for these differences in probability, each box is divided into a specific position class depending on attributes of the adjacent boxes as well as into a volume class depending on the magnitude of the precipitation depth (Güntner et al., 2001; Olsson, 1998). The four position classes are (i) starting boxes, (ii) enclosed boxes, (iii) ending boxes and (iv) isolated boxes (Olsson, 1998; Müller & Haberlandt, 2015; Güntner et al., 2001). The probabilites for P((1/3)/(1/3)/(1/3)), P((1/2)/(1/2)/0) and P(1/0/0) are the same for all position classes and the same goes for the different probabilities when b is equal to two (Müller & Haberlandt, 2015).

(i) starting boxes dry,wet,wet

(ii) enclosed boxes wet,wet,wet

(iii) ending boxes wet,wet,dry

(iv) isolated boxes dry,wet,dry

A higher and lower volume class was also chosen for level two-five (where b is equal to two) to account for the different probabilities related to the volume of the precipitation mentioned above. The volume threshold for level two-five was here chosen to the precipitation median. This is motivated because the median had the advantage of having equal number of boxes for each volume class. The parameters were estimated by counting the number of occurrences from the historical high-resolution time series for each combination of position-volume class. The approximation of $f(x)$ was done in a similar fashion (Müller & Haberlandt, 2015).

2.5 METHOD OF FRAGMENTS INTRODUCTION

The method of fragments model was programmed in Matlab by looking at model structures used in other studies (Pui et al., 2012; Westra et al., 2012; Westra & Sharma, 2010).

The method of fragments is categorized as a resampling and a non-parametric model. One constructive attribute of the method is that it does not require the assumption of a theoretical distribution related to the dataset. Instead, attributes from the data record of interest are used as criteria in the disaggregation process. The sequence of daily precipitation to be disaggregated may either come from historical records at the site of interest or from stochastic weather generation models. The method then samples fragments of sub-daily to daily ratios from historical hourly data at the same site or from multiple nearby sites conditional on properties of the daily precipitation at the site of interest. These conditions include comparison between total daily precipitation amounts between the day of interest and aggregated hourly data from historical records in a predefined time window to account for seasonal differences. Furthermore, a classification based on whether the wetness state of the antecedent and successive day is wet or dry is used as criteria in the sampling process (Westra & Sharma, 2010).

2.5.1 Method of fragments algorithm

The following algorithm gives a thorough description of the methodology used to disaggregate the daily precipitation time series into a time series of hourly resolution in this study. The steps are as follows:

(i) Obtain the daily time series R_t to disaggregate where t is the notation for the day to disaggregate. Use historical records of hourly data, $X_{i,m}$, to form daily time series R_i (equation 13) were m is the sub-daily time step and i denotes the day. Form a data series with sub-daily to daily ratios (equation 14).

$$R_i = \sum_{m}^{24} X_{i,m} \tag{13}$$

$$f_{i,m} = \frac{X_{i,m}}{R_i} \tag{14}$$

(ii) Form a moving window with l days centered around a particular day t of the daily precipitation time series, R_t, which you want to disaggregate. l is dependent on the length of historic time series and is in this example chosen to 15 days. If March the 16^{th} is to be disaggregated, the observational window spans from 1-31 March (subtracting the same year if one uses historical precipitation from the same station). By limiting the observational window, the model also accounts for seasonal differences. Similar to the microcanonical approach, four different classes are formed depending on the wetness state of neighboring days. This is done to account for continuity in precipitation events (Pui et al., 2012).

$Class$ (1) starting boxes $[dry, wet, wet]$, $R_j \geq 0 \mid (R_{j-1} = 0, R_{j+1} \geq 0)$

$Class$ (2) enclosed boxes $[wet, wet, wet]$, $R_j \geq 0 \mid (R_{j-1} \geq 0, R_{j+1} \geq 0)$

$Class$ (3) ending boxes $[wet, wet, dry]$, $R_j \geq 0 \mid (R_{j-1} \geq 0, R_{j+1} = 0)$

$Class$ (4) isolated boxes $[dry, wet, dry]$, $R_j \geq 0 \mid (R_{j-1} = 0, R_{j+1} = 0)$

Here j represents any day within the window centered around the specific date of the day t to be disaggregated.

(iii) Identify which class R_t belongs to.

(iv) Identify the number of nearest data observation neighbors k by $k = \sqrt{n}$, where n is the notation for the sample size of the days falling within the time window l and satisfying the class criteria. Classify these neighbors according to $|R_j - R_t|$ giving the day with the lowest absolute difference the smallest rank out of the numbers $j = 1, 2, \ldots, k$ so that the ranked daily precipitation time series is noted R_j. The lowest ranked neighbor will have the highest probability $p(j)$ to be picked from equation (15) using a uniformly distributed random number (0,1) (Pui et al., 2012).

$$p(j) = \frac{1/j}{\sum_{i=1}^{k} 1/i} \tag{15}$$

Use the date of picked day to find the corresponding fragments vector in $f_{i,m}$. Insert the fragments into day t using equation (16) to form the new disaggregated time series $R_{thourly}$.

$$R_{thourly} = R_t \; x \; f_{i,m} \tag{16}$$

(v) Repeat step (ii) to (iv) for each day until disaggregation is completed. All the steps and further reading can be found in (Pui et al., 2012; Westra et al., 2012; Westra & Sharma, 2010).

3 MATERIAL AND METHODS

3.1 STUDY AREA AND DATA

The study was conducted with data from three catchments in Tyrol around the 47°N 10-12°E in the Austrian Alps. Table 1 summarizes the available data where it can be seen that Kelchsauer and Gurglbach are similar with respect to elevation ranges m.a.s.l and potential evaporation (PET). Precipitation and temperature gauges for Ruetz are located at a higher altitude compared to the gauges in the Kelchsauer and Gurglbach catchments.

The catchments were chosen based on available data and catchment area. Moreover, the catchments had to be without the influence of hydropower stations to simplify analysis. These catchments also had the HBV required continuous hourly precipitation, temperature and discharge series (referred to as PTQ in HBV). The PTQ series ranged from 2000 - 2015 for Kelchsauer, 1997 - 2015 for Gurglbach and 2000 - 2015 for Ruetz. Continuous precipitation series were longer for Kelchsauer (1978 - 2015) and Gurglbach (1979 - 2015) in comparison with Ruetz (2000 - 2015). Ruetz was of particular interest since it was recently affected by a flood event (see Anon, 2015).

Table 1. Defining characteristics for the study catchments.

Catchment	**Krößbach Ruetz**	**Kelchsauer Ache**	**Gurglbach**
Area $[km^2]$	128.8	134.2	78.5
PET $[mm\ year^{-1}]$	284	470	510
Precipitation $[mm\ year^{-1}]$	1108	1373	966
Average discharge $[m^3 s^{-1}]$	5.3	5.4	1.9
Highest discharge peak $[m^3 s^{-1}]$	141.5	99.9	39.1
Elevation range $[m\ a.s.l]$	1095-3484	661-2450	804-2580
Rain gauge elevation $[m\ a.s.l]$	2308	815	854

3.1.1 Hydrology in the Alps

Typically, Alpine regions are affected by big differences in hydrology between summer and winter. During winter discharge is generally lower due to storage in ice or snow. Ice and snow significantly influence the hydrology in Alpine catchment. Runoff generation during winter is regulated by snow cover which results in a time lag between precipitation and discharge (Vanham et al., 2008).

3.2 MODELING FRAMEWORK

A conceptual sketch of the modeling framework seen in Figure (3) clarifies the different steps of the proposed framework. The weather generator extrapolates data points from precipitation and temperature observations while the disaggregation procedure increases the temporal resolution so that a design flood of wanted return period T on hourly timescale can be estimated. To estimate design floods beyond 50 years is problematic since environmental circumstances may change during such time intervals.

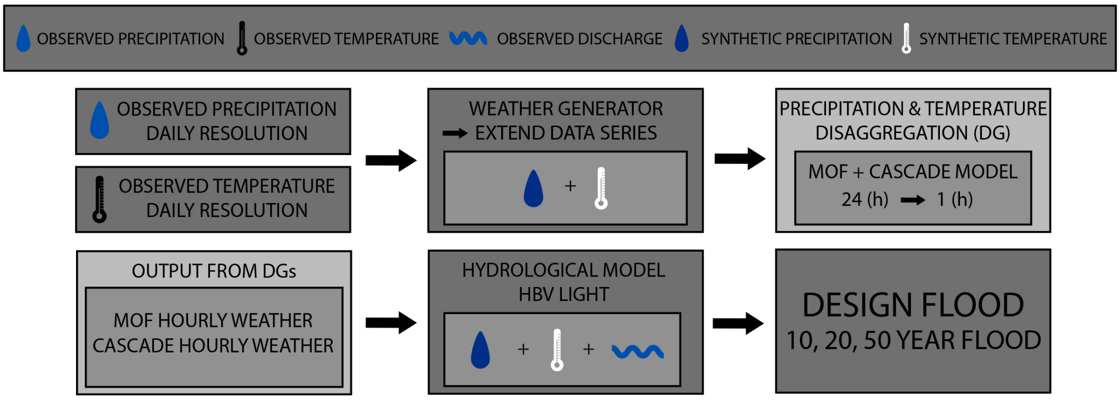


Figure 3. A conceptual sketch of the modeling framework.

3.3 FREQUENCY ANALYSIS METHODOLOGY

In this study, the annual maximum hourly flows were assessed to estimate design floods for the study catchments. The frequency analysis was done in Matlab by approximating parameters from the probability distributions (EV1, LP3, P3, GEV, LN) to minimize the error to the observed series of annual discharge peaks with the method of maximum likelihood (Okoli 2017, personal communication; Myung, 2003). The fitted distributions could then be used to estimate a 20- and 50-year flood.

3.4 HBV LIGHT METHODOLOGY

The HBV-light version requires a warm up period for the initial state variables of the model to take on appropriate values for the simulation based on meteorological conditions and parameter values. Time series of one year have turned out to be sufficient for the warm up process which was also used in this study (Seibert, 2012).

Potential evaporation (PET) had to be estimated to run the HBV model. It was calculated using the Thornthwaite equation (17) found in (Xu & Singh, 2001). T_a is the longterm mean monthly temperature in degrees celcius, $C = 16$ is a constant, I is the annual heat index, N the number of days in a month and d is the average monthly daylight in hours (Xu & Singh, 2001). Thornthwaite is a temperature based method to estimate PET. It has been shown that simple temperature based methods for approximating PET are as efficient as more intricate methods like the Penman model (Oudin, et al., 2005).

$$ET = C(\frac{10T_a}{I})^a(\frac{d}{12})(\frac{N}{30}) \tag{17}$$

The annual heat index I is calculated with equation (18). If the mean monthly temperature is zero or less than zero, T_a=0 (Xu & Singh, 2001).

$$I = \sum_{j=1}^{12}(\frac{T_a}{5})^{1.51} \tag{18}$$

The exponent a is calculated using equation (19) (Xu & Singh, 2001)

$$a = 67.5 \cdot 10^{-8} \cdot I^3 - 77.110^{-6} \cdot I^2 + 0.0179 \cdot I + 0.492 \qquad (19)$$

The hours of daylight, d, in equation (17) was estimated based on a table found in Yoo & Boyd (1994). The calculated values from the Thornthwaite equation (17) were adjusted to the long term mean potential evaporation in the study areas by looking at historic observations.

To account for differences in altitude in the different parts of the catchment, elevation zones were computed in a GIS software using natural breaks for classification as seen in a similar study by Breinl (2016). The use of elevation zones have also been seen to improve snow retention simulations in the HBV model (Breinl, 2016). However, to avoid over-parametrization the parameters CWH and CFR were fixed to 0.1 and 0.05 (Seibert, 1997, 2000; Breinl, 2016).

The elevations zones were extracted using the "extract by mask" tool in ArcGIS.

Two different objective functions were used in HBV for this study. The default objective function in the HBV model's interface is the Nash & Sutcliffe (1970) R_{eff}. It is the value of this objective function that was used to assess model efficiency. The other objective function used in HBV for this study was the LindstroemMeasure. The LindstroemMeasure was only used during automatic calibration to constrain parameter ranges (Seibert, 1999).

Some analyses were made to assess parameter uncertainty. This was done by plotting different parameter values against the values of R_{eff} (1). For a well-defined parameter, a distinct peak could be seen in the plot and abnormalities from the peak value reduced the fit in relation to the objective function. For parameters that are less well-defined, values on broad ranges may result in equally good fits resulting in a plateau in the plot (Seibert, 1997).

To assess parameter sensitivity, varying one or two parameters at a time was done and changes in the objective function R_{eff} were evaluated (Seibert, 2012). However, even if equally good results may be produced from a broad range of values of one parameter, this does not mean that the model is not sensitive to changes in this parameter. As stated above, changes in a sensitive parameter may be compensated by other parameters. Therefore, it is of importance to distinguish between uncertain parameters and insensitive parameters which may be set to some constant value (something seen in many HBV applications for the parameters CWH and CFR for example as described above) and sensitive parameters that have a big influence on the objective function when they are changed alone. The aforementioned is also important to avoid overparametrization (Seibert, 1997).

In this study, the HBV model was calibrated with Monte-Carlo and GAP simulations. The initial values of the parameters were set according to physically reasonable ranges used in other HBV studies (Bergström, 1990; Seibert, 1997). The parameters which depend on time step where adjusted accordingly. Thereafter, 5000 Monte-Carlo simulations were used to find which ranges of parameter values that resulted in sufficient model fits. The ranges of some parameters were extended somewhat for the simulations where the parameter values were close the maximum or minimum during the first phase of the calibrations (Seibert, 1999). To account for equifinality (Beven & Freer, 2001) 100 different but suitable parameter sets (sets with the smallest error according to R_{eff}) from 200 000 Monte-Carlo runs were determined to be used for the runoff simulations (Bergström, 1990; Breinl, 2016; Seibert, 1997, 2000). A validation was carried out on an independent time period to test the uncertainty of the calibrated parameter sets (Seibert, 2000). The reference values for the parameters in the HBV light model are listed in appendix 1 (Bergström, 1990).

3.5 WEATHER GENERATION METHODOLOGY

To account for many different scenarios, 25 realizations of 100 years of daily rainfall and temperature data were generated based on the observed time series for each catchment (Breinl 2017, personal communication). These time series were then used for analysis.

3.6 METHOD OF FRAGMENTS METHODOLOGY

There were big deviations between the daily observed precipitation time series and the observed hourly precipitation series for all three catchments used in the study. The Tyrolian meteorological office reported in a request, that the hourly values are measured from 06:00 to 07:00 on the consecutive day and that deviations between daily and aggregated hourly values are still possible in the region due to slightly different locations of the gauges over time or changes in the hourly measuring techniques over time. Aggregating hourly values according to this interval, confirmed some of these deviations between the aggregated hourly-to-daily values and the original daily values. However, this was confirmed at a late stage so that aggregation between 00:00-23:00 already had been conducted. The deviations seemed to be smaller for extremes. This may result in slightly offset disaggregated hourly values since they depend upon the aggregated hourly data.

The performance of the implemented method of fragments model had to be validated. As daily records were not available when this research was carried out and hourly time series in the study areas are sufficiently long with 14, 34 and 37 years for Ruetz, Kelchsauer and Gurglbach, hourly time series were aggregated to daily values. The daily series were then disaggregated again using the same hourly data series used to aggregate the daily values. The disaggregation was done as described in the theory section of the method of fragments algorithm. Since the aggregated daily series and hourly series came from the same station, the algorithm was computed so that the model could not choose precipitation ratios from days of the same year as the day to be disaggregated.

A parameter uncertainty analysis was conducted for each of the study catchments with the method of fragments model. The only two adjustable parameters in the method of fragments model are the length of the time window l and the parameter controlling the number of nearest neighbors k. The sensitivity analysis was made by varying one parameter while the other was kept constant. The length of the time window was tested for $l = 31$ days and $l = 37$ days. For each setting of the time window, l, the parameter k was given the values $k = 8 - 10 - 12$. For each different model set, 15 runs of the method of fragments model were made. The best model structure was determined by evaluating the model's performance to reproduce monthly maximas compared to the observed monthly maximum values. A similar validation criteria was used by Westra (2012) and Pui (2012) with the difference that the mentioned studies used annual maximas instead of monthly maximum values. I chose to look at the reproduction of monthly maximas to have more data points to evaluate the model performance from since the data series used in this study were smaller than all samples used in Westra (2012). Other validation statistics to assess the model's reproduction of hourly precipitation, were the total mean, the total variance and the proportion of days with zero rainfall (Westra et al., 2012; Pui et al., 2012). The mean values are expected to be well represented by the different disaggregation models (Pui et al., 2012).

Confidence boundaries were represented by the 5^{th} and 95^{th}percentile seen in the 15 realizations. This was also done by Westra (2012) with the difference that he used 100 simulations for each setting. On the other hand, Pui (2012) used 10 realizations to asses the model performance. I chose to do 15 initial runs with each model setting. If the difference between the 10^{th} and the 15^{th} run for a specific parameter setup was more than five percent, another five runs were made until the difference between monthly maximas from the X^{th} and the X^{th+5} runs was smaller than five percent.

In addition to plotting empirical intensity-frequency curves (probability of exceedance plots) to evaluate the reproduction of the monthly maximum values (Westra et al., 2012; Pui et al., 2012), an objective function was used to evaluate model performance. For this purpose, the R_{eff} objective function was used.

From the weather generator, 25 realizations of 100 years of rainfall data for each of the three study catchments were applied for disaggregation with the method of fragments model. The realizations are extrapolations out of the existing precipitation data series generated with the weather generator described earlier. Each realization was modeled two times for the best model structure determined from the validation procedure described above. Hence, 50 realizations were produced with 100 years of hourly data for each catchment.

3.7 MICROCANONICAL MODEL METHODOLOGY

As for the method of fragments model, 25 realizations of 100 years of weather generated data were used as input. The disaggregated precipitation data series from the microcanonical model was then used for analysis.

3.8 TEMPERATURE DISAGGREGATION METHODOLOGY

Temperature classes were defined based on the rainfall of neighboring days, similar to the procedure seen for the method of fragments algorithm. Temperature classes conditioned to the precipitation accounted for the underlying relations between meteorological processes. In addition to the precipitation classes presented under section 2.5.1, three more classes were added to give a criterion of uniformity for the days without precipitation which were not classified in the method of fragments algorithm. This was motivated since approximately 80 % of the days from the time series of the study sites are days without precipitation. Hence, without these extra classes, many sequences of days would not have been included in the original set of classes and the only criteria to distinguish the day to disaggregate would be from the smallest absolute difference in temperature from days within the time window. The three extra classes are listed below.

$Class$ (1) starting boxes $[dry, wet, wet]$, $R_j \geq 0 \mid (R_{j-1} = 0, R_{j+1} \geq 0)$

$Class$ (2) enclosed boxes $[wet, wet, wet]$, $R_j \geq 0 \mid (R_{j-1} \geq 0, R_{j+1} \geq 0)$

$Class$ (3) ending boxes $[wet, wet, dry]$, $R_j \geq 0 \mid (R_{j-1} \geq 0, R_{j+1} = 0)$

$Class$ (4) isolated boxes $[dry, wet, dry]$, $R_j \geq 0 \mid (R_{j-1} = 0, R_{j+1} = 0)$

Class (5) dry boxes $[dry, dry, dry]$, $R_j = 0 \mid (R_{j-1} = 0, R_{j+1} = 0)$

Class (6) end of dry boxes $[dry, dry, wet]$, $R_j = 0 \mid (R_{j-1} = 0, R_{j+1} \geq 0)$

Class (7) start of dry boxes $[wet, dry, dry]$, $R_j = 0 \mid (R_{j-1} \geq 0, R_{j+1} = 0)$

As with the method of fragments algorithm, j represents any day centered around day t to be disaggregated.

The time window was constructed as described for the method of fragments algorithm with 37 days, $l = +/-19$ centered around day t to be disaggregated. The parameter that regulates the number of possible days to choose from, k, was set to five. If more than five days fulfill the disaggregation criteria, the code considers only the five days with the least absolute difference in temperature. The absolute difference was pursued to account for the negative sign for temperature below zero.

Another difference from the method of fragments algorithm was how the daily values were summarized. To account for temperatures below zero when aggregating hourly values into daily, the absolute sum was used. This was done to give representative ratios of sub daily to daily temperatures and to have the right sign (minus/plus) for the day in the temperature ratio vector $T_{i,m}$ seen in equation (20) below.

$$TempRatio_{i,m} = \frac{T_{hourly,i,m}}{T_{absdaily_i}} \tag{20}$$

The daily temperature time series were disaggregated in a similar fashion seen for the precipitation disaggregation with the method of fragments model with the above differences. Residuals were calculated between the observed hourly temperature series and the simulated hourly temperature series, $\mid T_{hourly} - T_{simhourly} \mid$ compared to residuals between daily values and their daily averages, $\mid T_{hourly} - T_{dailyaverages} \mid$

4 RESULTS

4.1 FREQUENCY ANALYSIS

A conventional frequency analysis was made where the asymptotic distributions called GEV, Gumbel (EV1), Lognormal, Pearson type III and the Log Pearson type III were fitted to a series of annual maximum flows $[m^3s^{-1}]$ for each of the three study catchments. First of all it can be seen in Table 3 and 2 that there is a big difference in design flows for each of the study sites. Furthermore, the estimated design floods with return period T = 50 & T = 20 years seen in the tables vary greatly between the different distributions fitted to the sample. Note that Q_{obs} is not a representation of the 20 or 50-year flood, it is the highest flow seen in the data samples for each study site. For the Ruetz catchment, 20 annual maximum runoff peaks were constructed (from 20 full years of discharge data), while 37 and 34 maximum annual discharges were constructed for the Kelchsauer and Gurglbach catchment respectively.

Table 2. A summary of the results obtained with frequency analysis. The distributions were fitted to the sample with maximum likelihood to estimate a design flood $[m^3s^{-1}]$ with 20 year return period.

T = 20 $[y]$	$Q_{obs,max}$	GEV	EV1	P3	LP3	LogNorm	Unit
Ruetz	141.5	150.7	87.3	103.2	112.1	95.2	m^3s^{-1}
Kelchsauer	99.9	85.5	82.2	82.1	84.6	84.1	m^3s^{-1}
Gurglbach	39.06	20.5	18.1	19.9	20.4	19.7	m^3s^{-1}

Table 3. A summary of the results obtained with frequency analysis. The distributions were fitted to the sample with maximum likelihood to estimate a design flood $[m^3s^{-1}]$ with 50 year return period.

T = 50 $[y]$	$Q_{obs,max}$	GEV	EV1	P3	LP3	LogNorm	Unit
Ruetz	141.5	304.7	101.9	127.2	162.8	114.0	m^3s^{-1}
Kelchsauer	99.9	85.5	82.2	82.1	84.6	84.1	m^3s^{-1}
Gurglbach	39.06	26.9	21.4	23.7	26.2	24.1	m^3s^{-1}

The number of years with historical data means that Q_{20} is within the observations while Q_{50} is extrapolated. To give a visual representation of the uncertainty in conventional frequency analysis, graphs of the maximum monthly flows $[m^3s^{-1}]$ were also constructed. In the Ruetz catchment, the catchment with the least amount of data, it can be seen in Figure (4) that the spread between the different fitted distributions is wide. Especially GEV seems to diverge from the observed data points. For the Kelchsauer catchment with most observed data points, it can be seen in Figure (5) that the variation between the different fitted asymptotic distributions is less than the difference seen between the distributions for the other catchment. However, despite the smaller variation in estimated design flows for the fitted distributions, the residuals, $Y_{obs} - Y_{sim}$, are big, especially for flows with a return period T > 10 years.

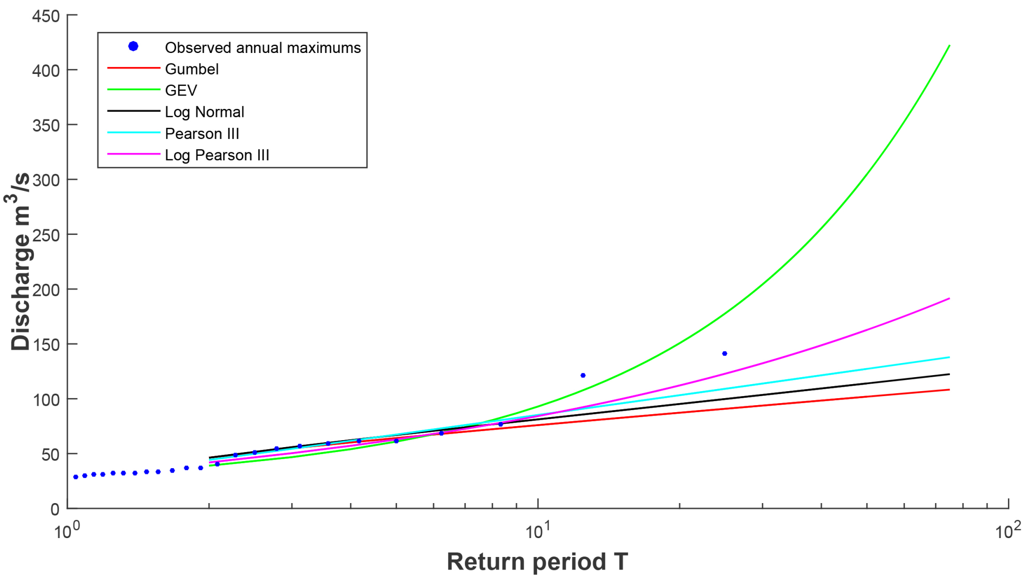


Figure 4. Probability of exceedence plots for Ruetz. Different asymptotic probability distributions have been fitted to observed annual maximum flows $[m^3s^{-1}]$ using the method of maximum likelihood. The blue points represent observed discharge and the colored lines represent different probability distributions.

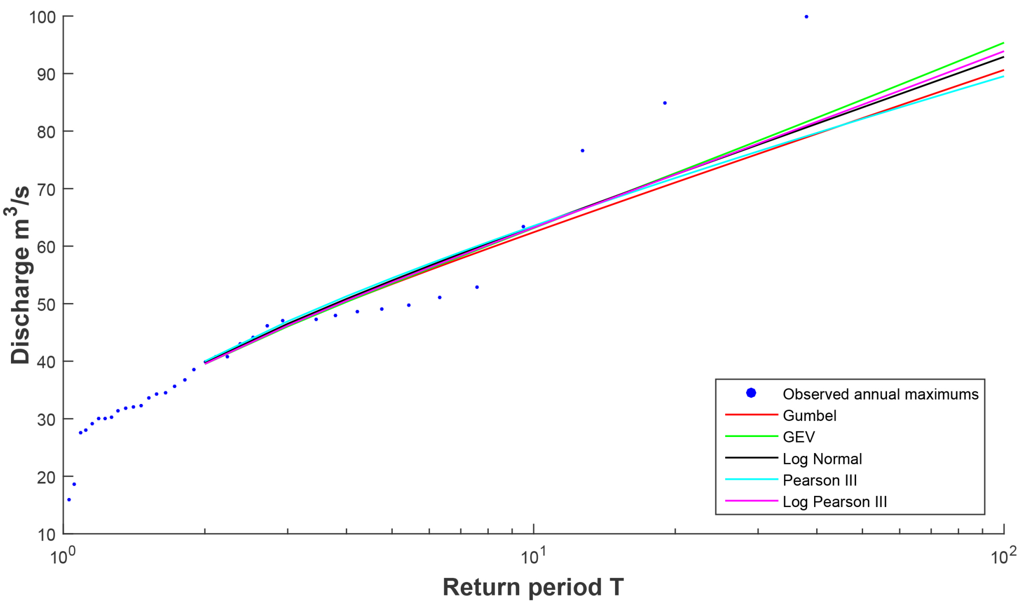


Figure 5. Probability of exceedence plots for Kelchsauer. Different asymptotic probability distributions have been fitted to observed annual maximum flows $[m^3s^{-1}]$ using the method of maximum likelihood. The blue points represent observed discharge and the colored lines represent different probability distributions.

In Figure (6) the probability distributions are constrained but the observed data points diverge from the fitted distributions, especially for flows with return period greater than 10 years.

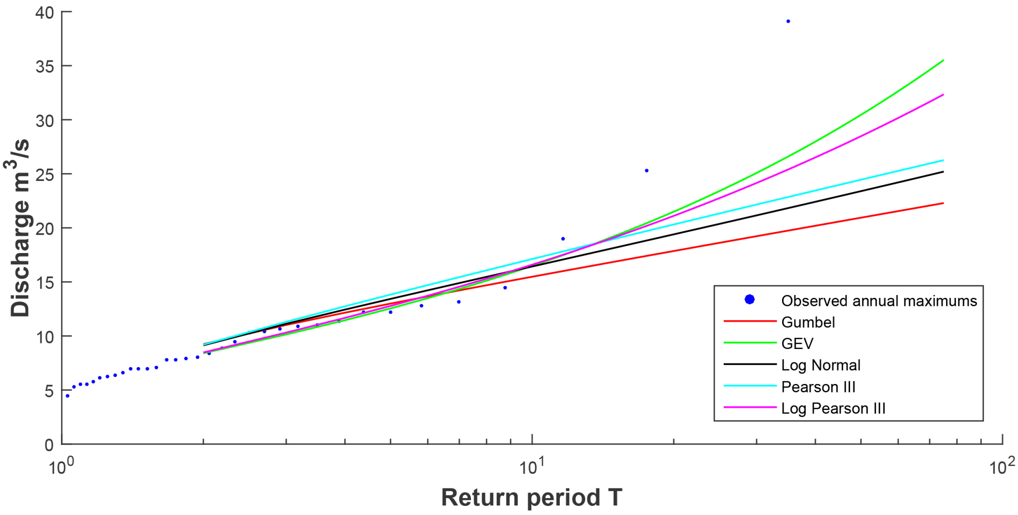


Figure 6. Probability of exceedence plots for Gurglbach. Different asymptotic probability distributions have been fitted to observed annual maximum flows $[m^3 s^{-1}]$ using the method of maximum likelihood. The blue points represent observed discharge and the colored lines represent different probability distributions.

4.2 HBV CALIBRATION AND VALIDATION

The calibration and validation plots in this section are computed from the parameter set with highest objective function value simulated from GAP-optimizations. In Figure (7a)/(7b),(8a)/(8b) and especially (9a)/(9b), it is noteworthy that peaks in general are underestimated and sometimes not simulated at all by the model with these parameter sets.

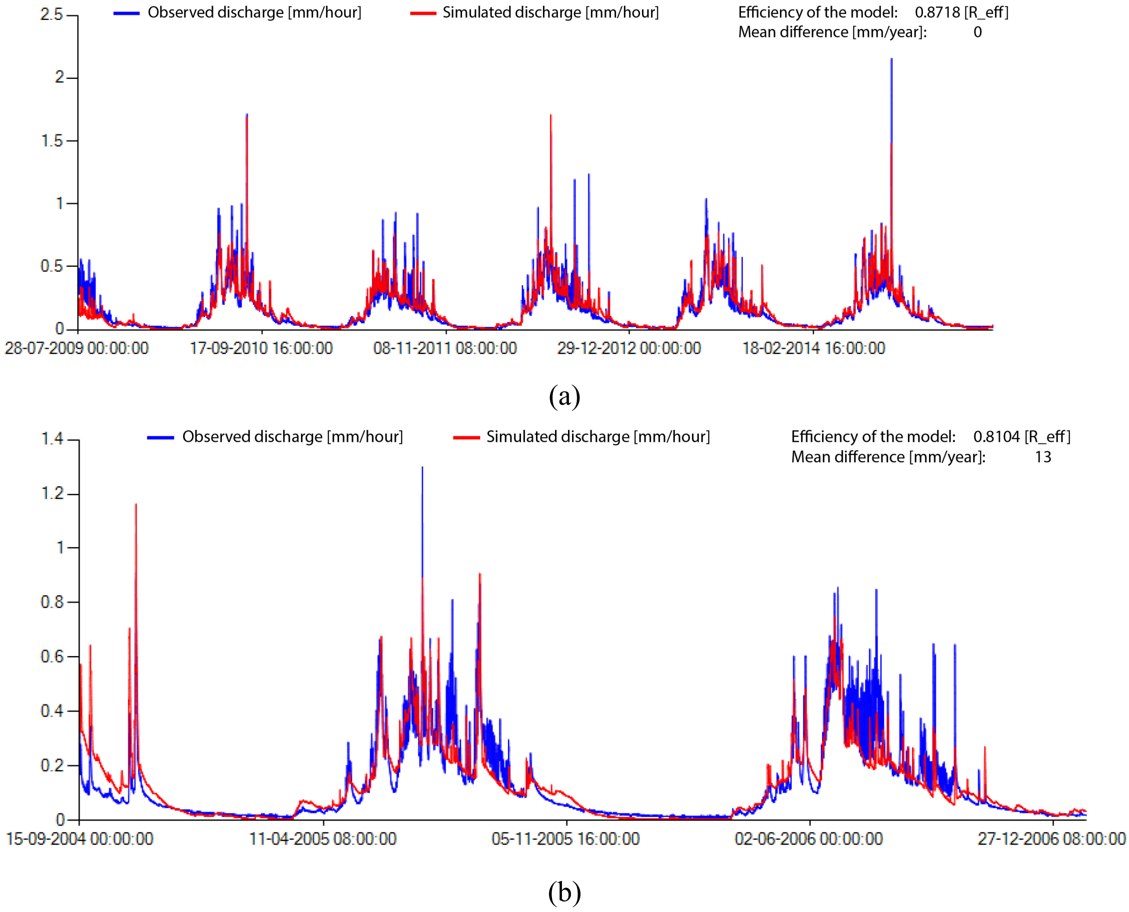
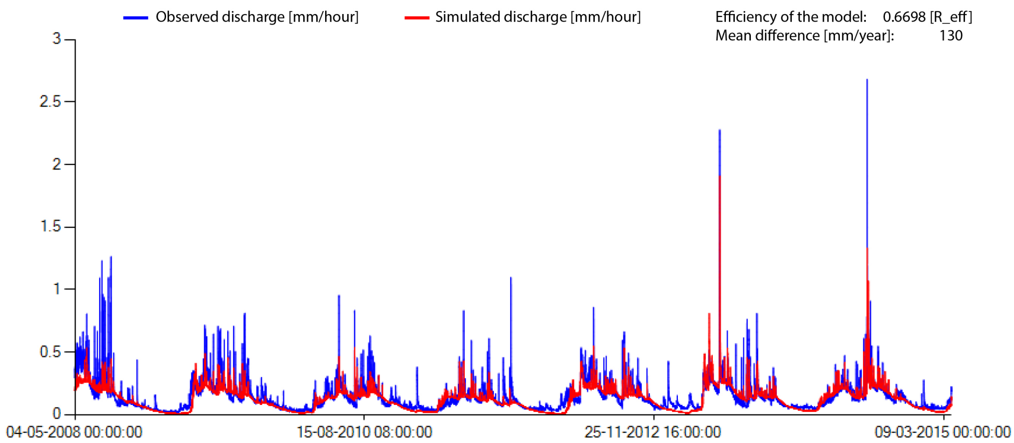
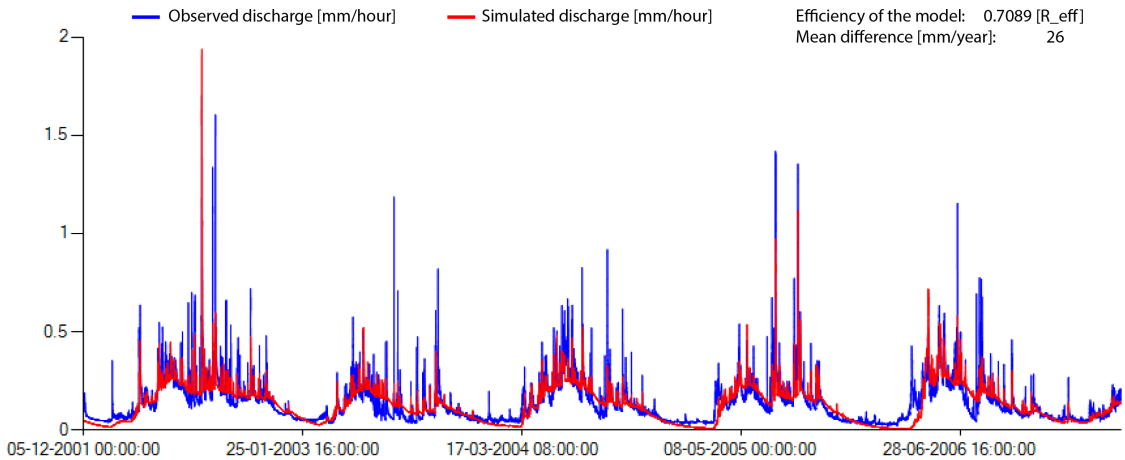


Figure 7. Calibration (a) and validation (b) plots for Ruetz. The simulations were based on historic observations between 2008/07/28 - 2015/03/31 for calibration and 2004/09/15 - 2007/01/24 for validation.

The final parameter ranges used during calibration and validation in the HBV model were estimated by plotting parameter values from 5000 Monte-Carlo simulations against the objective function. If a distinct peak or flat area was found in the plots, the ranges were not extended. Else, if the reference ranges from Table (4) did not catch the peak, ranges were slightly extended. These final ranges can be seen in Appendix A 1 were $PCALT$ was one of parameters with a displaced range in relation to reference values. Some example plots on parameter ranges can be found in Appendix B. These were used to get an understanding of how well defined the parameter was.

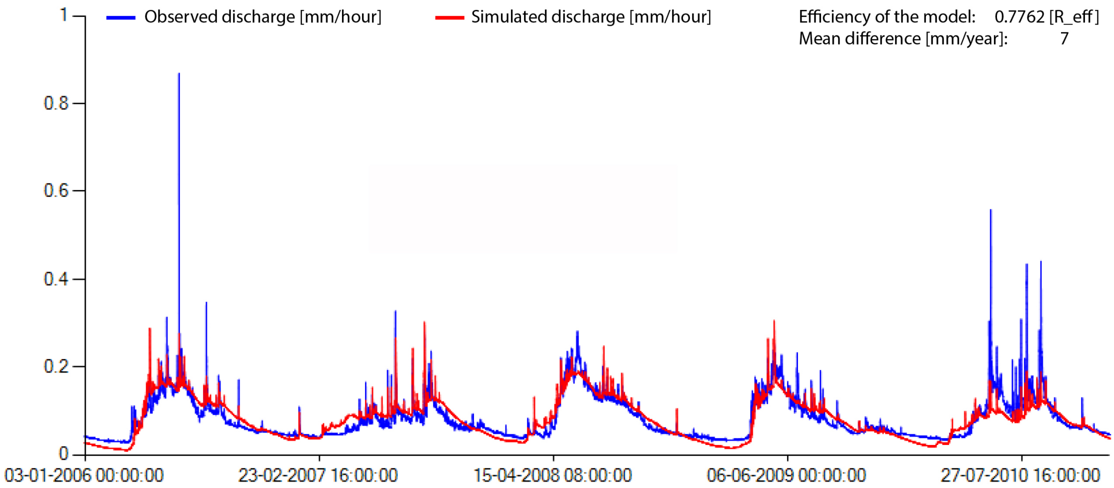


(a)

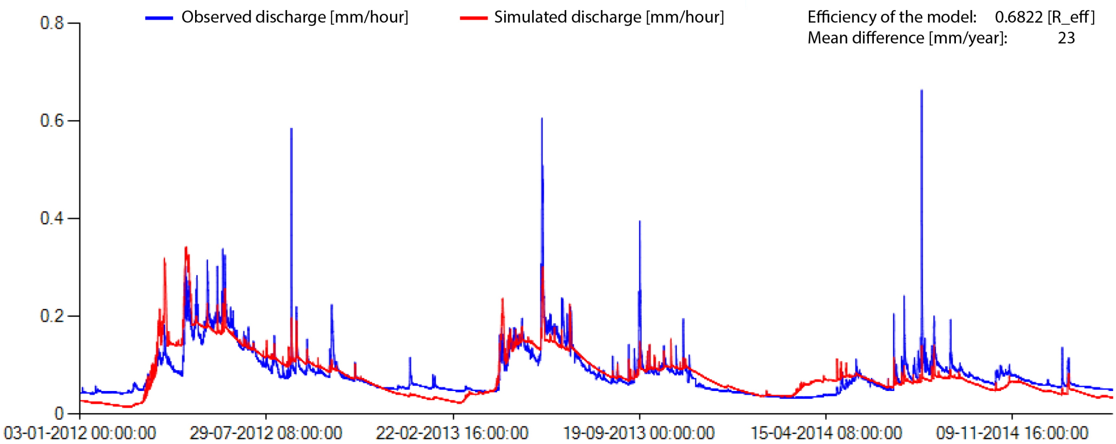


(b)

Figure 8. Calibration (a) and validation (b) plots for Kelchsauer. The simulations were based on historic observations between 2007/05/04 - 2015/03/31 for calibration and 2000/12/05 - 2007/04/29 for validation.



(a)



(b)

Figure 9. Calibration (a) and validation (b) plots for Gurglbach. The simulations were based on historic observations between 2005/03/31 - 2011/01/01 for calibration and 2011/01/02 - 2015/03/01 for validation.

Table 4. The modeled parameters and their ranges in the HBV model.

Parameter	Explanation	Min	Max	Unit
Snow Routine				
TT	Threshold temperature	-0.5	1.5	$^{\circ}C$
CFMAX	Degree-day factor	0.05	0.3	*mm* $^{\circ}C^{-1}\ h^{-1}$
SFCF	Snowfall correction factor	0.5	1.5	-
CFR	Refreezing coefficient	0.05	0.05	-
CWH	Water holding capacity	0.1	0.1	-
Soil& evaporation Routine				
FC	Maximum soil moisture	100	300	*mm*
LP	Soil moisture threshold for reduction of evaporation	0	1	-
Beta	Shape coefficient	0.01	4	-
Groundwater& response routine				
K_0	Recession coefficient	0.0033	0.02916	h^{-1}
K_1	Recession coefficient	0.00083	0.02083	h^{-1}
K_2	Recession coefficient	8.33E-05	0.01	h^{-1}
UZL	Threshold for K0-outflow	8	70	*mm*
PERC	Maximal flow from upper to lower GW-box	0	0.5	*mm* h^{-1}
MAXBAS	Routing, length of weighting function	1	30	h
Other				
CET	Potential evaporation correction factor	0	0.3	$^{\circ}C^{-1}$
PCALT	Change of precipitation with elevation	0	14	$\%\ 100^{-1}\ m^{-1}$
TCALT	Change of temperature with elevation	0.6	0.6	$\%\ 100^{-1}\ m^{-1}$

4.2.1 HBV sensitivity study

The sensitivity study was performed by changing one parameter at a time while the other were kept constant. Each parameter value was multiplied by 2 and 0.5 to see what influence the encroachment would have on the value of the objective function R_{eff}.

Changes in the snow correction factor, $SFCF$, had great influence on the value of the objective function for all the three study sites. This indicates the importance of the snow volume in the three catchments as this parameter compensates for errors in the ”missing” evaporation of snow as well as systematic errors in snowfall measurements (Seibert, 1999). The degree day factor, $CFMAX$, controls the amount of meltwater from the simulated snow pack which was also seen to have a big influence on the value of the objective function for all three catchments.

The change of precipitation with altitude, $PCALT$, was another parameter identified as sensitive for all the three catchments. Ruetz, whose measuring gauge is located on higher altitude than those in Kelchsauer and Gurglbach was not very sensitive to changes in any other parameter (a doubling/halving of the parameter value lead to changes smaller than 5 % for the value of the objective function).

For Gurglbach and Kelchsauer, some additional parameters were identified as important for the value of the objective function. $PERC$, the parameter controlling the maximum proportion of percolating water from the upper to the lower groundwater box, SLZ, as well as the parameter K_2, regulating the amount runoff from the lower groundwater box. Kelchsauer was also sensitive to changes in the parameter K_1 controlling runoff from the upper groundwater box.

Gurglbach was further sensitive to changes in LP, the parameter controlling evaporation.

4.3 PRECIPITATION DISAGGREGATION

4.3.1 Method of fragments

The adjustable values in the programmed method of fragments model are the length of the time window l and the number of nearest neighbors to choose precipitation rate from, k. To assess the parameter uncertainty of these values, different settings were tested by varying one of these parameters while keeping the other constant. The most efficient model structure for Ruetz was $l = 37$ d and $k = 10$ (Table 5). This structure was most effective at reproducing observed statistics according to three out of four criteria. The best model structure for Kelchsauer was also the model with the parameter values $l = 37$ d and $k = 10$ and for Gurglbach the best model structure was $l = 31$ d and $k = 8$.

Table 5. A summary of the results from a model evaluation procedure. By varying the length of the window l and the number of possible values to chsose the precipitation rate from k, the most efficient model structure could be determined for each study catchment.

Parameter	μ $[mm\ h^{-1}]$	$\Delta\ \mu$ %	σ^2 $[mm^2\ h^{-2}]$	$\Delta\ \sigma^2$ %	Dry ratio	R_{eff}
Obs Ruetz 31 d	0.1328	0	0.2759	0	0.7698	1
l = 31 d k = 8	0.1322	0.44	0.2990	-8.41	0.8003	0.9165
l = 31 d k = 10	0.1322	0.43	0.3037	-10.09	0.8009	0.9020
l = 31 d k = 12	0.1322	0.44	0.3023	-9.58	0.8008	0.9030
Obs Ruetz 37 d	0.1330	0	0.2763	0	0.7966	1
l = 37 d k = 8	0.1325	0.43	0.3016	-9.15	0.7993	0.9246
l = 37 d k = 10	**0.1325**	**0.42**	**0.2961**	**-7.17**	**0.7999**	**0.9390**
l = 37 d k = 12	0.1324	0.44	0.2991	-8.26	0.8000	0.9243
Obs Kelchsauer 31 d	0.1571	0	0.4196	0	0.8044	1
l = 31 d k = 8	0.1571	0	0.4421	-5.38	0.8051	0.9500
l = 31 d k = 10	0.1571	0	0.4401	-4.89	0.8053	0.9717
l = 31 d k = 12	0.1571	0	0.4368	-4.12	0.8057	0.9697
Obs Kelchsauer 37 d	0.1571	0	0.4197	0	0.8043	1
l = 37 d k = 8	0.1571	0	0.4441	-5.83	0.8049	0.9746
l = 37 d k = 10	**0.1571**	**0**	**0.4331**	**-3.20**	**0.8054**	**0.9747**
l = 37 d k = 12	0.1571	0	0.4333	-3.24	0.8053	0.9710
Obs Gurglbach 31 d	0.1101	0	0.2626	0	0.8754	1
l = 31 d k = 8	**0.1100**	**0**	**0.2646**	**-0.77**	**0.8769**	**0.9893**
l = 31 d k = 10	0.1100	0	0.2779	-5.84	0.8770	0.9877
l = 31 d k = 12	0.1100	0	0.2812	-7.11	0.8770	0.9807
Obs Gurglbach 37 d	0.1101	0	0.2626	0	0.8754	1
l = 37 d k = 8	0.1101	0	0.2769	-5.44	0.8764	0.9879
l = 37 d k = 10	0.1101	0	0.2794	-6.40	0.8765	0.9821
l = 37 d k = 12	0.1101	0	0.2796	-6.46	0.8769	0.9872

For each study site, the means of the highest simulated precipitation depths $[mm\ h^{-1}]$ simulated from the best model structure as well their 5^{th} and 95^{th} percentiles were plotted against their corresponding probability of exceedance. The same was done for the observed values. The observed values were within confidence bounds for Kelchsauer and Gurglbach (Figure 11 and 12, while many were found just outside for the Ruetz catchment (Figure 10).

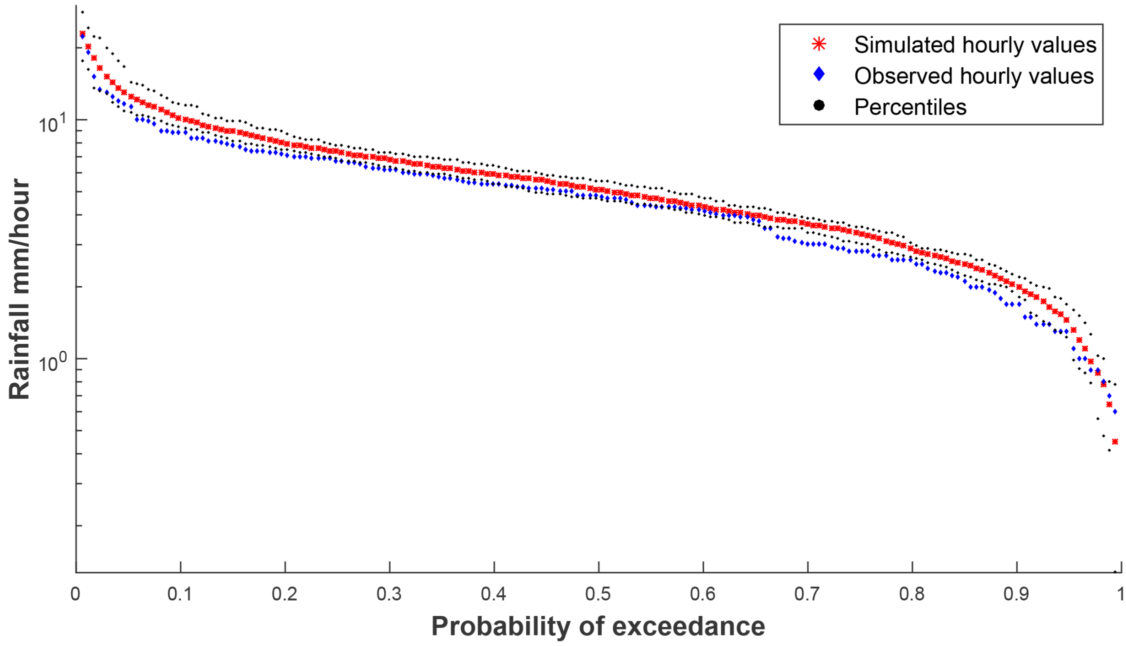


Figure 10. The highest precipitation depths $[mm\ h^{-1}]$ from each month for Ruetz plotted against its probability of exceedance. Simulated values are represented by a red '*', and their 5^{th} and 95^{th} are represented by a black dot 'o' while observed hourly values are visualized as blue diamonds.

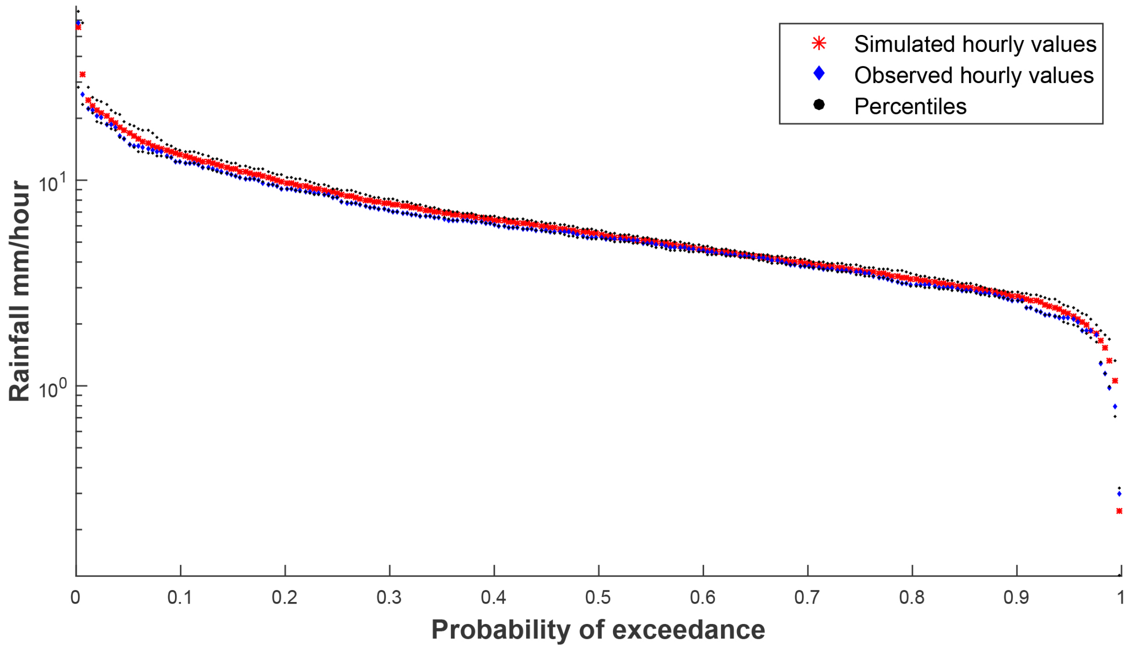


Figure 11. The highest precipitation depths $[mm\ h^{-1}]$ from each month for Kelchsauer plotted against its probability of exceedance. Simulated values are represented by a red '*', and their 5^{th} and 95^{th} are represented by a black dot 'o' while observed hourly values are visualized as blue diamonds.

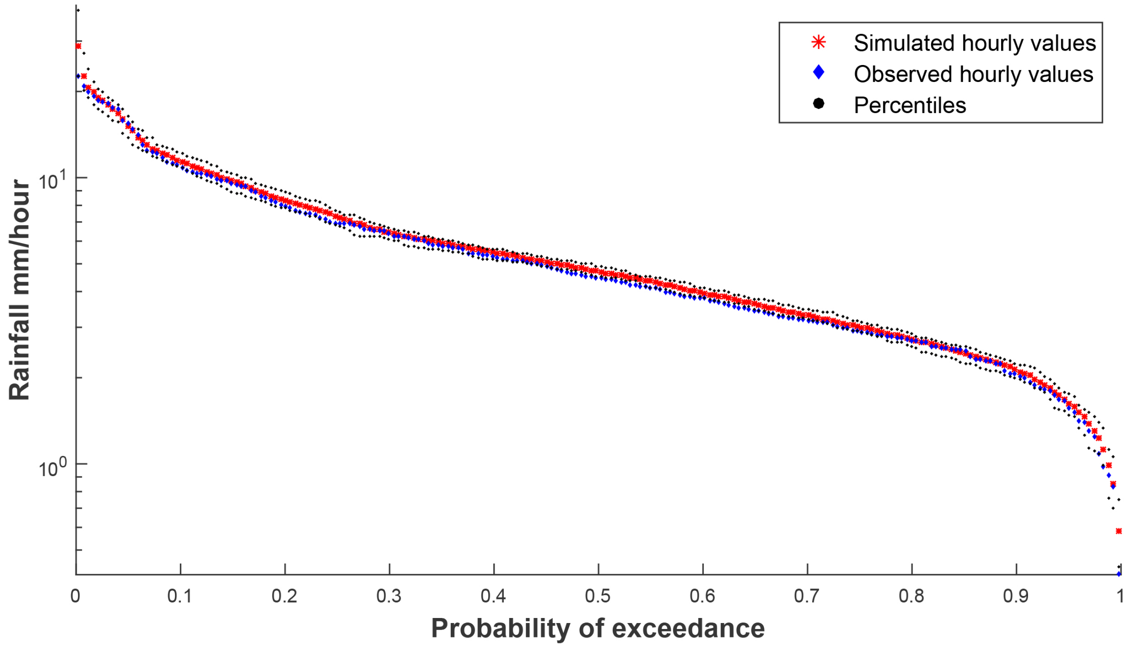


Figure 12. The highest precipitation depths [$mm\ h^{-1}$] from each month for Gurglbach plotted against its probability of exceedance. Simulated values are represented by a red '*', and their 5^{th} and 95^{th} are represented by a black dot 'o' while observed hourly values are visualized as blue diamonds.

4.3.2 Disaggregation models with weather generated data as input

The most efficient model structures of the method of fragments models were used to estimate design precipitation depths with return period T=10, 20 and 50 years with the weather generated data as input. These values were estimated using 100 years of extrapolated synthetic data to evaluate the reproduction of unobserved extremes. The same procedure was done for the microcanonical model with uniform splitting.

The two different disaggregation models simulate different design precipitation depths (Figure 13). The method of fragments model is most similar to the observed extremes. Another big difference between the two models is the variance which is bigger for the microcanonical model (Table 6).

Table 6. A comparison of the reproduction of conventional statistics for Ruetz simulated by the two dissaggregation models.

Statistics	Observed	MOF	MICRO	Unit
Mean μ	0.133	0.135	0.135	$mm\ h^{-1}$
Variance σ^2	0.276	0.311	0.400	$mm^2\ h^{-2}$
Dry ratio	0.797	0.830	0.836	-

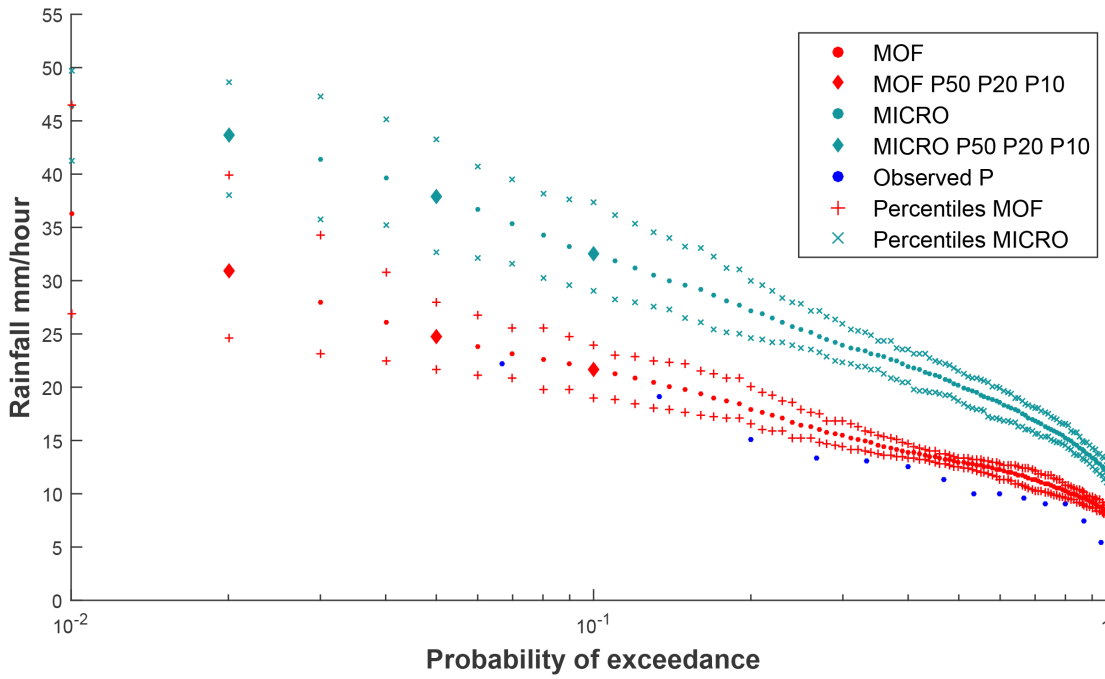


Figure 13. A probability of exceedance for Ruetz. Values in the color cyan represent simulated data from the microcanonical model while the red data points are simulations from the method of fragments model.

The simulated design depths from the two disaggregation models are more alike for Kelchsauer, the catchment with most available precipitation data. The two models simulated P50 is approximately the same. However, the observed precipitation depth for P20 is higher than the P50s estimated by the two models (Figure 14). Again, the biggest difference between conventional statistical parameters is the difference between the two models' simulated variance which is bigger for the microcanonical model (Table 7).

Table 7. A comparison of the reproduction of conventional statistics for Kelchsauer simulated by the two dissaggregation models

Statistics	**Observed**	**MOF**	**MICRO**	**Unit**
Mean μ	0.1571	0.1566	0.1566	$mm\ h^{-1}$
Variance σ^2	0.4197	0.4444	0.4706	$mm^2\ h^{-2}$
Dry ratio	0.8043	0.8080	0.8223	-

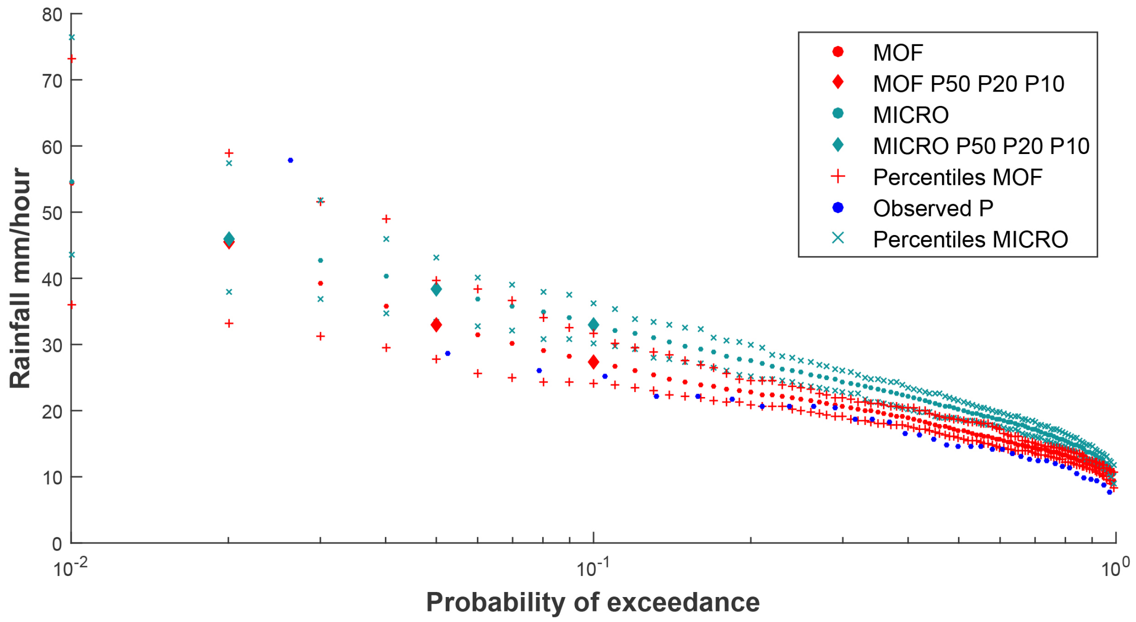


Figure 14. A probability of exceedance for Kelchsauer. Values in the color cyan represent simulated data from the microcanonical model while the red data points are simulations from the method of fragments model.

For Gurglbach, the simulated precipitation depths from the two disaggregation models are most alike compared to the results of the other catchments (Figure 15).

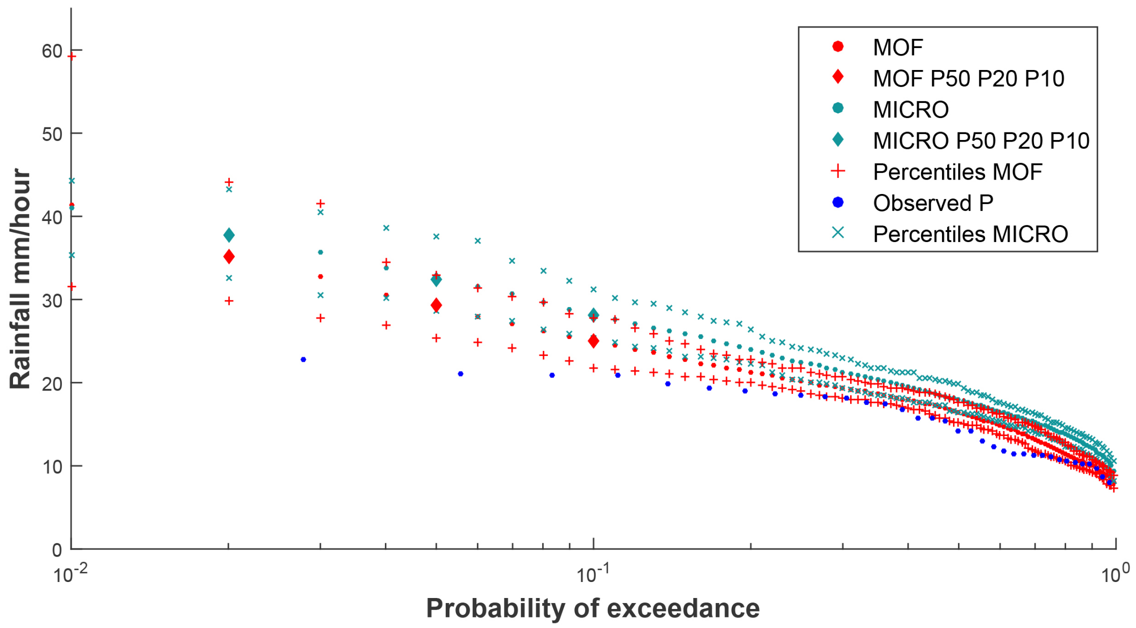


Figure 15. A probability of exceedance for Gurglbach. Values in the color cyan represent simulated data from the microcanonical model while the red data points are simulations from the method of fragments model.

Also here, the biggest difference in the conventional statistics is the difference between the two models simulated variance which is bigger for the microcanonical model (Table 8).

Table 8. A comparison of the reproduction of conventional statistics for Gurglbach simulated by the two dissaggregation models

Statistics	**Observed**	**MOF**	**MICRO**	**Unit**
Mean μ	0.1101	0.1100	0.1100	$mm\ h^{-1}$
Variance σ^2	0.2626	0.2835	0.3283	$mm^2\ h^{-2}$
Dry ratio	0.8754	0.8769	0.8840	-

4.4 TEMPERATURE DISAGGREGATION

4.4.1 Temperature Disaggregation With Observed Data

The sum of the residuals from the mean of 15 simulations with the temperature disaggregation algorithm shows that the disaggregation of temperature decrease the residuals, $\mid T - T' \mid$, by 39% compared to residuals between observed and daily averaged values, $\mid T - T_\mu \mid$ for Gurglbach. The average temperatures were calculated as the sum of 24 observed aggregated hourly values which were then divided by 24 to get the average. The sum of the residuals for the daily averages shows that this leads to an average error of 2.98 °C for every individual hourly value while the error for the disaggregation model was 1.82 °C for every single hourly value for the Gurglbach catchment. Furthermore, temperature disaggregation for Ruetz was not as successful. In fact, the two percent decrease in the sum of the residuals for the Ruetz catchment, were overshadowed by big differences for individual days (something compensated for by running the model 10 times for the Gurglbach and Kelchsauer catchments) (Table 9).

Table 9. A summary of the results obtained when disaggregating observed temperature (°C on hourly resolution) series.

Catchment	μ_{obs} $[°C]$	μ_{sim} $[°C]$	μ_{ave} $[°C]$	σ^2_{obs}	σ^2_{sim}	σ^2_{ave}	1-$\frac{\sum\mid T-T'\mid}{\sum\mid T-T_\mu\mid}$
Ruetz	0.72	0.73	0.72	57.77	57.83	53.42	+0.02
Kelchsauer	6.64	6.64	6.64	76.01	76.10	62.71	+0.42
Gurglbach	6.91	6.92	6.91	80.35	78.42	65.63	+0.39

In Figure 16 it can be seen that the temperature disaggregation model outperforms hourly averages of temperature leading to smaller residuals (bottom graph) for the Kelchsauer and the Gurglbach study site. The temperature disaggregation did not decrease the sum of the residuals for Ruetz at the same magnitude seen for the other two catchments. The two graphs are a visualization from the results for the Gurglbach catchment. Similar results were observed when disaggregating temperature for Kelchsauer.

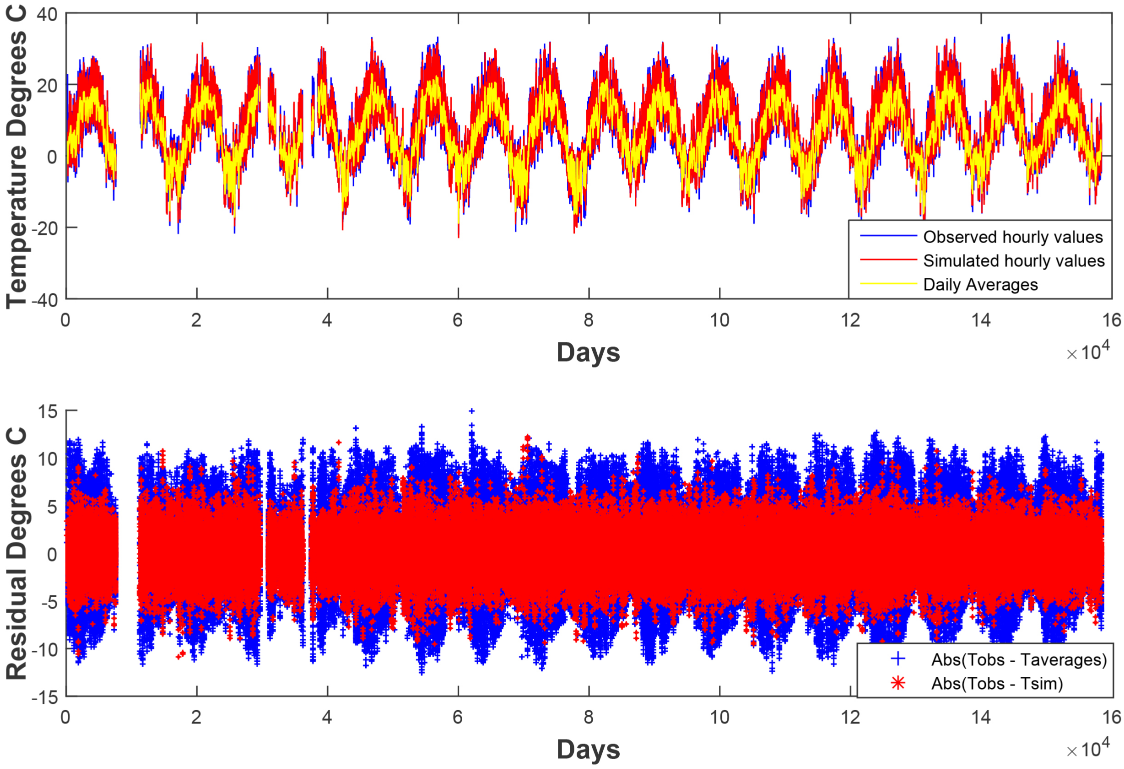


Figure 16. Temperature disaggregation with observed data series for Gurglbach. The blue data points represent observed hourly values, while yellow are daily averages and red data points represent simulated values.

4.4.2 Temperature disaggregation influence in the HBV model

To evaluate the need for temperature disaggregation, daily averages and disaggregated temperature series for Gurglbach and Kelchsauer were compared with the results seen with observed hourly values. In Ruetz, the only performed comparison was between observed hourly temperature (the original input for the validation period) and the hourly averaged temperature series. The impact on the value of the objective function was rather small despite the big decrease in the sum of the residuals for Kelchsauer and Gurglbach (Table 10).

However, the impact of temperature disaggregation and the use of mean values on the discharge peaks in the HBV model, was bigger than the influence of the objective function. As can be seen in Table (10), the impact on the highest simulated flow Q_{1Td}, as well as on the 5^{th} highest simulated flow Q_{5Td} was considerable with differences up to almost 16 % for a temperature series with daily averages compared to a difference of 5 % when using disaggregated data for the same period. The simulations were applied on the validation period for each catchment in the HBV model.

Table 10. The results from historic observations were compared with the results from using disaggregated temperature data, T_d, and hourly temperature mean values, T_μ, as input in the HBV model. The values in the table are unitless since the represent percentual differences from the value of the objective function and peak discharges.

Statistics	Ruetz	Kelchsauer	Gurglbach
$R_{eff}\ T_{obs}$	0.8104	0.7089	0.6908
$T_d\ \Delta\%$	-	0.2 %	0.35%
$T_\mu\ \Delta\%$	0.83%	9.11%	0.90%
$Q_{1Td}\ \Delta\%$	-	-1.19%	+5.1%
$Q_{1T\mu}\ \Delta\%$	-10.48%	+7.12%	+15.57%
$Q_{5Td}\ \Delta\%$	-	-1.23%	+4.49%
$Q_{5T\mu}\ \Delta\%$	-12.05%	+7.26%	+13.77%

4.4.3 Temperature disaggregation with weather generated data

Temperature disaggregation was further attempted for Gurglbach with the weather generated temperature series. These temperature series were conditioned to rainfall in the weather generation process and comprised equal amounts of data, i.e., 25 realizations of 100 years of daily temperature.

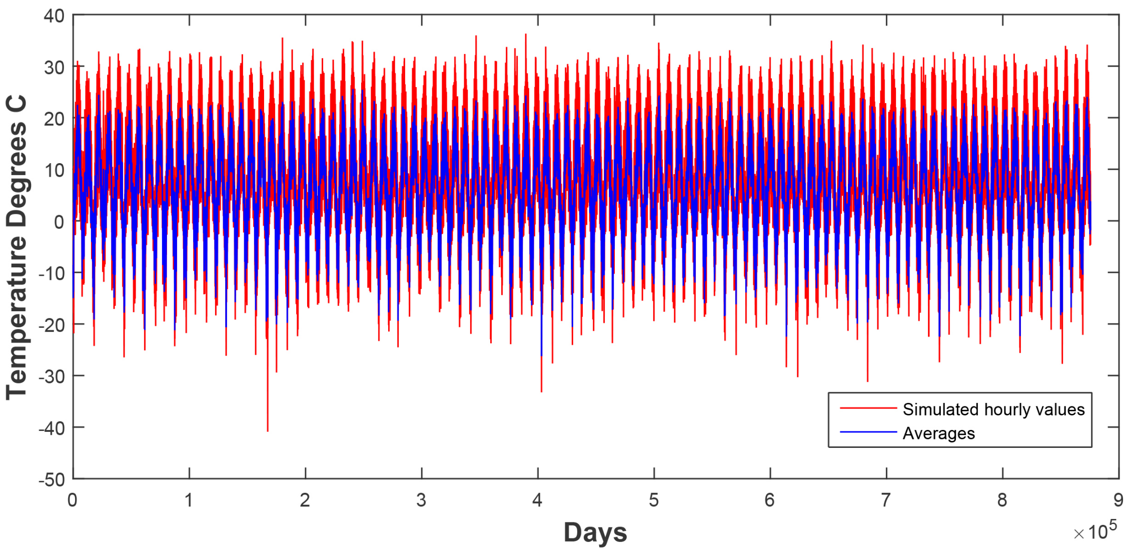


Figure 17. Temperature disaggregation with observed data series for Gurglbach. The blue data points represent daily averages and red data points represent simulated values.

5 DISCUSSION

5.0.1 Frequency analysis

As been evidenced in earlier studies, frequency analysis is practical and simple to do, but is limited when it comes to estimate and model extremes beyond the conditions of the sample series. Moreover, it does not provide any information on the processes that gave rise to the observed maximum flows (Merz & Blöschl, 2003; Moran, 1958; Chow et al., 1988). The same results were obtained in this study. In Figure (4) it can be seen that the estimated 50 year design flood for Ruetz varies between approximately $100 - 300$ $[m^3s^{-1}]$ for the different applied distributions. Furthermore, the results are not improved with more data (see the results for Kelchsauer and Gurglbach). In fact, even if the confidence bounds between different distributions are constrained for Kelchsauer and Gurglbach in comparison with Ruetz, the observed data points for bigger return periods are far from the simulations. These results strengthen the need for more physical methods that are closer to understanding the physical causes of floods so that more reliable projections can be made.

5.0.2 HBV

As thoroughly described by Seibert (1997) there are many uncertainties related to calibration of a hydrological conceptual model. Due to errors in model structure, observed variables and interactions between different parameters, it is hardly possible to find a unique, "true", parameter set. This was confirmed in this thesis where most parameters were found to be uncertain i.e. not well-defined (something seen in other HBV studies as well)(Seibert, 1997). The aim with a fuzzy measure (the LindstroemMeasure) when using the automatic calibration tools available in HBV, was to constrain, hence decrease the big uncertainty of the simulated parameters (Seibert, 1997). But even when using a fuzzy measure, many of the calibrated parameters were seen to be highly uncertain (see example of an uncertain parameter in Figure 2 Appendix B).

Many different parameter sets gave equally good fits according to R_{eff} (parameters from 200 000 Monte Carlo simulations). However, it is not certain that the different parameter sets would have resulted in the same runoff projections. Therefore, as was planned, it is very much motivated to use dissimilar parameter sets to account for the equifinality to justify design flood modeling with the HBV model (Beven & Freer, 2001; Seibert, 1997).

One parameter that in general was different from the reference values (Bergström, 1990; Seibert, 1997), was the parameter $PCALT$, whose ranges had to be extended considerably. This is most certainly related to the altitude of the gauges and the mountainous environment from which they collect precipitation and can therefore be justified. However, this parameter is not always displayed in studies referred to, something complicating a comparison (Breinl, 2016). Furthermore, the sensitivity analysis shows that changes in parameters related to the snow routine had a big impact on the objective function. This is also motivated since snow is a great determining factor of hydrology in the Alps (Vanham et al., 2008).

In the calibration and validation plots simulated in the HBV light model (Figure 7, 8 and especially Figure 9), the discharge peaks are constantly underestimated for all three study catchments. This highlights the complex process of calibration, where the fit according to an objective function can be sufficient even though, as in this case, peaks are not well defined.

To get an understanding on why these peaks were underestimated, the raw data was analyzed for the dates of some discharge peaks without finding any general pattern. It is much like Merz & Blöschl (2003) put it, understanding the physical processes behind a flood of a given probability may be one of the most intriguing questions of catchment hydrology and not an easy one to answer. The underestimation error would have influenced the projections. However, even though the errors are considerable, the error was to some extent assessed, since validation was performed on an independent time interval (Seibert, 1999a). It is thus suggested to try other objective functions more focused on the discharge peaks for future studies to increase confidence in the modeling.

The choice of the Thornthwaite method for estimating PET may also introduce uncertainties. However, as concluded by Oudin et al (2005), it is not assured that more complex methods improve projections. The use of elevation zones in the HBV model is another area where errors can be introduced. Here, it was considered motivated since it improved the fit to the objective function considerably.

5.0.3 Precipitation disaggregation

The method of fragments model was assessed on its ability to reproduce observed extremes and conventional statistics. This first test was done with observed hourly time series as well as daily time series constructed from the hourly. For the study areas with more data available, i.e., Kelchsauer and Gurglbach, model results were seen to have most resemblance with their observed counterparts. The majority of data points in the probability of exceedance plots (Figure 10), i.e., the extremes for Kelchsauer and Gurglbach, were within confidence bounds. On the other hand, many of the simulated precipitation extremes for Ruetz were just outside the confidence ranges. In general, these results seem to be in line with other studies where similar findings have been published (Pui et al., 2009; Westra et al., 2012).

These results were then used to assess the performance when using 100 years of generated weather data as input into the models. The microcanonical model produced bigger precipitation amounts in comparison with the method of fragments model. This was to be expected, since the method of fragments model tries to reproduce observed precipitation (Pui, et al., 2009). Some of the observed probability of exceedances fell within the confidence boundaries for the method of fragments model, while all were outside the boundaries when simulated with the microcanonical model (Figure 13, 14, 15). These patterns are strenghtened in Müller & Haberlandts' study (2015), who found that their microcanonical model slightly overestimated precipitation extremes.

The performances of the two models seemed to be more alike for the two catchments with more data (Kelchsauer and Gurglbach). For these two catchments, the confidence ranges are constrained in comparison with the results for Ruetz. Moreover, the estimated precipitation depths are comparable between the two model approaches for Kelchsauer and Gurglbach. This is especially demonstrated with the simulated precipitation depth with 50 a year return period (P50) for Kelchsauer (Figure 14) where the two models estimated approximately the same depths. A difference between the two models for the simulated P50 in all catchments, is that the confidence ranges seem slightly more constrained for the microcanonical model.

The two models also seemed to preserve conventional statistics seen in the observed time series (Müller & Haberlandt, 2015; Pui et al., 2009) even if 100 years of extrapolated weather data were used as input. The method of fragments model simulated precipitation with statistics that are more alike observed events in comparison to the microcanonical model apart from the mean, where the two models had similar results.

Determining the most efficient structure for the method of fragments model was done by looking at the score out of four different criterions. Many structures produced results with similar statistics, especially if the variance is excluded. The biggest difference between variables simulated with the most efficient and least efficient structure was 6.3 % for the variance in Gurglbach. However, in general differences were much smaller even for the variance. As the parameters for the model is determined to account for seasonality and to avoid disproportionate biases when disaggregating extreme events (Westra et al., 2012; Pui et al., 2012), it is not motivated to try more structures.

As described in the theory section of the method of fragments algorithm, the model does not necessarily maintain dependence across days for storm events. The correlation in rainfall between the 24^{th} hour of the previous day and the first hour of the current day may therefore be compromised (Pui et al., 2012). As this affects antecedent conditions, it is noteworthy and something to consider if disaggregated data are to be used for flood modeling.

A big uncertainty related to the disaggregation procedure is the number of model runs, i.e., the number of realizations to simulate for every single weather generated data series. Disaggregation is a random process since each run of the model allows choosing up to 10 different nearest neighbors. To account for this random behavior, a sufficient amount of disaggregation runs need to be performed to get stable values (Müller & Haberlandt, 2015). The biggest uncertainty was related to the greatest precipitation depths simulated by the model. It was shown that 15-25 runs of the model seemed to make mean values of the produced statistics converge, however, no significance test was performed to assess these differences. This is a limit that is motivated to account for in future studies.

5.0.4 Temperature disaggregation

Temperature disaggregation was shown to reduce residuals considerably in comparison with using daily averages as hourly values for Kelschauer and Gurglbach. For Ruetz, the same results could not be achieved for unknown reason, especially since the length of the temperature series for the three catchments were similar (16-18 years). Using a temperature disaggregation approach conditioned to rainfall strengthened by relationships between the two atmospheric variables (Richardson, 1981), seems promising according to the results of this study. However, due to time constraints, a proper literature study, assessment and analysis of the temperature disaggregation model was not feasible.

Some temperature disaggregation simulations were made with the generated weather data. These preliminary test scores seem optimistic even though there are problems with the temperature of individual days being unrealistically high in comparison with observed temperatures of that time interval. This is probably the result of big differences between the nearest neighbors $|T_j - T_t|$ leading to big differences when multiplied with the $TempRatio_{i,m}$

vector. A tool to cut outliers is suggested to be implemented in the temperature disaggregation to disable the possibility for unrealistic temperatures.

In Table 10 it is demonstrated that the highest discharge peak modeled with disaggregated temperature data for Gurglbach was 10.47% closer to simulations performed with historic data in comparison with using temperature averages as input. An improvement, 5.93%, was also seen for the simulations performed in Kelchsauer with disaggregated temperature data. Simulations depend on the quality of the input data. If improvements can be made so that the temperature data is more similar to observed values (approximately 40 % smaller residuals in comparison with daily averages, see Table 9), this is motivated since the framework as a whole is associated with many uncertainties. The aim must be to reduce all the errors so that the framework is as realistic as possible.

5.0.5 Modeling Framework

To test the modeling framework as a whole, one disaggregated precipitation realization together with the corresponding hourly temperature series with hourly averages were used as input in the HBV model. The long term mean monthly temperature and potential evaporation used for the calibration and validation procedure in the HBV model were also used here to enable comparison between the results. 100 years of hourly data equals 876,000 data points for each variable (T, P). To account for equifinality, all of the 100 realizations from each study area (50 realizations produced by the method of fragments model and 50 from the microcanonical model) were going to be simulated with 100 suitable but different parameter sets simulated from 200 000 Monte Carlo runs. In HBV light, simulations with many different parameter sets is possible with the tool ”batch runs”. However, when this was practiced in the HBV model, it almost immediately crashed due to a transcendence of simulated data points (87,600,000 runoff simulations). When batch runs were tried for a smaller interval, just to test the model framework, the program displayed an error with the input files. This error was not solved during the end phase of the thesis.

A big advantage with the proposed modeling framework is as mentioned before; that coupling of weather generation with hydrological modeling allows for impact assessment on changing land use and climate conditions (Bergström et al., 2001; Booij, 2005; Breinl, 2016), something that is not possible with frequency analysis. The framework also attempts to tackle challenges with inadequate data series (too short and with insufficient resolution) (Müller & Haberlandt, 2015; Westra et al., 2012; 2013; Kobold & Brilly, 2006; Pui et al., 2009).

Another proposed benefit of using a conceptual runoff model as an alternative to frequency analysis with distribution fitting, is the closer understanding of the physical processes of the catchment hydrology (Lindström et al., 1997; Seibert, 1999a, 1999b, 2012). As seen here and elsewhere, extrapolation with distribution fitting performs badly beyond the conditions of the sample (Merz & Blöschl, 2003; Moran, 1958; Chow et al., 1988).

However, there are many uncertainties related to the modeling framework as well. First of all, the weather generation process involves curve fitting (Breinl et al., 2013). The same goes for the microcanonical model with uniform splitting when precipitation volumes are distributed according to $W_{x/x}$, where x is associated with a certain probability distribution (Olsson, 1998; Müller & Haberlandt, 2015; Güntner et al., 2001). The microcanonical model has also been seen to overestimate extremes (Müller & Haberlandt, 2015). Another uncertainty is that storm events crossing days may be redistributed so that these conditions are altered. There are also uncertainties related to measurements of precipitation intensity and duration. Since precipitation data is one of the main inputs in the HBV model, this may have a significant effect on the flood modeling (Habib et al., 2001). The preliminary tests of temperature disaggregation show that disaggregating temperature results in simulations of discharge peaks with more resemblance to historic observations in comparison with using daily temperature averages as input in the HBV model.

The HBV model, the last model of the framework, may introduce big errors. The choice of model structure determines the number of parameters, where a more complex model structure introduces more uncertainty. The use of elevation zones adds complexity to the model and the method to calculate potential evapotranspiration can lead to different estimations which may influence runoff simulations. Many parameters are highly uncertain, hence a big range of values was seen to result in equally good fits according to the objective function meaning that many parameter sets may be used for simulation which could lead to different estimations of design floods. Different choices of objective function may also lead to dissimilar results.

The many elements of the modeling framework is another source of complexity that makes it hard to quantify the uncertainties of the whole framework. However, the framework as a whole attempts to understand the physical causes leading to a flood of certain return period as well as to understand how changes in weather data and changes in land-use and climate can alter runoff. It also tackles challenges with insufficient data amounts and resolution. The trade off between introduced uncertainties in comparison with the many applications provided by the framework motivates further research, especially as an alternative to frequency analysis.

6 CONCLUSIONS

- Frequency analysis for design floods estimation is straightforward but shown to be uncertain in this study. It does not give any information on antecedent conditions or on causes of the floods. It is not applicable for land use and climate change studies. It does not solve issues with scarce data series or data series at coarse temporal resolution.

- The method of fragments model is non-parametric, has few parameters and is shown to reproduce observed precipitation extremes within confidence boundaries for Kelchsauer and Gurglbach. The results are better when more data is available, indicating that the model is data driven. The model can be used for temperature disaggregation with slight changes, something that from preliminary tests seems to improve hydrological simulations in comparison with using temperature averages.

- The two disaggregation models used in this study can produce high resolution precipitation data which can be used as input in the HBV model to estimate design floods. The difference between the disaggregation models was bigger when smaller amounts of historic high resolution data was available. Simulations with the method of fragments model reproduce precpitation extremes with more resemblance to historic observations in comparison with the microcanonical model. The reproduction of precipitation variance and the proportion of wet spells simulated with weather generated precipitation is also closer to observations with the method of fragments model in comparison with the microcanonical model.

- There are many uncertainties related to the modeling framework and it is complex in comparison with frequency analysis. The weather generator uses theoretical probability distributions for simulating precipitation amounts and occurrence, the disaggregation models are data driven, the parameters of the HBV may be uncertain and difficult to estimate, which in combination with the choice of objective function influence the modeling of discharge peaks. However, the framework is closer to a physical understanding of the causes behind floods and it is applicable for land-use and climate change studies. It is also positive since it attempts to solve problems related to data scarcity.

7 REFERENCES

Allan, R. P. & Soden, B. J. (2008). Atmospheric Warming and the Amplification of Precipitation Extremes. *American Association for the Advancement of Science*, vol. 321, pp. 1481-1484.

Anon., 2015. ORF.at. [Online] Available at: http://tirol.orf.at/news/stories/2725468/ [Accessed 23 05 2017].

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: model development. *Journal of the American Water Resources Association*, vol. 34(1), pp. 73-89.

Baldassare, B. & Ranzi, R. (2003). Hydrological and meterological aspects of floods in the Alps: an overview. *Hydrology and Earth System Sciences*, vol. 7(6), pp. 785-798.

Bergström, S. (1976). *Development and application of a conceptual runoff model for Scandinavian catchments*, Norrköping: Swedish Meteorological and Hydrological Institute (SMHI).

Bergström, S. (1990). *Parametervärden för HBV-modellen i Sverige, Erfarenheter från modellkalibreringar under perioden 1975-1989*. 28 ed. Norrköping: SMHI Hydrologi.

Bergström, S. (1992). *The HBV Model - its structure and applications*, SMHI Hydrology, RH No.4, Norrköping, 35 pp.: s.n.

Bergström, S., Carlsson, B., Gardelin, M., Rummukainen, M. (2001). Climate change impacts on runoff in Sweden‹ assessments by global climate models, dynamical downscaling and hydrological modelling. *Climate Research*, vol. 16(2), pp. 101-112.

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

Booij, M. J. (2005). Impact of climate change on river flooding assessed with different spatial model resolutions. *Journal of Hydrology*, vol. 303, pp. 176-198.

Breinl, K. (2016). Driving a lumped hydrological model with precipitation output from weather generators of different complexity. *Hydrological Sciences Journal*, vol. 61(8), pp. 1395-1414.

Breinl, K., Strasser, U., Bates, P., Kienberger, S. (2015a). A joint modelling framework for daily extremes of river discharge and precipitation in urban areas. *Journal of Flood Risk Management*, vol. 10(1),pp. 97-114.

Breinl, K., Turkington, T., Stowasser, M. (2013). Stochastic generation of multi-site daily precipitation for applications in risk management. *Journal of Hydrology*, vol. 498, pp. 23-

35.

Breinl, K., Turkington, T., Stowasser, M. (2015b). Simulating daily precipitation and temperature: a weather generation framework for assesing hydrometeorological hazards. *Meteorological Applications*, vol. pp. 334-347.

Chow, V. T., Maidment, D. R., Mays, L. W. (1988). *Applied Hydrology*. Singapore: McGraw Hill Company.

Das, T., Bárdossy, A., Zehe, E. & He, Y. (2008). Comparison of conceptual model performance using different representations of spatial variability. *Journal of Hydrology*, vol. 356, pp. 106-118.

Ding, J., Wallner, M., Mülle, H. & Haberlandt, U. (2016). Estimation of instantaneous peak flows from maximum mean daily flows using the HBV hydrological model. *Hydrological Processes*, vol. 30(9), pp. 1431-1448.

Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J., Stoffel, M. (2014). 21st century climate change in the European Alps - A review. *Science of the Total Environment*, vol. 493, pp. 1138-1151.

Güntner, A., Olsson, J., Calver, A. & Gannon, B. (2001). Cascade-based disaggregation of continuous rainfall time series: the influence of climate. *Hydrology and Earth System Sciences*, vol. 5(2), pp. 145-164.

Haberlandt, U., Hundecha, Y., Pahlow, M. & Schumann. A, H. (2011). Rainfall Generators for Application in Flood Studies.. In: H. Schumann. A, ed. *Flood Risk Assessment and Management: How to Specify Hydrological Loads, Their Consequences and Uncertainties*. New York: Springer, pp. 117-147.

Habib, E., Krajewski. F, W., Kruger, A. (2001). Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrological Engineering*, vol. 6(2), pp. 159-166.

Kobold, M. & Brilly, M. (2006). The use of HBV model for flash flood forecasting. *Natural Hazards and Earth System Sciences*, vol. 6, pp. 407-417.

Kottegoda, N. T. & Rosso, R. (2008). Frequency Analysis of Extreme Events. In: *Applied Statistics for Civil and Environmental Engineers*. 2 ed. s.l.:Blackwell Publishing Ltd, pp. 405-478.

Kundzewicz, Z. W., Pińskwar, I., Brakenridge, G. R. (2012). Large floods in Europe, 1985–2009. *Hydrological Sciences Journal*, vol. 58(1), pp. 1-7.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, vol. 201, pp. 272-288.

Merz, R. & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, vol. 39(12).

Molnar, P. & Burlando, P. (2005). Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model. *Atmospheric Research*, vol. 77(1-4), pp. 137-151.

Moran, P. (1957). The Statistical Treatment of Flood Flows. *Eos, Transactions American Geophysical Union*, 38(4), pp. 519-523.

Moran, P. (1958). Discussion of ”The Statistical Treatment of Flood Flows”. *Eos, Transactions American Geophysical Union*, vol. 39, pp. 732-736.

Müller, H. & Haberlandt, U. (2015). Temporal Rainfall Disaggregation with a Cascade Model: From Single-Station Disaggregation to Spatial Rainfall. *Journal of Hydrologic Engineering*, vol. 20(11).

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, vol. 47, pp. 90-100.

Nash, J. E. & sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology*, vol. 10, pp. 282-290.

Olsson, J. (1998). Evaluation of a scaling cascade model for temporal rainfall disaggregation. *Hydrology and Earth System Sciences*, vol. 2(1), pp. 19-30.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C. (2005). Which potential evapotranspiration input for a lumped Part 2—Towards a simple and efficient potential evapotranspiration. *Journal of Hydrology*, vol. 303, pp. 290-306.

Pui, A., Sharma, A. & Mehrotra, R. (2009). A comparison of Alternatives for Daily to Sub-Daily Rainfall Disaggregation. Cairns, School of Civil and Engineering University of New South Wales, Sydney Australia.

Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of Hydrology*, vol. 470-471, pp. 138-157.

Richardson, C. W. (1981). Stochastic Simulation of Daily Precipitation. *Water Resources Research*, vol. 17(1), pp. 182-190.

Seibert, J. (1997). Estimation of Parameter Uncertainty in the HBV Model. *Nordic Hydrology*, vol. 28(4/5), pp. 247-262.

Seibert, J., (1999a). *Conceptual runoff models - fiction or representation of reality?*, Uppsala: Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 436.

Seibert, J. (1999b). Regionalisation of parameters for a conceptual rainfall-runoff model. *Agricultural and Forest Meteorology*, vol. 98-99, pp. 279-293.

Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences*, vol. 4(2), pp. 215-224.

Seibert, J. (2001). On the need for benchmarks in hydrological modelling. *Hydrological Processes*, vol. 15, pp. 1063-1064.

Seibert, J. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, vol. 16, pp. 3315-3325.

Vanham, D., Fleischhacker, E. & Rauch, W. (2008). Technical Note: Seasonality in alpine water resources management – a regional assessment. *Hydrology and Earth System Sciences*, vol. 12, pp. 91-100.

Weingartner, R., Barben, M. & Spreafico, M. (2003). Floods in mountain areas—an overview based on examples from Switzerland. *Journal of Hydrology*, vol. 282(1-4), pp. 10-24.

Westra, S., Evans, P. J., Mehotra, R., Sharma, A. (2013). A conditional disaggregation algorithm for generating fine time-scale rainfall data. *Journal of Hydrology*, vol. 479, pp. 86-99.

Westra, S., Mehrotra, R., Sharma, A., Srikanthan, R. (2012). Continuous Rainfall Simulation: 1. A regionalized subdaily disaggregation approach. *Water Resources Research*, vol. 48.

Westra, S. & Sharma, A. (2010). *Australian Rainfall and Runoff Revision* , s.l.: Engineers Australia Water Engineering.

Xu, C. Y. & Singh, V. P. (2001). Evaluation and generalization of temperature-based methods for calculating evaporation. *Hydrological Processes*, vol. 15, pp. 305-319.

Yoo, K. H. & Boyd, C. E. (1994). *Hydrology and Water Supply for Pond Aquaculture*. New York: Chapman & Hall.

7.0.1 Personal communication

Breinl, K. (2017). Postdoctoral at the Department of Earth Sciences, Program for Air, Water and Landscape Sciences.

Okoli, K. (2017). Doctoral student at the Department of Earth Sciences, Program for Air, Water and Landscape Sciences.

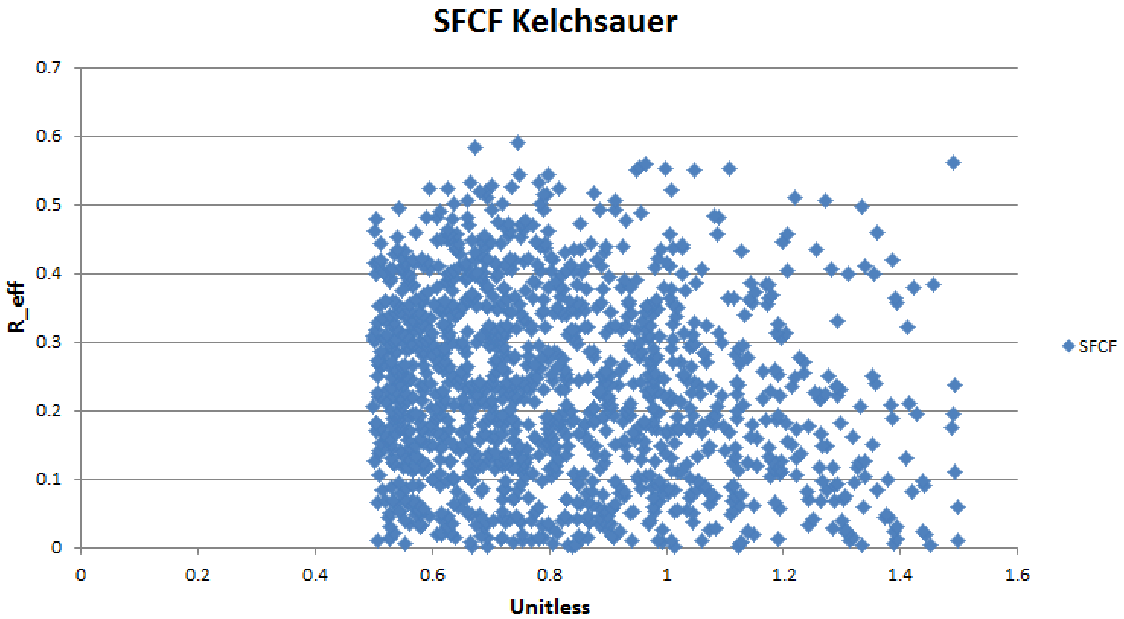
APPENDIX
APPENDIX A: HBV PARAMETER REFERENCE VALUES

A 1. The initial parameters and their ranges used for the Monte Carlo simulations used in the HBV model

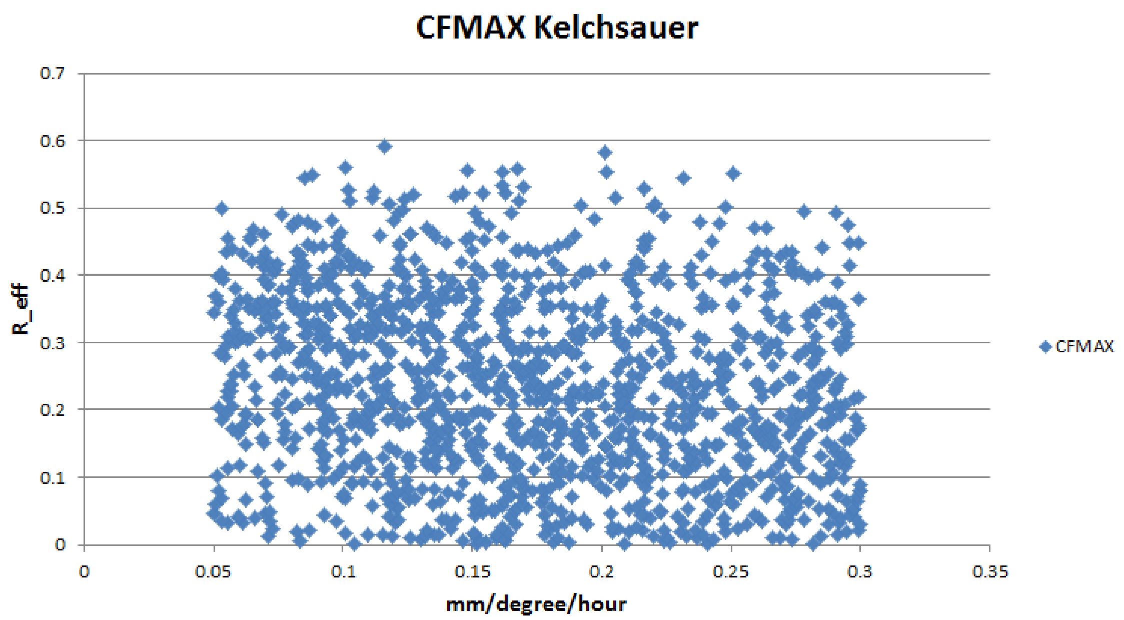
Parameter	Explanation	Min	Max	Unit
Snow Routine				
TT	Threshold temperature	-0.5	1.5	$°C$
CFMAX	Degree-day factor	0.05	0.2083	$mm\,°C^{-1}\,h^{-1}$
SFCF	Snowfall correction factor	0.6	1.5	-
CFR	Refreezing coefficient	0.05	0.05	-
CWH	Water holding capacity	0.1	0.1	-
Soil& evaporation Routine				
FC	Maximum soil moisture	100	300	*mm*
LP	Soil moisture threshold for reduction of evaporation	0	1	-
Beta	Shape coefficient	0.01	4	-
Groundwater& response routine				
K_0	Recession coefficient	0.0033	0.02916	h^{-1}
K_1	Recession coefficient	0.00083	0.02083	h^{-1}
K_2	Recession coefficient	8.33E-05	0.00083	h^{-1}
UZL	Threshold for K0-outflow	8	70	*mm*
PERC	Maximal flow from upper to lower GW-box	0	0.2083	*mm* h^{-1}
MAXBAS	Routing, length of weighting function	1	60	*h*
Other				
CET	Potential evaporation correction factor	0	0.3	$°C^{-1}$
PCALT	Change of precipitation with elevation	10	18	$\%\,100^{-1}\,m^{-1}$
TCALT	Change of temperature with elevation	0.6	0.6	$\%\,100^{-1}\,m^{-1}$

APPENDIX B SENSITIVITY ANALYSIS

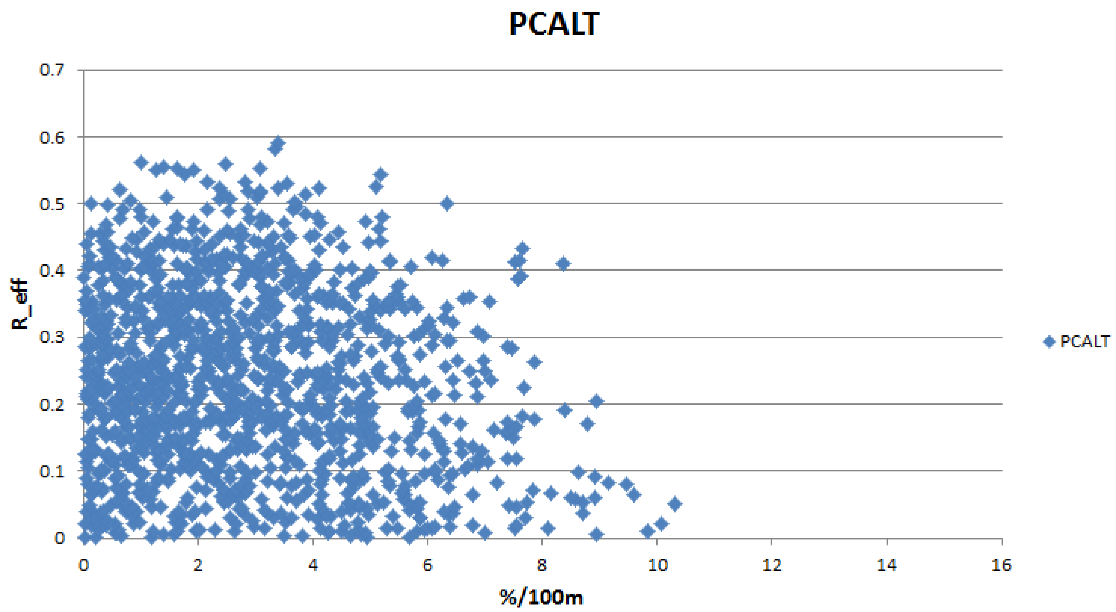
Here are a few example plots from the parameter uncertainty test. For well defined parameters, a distinct peak can be found (Figure B 3) while more uncertain parameters looks more like the results for the parameter $CFMAX$ where many parameter values give the same fit (Figure B 2).



B 1. Uncertainty plot for the parameter SFCF, the snow correction factor and the Nash objective function.



B 2. Uncertainty plot for the parameter CFMAX, the snow degree day factor and the Nash objective function.



B 3. Uncertainty plot for the parameter PCALT which control differences in precpitation in relation to altitude and the Nash objective function.