

# Spatially distributed modelling of hydro-meteorological hazards for a physical deterministic analysis of flash flood occurrence

A case study in Khao Yai National Park, Thailand

Eva C. Potthoff





# Spatially distributed modelling of hydro-meteorological hazards for a physical deterministic analysis of flash flood occurrence

**A case study in Khao Yai National Park, Thailand**

by

Eva C. Potthoff

to obtain the degree of Master of Science  
at the Delft University of Technology,  
to be defended publicly on 23-09-2022

Student number: 4369068  
Thesis committee: Dr. T. A. Bogaard, TU Delft, supervisor  
Dr. L. P. H. van Beek, Universiteit Utrecht  
Dr. P. P. Mapiam, Kasetsart University  
Dr. Ir. J. A. E. ten Veldhuis, TU Delft

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.



# Acknowledgements

Writing this thesis has been a roller coaster of a journey, which I would not have been able to do without the help and support of the people around me. This past year has been full of ups and downs, including lockdowns, a model that did not do what I wanted it to do and of course a wonderful time in Thailand, working together with the people there.

I would like to take this moment to thank the people who made it possible for me to enjoy the ups and get through the downs. First of all I would like to thank Thom, for guiding me throughout this whole journey. Whenever I was losing track, you always knew to motivate me again with your undying enthusiasm. Thank you for providing me with the opportunity to go to Thailand and helping me get approval from the TU. I would also like to thank Punpim, for welcoming me with open arms in Thailand. The scarf and wooden elephant you gave me have a special spot in my bedroom now! I had a wonderful time in Bangkok, which is for a big part thanks to you and your family. I hope we will meet again soon, maybe during my travels through Asia this winter. Next, I would like to thank Rens for all the energy and time spent on me and my struggles with understanding and running the model. I hope I didn't give you too much late night work... Also I want to thank Marie-Claire for agreeing to join my graduation committee and giving me some good questions and suggestions during my green-light meeting. I would also like to thank Poom. I really appreciate all your efforts to make my time in Bangkok comfortable and welcoming. You made sure I would taste all the tasty dishes served in the KU canteen. Luckily we both agreed the crispy pork was the best dish on the menu! I would also like to thank the people reading through my report and providing me with comments and suggestions to improve it. Your feedback was invaluable.

I want to thank my parent for being patient with me throughout this period and for their support throughout my studies at the TU and my entire life. Special thank yous to: My dear friends from het Afs-tudeerhok, with whom I had a lot of laughs, good talks and heated ping-pong matches; my friends at home, who helped me get my mind of my thesis work by going out and having good times. A special shout out to Eefje, who was my buddy in all this, especially the last few weeks. I hope that after this, we will remember how to have regular conversations again and not just talk about our research projects.

Thanks to everyone who helped me in whatever way possible and I hope you will enjoy reading this thesis!

*Eva Potthoff*  
*Delft, September 16, 2022*



# Abstract

Due to rising temperatures, rainfall patterns around the world are being affected, causing extreme precipitation events to become more frequent and intense, resulting in an increased probability of severe flash floods. Thailand is no exception to the increased risk of these hazards, which is why Early Warning Systems are being set up. Since flash floods occur within a few hours after the triggering precipitation event, timely and accurate precipitation observations are critical, to enable timely warning and evacuation of inhabitants to mitigate the risk. One solution for this is the use of rain radar, which provides rain data with high spatial and temporal resolutions.

An analysis of this technique was chosen to form the basis of this research in Northeastern Thailand's Lam Takhlung basin. The objective was to investigate the importance of using distributed precipitation data. In this research, the hydrological response of the catchment of interest was studied. The ability of the physically-based, conceptual and distributed CALEROS model to capture this response was assessed. Different modelling strategies and sources of precipitation input were analysed, and the additional value of using distributed precipitation data was determined. The examination of multiple hydrological characteristics in the study area shows great heterogeneity of the catchments response to precipitation. A clear differentiation can be made between the hydrological response of the upstream and downstream part of the catchment. Additionally, the results from the catchment characterisation indicate great spatial variability of the precipitation patterns in the study area. This is confirmed when using the CALEROS model to recreate the catchments hydrology.

Four modelling strategies were used, varying by spatial and temporal constancy. Calibration performed using uniform precipitation showed to be incapable of capturing the discharge trends in the catchment, failing to properly catch the observed peaks as well as the base flow. Runs performed using distributed precipitation maps obtained by Inverse Distance interpolation of rain gauge data showed significantly better results, adequately capturing the trend of the discharge observed in the catchment. Differences in parameterisation only had limited effect on the outcome, making the precipitation input the most important parameter. Runs using synthetic precipitation data demonstrated the importance of properly capturing the precipitation pattern and movement across the catchment, as it greatly influences the timing of occurring discharge peaks.

A comparison between rain gauge data and rain radar data shows that, although data quality of the rain gauges seems acceptable, rain gauges are not capable of properly capturing rainfall patterns, while rain radar does. Precipitation patterns were found to be the most crucial parameter for modelling of flash floods in the area of interest, signifying the importance of using rain radar data for accurate flash flood forecasting.

**Keywords:** flash floods, physically-based model, distributed precipitation, rain gauges, hydrological characterisation, radar, Thailand



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.1.1	Floods . . . . .	1
1.1.2	Landslides . . . . .	2
1.1.3	Early Warning Systems . . . . .	3
1.1.4	Modelling . . . . .	3
1.2	Relevance . . . . .	6
1.3	Research Goal . . . . .	7
1.4	Thesis outline . . . . .	7
<b>2</b>	<b>Hydrological Characterization of the Catchment</b>	<b>9</b>
2.1	Study Area . . . . .	9
2.1.1	Climate . . . . .	9
2.1.2	Topography, soil type and land cover . . . . .	10
2.1.3	Hazards . . . . .	11
2.2	Data . . . . .	12
2.2.1	Climate data . . . . .	12
2.3	Methodology hydrological characterisation . . . . .	14
2.3.1	Base Flow . . . . .	14
2.3.2	Wave celerity . . . . .	15
2.3.3	Runoff . . . . .	15
2.3.4	Recession coefficient . . . . .	16
2.3.5	Individual event analysis . . . . .	17
2.3.6	Water balance . . . . .	18
2.4	Results . . . . .	18
2.4.1	Base Flow . . . . .	19
2.4.2	Wave celerity . . . . .	19
2.4.3	Runoff . . . . .	19
2.4.4	Recession coefficient . . . . .	20
2.4.5	Individual event analysis . . . . .	20
2.4.6	Water balance . . . . .	23
2.5	Discussion . . . . .	23
<b>3</b>	<b>Modelling</b>	<b>27</b>
3.1	Model Description . . . . .	27
3.1.1	Data input . . . . .	27
3.1.2	Data output . . . . .	29
3.2	Method modelling . . . . .	31
3.2.1	Sensitivity Analysis . . . . .	31
3.2.2	Uniform and static precipitation input . . . . .	33
3.2.3	Uniform and dynamic precipitation input . . . . .	34
3.2.4	Spatial and static precipitation input . . . . .	36
3.2.5	Spatial and dynamic precipitation input . . . . .	36
3.3	Results modelling . . . . .	38
3.3.1	Sensitivity Analysis . . . . .	38
3.3.2	Uniform and static precipitation input . . . . .	39
3.3.3	Uniform and dynamic precipitation input . . . . .	39
3.3.4	Spatial and static precipitation input . . . . .	44
3.3.5	Spatial and dynamic precipitation input . . . . .	44

3.4 Discussion . . . . .	47
<b>4 Radar data analysis</b>	<b>51</b>
4.1 Methodology Radar Data. . . . .	51
4.1.1 Radar pixel comparison . . . . .	51
4.1.2 Radar image comparison . . . . .	51
4.2 Results Radar Data . . . . .	52
4.2.1 Radar pixel comparison . . . . .	52
4.2.2 Radar image comparison . . . . .	53
4.3 Discussion . . . . .	54
<b>5 Conclusion and Recommendations</b>	<b>57</b>
5.1 Conclusion . . . . .	57
5.2 Recommendations . . . . .	59
<b>A Station Locations</b>	<b>65</b>
<b>B Data consistency</b>	<b>67</b>
B.1 Missing Data . . . . .	67
B.2 Double Mass Curves . . . . .	68
<b>C Base Flow</b>	<b>75</b>
<b>D Lag Time and Runoff Coefficient</b>	<b>77</b>
D.1 Event overview . . . . .	77
D.2 Lag time and runoff correlations . . . . .	78
<b>E Pattern Description</b>	<b>81</b>
<b>F Model Description</b>	<b>85</b>
F.1 Data Input. . . . .	85
F.2 Data output . . . . .	88
<b>G Initial input variables</b>	<b>91</b>
<b>H Model variable list</b>	<b>95</b>
<b>I Sensitivity Analysis</b>	<b>97</b>
<b>J Parameterisation</b>	<b>99</b>
<b>K Synthetic spatial and dynamic precipitation input</b>	<b>101</b>
<b>L Radar data comparison</b>	<b>103</b>
L.1 Data images . . . . .	103
L.2 Error results. . . . .	107
L.3 Difference map evaluation . . . . .	107

# List of Figures

1.1	<i>Components of an effective Early Warning System (Budimir et al., 2021).</i>	3
1.2	<i>Illustration of the different options for model distribution</i>	4
2.1	<i>Location of the study area.</i>	10
2.2	<i>The (a) Digital Elevation Map (DEM), (b) land cover map and (c) soil type map of the study area.</i>	11
2.3	<i>Slope of the river bed.</i>	11
2.4	<i>Hazards occurring as a result of the heavy rainfall in Khao Yai National Park.</i>	12
2.5	<i>Location of the different stations in and near the study area.</i>	13
2.6	<i>(a) The two hydrological response units and (b) the (sub) catchments in the study area.</i>	13
2.7	<i>Illustration of base flow and direct flow in a discharge hydrograph.</i>	15
2.8	<i>Peaks at (a) discharge station M43a and (b) discharge station M89, used to determine the wave celerity.</i>	16
2.9	<i>Used rain gauges for lag time determination of precipitation events based on Thiessen polygons and distance to downstream discharge station.</i>	17
2.10	<i>Illustration of lag time.</i>	18
2.11	<i>An overview of (a) the determined BFI during the three seasons for the different years and (b) the overall annual BFI.</i>	19
2.12	<i>Correlation between lag time and various hydrological characteristics of the corresponding precipitation event.</i>	21
2.13	<i>(a) The hyetograph of three similar precipitation events and (b) the resulting hydrograph of the direct discharge.</i>	22
2.14	<i>The water balance according to the different hydrological characteristics determined at the two different discharge stations.</i>	25
3.1	<i>Schematisation of the calculations done for each time step.</i>	28
3.2	<i>Simplified representation of the water fluxes within the catchment determined by the model.</i>	30
3.3	<i>Overview of the four different types of modelling strategies performed based on the spatial and temporal variation of the precipitation input.</i>	32
3.4	<i>Schematic of the relation between an output variable <math>y</math> and an input parameter <math>x</math> (Lenhart et al., 2002).</i>	32
3.5	<i>Schematic illustration of the boundary conditions of the different water balances and their corresponding in- and outfluxes.</i>	33
3.6	<i>Identified isolated discharge peaks with a peak larger than <math>50 \text{ m}^3/\text{s}</math>.</i>	34
3.7	<i>Flow Duration Curve of the discharge observed in 2018.</i>	35
3.8	<i>Observed discharge at the two discharge station for the year 2018.</i>	36
3.9	<i>The four sub areas for the spatial and static precipitation input runs.</i>	37
3.10	<i>Design storm used for the spatial, dynamic precipitation modelling.</i>	37
3.11	<i>The precipitation event moving in different direction using distribution according to polygons and Inverse Distance.</i>	38
3.12	<i>(a) Modelled total annual discharge, (b) modelled annual actual evaporation and (c) modelled annual deep percolation for the <math>10 \text{ mm/day}</math> run using the initial parameters.</i>	39
3.13	<i>Test precipitation used for the partial static precipitation run and the resulting modelled discharge and determined base flow.</i>	40
3.14	<i>Modelled discharge after parameter tuning based on an isolated individual discharge event.</i>	40
3.15	<i>Results of the validation run for Parameter Set B and Parameter Set C.</i>	41
3.16	<i>The Flow Duration Curves of the modelled discharge using different parameterisations.</i>	41

3.17	<i>Observed precipitation of the year 2018 at rain gauge STN0551 and the resulting modelled discharge for the five parameter sets.</i>	43
3.18	<i>Observed and modelled discharge using uniform precipitation zoomed in to May 2018.</i>	43
3.19	<i>Observed and modelled discharge using uniform precipitation zoomed in to August and September 2018.</i>	43
3.20	<i>Modelled discharge for the spatial, static precipitation input.</i>	44
3.21	<i>Modelled discharge for the spatial, dynamic precipitation input, with precipitation only occurring at one sub area of the catchment.</i>	45
3.22	<i>Modelled discharge for the spatial, dynamic precipitation input, with precipitation moving across the catchment using polygon distribution.</i>	45
3.23	<i>Modelled discharge for the spatial, dynamic precipitation input, with precipitation moving across the catchment using Inverse Distance distribution.</i>	45
3.24	<i>Modelled discharge of 2018 for the five parameter sets using distributed precipitation obtained from rain gauge data.</i>	46
3.25	<i>Modelled discharge using distributed precipitation interpolated from rain gauge data zoomed in to August and September 2018 using the obtained parameter sets.</i>	46
4.1	<i>Hourly precipitation observations from multiple rain gauges in the study area and the radar pixels corresponding to those locations.</i>	52
4.2	<i>Daily precipitation observations from multiple rain gauges in the study area and the radar pixels corresponding to those locations.</i>	52
4.3	<i>Overview of Inverse Distance of rain gauge data, Radar data and difference between rain gauge data and radar data on August 15<sup>th</sup> until August 17<sup>th</sup>.</i>	53
4.4	<i>Percentage of overestimated (positive) and underestimated (negative) precipitation values for each day when using interpolated rain gauge data opposed to radar data.</i>	54
A.1	<i>Locations of stations in and around the study area.</i>	66
B.1	<i>Double Mass Curves of the rain gauges for the different years.</i>	73
C.1	<i>Total flow measured at discharge station M43a and the corresponding baseflow as determined with the Eckhardt Algorithm.</i>	75
C.2	<i>Total flow measured at discharge station M89 and the corresponding baseflow as determined with the Eckhardt Algorithm.</i>	76
D.1	<i>Correlation between lag time, runoff coefficient and various characteristics of the precipitation event.</i>	79
F.1	<i>The potential evaporation as determined by the model using Hamon, as determined using the Penman-Monteith equation and as determined using Hamon, after scaling it to the Penman-Monteith results.</i>	89
K.1	<i>Modelled discharge for the synthetic precipitation input using both Thiessen and Inverse distance interpolation for the movement of the precipitation event in different directions.</i>	101
L.1	<i>Overview of the interpolated rain gauge map, the rain radar map and the difference map between rain gauge data and radar data.</i>	106
L.2	<i>Boxplots of the pixel values of the difference maps for each day.</i>	107

# List of Tables

2.1	<i>Seasonal temperatures (Celsius) in Northeastern Thailand.</i>	10
2.2	<i>Seasonal rainfall (mm) in Northeastern Thailand.</i>	10
2.3	<i>Summary of obtained data, including information on the data type, sources, resolution and period of the data set.</i>	12
2.4	<i>Coefficients of determination determined from yearly Double Mass Curves, created with daily data.</i>	14
2.5	<i>Distance between discharge stations and rain gauges.</i>	17
2.6	<i>Mean monthly BFI determined over the years 2012-2020.</i>	19
2.7	<i>Routing time of discharge peaks between station M43a and station M89.</i>	20
2.8	<i>Seasonal and annual runoff coefficients for the different years.</i>	20
2.9	<i>Base flow recession coefficient for the different years.</i>	20
2.10	<i>Number of precipitation events meeting the set criteria, but unsuitable for lag time and runoff determination.</i>	21
2.11	<i>Characteristics of the three similar precipitation events and the lag time and runoff coefficient determined from the resulting direct runoff peak.</i>	22
2.12	<i>Annual total evaporation as determined with the water balance.</i>	23
3.1	<i>Minimum, maximum and mean monthly temperatures (°C) from Pakchong Agromet weather station determined over the years 2006-2021.</i>	28
3.2	<i>Sensitivity classes.</i>	33
3.3	<i>Overview of determined the sensitivity class per parameter.</i>	38
3.4	<i>Errors calculated for the FDC of modelled discharge compared to observed discharge.</i>	42
3.5	<i>Errors calculated for the annual model runs using uniform precipitation input.</i>	44
3.6	<i>Errors calculated for the period August and September of the annual model runs using uniform precipitation input.</i>	44
3.7	<i>Errors calculated for the annual model runs using distributed precipitation input.</i>	46
3.8	<i>Errors calculated for the period August and September of the annual model runs using distributed precipitation input.</i>	46
3.9	<i>Base Flow Index values determined from modelled discharge.</i>	47
3.10	<i>Runoff coefficients determined from modelled discharge.</i>	47
B.1	<i>Missing data per station per year.</i>	67
D.1	<i>Information on the precipitation events used to determine the lag time for direct runoff peaks at discharge station M43a.</i>	77
D.2	<i>Information on the precipitation events used to determine the lag time for direct runoff peaks at discharge station M89.</i>	77
F.1	<i>Minimum, maximum and mean monthly temperatures (°C) from Pakchong Agromet weather station determined over the years 2006-2021.</i>	85
F.2	<i>Minimum, maximum and mean groundwater levels (in meters from the ground surface) observed by the groundwater stations in the study area.</i>	87
I.1	<i>Overview of the sensitivity indices for different input parameters and output variables.</i>	97
L.1	<i>Overview of some specifics of the different maps obtained.</i>	107



# Acronyms

<b>BFI</b>	Base Flow Index
<b>CALEROS</b>	Calabrian Erosion
<b>DEM</b>	Digital Elevation Map
<b>DMC</b>	Double Mass Curve
<b>DMR</b>	Department of Mineral Resources
<b>EWS</b>	Early Warning System
<b>FAO</b>	Food and Agricultural Organization of the United Nations
<b>FDC</b>	Flow Duration Curve
<b>GLAD</b>	Global Land Analysis and Discovery
<b>HII</b>	Hydro-Informatics Institute
<b>HRU</b>	Hydrological Response Unit
<b>ISRIC</b>	International Soil Reference and Information Centre
<b>KGE</b>	Kling-Gupta Efficiency
<b>LAI</b>	Leave Area Index
<b>NSE</b>	Nash-Sutcliffe Efficiency
<b>RMSE</b>	Root Mean Square Error
<b>TMD</b>	Thai Meteorological Department
<b>UN-OCHA</b>	United Nations Office for the Coordination of Humanitarian Affairs



# Introduction

In the last decades, the Earth has undergone major changes as a result of climate change. Temperatures have risen immensely, affecting among others the rainfall patterns around the world. This manifests itself in more extreme weather in the form of severe droughts and more intense precipitation. Current climate models indicate that temperatures will continue to rise, intensifying the Earth's water cycle even more. As a result, storm affected areas around the world will experience more frequent and intense precipitation and consequently, an increased probability of severe floods and landslides (Merzdorf, 2020; United Nations, n.d.-b). In addition, changes in worldwide demography, such as expanding urbanisation, result in a greater portion of the population living in areas prone to these hazards, increasing the resulting risk (Hapuarachchi et al., 2011).

## 1.1. Background

Landslides and flash floods are two hydro-meteorological hazards causing extreme damage to life and property worldwide. A flash flood is the rapid flooding of an area and a landslide is the gravity driven movement of rock, debris or soil masses down a slope, both often as a result of heavy rainfall .

### 1.1.1. Floods

Floods are the most common natural hazard worldwide, accounting for roughly 40% of all natural disasters. They are also among the most destructive ones (Chen et al., 2020). In the period between 1994 and 2013, more people were affected by floods than by all other natural hazards combined (UNISDR et al., 2015a). The region of Southeast Asia especially is very prone to frequent and severe floods (Cheng, 2021), being the most flood vulnerable region in the world and consequently having the highest fatality rate (Chen et al., 2020; Shah et al., 2020). In recent decades both frequency and severity have been rising as a result of climate change and human interventions (Shah et al., 2020; Torti, 2012). Most of Asia's floods occur following the heavy rainfalls brought by the summer monsoon. For centuries, Asia has flourished under this abundant rainfall during its wet seasons, as communities relied on them for the fertility of agricultural lands and the generation of hydropower. Climate change, however, is making the monsoon rains more unpredictable, exposing an increasing number of communities to risk (Cheng, 2021; Shah et al., 2020). As the monsoons arrive increasingly later, while the amount of rainfall has stayed roughly the same, precipitation falls over a shorter period of time resulting in more intense rainfall and more severe floods (Cheng, 2021). Since communities located in Southeast Asia often do not have the essentials to handle these floods and restore the resulting damage. Floods threaten the lives of millions of people in that area (Chen et al., 2020; Torti, 2012).

When looking at Thailand specifically, floods are a serious issue. Of the large number of people affected by floods in Southeast Asia, the majority lives in Thailand. During the wet season of 2011, the United Nations Office for the Coordination of Humanitarian Affairs (UN-OCHA) estimated that 9.5 million people in Southeast Asia were affected by floods that year, of which 5.3 million people living in Thailand. The floods in Thailand during that year were so severe that they were labelled to be the most damaging floods in over 50 years (Bidorn et al., 2015; Chen et al., 2020). Amongst the people affected, many lost

their jobs, their homes, and some even their lives (DisasterAWARE, 2011). While 2011 was extreme, it is common to have multiple cases of flooding a year in Thailand, affecting the lives of tens to hundreds of thousands of people each year and causing numerous fatalities per event. The majority of these floods are riverine floods, but also flash floods and landslides are quite common as a result of extreme precipitation (Floodlist, n.d.).

### **Flash floods**

There are several types of floods, of which flash floods are an extremely dangerous one, developing very quickly. Where most floods take days to weeks to develop, flash floods only take six hours or less. They are caused by extremely intense rainfall in a short period of time over a relatively small area (Campbell, 2017) and are often accompanied with other hazards such as landslides or mud flows (Hapuarachchi et al., 2011). The energy of the fast moving water mass during a flash flood is high enough to sweep entire houses away and, due to their sudden onset, there is little to no advance warning. This makes them extremely difficult to predict and evacuate people in time (Borga et al., 2011; Campbell, 2017; Cheng, 2021; Shah et al., 2020).

Flash flood forecasting is becoming increasingly important worldwide. For appropriate warning and real time forecasting of flash floods, both hydrologic and meteorologic expertise is required. As flash floods are a local phenomenon with a short response time, local information of small temporal scale is required for effective warning. Warning strategies used for river floods, that build up over time, are often not effective for flash floods, as flash floods have a quick response to rainfall input. River flood forecasting is typically solely hydrological based, while flash flood forecasting is a hydro-meteorological problem (Campbell, 2017), as flash floods are generated by the interaction between the meteorology and the hydrology of a location. Besides the intensity of a rainfall event, also the antecedent conditions, land use, terrain characteristics and geomorphology affect the likelihood of a flash flood event to occur (Campbell, 2017; Chen et al., 2020). This adds complexity to predicting flash floods as the interaction between multiple aspects needs to be considered.

### **1.1.2. Landslides**

Landslides occur all around the world and are considered a significant global hazard. Although they occur significantly less frequent than floods, still 5% of all natural disasters between 1995 and 2015 were landslides (UNISDR et al., 2015b) and approximately 14% of all casualties resulting from natural hazards can be attributed to landslides (Froude and Petley, 2018). A landslide event can have multiple causes: geological, morphological or human, but only one trigger (Wieczorek, 1996). The triggers for landslides can be diverse, ranging from natural factors, such as intense rainfall, snow melt, freezing, earthquakes or volcanic eruptions, to anthropogenic activities, such as hill cutting, mining or construction (Froude and Petley, 2018; Zhao and Lu, 2018). One of the most common triggers for landslides is excessive water, often as a result of heavy rainfall (Connors, 2019). Especially in Asia, this is an important trigger as a result of the intense precipitation events during the wet season. Making Asia the most landslide prone region in the world, with more than half of all events worldwide occurring there (Froude and Petley, 2018).

In Thailand, landslides pose an increasing risk. Although the dominating natural hazards in the country are floods, landslides also lead to considerable casualties (Department of Mineral Resources, n.d.), especially in the more mountainous regions. Compared to floods, landslides are considered to be less harmful, but the number of people affected by landslides is rapidly increasing (Schmidt-Thomé et al., 2018). The annual number of landslides has been increasing, as a result of changing rainfall patterns due to both climate change and anthropogenic changes (Schmidt-Thomé et al., 2018). Additionally, the population in Thailand is increasing. In many of the rural, mountainous areas, forests have been converted to agricultural lands and plantations. According to the Department of Mineral Resources (DMR), there are 1084 communities, comprising 6563 villages in 54 provinces located in a landslide hazard zone (Schmidt-Thomé et al., 2018). Most of the areas in Thailand that are prone to landslides are located in remote and mountainous areas which are difficult to reach and thus far from help when a landslide strikes (Schmidt-Thomé et al., 2018). Early warning systems that forecast landslides are crucial in mitigating their effects and damage.

### 1.1.3. Early Warning Systems

In order to mitigate the effects of both hazards, proper preparation and response is required. To ensure this, often Early Warning Systems (EWS) are put in place. Early Warning System (EWS) are complex systems, in which multiple components are interconnected with each other. An effective EWS is comprised of four inter-related key elements as are illustrated in Figure 1.1. These four key elements, as defined during the Third International Conference on Early Warning, are risk knowledge, monitoring and warning, dissemination and communication, and response capability (UN/ISDR, 2006).

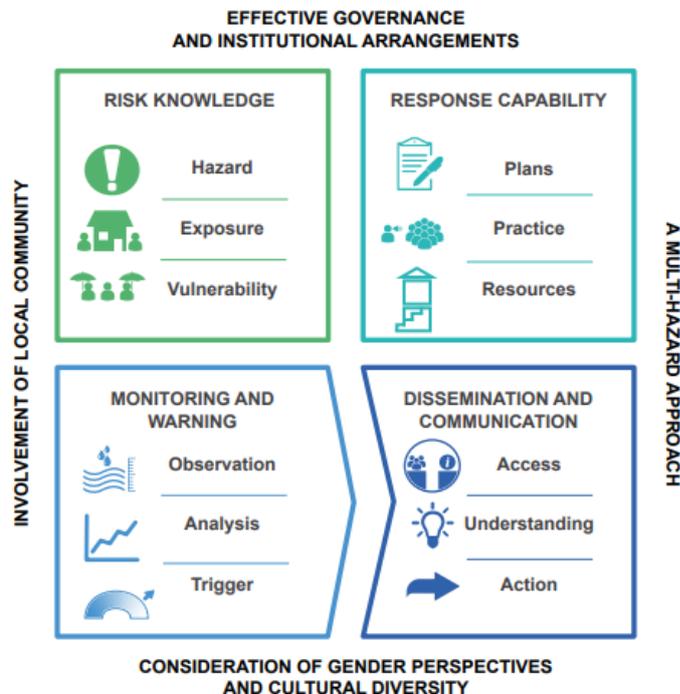


Figure 1.1: Components of an effective Early Warning System (Budimir et al., 2021).

Knowledge on the risks, including the hazards, exposure and vulnerabilities, is important in order to understand how these risks could affect hazard prone communities. To assess the risks, systematic collection and analysis of data is required, considering the dynamic nature of these aspects. Prediction and forecasting of the potential hazard should be timely and reliable and should operate on a 24-hour basis. A continuous monitoring of the hazard parameters is essential to generate accurate and timely warnings. As warnings need to reach the people at risk, messaging should be clear and simple. It is crucial that communities understand the risks and know how to respond in the case of a hazard. To accomplish this, education and preparedness programs are fundamental (Budimir et al., 2021; UN/ISDR, 2006). This thesis will focus on the first two key elements.

### 1.1.4. Modelling

For EWS to be put in place, a proper model predicting the potential hazards must be available. A model can be seen as a simplified representation of the real world, with the best model giving results closest to reality, optimally using as little parameters as possible. Hydrological models can be classified based on multiple aspects, such as the model input and parameters and the extent of physical principles applied in the model (Devia et al., 2015).

When classifying a model based on the model parameters as a function of space and time, the hydrological model can be categorised into three different groups; lumped, semi-distributed and distributed (Figure 1.2). The classification of these groups is based on their spatial variability representing the hydrological processes of the catchment of interest (Kumar and Bhattacharjya, 2020). A lumped model is often used for regional models. In a lumped model, the whole catchment of interest is represented as one single model entity. The behaviour of spatially distributed physical systems within the catchment are simplified, assuming the parameters to be constant throughout space and only being dependent

of time. Because of this, lumped models require little data, and thus use little computational power. A variation on the lumped modelling method is semi-distributed modelling. Here, the catchment is divided into multiple smaller sub-catchments or Hydrological Response Unit (HRU). By doing this, the variation of runoff processes and pathways within a catchment is considered, as for each sub-catchment or HRU specific model parameters are set. In semi-distributed modelling, factors such as runoff, elements of the water balance, and the interaction between them are simulated separately for each sub unit and subsequently combined to obtain the overall catchment response (Kása et al., 2017). In a fully distributed model, the processes in catchments are represented in a high resolution spatial scale. The catchment of interest is divided into numerous individual entities, most commonly grid cells. This allows for a detailed prediction at many points within the catchment as the parameters are a function of both space and time. Distributed models allow for characterising different parts of the landscape by different hydro-meteorological conditions, vegetation and subsurface structures.

In large systems or systems with much variability, lumped models are not desired as they oversimplify the situation, misrepresenting the system and reducing the predictive capability of the model. Distributed models have the potential to be more accurate but also pose some challenges, as significantly more parameters need to be quantified. These can be problematic to obtain when modelling a large region.



Figure 1.2: Illustration of the different options for model distribution which are (from left to right) lumped, semi-distributed or distributed models.

When classification is done based on the variability of the output, a model can be classified as either deterministic or as stochastic. A deterministic model will always give an identical output for a given set of input values, while a stochastic model can produce a different output for the same set of input values each time you run it (Devia et al., 2015).

The last classification is based on the extent of physical principles applied in the model, making the model either empirical, conceptual or physically-based (Devia et al., 2015). Physically-based models, or mechanistic models, are based on observed parameters, imposing physical laws such as the conservation of mass and energy (Aghakouchak and Habib, 2010). They allow for a high degree of detail, making them very suitable for small-scale applications where parameters can be observed, or can at least be reasonably assumed, with sufficient detail in both time and space. At too large scale however, parameters cannot be sufficiently well observed, making a physically-based model an inconvenient method. Physically-based models generally do not need a lot of data for calibration, due to the large number of parameters used to describe the physical characteristics of the catchment. They give information on aspects such as soil moisture, water depth and topography (Devia et al., 2015). The spatial resolution of physically-based models is generally high, making them appropriate to use for distributed modelling.

Empirical models are based on limited physical process knowledge, distinguishing a relationship between different variables without considering the laws of physics. These models are based on observations, taking only information of existing data into account without regarding hydrological processes and features of the catchment. The mathematical equations used in empirical models are derived from input and output time series instead of from the physical processes within the catchment (Devia

et al., 2015). These models typically have a low spatial resolution, making them suitable for lumped modelling.

In conceptual models, a hydrological system is still simplified, but contrary to the empirical model, some laws of physics are considered. Semi-empirical equations are used, assessing the model parameters not only from field data but also through calibration (Devia et al., 2015). Simple mathematical equations describe the main hydrological processes, such as evaporation, percolation, base flow and runoff. Conceptual models are less complex than physically based models, as the processes in the system are described with simplified equations instead of partial differential equations. To compensate for the simplification of the calculations, however, model parameters need to be used, that may not have any physical meaning (Aghakouchak and Habib, 2010).

### **Flash flood modelling**

In the past decades, efforts to improve flash flood forecasting have been made and new modelling techniques have been investigated. Hapuarachchi et al., 2011, did a review on the different approaches in flash flood modelling and forecasting. Originally mainly empirical and lumped models were used in flash flood modelling, as these models are relatively simple to set up and only limited information is required in order to obtain plausible results (Hapuarachchi et al., 2011). However, as empirical models require long-term data sets for calibration and are rather site specific, popularity is shifting towards physically-based models. Distributed models are gaining popularity over lumped models, due to the increased availability of new data. Lumped models have a coarse resolution and, like empirical models, require long-term data sets for calibration. This in combination with poorly gauged catchments makes the use of lumped models for flash flood forecasting generally inconvenient. Because of this, more models are becoming distributed or semi-distributed, providing similar or better performance, while being capable of representing the physical characteristics of a catchment (Hapuarachchi et al., 2011).

An increasing number of physically-based distributed models have been developed, having high temporal and spatial resolution. Various tests showed that they give realistic simulations with only limited training data. Because of the spatial scale of these models, they are very suitable for flash flood forecasting and due to the high number of physically based parameters, no long data records are required for parameter calibration. Not only does this result in a relatively short computation time, making it an appropriate tool for real-time flash flood forecasting, the high number of physical parameters increases the potential of the models being successful in catchments where gauging is poor (Hapuarachchi et al., 2011). Additionally, Moore et al., 2006, suggested that distributed models often give more plausible results than lumped models, especially for extreme weather events. A great disadvantage of physically-based distributed models is the fact that they require high quality data of the catchment, such as a DEM, land use maps and soil characteristics. When deciding which type of model is suitable, it is important to consider the application of the model as well as factors such as data availability, process complexity and temporal and spatial scale. But overall, in flash flood modelling, physically-based distributed models are deemed to be more suitable than lumped and empirical models, as flash floods generally occur in small catchments with limited rain gauges, making it difficult to calibrate those models (Hapuarachchi et al., 2011).

### **Landslide modelling**

As rainfall is seen as one of the main initiators of slope instability, rainfall thresholds are often used to evaluate the risk that a landslide will occur. Singh and Kumar, 2021, did a review on different rainfall threshold techniques for determining and forecasting landslides. The rainfall threshold is defined as the amount of rain which, when reached or exceeded, is expected to cause slope failure at a specific location. It can be determined either empirically, in which case it is dependent on historical analysis of landslide triggering rainfall events, or physically-based, in which case, it is based on the physical interaction of rainfall and slope stability. For the latter, the general aim is to simulate the physical processes related to water infiltration as a result of a precipitation event. This method does not only consider rainfall, but also additional aspects that could cause destabilisation of the slope. Such aspect could be either static or dynamic, including topography, geology, hydrology and geo-technical parameters (Singh and Kumar, 2021).

Physically based rainfall thresholds are very popular to apply in EWS for landslide predictions, as both location and timing of the forecasted landslide can be predicted, considering site-specific parameters

in combination with dynamic factors such as rainfall and soil characteristics. However, at the same time this is one of the disadvantages of these types of thresholds, as detailed information related to hydrology, morphology, lithology, and soil are required in order to correctly predict slope instability, being difficult to collect, especially on larger scale. Therefore, physically- based rainfall thresholds are mainly effective for small scale catchments (Singh and Kumar, 2021).

Empirical thresholds, on the other hand, do not require much additional information, aside from historical data on rainfall and landslides. Different types of empirical rainfall thresholds have been developed, such as Intensity-Duration thresholds, Total Event rainfall thresholds, Rainfall Event-Duration thresholds, Rainfall Event-Intensity thresholds and antecedent rainfall thresholds. However, only few models are based on the latter, although antecedent rainfall can play a significant role in slope destabilisation. Empirical rainfall thresholds are most useful at regional, national or global scale, but it is disadvantageous that these thresholds are very site specific and can not be applied to other areas. Additionally, the temporal accuracy of models based on these thresholds is not sound. Furthermore, historical records of both rainfall events and landslides in an area are necessary, which can be a problem, especially in developing countries (Singh and Kumar, 2021).

### **Radar now-casting**

As both flash floods and rainfall induced landslides are triggered by heavy rainfall events, it is desirable to predict these events ahead of time, for timely forecasting of both hazards. To do so, rainfall input of high spatial and temporal resolution is required (Liechti et al., 2013; Mapiam et al., 2022). As landslides and flash floods typically occur within a short time span after the rainfall event, accurate and fast rainfall information is desired.

In developing countries rain gauges are often sparse and unreliable. Classic weather prediction models often lack the ability to capture convective rainfall patterns. One solution could be offered by using radar now-casting. Where many of the classic numerical weather prediction models are mesoscale and thus incapable of capturing smaller scale convective storms, radar observations grant the possibility to represent rainfall patterns at a significantly higher resolution (Prudden et al., 2020). Its capability of detecting precipitation at both finer spatial and temporal resolution makes it a promising option in forecasting models, due to its ability of accurately capturing spatial patterns. Especially for short duration convective storms, weather radar observations could be beneficial for deriving rainfall events triggering potential flash floods or landslides (Marra et al., 2014).

### **Chosen model type**

In this research a distributed, deterministic and physically based model will be used to model the discharge and consequently flash floods in the study area. This type of model has been chosen as it allows predictions at local and regional scale, which is appropriate since flash floods, and landslides generally occur at small spatial scale. Additionally, this type of modelling does not require a lot of calibration, and with the sparse rain gauging in the area, i.e. limited data records, not much calibration is possible. Furthermore, the spatial and physical nature of the model allows for consideration of additional hydrological factors, which can be important in the modelling of both flash floods and landslides. Lastly, a distributed model is desired when eventually working with radar input data. Since weather radar provides data of a fine spatial resolution, a model that allows for these spatial differences is most suitable.

## **1.2. Relevance**

As discussed, flash floods and landslides are two major natural hazards bringing damage to Thailand on a yearly basis. As the number of vulnerable areas in the country is high, it is expected that the damage caused by natural hazards will continue to grow each year, substantially affecting the economy of Thailand and the lives of its inhabitants (Ikeda and Palakhamarn, 2020). In order to mitigate these hazards, EWS are being set up, helping communities to prepare for the possibly hazardous events. Successful EWS could save lives, land and infrastructure and protect the economy in the long term (United Nations, n.d.-a). One of the key elements of an EWS is proper monitoring of the hazard (UN/ISDR, 2006), for which understanding of the events and their potential triggers and causes is of high importance. For both landslide and flash flood initiation intense rainfall is recognised as one of the

most critical factors, however also additional hydrological information can be an important dynamic factor causing slopes to be susceptible for landslide or flash flood initiation during triggering storm events. Despite the fact that both landslides and flash floods are triggered by the same factors, little research has been conducted on the relationship between the two. Additionally, most warning models are solely based on rainfall data, neglecting further hydrological information, while addition of such information in hazard assessment can improve the prediction of landslide and flash flood initiation (Bogaard and Greco, 2016; Marino et al., 2020). Considering the relationship between landslides and flash floods as well as addition of hydrological information could improve monitoring of both hazards, especially when combined, and thus becomes a valuable tool for a good functioning EWS.

Another important factor in proper monitoring and forecasting of hydro-meteorological hazards is accurate and high resolution precipitation data. As landslides and flash floods typically develop as the result of small-scale, heavy rainfall within a short period of time, it is difficult to monitor these hazards timely, thus making it complicated to accurately issue a warning. The conventional rain gauge network in Thailand is of low density, and thus insufficient for capturing these precipitation events. The use of weather radar could pose a solution to this, as it is capable of detecting precipitation at a higher temporal and spatial resolution (Liechti et al., 2013). Although capable of detecting precipitation at the desired resolution, the accuracy of the weather radar is often lacking as errors in the reflectivity measurements and reflectivity rainfall conversions often affect the radar rainfall estimates. Still, weather radar could be a valuable addition, as methods have been found to correct for the radar reflectivity with a combination of conventional rain gauge and additional low-cost rain gauge networks. Research by Mapiam et al., 2022, used this method to improve the bias correction of radar data, which is crucial for increasing the accuracy of radar rainfall estimates and consequently forecasting of potential flash floods and landslides.

### 1.3. Research Goal

This research is part of a bigger project, with an overall goal to construct a coupled near real-time prediction model for landslides and flash floods, which can be used for EWS in Thailand. In order to reach this goal, the bias correction technique for weather radar as discussed by Mapiam et al., 2022, will be used to obtain optimised radar data. This data can then be used for rainfall predictions, which will be applied for forecasting of potential flash floods and landslides, with a lead time of 6 hours. These predictions will then be used in a coupled EWS for both natural hazards.

This thesis will focus on the flash flood aspect of this project. A model will be set up, with an aim to make a sound conceptual representation of the hydrological processes in the study area. With this model, various rain events will be analysed, to understand how the study area reacts to different types of precipitation events. Additionally, the importance of distributed rainfall information for the flash flood forecasting will be assessed.

This all has resulted in the following research question for this thesis:

*"How does distributed precipitation data improve the accuracy of flash flood modelling?"*

In order to answer this question, three sub-questions have been defined:

1. How does the catchment respond to precipitation and vary throughout the area?
2. To what extent is the model capable of recreating the hydrological response of the study area and how is this affected by the spatial resolution of the input data?
3. What will be the effect of using radar data as input for the model?

### 1.4. Thesis outline

This report consists of three main parts, all focused on answering one of the sub questions of this research, and will be closed with a conclusion and recommendations, where the overall research question will be answered.

Chapter 2 starts of with an introduction of the study area, after which an overview will be given of the different data sets used, and the different institutions in Thailand the data was obtained through.

Next, a catchment characterisation will be done using the observed data, analysing the hydrological response of the study area to precipitation and whether this differs for different parts of the catchment. First the methods used for the catchment characterisation will be explained, after which the results will be presented, ending with a discussion.

In chapter 3, the response of the catchment to incoming precipitation will be modelled, using both uniform and distributed input data, in order to analyse the effect of distributed precipitation data. First a model description will be given, after which the different methods used will be explained. Lastly, the results will be given and discussed.

In chapter 4, radar images will be compared to distributed precipitation data created from the rain gauges in the study area, to evaluate what the added value of using radar input data could be for modelling of discharge.

Lastly, a conclusion of the different parts will be given in chapter 5, elaborating on the different sub questions, and finally answering the research question of this report. Recommendations will be given to optimise future research.

# 2

## Hydrological Characterization of the Catchment

In this chapter the hydrological characteristics of the catchment will be discussed. To start with, a description of the study area will be given, including information on the climate, topography, soil types and land cover, as well as a brief example of the hazards occurring in the area. Next, the data used will be discussed and consistency will be tested. After that, the methods used to determine multiple hydrological characteristics of the catchment will be discussed, and the results will be presented. Lastly, the results will be discussed.

### 2.1. Study Area

The region of interest for this research is a catchment located in the Lam Takhlung basin in Northeastern Thailand in the province of Nakhon Ratchasima, on the north side of the Khao Yai National Park (Figure 2.1). Khao Yai National Park is the third largest National Park of Thailand, and is a mountainous area, mainly consisting of igneous rocks overlain by sedimentary rocks, with an average elevation between 400 and 1000 meters above sea level. Due to the steep terrain of the National Park and the heavy rainfall during the wet season, every few years landslides and flash floods occur, causing obstruction of the road and damage to nearby districts. Especially the northern part of the National Park is susceptible to these hazards, which is why it has been chosen as the study area for this research. The catchment of interest has a total area of 703  $km^2$ .

#### 2.1.1. Climate

The climate in Thailand is classified as a humid tropical climate. Thailand can be divided into five climate zones, based on climate patterns and meteorological conditions. These different climate zones are displayed in Figure 2.1b. Thailand's general climate is influenced by two monsoon winds of seasonal character; the southwest monsoon, which is from May to October, and the northeast monsoon, being from November to February. The monsoons are caused by seasonal temperature changes between the land mass and the Gulf of Thailand, bordering the south of the country. The southwest monsoon coincides with the wet season in the country, with the wettest months of the year being August and September (Land Development Department, 2011c; Thai Meteorological Department, 2015). Warm moist air is transported from the Indian Ocean towards the mainland of Thailand, resulting in abundant rainfall, especially in the western part of the country. However, as the mountain ranges in this western region block the winds from travelling further into the country, the southwest monsoon is not a major source of rainfall in the Northeastern part of Thailand. Here, other factors contribute to the rainfall during the wet season, such as monsoon troughs, low pressure cells and tropical cyclones (Land Development Department, 2011a; Thai Meteorological Department, 2015). Especially the latter two are important contributors to the rainfall in the study area. In low pressure cells, the warm air rises and as it cools down, rainfall can occur. Similar to this, the tropical cyclones are initiated by low pressure systems over the Pacific Ocean, and can either result in tropical depressions, tropical

storms or typhoons, which differ by wind speed. However, tropical storms and tropical typhoons are rare in Thailand, as the land buffer of Vietnam and Laos decreases the wind speed before drifting into Thailand (Thai Meteorological Department, 2015). The Northeast monsoon coincides with the winter season, with December to January being the coldest period, where cold and dry air is transported from the Chinese mainland over to the majority of Thailand (Land Development Department, 2011a, 2011d). The temperatures are mild and little to no precipitation occurs over the mainland of Thailand (Land Development Department, 2011c). The transition period between the northeast monsoon and the southwest monsoon is the pre-monsoon, or summer season. This season is characterized by increasing temperatures and is the warmest period of the year, with April being the warmest month (Land Development Department, 2011c, 2011d). Rainfall gradually increases during this season (Land Development Department, 2011b).

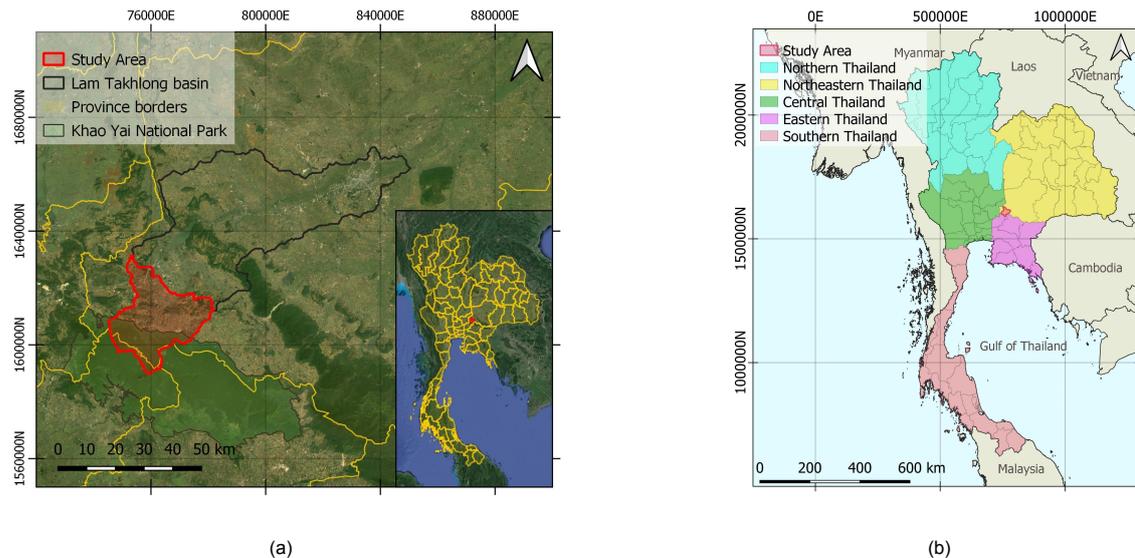


Figure 2.1: Location of the study area relative to (a) the Lam Taklong basin in Nakhon Ratchasima province and (b) the regions in Thailand based on climate patterns and meteorological conditions. The study area is located in northeastern Thailand.

An overview of the mean temperatures and rainfall in Northeastern Thailand for the different seasons can be found in Table 2.1 and Table 2.2 respectively (Thai Meteorological Department, 2015). Most areas in Thailand receive between 1200 - 1600 mm of rainfall per year, for Northeastern Thailand this is around 1400 mm per year.

	Winter season	Summer season	Wet season
<b>Min</b>	18.7	32.2	24.4
<b>Max</b>	30.6	35.2	32.6
<b>Mean</b>	24.2	28.6	27.6

Table 2.1: Seasonal temperatures (Celsius) in Northeastern Thailand.

Winter season	Summer season	Wet season	Annual rainy days
76.3	224.4	1103.8	116

Table 2.2: Seasonal rainfall (mm) in Northeastern Thailand.

### 2.1.2. Topography, soil type and land cover

The study area is partly located in the mountainous region of Khao Yai National Park ( $243 \text{ km}^2$ ) and partly in the flatter provincial landscape on the north side of the National Park ( $460 \text{ km}^2$ ), resulting in a clear distinction between the geomorphological characteristics within the study area.

Elevation ranges between 380 and 1350 metres in the National Park region and between 299 and

769 metres in the provincial area, with mean elevations of 751 metres and 419 metres respectively (Figure 2.2a). This results in significantly steeper slopes in the National Park than in the provincial area, which can be seen by the simplified visualisation of the slope of the river bed in Figure 2.3. Moreover, in the National Park 99.7% of the land is covered by forest, while in the provincial area this is only 43.7%. Besides forests, a large part of the provincial area is covered by croplands (48.6%) and built up areas (8.9%) as shown in Figure 2.2b.

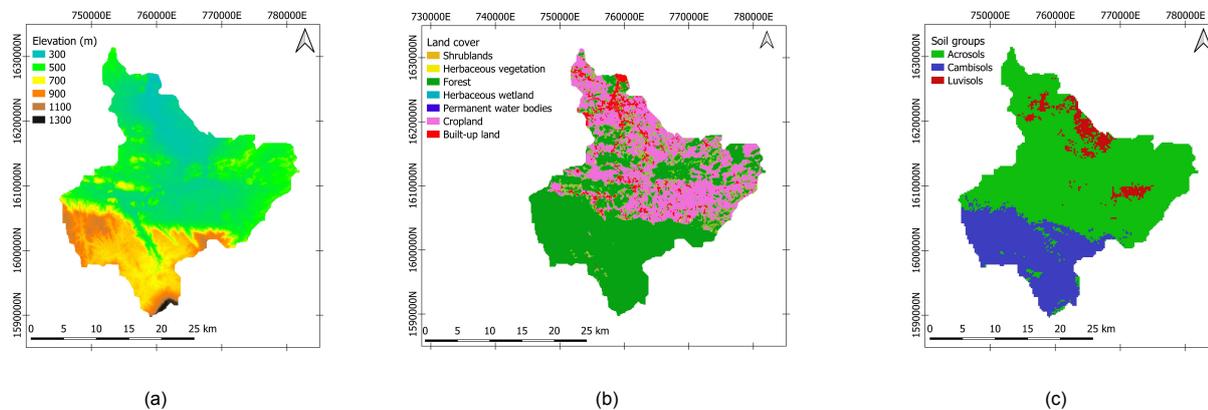


Figure 2.2: The (a) DEM, (b) land cover map and (c) soil type map of the study area.

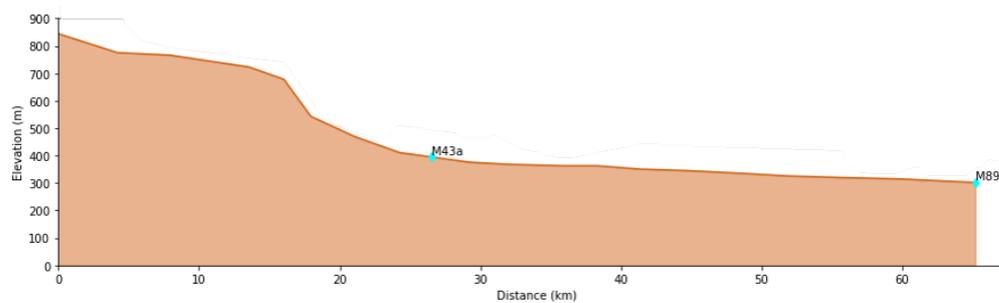


Figure 2.3: Slope of the river bed with the origin of the graph coinciding with the origin of the river and the end coinciding with the outlet point of the catchment.

Furthermore, when considering the soil groups as defined by the classification system of the Food and Agricultural Organization of the United Nations (FAO) (Figure 2.2c) the soil in the National Park predominantly consists of cambisols (75.6%). Cambisols are characterised by slight weathering of the parent material, and are often medium or fine textured, young soils. They often appear in mountainous regions due to the erosion and depositions cycles occurring there (IUSS Working Group WRB, 2015). In the provincial area on the other hand, acrisols are dominant (91.9%) and some luvisols can be found (8.1%). They both show low clay content in the top soil, with higher content in the subsoil. Acrisols often occur in regions with a tropical or monsoonal climate, and are the result of the weathering of acid rocks. The combination of humid climate and the acidic parent material leads to leaching, resulting in advanced degrees of weathering of the soil. This applies less to luvisols, which are characterised by high contents of silt (IUSS Working Group WRB, 2015).

The differences between the northern part and the southern part of the study area result in the distinction of two HRU during this research, as visualised in Figure 2.6a.

### 2.1.3. Hazards

As discussed, heavy precipitation events during the wet seasons causes obstruction and damage to roads and districts in the study area on an annual basis. An example of such events can be found in 2020, when a lot of damage was caused by landslides and flash floods in the park. At the end of September landslides around the side slope of the road caused the road to be closed as can be seen in Figure 2.4a. A few days later the road was closed again as heavy rainfall resulted in flash floods

in the park, causing blockage of the road (The Nation, 2020). Additionally, the province of Nakhon Ratchasima encountered the worst flood in 15 years (Figure 2.4b), affecting over 13000 households (Floodlist, 2020), inundating markets, the Pak Chong district office, police stations, temples and hotels (Bangkok Post, 2020).



Figure 2.4: Hazards occurring as a result of the heavy rainfall in Khao Yai National Park. In 2020 rainfall triggered (a) landslides in the National Park (28 September) and (b) flash floods in the Pak Chong district in Nakhon Ratchasima province (10 October).

## 2.2. Data

Table 2.3 gives an overview of the various data sets used, including additional information. For spatial data sets with data of multiple years, the data set of the most recent year has been used.

Type	Data	Source	Resolution	Year											
				2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	
Climate	Precipitation	TMD	Hourly												
		HII	Hourly												
	EWS	15 min													
	Discharge	RID	15 min												
	Groundwater	DGR	Monthly												
Elevation	DEM	NASA	30 meter												
Soil	Land use map	Copernicus	100 meter												
	Soil content	ISRIC	250 meter												
Vegetation	Cover fraction	Copernicus	300 meter												
	LAI	Copernicus	300 meter												
	Canopy height	GLAD	30 meter												

Table 2.3: Summary of obtained data, including information on the data type, sources, resolution and period of the data set.

### 2.2.1. Climate data

The climate data includes information on the precipitation and discharge in and around the study area. An overview of the different climate stations in the study area can be found in Figure 2.5. Precipitation information is obtained through 10 rain gauges located within the study area, of which 1 is an hourly gauge from the Hydro-Informatics Institute (HII) and 9 are quarter hourly gauges from the EWS. Downstream of the study area, still within the Lam Takhlung basin, 5 hourly rain gauges from the Thai Meteorological Department (TMD) are located (see Figure A.1), which will (partly) be used for the hydrological analysis of the study area.

Besides rain gauges, 2 discharge stations are located in the study area. One on the edge of the National Park (M43a) and one in the provincial area (M89). The size of the catchment above M43a is  $151 \text{ km}^2$ , which is about 20% of the total catchment (see Figure 2.6b). Station M89 has been set to be the outlet point of the catchment for this research. Discharge data of station M89 was provided in  $\text{m}^3/\text{s}$  and was converted to  $\text{mm}/\text{hr}$  using the area upstream of the discharge station. Initially data of discharge station M43a was only available between 29-10-2020 and 25-08-2021, also in  $\text{m}^3/\text{s}$ . Identically as for

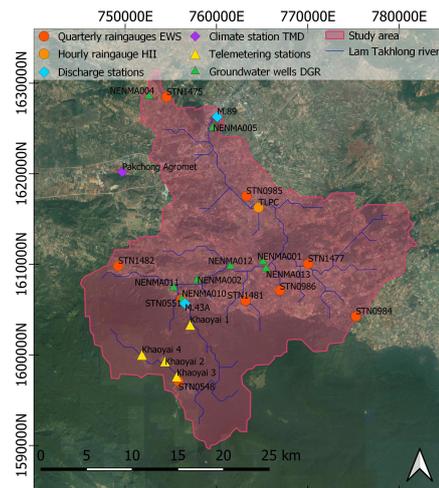


Figure 2.5: Location of the different stations in and near the study area.

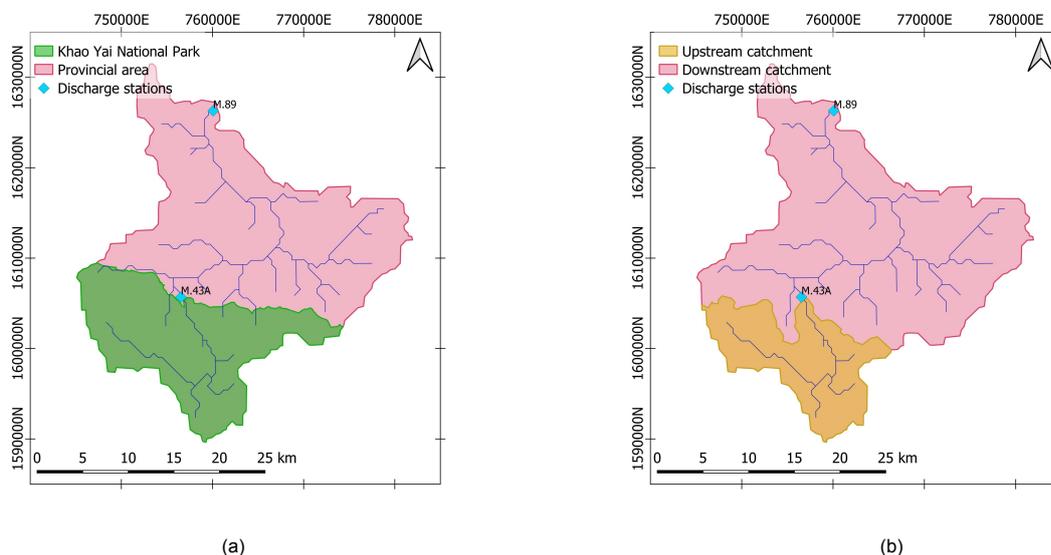


Figure 2.6: (a) The two hydrological response units and (b) the (sub) catchments in the study area.

M89, this has been converted to  $mm/hr$  using the area upstream of the station. Eventually, more data at gauge M43a became available, consisting of the measured water levels at the station, which was converted to discharge in  $m^3/s$  using the corresponding rating tables and again converted to  $mm/hr$ . Some of the calculations have been done using the initial discharge data from station M43a, however for most calculations the latter obtained data was used. When the earlier obtained data has been used, this will explicitly be mentioned.

### Data consistency

As rain gauge data in Thailand is generally of poor quality some checks were performed to increase the reliability of the data used. To check the consistency of the meteorological data, the Double Mass Curve (DMC) was used. The Double Mass Curve (DMC) checks the consistency of the data by comparing data of a single station to that of a pattern composed from data from multiple stations in the area (Searcy and Hardison, 1960). Rain gauges with more than 50% missing data values in a year were removed before making the DMC. An overview of the missing data per year for each station can be found in Table B.1 of Appendix B. Also the DMC for each station for different years can be found in Appendix B.

For all DMC the coefficient of determination ( $R^2$ -value) was calculated to determine how well the data

of each stations fits the overall pattern. In all further calculations and analysis, only stations with an  $R^2$  value of 0.985 or higher will be included (the green cells in Table 2.4). This boundary condition has been determined based on an initial reliability check of the data. For each rain gauge the total precipitation amount of all available years was determined. Some years gave unrealistically low annual precipitation values, but relatively high coefficients of determination. The  $R^2$  value has been chosen to be such that these years would be excluded.

Some annual precipitation values still seemed too low compared to average annual precipitation numbers given by literature and the TMD, however, due to lack of more accurate data, it has been decided to still use it for this research. Additionally, some stations showed unrealistically high values for certain time steps. For these cases, the values have either been corrected or set to 0.

	TLPC	4310024	4310021	4310006	4310042	4310040	STN0548	STN0551	STN0984	STN0985	STN0986	STN1475	STN1477	STN1481	STN1482
2011	-	0.997	0.997	0.999	0.994	0.995	-	-	-	-	-	-	-	-	-
2012	-	0.996	0.999	0.972	-	0.985	-	-	-	-	-	-	-	-	-
2013	0.998	0.992	0.999	0.999	-	0.990	-	-	-	-	-	-	-	-	-
2014	0.986	0.998	0.992	-	-	0.996	-	-	-	-	-	-	-	-	-
2015	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2016	0.960	0.966	0.970	0.838	-	0.982	-	-	-	-	-	-	-	-	-
2017	0.997	0.898	-	0.995	-	0.293	-	-	-	-	-	-	-	-	-
2018	0.870	-	-	-	-	-	0.990	0.997	0.997	0.982	0.995	0.997	0.995	0.987	0.998
2019	0.999	0.898	0.899	0.969	-	0.990	0.993	0.979	0.991	0.977	-	0.990	0.996	0.997	0.999
2020	0.986	-	-	-	-	-	0.997	0.990	-	-	-	-	-	-	-

Table 2.4: Coefficients of determination determined from yearly Double Mass Curves, created with daily data. Red cells indicate years with 50% or more missing values, and are thus excluded from the DMC. Orange cells indicate an  $R^2$  value lower than 0.985, and will thus be excluded from further calculation. Green cells indicate an  $R^2$  value of 0.985 or higher.

## 2.3. Methodology hydrological characterisation

For the hydrological characterisation of the study area, multiple aspects of the catchment have been evaluated using the precipitation and discharge data from the locations indicated in Figure 2.5. As explained in subsection 2.1.1, three climates can be distinguished in Thailand. For this research, the summer season is assumed to be during the months March and April, the wet season from May until October and the winter season from November until February.

### 2.3.1. Base Flow

Discharge consists of two components, the base flow and the direct flow, as is illustrated in Figure 2.7. Base flow is the result of the systems slow response to precipitation and is mainly maintained by groundwater. Direct flow on the other hand is the result of the system quickly responding to precipitation, with rainfall running off into the stream, mainly through overland flow and fast lateral flow. Knowledge on base flow and direct flow is important to get an understanding of the catchments' response to precipitation events. Additionally, it can provide useful information when looking into flash floods. Base flow results predominantly from groundwater recharge, contributing to the soil moisture and pore pressure in the subsurface before entering the streamflow. The part of the streamflow that does not result from base flow is a result of the direct flow, which, when abundant, can cause flash floods to occur.

For this study, the base flow is determined using the Eckhardt Algorithm, which is a two-parameter recursive filtering algorithm for base flow separation. The algorithm is known for its clear physical basis and easy operation (Yang et al., 2021) and additionally models a nicely balanced base flow, where the flow is not too low but also not too much of the discharge peaks is included. The Eckhardt Algorithm is given in Equation 2.1 (Eckhardt, 2005).

$$b_k = \frac{(1 - \beta_{max})ab_{k-1} + (1 - a)\beta_{max}y_k}{1 - a\beta_{max}} \quad (2.1)$$

where  $b_k \leq y_k$ .

$b_k$  = base flow at time step  $k$  (mm/hr),  $y_k$  = total streamflow at time step  $k$  (mm/hr),  $a$  = filter parameter (-),  $\beta_{max}$  = long-term maximum value of the base flow index that can be modelled by the algorithm (-).

The values for  $a$  and  $\beta_{max}$ , which are both unit-less, have been determined using the software BFI+.

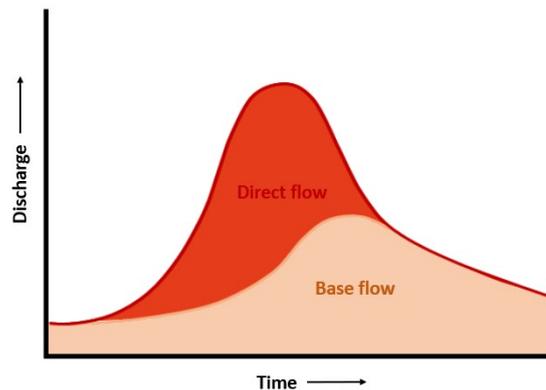


Figure 2.7: Illustration of base flow and direct flow in a discharge hydrograph.

For both stations parameter calibration resulted in a value of 0.999 for  $\alpha$  and the resulting values of  $\beta_{max}$  were 0.7 and 0.8 for station M43a and station M89 respectively. The resulting base flow at stations M43a and M89 of the different years can be found in Appendix C.

During dry seasons, the base flow is similar to the total discharge, as a lack of precipitation results in the absence of direct flow. During wet seasons, total flow often exceeds the base flow, as heavy rainfall leads to large peaks in the discharge, as a result of direct runoff. In this research, a relatively high base flow has been modelled. This decision has been made to obtain a smaller direct flow, primarily comprising quick runoff as a result of overland flow.

With the base flow, the Base Flow Index (BFI) can be determined. The BFI is the ratio between the long-term mean base flow and the long-term mean total streamflow. A high BFI indicates a large portion of the total streamflow being base flow while a low BFI indicated a large portion of the total stream flow being direct flow. The BFI has been determined for the years 2012-2020. Additionally, the BFI has been determined for the three different seasons.

### 2.3.2. Wave celerity

Flood routing has been applied to determine the development of the discharge peaks throughout the catchment. Peaks in the discharge measured at station M43a have been compared to corresponding discharge peaks measured at the downstream station M89. For the flood routing, the initial data set of station M43a has been used. In Figure 2.8, the peaks that have been compared are marked. By comparing the timing of these peaks the travel time, and consequently the wave celerity can be determined. The length of the river course between the two discharge stations is approximately 37 km.

For the peak event occurring in May, one peak has been measured at discharge station M43a whereas three peaks have been measured at discharge station M89. The time difference of all three peaks has been compared to that of other peak events to determine which of the three corresponds to the peak measured at M43a.

### 2.3.3. Runoff

The runoff coefficient relates the amount of direct flow to the amount of rainfall received by the system, giving information on the infiltration capacity and evaporation within the catchment. The runoff coefficient can be calculated using Equation 2.2. High values indicate a high amounts of direct runoff and thus lower infiltration, while low values indicate the opposite. The annual runoff coefficient for the different years has been determined as well as for the three seasons individually. Runoff values have been determined for both discharge stations, in combination with the rain gauge located within the same Thiessen polygon, which is STN0551 for station M43a and STN1475 for station M89 (see Figure 2.9).

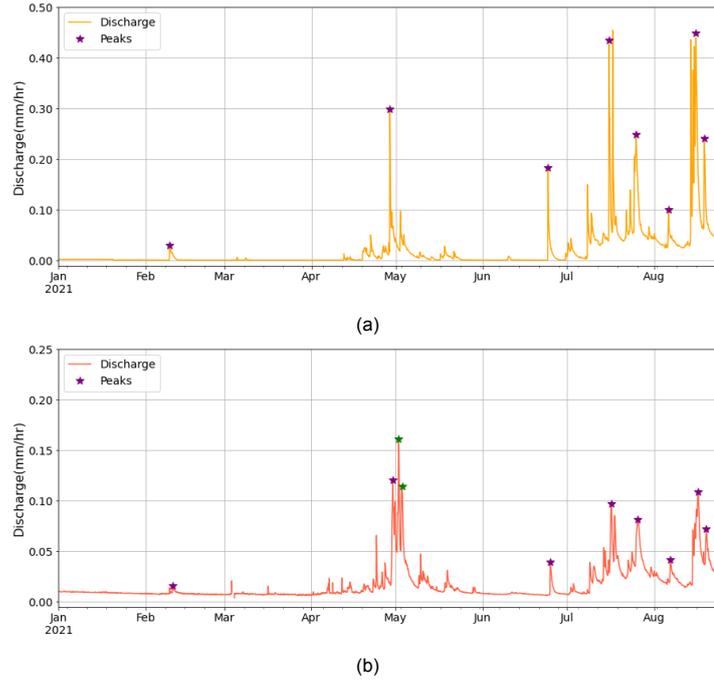


Figure 2.8: Peaks at (a) discharge station M43a and (b) discharge station M89, used to determine the wave celerity.

$$C = \frac{Q_{direct}}{P} \quad (2.2)$$

$C$  = runoff coefficient (-),  $Q_{direct}$  = direct flow (mm),  $P$  = precipitation amount (mm).

### 2.3.4. Recession coefficient

Periods with little or no precipitation cause a gradual depletion of discharge, constituting to the recession of discharge. This period of recession lasts until subsequent precipitation causes the discharge to increase again. The analysis of the recession curve of the discharge in a system yields information on the retention characteristics of that system and of groundwater storage and depletion. The shape of the recession curve primarily depends on the topography and geology of the catchment (Jain, 2011). The recession coefficient consists of roughly three components, the component for surface storage, inter flow storage and groundwater storage (Goel, 2011). For this analysis only the recession coefficient for the groundwater storage will be regarded, thus looking at the base flow of the system. The recession coefficient can be calculated using Equation 2.3.

$$Q_t = Q_0 e^{-t/k} \quad (2.3)$$

$Q_0$  = initial discharge at the start of the recession,  $Q_t$  = discharge after time  $t$ ,  $k$  = recession constant (days).

In Figure 2.7,  $Q_0$  would be the discharge at the point in time where the base flow and the direct flow become equal. This is the point on the recession curve of the hydrograph where the graph begins to decline and thus the base flow becomes more important to the total flow than the direct flow.

The recession constant was obtained by plotting the natural logarithmic discharge during the recession against time and determining the slope of the trendline, by rewriting Equation 2.3 to Equation 2.4. This was done for each year by looking at the last discharge peak of the wet season.

$$\ln(Q_t) = -t/k + \ln(Q_0) \quad (2.4)$$

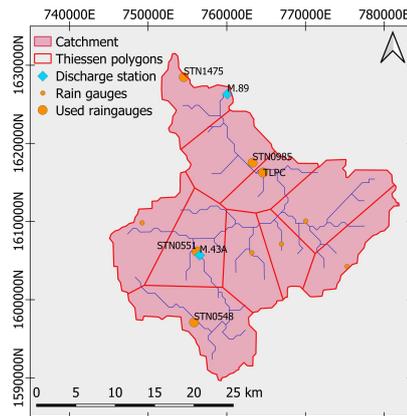


Figure 2.9: Used rain gauges for lag time determination of precipitation events based on Thiessen polygons and distance to downstream discharge station.

### 2.3.5. Individual event analysis

Besides looking at the general response of the catchment, also individual events have been compared and analysed. As flash floods are an event based hazard, insight on the catchments response to individual events is important to have.

Lag time and runoff coefficient were determined looking at direct runoff obtained from both discharge stations and comparing this to observed precipitation amount from rain gauges either in the same Thiessen polygon as the discharge station or within a range of 20 km upstream of the discharge station (see Figure 2.9). An overview of the distance between the rain gauges and the discharge station can be found in Table 2.5. Negative values indicate the rain gauge to be downstream of the discharge station and positive values indicate the rain gauge to be upstream of the discharge station.

Precipitation events with a cumulative rainfall amount of 40 mm or more were identified and the lag time and runoff coefficient were calculated for these events. This boundary condition was set as only large precipitation events are expected to result in flash floods. The end of a precipitation event was marked when less than 0.5 mm of rainfall had fallen in four consecutive hours. The start and end of a discharge peak were set at the points where the direct flow was smaller than or equal to 5% of the total discharge.

	M43a		M89		
	STN0551	STN0548	STN1475	STN0985	TLPC
<b>Shortest distance (km)</b>	-0.6	8.7	5.9	9.3	11
<b>River trajectory distance (km)</b>	-0.8	16.1	-	11.3	13.3

Table 2.5: Distance between discharge stations and rain gauges. Negative distance means the rain gauge is downstream of the discharge station and positive distance means the rain gauges is upstream of the discharge station.

#### Lag time

The lag time is the response time between a precipitation event and the resulting increase in direct flow. With direct flow as a result of overland flow, a shorter lag time typically indicates the overland flow to be Hortonian whereas the discharge pulse of saturated overland flow is generally more delayed, resulting in a longer lag time.

For this study, the lag time has been defined as the time between the centre of mass of a precipitation event and the centre of mass of the corresponding peak in the direct flow, as illustrated in Figure 2.10. The centre of mass for the discharge and precipitation peaks was determined using Equation 2.5 (Sapač

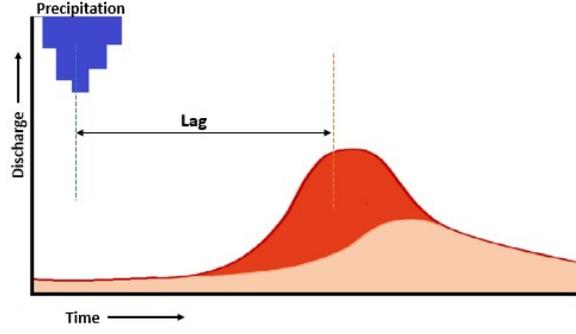


Figure 2.10: Illustration of lag time.

et al., 2020). Next the lag time was determined according to Equation 2.6.

$$M_x = \frac{\sum_{i=1}^n Y_i t_i}{\sum_{i=1}^n Y_i} \quad (2.5)$$

$Y_i$  = total amount of discharge or precipitation for  $i = 1, 2, \dots, n$  time periods of equal length,  $t$  = time of from the beginning of the event to halfway through period  $i$  (min). As the data set used has a resolution of 60 minutes,  $t_i$  is  $i \times 60 - 30$ .

$$T_{lag} = M_Q - M_P + \Delta t \quad (2.6)$$

$M_Q$  = centre of mass of the discharge,  $M_P$  = centre of mass of the precipitation,  $\Delta t$  = time difference between the start of the precipitation event and the start of the discharge peak.

### Distributed patterns of individual events

For the identified rainfall events, the spatial rainfall patterns have also been analysed. Spatially distributed precipitation maps have been created using data from the different rain gauges throughout the area. The rain gauge data has been interpolated using the inverse distance method. Patterns of events with similar lag times or runoff coefficients were compared to see if there was any correlation in the course of the precipitation throughout the area.

### 2.3.6. Water balance

For the years with adequate data on both precipitation and discharge an estimate of the total actual evapotranspiration was made using the simplified water balance in Equation 2.7.

$$\Delta S = P - Q - E_a \quad (2.7)$$

$\Delta S$  = storage change in the soil,  $P$  = total annual precipitation,  $Q$  = total discharge,  $E_a$  = yearly evapotranspiration, all in  $mm$ .

For the estimation of the annual evapotranspiration the ground water storage is neglected. Consequently  $\Delta S$  is assumed to be zero and the evapotranspiration can be determined using Equation 2.8.

$$E_a = P - Q \quad (2.8)$$

This was done for both discharge stations in combination with the rain gauge located in the same Thiessen polygon as the discharge station.

## 2.4. Results

In this section the results of the multiple characteristics determined as discussed above will be presented.

### 2.4.1. Base Flow

In Figure 2.11 the BFI values for the different seasons in each year are plotted, as well as the annual mean. The mean BFI values determined for the discharge at station M43a for the winter season, summer season and wet season are 0.85, 0.63 and 0.57 respectively, with an overall BFI of 0.63. Station M89 has an overall BFI of 0.74, with values of 0.89 for the winter season, 0.76 for the summer season and 0.63 for the wet season.

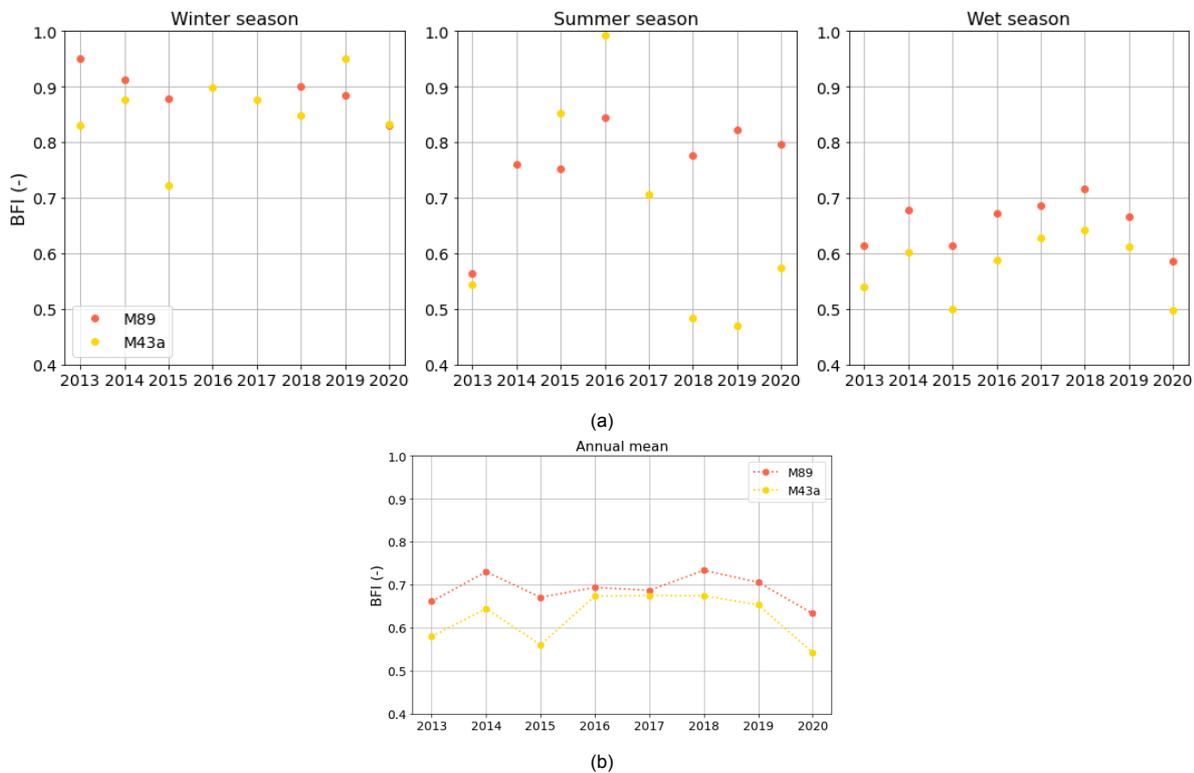


Figure 2.11: An overview of (a) the determined BFI during the three seasons for the different years and (b) the overall annual BFI.

An overview of the mean monthly BFI values can be found in Table 2.6. Months with 50% or more missing data were excluded. The monthly BFI values show the larger contribution of base flow in the downstream area compared to the upstream area, except for the months September, October and November, where the BFI for both stations is similar. Especially during the summer months and the start of the wet season, a big difference between upstream and downstream can be seen. Additionally, the monthly BFI values show a clear difference between months in the dry season and months in the wet season.

	Station	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
BFI	M43a	0.81	0.82	0.69	0.43	0.33	0.44	0.56	0.58	0.59	0.70	0.89	0.85
	M89	0.86	0.83	0.78	0.72	0.74	0.73	0.66	0.63	0.60	0.69	0.92	0.93

Table 2.6: Mean monthly BFI determined over the years 2012-2020.

### 2.4.2. Wave celerity

Analyses of the 8 peaks shown in Figure 2.8, resulted in the travel times as shown in Table 2.7. The travel time ranges between 18 and 25 hours, with a mean travel time of 21.1 hours. Over a course of 37 km, this corresponds with a wave celerity of 0.49 m/s.

### 2.4.3. Runoff

The annual and seasonal runoff coefficients determined for discharge measured at stations M43a and M89 can be found in Table 2.8. The runoff coefficient measured at station M43a is substantially higher

Peak	1	2	3	4	5	6	7	8
Travelling time (hr)	25	25	21	22	18	18	19	21

Table 2.7: Routing time of discharge peaks between station M43a and station M89.

than the runoff coefficient measured at station M89, especially for the winter and wet season. During the summer season the runoff coefficient of the two stations is quite similar.

Both stations show the highest runoff during the wet season and the lowest runoff during the summer season. At station M43a however, the differences in runoff between the seasons are significantly larger than at station M89. The runoff coefficients determined for station M89 are quite similar for both years, for station M43a the difference is larger, especially for the winter season.

	Winter season		Summer season		Wet season		Annual runoff	
	M43a	M89	M43a	M89	M43a	M89	M43a	M89
2018	0.12	0.01	0.03	0.02	0.40	0.08	0.25	0.05
2019	-	0.06	-	0.01	-	0.06	-	0.05
2020	0.27	-	0.03	-	0.49	-	0.34	-
Mean	0.20	0.04	0.03	0.02	0.45	0.07	0.30	0.05

Table 2.8: Seasonal and annual runoff coefficients for the different years.

#### 2.4.4. Recession coefficient

The results of the determined recession coefficient for the base flow at the two discharge stations can be found in Table 2.9. The recession coefficients at both stations are quite regular. At station M43a two outliers can be identified in 2016 and 2017. When excluding these outliers,  $k$  ranges between 17 days and 30 days, with a mean value of 23 days and a standard deviation of 5 days. When including 2016 and 2017, the range is between 17 and 49 days, with an average of 27 days and a standard deviation of 10 days. For station M89,  $k$  ranges between 6 and 19 days, with an average of 12 days and a standard deviation of 4 days.

The recession coefficients at station M43a are generally between 2 and 3.4 times higher than those at station M89, with two exceptions in 2015, where  $k$  is slightly higher at station M89, and 2020, where  $k$  is only 1.3 times higher at station M43a than at station M89. On average,  $k$  is 2.2 times higher at M43a than at M89.

	Station	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
$k$ (days)	M43a	19	19	30	27	17	49	38	27	28	21
	M89	6	9	10	13	19	-	-	12	11	17

Table 2.9: Base flow recession coefficient for the different years.

#### 2.4.5. Individual event analysis

Of the observed precipitation events meeting the set criteria, multiple events were unsuitable for determination of the lag time and runoff coefficient, as either no significant peak in direct runoff appeared, the timing of the direct runoff peak was before the start of the precipitation event, or the amount of direct runoff was higher than the total amount of precipitation. An overview of the number of suitable and unsuitable events for the different rain gauges is given in Table 2.10. An overview of the resulting suitable events, some of the event characteristics and their determined lag time and runoff coefficient can be found in Appendix D.

The lag times determined at station M43a range between 7.1 to 60.9 hours. When only looking at the ones determined from precipitation events measured at STN0551 the range is between 23.4 and 42.4 hours. Event 14 gave a negative lag time and is therefore no further considered. The lag times determined from the discharge measured at station M89 gave results between 12.3 and 86.3 hours. For rain gauge STN1475 the range is between 38.0 and 86.3 hours and for rain gauge TLPC the range is between 12.3 and 66.1 hours. For the runoff coefficients obtained from direct discharge at station

Gauge	Number of events	No peak	Wrong timing	Too high runoff	Suitable events
STN0551	15	5	6	1	3
STN0485	21	6	3	0	12
STN1475	7	2	2	0	3
STN0985	5	3	1	0	1
TLPC	14	4	2	0	8

Table 2.10: Number of precipitation events meeting the set criteria, but unsuitable for lag time and runoff determination.

M43a the values range between 0.04 and 0.76, with higher runoff coefficients from precipitation events measured at STN0551 than at STN0548. At station M89 the runoff coefficient ranges between 0.02 and 0.36, with the runoff coefficients determined from precipitation data of STN1475 ranging between 0.02 and 0.05 and runoff coefficients determined from data of TLPC ranging between 0.02 and 0.36.

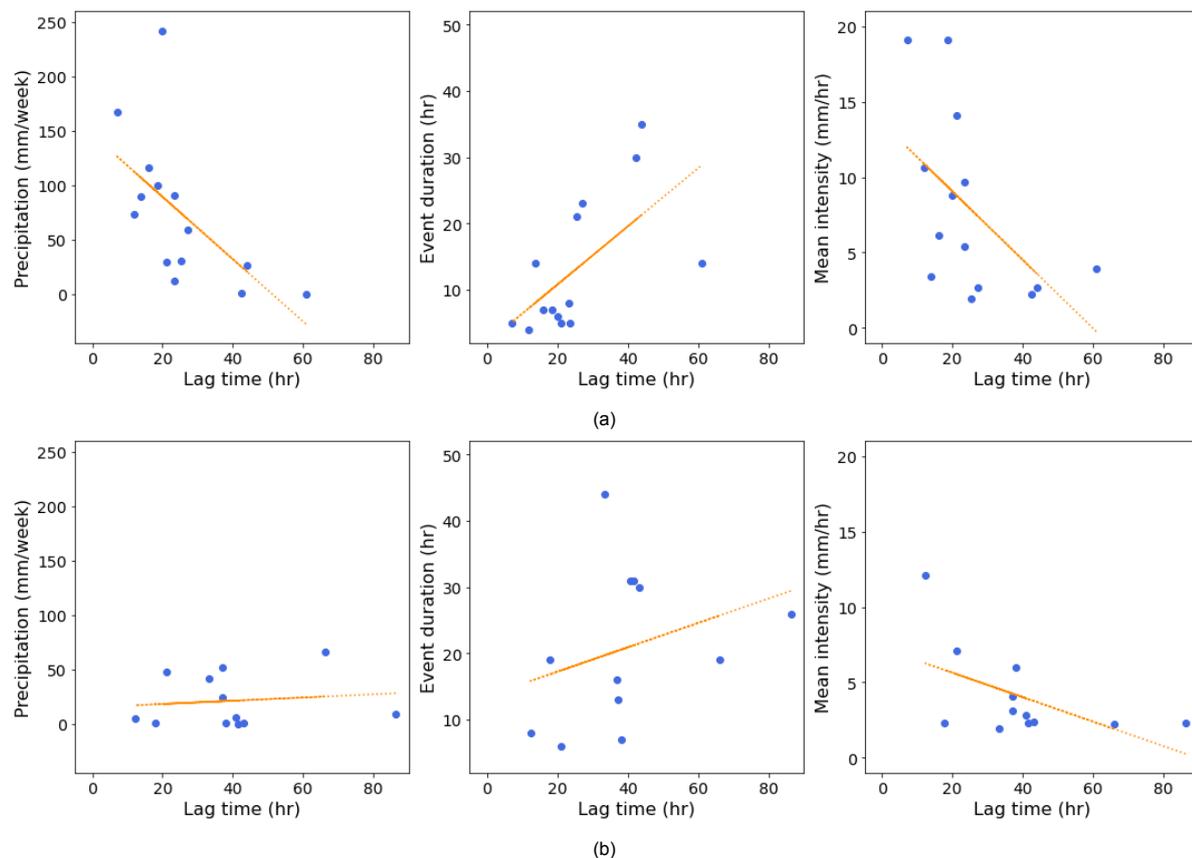


Figure 2.12: Correlation between lag time and (from left to right) antecedent conditions, the duration of the precipitation event and the mean intensity of the precipitation event for (a) station M43a and (b) station M89.

As the range of lag times and runoff coefficients is very wide, they both have been compared to multiple characteristics of the precipitation event, such as intensity, precipitation amount, duration and antecedent conditions, to see if there is any correlation. At station M43a there is a weak negative correlation between the lag time and antecedent conditions as well as mean intensity. Additionally, there is a slight positive correlation between the lag time and the duration of a precipitation event (Figure 2.12a). For station M89, the correlations between lag time and duration, and lag time and mean intensity can also be seen, however this correlation is less significant than for station M43a. The lag times determined with the discharge measured at station M89 do not seem to have any correlation with the antecedent conditions (Figure 2.12b). All correlation plots can be found in Appendix D. For the runoff coefficient, no distinct correlations were found, except for a weak positive correlation between the runoff coefficient measured at station M89 and the antecedent conditions, indicating that a higher amount of precipitation in the week prior to the event causes higher runoff.

The lacking correlation between lag time, runoff and event characteristics is also visible when examining the resulting direct runoff of three comparable events measured at gauge STN0548. The events have comparable hyetographs, as visualised in Figure 2.13a, and also the characteristics of the precipitation events are very alike. However, the resulting hydrographs of the direct runoff, as shown in Figure 2.13b, show large differences in magnitude and amount, consequently the resulting lag time and runoff coefficient of the three precipitation events is very different. An overview of this can be found in Table 2.11 together with the characteristics of the precipitation events.

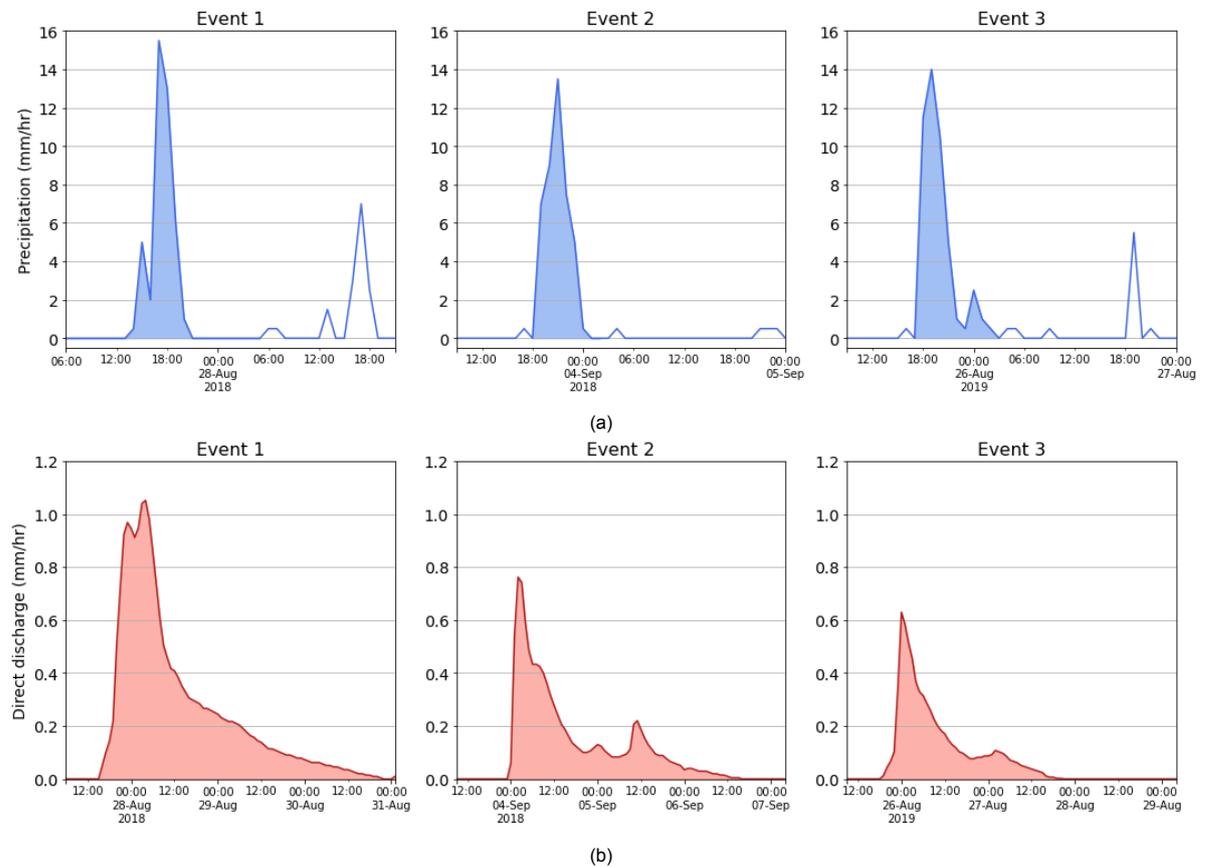


Figure 2.13: (a) The hyetograph of three similar precipitation events and (b) the resulting hydrograph of the direct discharge.

	Event 1	Event 2	Event 3
<b>Amount (mm)</b>	43	43	48
<b>Duration (hr)</b>	7	8	10
<b>Maximum intensity (mm/hr)</b>	15.5	13.5	14.0
<b>Mean intensity (mm/hr)</b>	6.2	5.4	4.2
<b>Antecedent precipitation (mm/week)</b>	116	91	89.5
<b>Lag time (hr)</b>	16	23.3	13.7
<b>Runoff coefficient (-)</b>	0.46	0.24	0.15

Table 2.11: Characteristics of the three similar precipitation events and the lag time and runoff coefficient determined from the resulting direct runoff peak.

Additionally, the timing of the measured precipitation events at the different rain gauges was compared. There were only two moments at which a total precipitation larger than 40 mm was measured at more than one station. On 02-07-2019 a high amount of precipitation was measured at all stations, and on 29-10-2020 a high amount of precipitation was measured at both STN0548 and TLPC, which are located in the upstream and downstream part of the catchment respectively. The lag times determined from the different rain gauges on 02-07-2019 give similar results ranging from 41 to 44 hours, although it must be considered that the distances to the discharge station are varying. When comparing the runoff

coefficients from 02-07-2019 the three downstream stations showed similar results ranging between 0.02 and 0.03, however the upstream stations showed significantly higher and more varying results (0.16 and 0.23).

### Distributed patterns of individual events

In Appendix E a brief explanation of the spatial pattern of the different precipitation events can be found. There did not seem to be any clear correlation between the pattern of the events as observed using the interpolated rain gauge data and lag time or runoff coefficient.

### 2.4.6. Water balance

In Table 2.12 the determined evaporation of the years 2018, 2019 and 2020 can be found for the two locations in the study area. For the year 2020 the total precipitation measured at gauge STN0551 was lower than the total annual discharge, indicating an incomplete data set, therefore the precipitation measured at station STN0548 was used instead.

	STN0551/ STN0548	M43		STN1475	M89	
	<i>P (mm)</i>	<i>Q (mm)</i>	<i>E<sub>a</sub> (mm)</i>	<i>P (mm)</i>	<i>Q (mm)</i>	<i>E<sub>a</sub> (mm)</i>
2018	1029.50	588.7	440.9	976.5	224.6	751.9
2019				712.5	132.4	580.1
2020	1183.5	774.8	424.8			

Table 2.12: Annual total evaporation as determined with the water balance.

The percentage of precipitation leaving the system through evaporation at the outlet of the study area for 2018 and 2019 was 77% and 81% respectively. However, at station M43a the amount of discharge is higher relative to the total precipitation, resulting in lower estimates of the total annual evaporation, between 34% and 43%.

### Add results for same rain gauge

## 2.5. Discussion

### Limitations and uncertainties

All calculations have been conducted based on observed precipitation and discharge data, which are both subject to uncertainties. Local sources confirm that rain gauge data in Thailand is of poor quality and has high levels of inaccuracy. After checking of the precipitation data using the DMC methods, some stations with unrealistically low annual precipitation amounts still gave high  $R^2$  values, resulting in high boundary conditions for the stations to pass the DMC check. However, even after applying these conditions, there were still stations of which the annual precipitation amount seemed too low. Besides the rain data, discharge data comes with uncertainties. Discharge values are obtained through conversion of water level measurements, using a rating table. For the observations at station M43a this conversion still had to be performed. Besides the possible errors in the measurements of the water levels in the river, the rating curve comes with uncertainties, resulting in a water level-discharge relation being inconsistent with the reality (Habert et al., 2016). Not only the observed data comes with uncertainties, the amount of data is quite limited. The distribution of rain gauges in the study area is sparse, and the gauges that are available often only have a few years of reliable data. The sparsity of the gauge network results in an inability to properly capture information such as precipitation amounts in a specific catchment, as the gauges only give point information.

Additionally, the limited length of data records available for the majority of the rain gauges increases the probability of part of the calculated catchment characteristics giving biased results. As only a couple of years are considered, this could lead to misrepresentation of the meteorology in the catchment, when the available years do not represent average years.

Besides the data uncertainties, the methods used to calculate the various characteristics also come with uncertainties. For the determination of the base flow, calibration on discharge data was performed to obtain satisfactory parameters for the Eckhardt algorithm. Different decisions made in the calibration

process would have resulted in either different parameters for the Eckhardt equation, or a different method to determine the base flow altogether. This would both have resulted in a different base flow and with that a different direct runoff, affecting the calculated BFI, runoff coefficient and recession coefficient. Additionally, simplified equations have been used, giving an oversimplified representation of reality. In subsection 2.3.6 only three processes are considered in the water balance, while in reality more processes contribute to this. Therefore results derived from this give merely an indication and are far from reality.

## Results

The two sub areas as divided in section 2.1 are also reflected in the catchment characteristics. In BFI, runoff coefficient, recession coefficient and the water balance differences between the two areas can be distinguished. The BFI values determined from discharge observed at station M43a show to be lower than those observed at station M89, indicating a smaller contribution of the base flow to the total flow. Especially during the wetter season this difference is apparent. Although findings in literature show varying conclusions on the relation between the BFI and topography (Beck et al., 2013; Mwakalila et al., 2002; Price, 2011), an explanation in the lower BFI determined for the upstream catchment can be found in the differing topography of the both sub areas, as topographic gradients affect the rate at which water moves down a slope, thus regulating whether the water becomes runoff or is retained in the soil (Price, 2011). As slopes are steeper in the upstream area, water will be drained more rapidly, resulting in a higher direct runoff and thus a lower contribution of the base flow, while in the downstream area the water has more time to infiltrate into the soil. Similar results were found for the runoff coefficient for both sub catchments, with significantly higher values for the runoff coefficient in the steeper, upstream area, which has also been confirmed by local sources. This can again be explained by the steep terrain in that area and possibly differences in soil characteristics.

Also the recession coefficient shows clear differences between both areas, with a more rapid depletion of the base flow in the downstream area than in the upstream area, although it being quite rapid in both. Recession of the base flow is often related to the transmissivity of the soil and its storage capacity (Hall, 1968), which could be an explanation for the rapid depletion. The reason for the slower depletion of base flow in the National Park area could be due to larger storage of the groundwater or differences in soil properties, resulting in a lower infiltration capacity in the provincial area but with that also a slower lateral groundwater flow, producing less base flow and a slower depletion of the groundwater. Another explanation could be water being stored in and around the river bed contributing to the base flow. As there is abundant vegetation in the upstream area, this could facilitate storage of the water and thus a delay in the runoff.

Besides differences between the two study areas, both the determined BFI values and runoff coefficient show clear differences between the different seasons. BFI values are highest in winter season, where little to no precipitation occurs, so streamflow mainly results from base flow. Values are lowest during the wet season, as abundant precipitation results in more direct runoff and thus a lower contribution of base flow (Padiyedath et al., 2017). The same explanation can be used for the high runoff coefficient in the wet season compared to the dry season and summer season.

When looking at the individual events examined, wide ranges in both lag time and runoff coefficient are found. The big differences in lag time and runoff coefficient show the heterogeneity of both the study area and the precipitation patterns within the study area, making it impossible to describe the catchment using either of the two. Especially the spatial differences in the precipitation within the study area become apparent. Of the selected precipitation events, only two days can be identified during which multiple rain gauges observed a precipitation event meeting the set criteria, showing the heterogeneous nature of the precipitation events and the fact that most events occur locally.

This can also be seen in the different hydrological response to the three similar precipitation events examined, probably due to deviating precipitation patterns upstream of the discharge station which are not captured by the rain gauge. Additionally, the high number of situations where the start of the discharge peak occurs before the start of the precipitation event indicates the spatial variation in precipitation, indicating the occurrence of mostly local and convective storms. As these local events are difficult to capture with a sparse rain gauge network as the one in the study area, properly characterising the catchment is challenging. The heterogeneity of the precipitation events in the study area thus result

in a heterogeneous hydrological response.

Just as the previous discussed catchment characteristics also the simplified water balance shows different results for both sub catchments. Observations at discharge station M89 show roughly 20% of the precipitation to leave the system through discharge, meaning roughly 80% leaves through either evaporation, percolation or is stored in the ground. For station M43a, 60% of the precipitation leaves through discharge, indicating a stronger precipitation-discharge relation in the upstream area. From different sources of literature annual actual evaporation values between 66.5% and 81.5% of the total annual precipitation were found, depending on year, location and estimation method (Hirota, 2001; Jha et al., 2000; Masaki et al., 2011; Zheng et al., 2019), corresponding to the estimated actual evaporation for station M89. The estimated evaporation at station M43a is far off from findings in literature, so the accuracy is plausible, but lower evaporation values in the National Park could be explained by the steeper terrain, causing the water to leave the system more quickly. The higher fraction of precipitation leaving as discharge in the upstream area shows the clear difference in precipitation-discharge relationship between the two sub catchments.

In Figure 2.14 the relation between the water balance, BFI and runoff coefficient for both locations is illustrated. The determined BFI at station M89 showed that of the total discharge, approximately 25% is due to direct flow and the other 75% due to base flow, which would mean 5% of the total precipitation reaches the stream through direct runoff, while 15% reaches the stream through base flow, which corresponds to the determined runoff coefficient at this location. At station M43a according to the overall runoff coefficient around 30% of the annual precipitation turns into direct flow, which would be 50% of the total discharge, meaning also 50% of the discharge is from base flow. According to the BFI however, a little over 60% of the total discharge is a result of base flow.

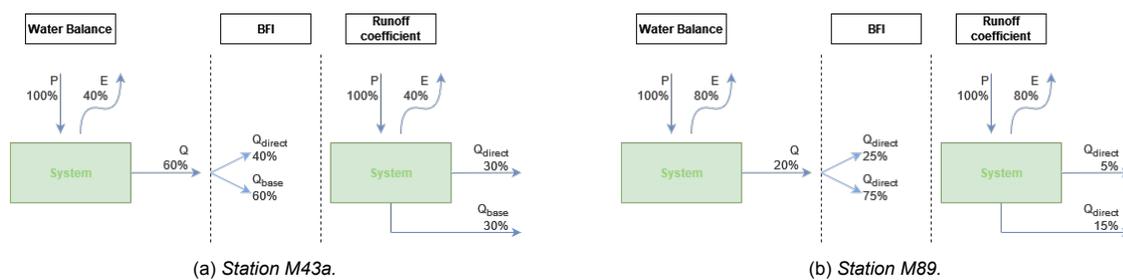


Figure 2.14: The water balance according to the different hydrological characteristics determined at the two different discharge stations.

This difference could be a result of incomplete or inaccurate data, but could also be attributed to the chosen method to determine the base flow, resulting in too high estimations of the base flow. A different method, or different values for the calibration parameters could have shown a lower base flow, thus a lower BFI and with that a better correspondence to the calculated runoff coefficient.



# 3

## Modelling

In this chapter the modelling aspects of this research will be discussed. To start with, a description of the model used in this research, the Calabrian Erosion (CALEROS) model, will be given. This includes some background information on the model, as well as information on the required input data and the used data sets. Furthermore, the most important output sets for this research will be discussed. Next, an elaboration on the used methods will be given, after which the results will be presented and discussed.

### 3.1. Model Description

The CALEROS model (Feiken, 2014) is a physically-based, conceptual, distributed model, coupling hydrology to slope stability in a dynamic and process based manner. It runs within the PCRaster software, which is software for environmental modelling developed by the University of Utrecht. CALEROS has been built as an erosion-deposition model for landscape development. The three main components on which the model is based are geology, climate and human activity.

The required input of the model consists of the initial state of various parameters describing the system and climatological time series. The model is conditioned by initial states that describe the topography in the form of a digital elevation model (DEM), the underlying lithology and the distribution of soil materials, including unsaturated and saturated water content and vegetation. CALEROS output consists of a series of maps and tables on basic characteristics of the area.

In the model, both spatial and temporal resolution is considered. During a model run, each time step consists of a series of multiple calculations of which a simplified schematisation is illustrated in Figure 3.1. First, the initial conditions, such as the soil moisture content and water levels in the soil are read. Secondly, the upper and lower boundary conditions are evaluated, consisting of the input precipitation, the actual evaporation calculated by the model and the base losses determined by the model. Thirdly, the vertical fluxes are calculated and the storage is updated. Next, the lateral fluxes are evaluated and the storage is updated again. This updated storage sets the new initial conditions for the next time step. The boundary conditions at the top are determined by precipitation and evaporation. The boundary conditions at the bottom are determined by percolation into the underlying bedrock reservoir.

The spatial resolution of the model used for this research is 100x100 meter. CALEROS gives the possibility for finer spatial resolutions, however this has not been done for computational convenience. The temporal resolution of the model is daily scale.

#### 3.1.1. Data input

CALEROS requires multiple sets of input, which can be divided into the categories meteorology, topography, soil, surface hydrology, soil hydrology and vegetation stand. In this section a brief motivation will be given on how and why different input values were chosen. In Appendix F a more elaborate explanation on the different input parameters is given as well as a more detailed motivation of the chosen values.

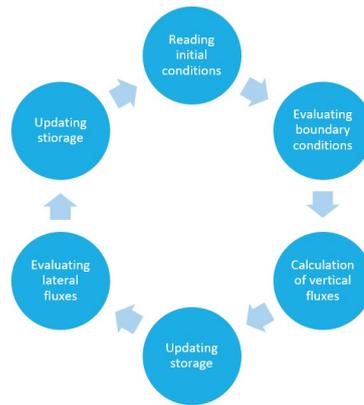


Figure 3.1: Schematisation of the calculations done for each time step.

### Meteorology

The meteorological input consists of precipitation and temperature data, both with a daily temporal resolution. The precipitation data includes information on the amount of precipitation ( $kg/m^2$ ), the maximum intensity of the rain event ( $kg/m^2/s$ ), the duration of the event ( $s$ ) and the timing of the peak ( $s$ ). The temperature data gives information of the temperature in Kelvin.

In the study area, there are multiple rain gauges providing information on precipitation amount and intensity (see section 2.2), either with a quarterly or hourly resolution. These data sets have been re-sampled to daily precipitation to fit the temporal resolution of the model. The hourly data sets have been used to find the maximum intensity of the rain event for each day. With the maximum intensity and the total daily precipitation amount, an estimation of the rainfall duration was made. The timing of the rainfall peak was assumed to be halfway each event.

For the temperature data, climatological information from the TMD weather station Pakchong Agromet was used. This weather station is located near the study area (Figure 2.5) and provides information on monthly minimum, maximum and mean temperature obtained in the period between 2006 and 2021. In this research the monthly temperature means have been used as input, and were assumed to be constant for each year. In Table 3.1 these monthly means can be seen, together with the average monthly minimum and maximum measured at the climatological station.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<b>Minimum</b>	18.0	19.6	21.5	22.9	23.7	24.1	23.8	23.5	22.7	22.0	20.7	18.5
<b>Maximum</b>	29.9	32.1	33.5	33.8	32.9	32.2	31.2	30.7	30.3	30.1	29.8	29.0
<b>Mean</b>	23.5	25.4	27.0	27.6	27.6	27.5	26.9	26.5	25.8	25.3	24.7	23.3

Table 3.1: Minimum, maximum and mean monthly temperatures ( $^{\circ}C$ ) from Pakchong Agromet weather station determined over the years 2006-2021.

As there is only one climatological station near the study area, for this research the temperature is assumed to be spatially homogeneous over the entire catchment. For the precipitation, both uniform and spatially distributed data has been used.

### Topography

The topography of the catchment is characterized by a DEM, with a spatial resolution of 100 m (Figure 2.2a). The data was obtained from the NASA SRTM (NASA JPL, 2013) with an initial resolution of 30m. This resolution was lowered for computational convenience.

### Soil

The CALEROS model assumes a soil distribution which is steady state. Information is required on the content of silt, sand, clay, organic matter and coarse fragments. Soil data has been obtained from the International Soil Reference and Information Centre (ISRIC) World Soil Information database (Poggio et al., 2021) and was converted to match the required units of the model. The ratio of the

soil types obtained through the ISRIC database did not match results found in literature (Trelo-ges and Sriboonlue, 2002) or results obtained from soil samples collected in the National Park at telemetering station Khaoyai 1 (Figure 2.5), therefore the maps were scaled to better match these results. The content of coarse fragments and organic matter has been set to zero, as no data was found on this.

### Surface hydrology

For this research, the ponded water mass was set to zero throughout the entire catchment, as there are no significant lakes or other water bodies in the study area.

### Soil hydrology

The soil hydrology includes information on the water content in the saturated as well as the unsaturated zone. Additionally, the saturated hydraulic conductivity of the ground layer below the bedrock layer is specified.

The water content in the saturated zone is set to zero for this research, as groundwater wells throughout the study area (Figure 2.5) show the water depth to be significantly deeper (Table F.2) than the base of the soil layer. To obtain the initial value for the unsaturated water content multiple-year test runs have been performed, as will be further elaborated in subsection 3.2.2. The results were compared with the field observations from the telemetering station installed in the National Park, giving information on the volumetric water content at different depths, being around 30%. The saturated conductivity of the bedrock layer is specified as a fraction of the hydraulic conductivity of the top layer and is obtained through an iterative process of multiple test runs.

### Vegetation stand

In the section on vegetation stand, multiple aspects of the vegetation are specified being initial interception, soil albedo, canopy albedo, canopy height, cover fraction, Leave Area Index (LAI), soil cropfactor, cropfactor, interception capacity, root depth, saturated hydraulic conductivity, root cohesion and surcharge. For some vegetation parameters constants have been used, while others have a spatially distributed input.

Distributed data of the vegetation stand was obtained through various sources. Potapov et al., 2021, developed a global forest canopy height map with the Global Land Analysis and Discovery (GLAD), which was used as input for the model. The data for both the cover fraction and the LAI was obtained from data sets from the Copernicus Global Land Service (Fuster et al., 2020). From the LAI map, the interception capacity was calculated according to Equation 3.1 (De Roo et al., 1996).

$$S_{max} = 0.935 + 0.498 * A_{leaveindex} - 0.00575 * A_{leaveindex}^2 \quad (3.1)$$

$S_{max}$  = interception capacity (mm),  $A_{leaveindex}$  = LAI.

Root depth has been determined based on the land cover types in the study area (see Figure 2.2b). For water bodies, herbaceous wetlands and build up areas the root depth was assumed to be zero. Croplands, shrub lands (Silva and Rego, 2003) and herbaceous vegetation were assumed to have a root depth of 30 cm and for forest covers the root depth was set to be equal to the soil thickness.

Initial interception was set to zero, as it can be assumed that with the high temperatures, intercepted water will rapidly evaporate. The canopy albedo was set to 0.14, as Pinker et al., 1980, found this to be the albedo of a tropical evergreen forest at the Khorat Plateau, located around 200 km northeast of the study area. The soil albedo was set to 0.2, after findings of An et al., 2017, and Dobos, 2003, on the albedo of different soil types. For the cropfactor a value of 1.00 has been used, based on the general crop factors determined by the FAO and for the soil cropfactor a value of 0.2 was chosen (Allen et al., 1998). Hydraulic saturated conductivity, root cohesion and surcharge are aspects determined by the model using pedotranfer-functions, based on the soil content of the different soil classes in the system.

### 3.1.2. Data output

There are a large number of variables that can be written as output in this model. An overview of all possible output variable can be found in Appendix H. In this chapter only the most relevant output variables for this research will be discussed, consisting of the different water fluxes in the catchment. In

Figure 3.2 simplified schematisations of the different water fluxes in the model are shown. As illustrated in the figure, the upper boundary conditions consist of precipitation entering the system and evaporative fluxes leaving the system. The lower boundary conditions consist of deep percolation leaving the system. Between the upper and lower boundary conditions the water is moved through various fluxes in and on the soil.

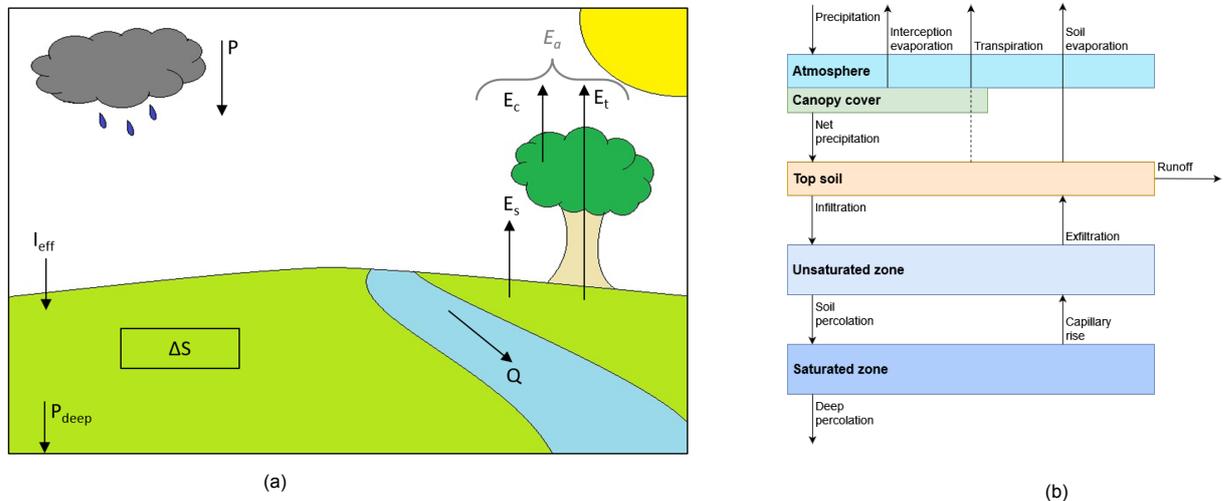


Figure 3.2: Simplified representation of the water fluxes within the catchment determined by the model. (a) Illustrated the boundary fluxes used to check the water balances and (b) schematises the same, but with more detail in the soil water fluxes.

### Evaporative fluxes

One of the boundary conditions of the model is evaporation. The actual evaporation is based on the determined potential evaporation.

**Potential evaporation** The potential evaporation in the model is determined using the Hamon equation, which is a convenient method to use when only temperature data is available. It is often used in hydrological studies as limited data is required and it has shown to be closely correlated to the actual evaporation (Lu et al., 2005; Rao et al., 2011).

The potential evaporation estimated with this method is based on air temperature and day length, according to Equation 3.2.

$$E_{hamon} = 0.1651 * L_d * \rho_{sat} * k \quad (3.2)$$

$E_{hamon}$  = potential evaporation (mm/day),  $L_d$  = amount of sunshine hours from sunrise to sunset (fraction of 12 hours),  $\rho_{sat}$  = saturated vapour density at the daily mean temperature ( $g/m^3$ ),  $k$  = calibration factor (-).

The potential evaporation determined by the model was checked by calculating the monthly potential evaporation using the Penman-Monteith equation, with data obtained from the climatological TMD station Pakchong Agromat (Figure 2.5). The potential evaporation determined with Penman-Monteith gave higher values than that determined by the model, so the calibration factor,  $k$ , in the model was set to 1.4 to better match the calculated potential evaporation.

**Actual evaporation** The actual evaporation is determined based on the combination of canopy or interception evaporation, soil evaporation and canopy transpiration. Canopy evaporation is equal to the interception stored on the canopy. The soil evaporation and canopy transpiration are determined by multiplying the potential evaporation with the soil cropfactor and cropfactor respectively. After this, a scale factor is applied to limit the maximum evaporation rate. This limiting rate can be e.g. the atmospheric demand or the allowable evaporation from the soil.

### Vertical groundwater fluxes

In the unsaturated zone of the soil, water flows vertically through percolation as a result of gravitational forces. This flow is controlled by the saturated hydraulic conductivity of the soil and the bedrock layer. The saturated hydraulic conductivity ( $k_{sat}$ , in  $m/d$ ) is estimated on the basis of soil properties using the relationship of Brutsaert, in Equation 3.3.

$$k_{sat} = a * \frac{(\theta_{sat} - \theta_{res}) * \lambda_{bc}^2}{\psi_{sat_{bc}}^2 * ((\lambda_{bc} + 1) * (\lambda_{bc} + 2))} \quad (3.3)$$

$\theta_{sat}$  = saturated volumetric water content ( $m^3/m^3$ ),  $\theta_{res}$  = residual volumetric water content ( $m^3/m^3$ ),  $\psi_{sat_{bc}}$  = air entry value of the Brooks & Corey soil water retention curve ( $m$ ),  $\lambda_{bc}$  = pore size distribution of the Brooks & Corey soil water retention curve (-),  $a$  = optional constant (default value is 3.01124)

A higher value of  $k_{sat}$  allows more water to infiltrate into the subsurface and thus reduces the runoff in the system. The saturated hydraulic conductivity of the bedrock layer,  $k_{sat_{bc}}$ , is a fraction of the saturated hydraulic conductivity of the soil. When set to zero, no water leaves the system through deep percolation, indicating a closed system.

### Lateral fluxes

Between the cells, also lateral exchange of water occurs, either as surface water flux (runoff) or groundwater flux.

**Runoff** The runoff in this model is determined using the PCRaster function *pcr.accuthresholdflux*. This function describes the accumulation of water masses in a drainage network and limits its transport by thresholds. Transport of the water will only occur when a particular threshold of losses has been attained. When the threshold is not reached, the water will be stored. For overland flow, the threshold is equal to the storage capacity of the soil, so that flow will only develop once the soil is saturated.

**Groundwater flow** Lateral groundwater flow occurs across the saturated zone due to differences in total head. The lateral flow is dependant on the saturated hydraulic conductivity of the soil, the water level and the width of a cell. It is calculated using Darcy's law given in Equation 3.4.

$$Q_{gw} = -k_{sat} \frac{\Delta h}{\Delta x} \quad (3.4)$$

$Q_{gw}$  = Lateral groundwater flow ( $m/d$ ),  $k_{sat}$  = saturated hydraulic conductivity ( $m^3/d$ ),  $\Delta h$  = head difference ( $m$ ),  $\Delta x$  = distance ( $m$ )

## 3.2. Method modelling

In this section the different modelling methods will be discussed. All methods are based on iterative processes. First a sensitivity analysis was conducted in order to find out which parameters are most critical for calibration of the model. Next, tuning of the model parameters was performed, using modelling strategy 1 and 2 as schematised in Figure 3.3. Both uniform calibration and spatial calibration was performed, where the spatial calibration was based on the HRU as illustrated in Figure 2.6a. Lastly, model runs were conducted according to modelling strategy 3 and 4 in Figure 3.3. Due to lack of time, no calibration was performed for strategy 3 and 4, so calibration results obtained from strategy 2 were used to run the model.

### 3.2.1. Sensitivity Analysis

Parameter sensitivity analysis is an important step when working with a model. A sensitivity analysis identifies which input parameters have a significant impact on the different output parameters and is crucial for reduction of the number of parameters required in model calibration and validation. Sensitivity analyses can roughly be divided into two different types, the local and the global analysis (Xu et al., 2016). In local sensitivity analyses the response of a model is obtained by changing the parameters one at a time, while keeping the other parameters at a fixed value. This method is simple to apply,

		Space	
		Uniform	Spatial
Time	Static	1	3
	Dynamic	2	4

Figure 3.3: Overview of the four different types of modelling strategies performed based on the spatial and temporal variation of the precipitation input.

however, there are some limitations of linearity and normality assumptions, and local variations (Xu et al., 2016). To overcome these limitations, global sensitivity analysis methods can be applied, which are able to provide relevant information of the sensitivity of model outputs to the whole range of model parameters (Xu et al., 2016), however, these are more complex and often time consuming. For this research a local sensitivity analysis has been conducted using the sensitivity index as described by Lenhart et al., 2002.

### Sensitivity Index

Mathematically, the dependency of output variable  $y$  on input parameter  $x$  can be approximated with Equation 3.5, given that  $y_0$  is the modelled output of a certain variable with an initial value of  $x_0$ , which is varied by both increasing and decreasing it with  $\Delta x$ , resulting in modelled values  $y_1$  and  $y_2$  as schematically illustrated in Figure 3.4 (Lenhart et al., 2002). When normalising this equation, the expression for the sensitivity index is obtained according to Equation 3.6.

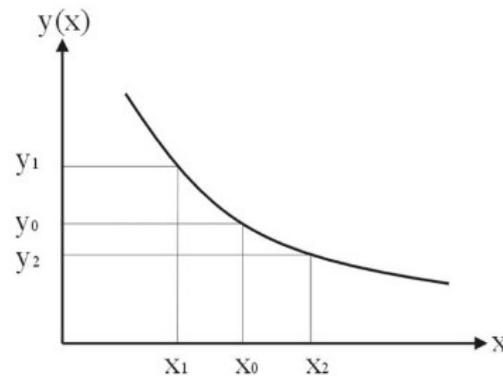


Figure 3.4: Schematic of the relation between an output variable  $y$  and an input parameter  $x$  (Lenhart et al., 2002).

$$I' = \frac{y_2 - y_1}{2\Delta x} \quad (3.5)$$

$$I = \frac{(y_2 - y_1)/y_0}{2\Delta x/x_0} \quad (3.6)$$

Four sensitivity classes can be distinguished based on the sensitivity index, indicating how sensitive the output variable is to the input parameter. An overview of these classes is given in Table 3.2. The sign of the sensitivity index shows whether the output variable reacts positively or negatively proportional to the change in input parameter.

Class	Index	Sensitivity
I	$0.00 \leq     < 0.05$	Small to negligible
II	$0.05 \leq     < 0.20$	Medium
III	$0.20 \leq     < 1.00$	High
IV	$    \geq 1.00$	Very high

Table 3.2: *Sensitivity classes.*

The sensitivity index was determined for all different combinations of  $x$  being one of the input parameters of the model and  $y$  being one of the fluxes in the water balances as will be discussed in subsection 3.2.2. For  $x_0$  either the constant value chosen in subsection 3.1.1 or the mean of the input map as obtained in subsection 3.1.1 was used. For some parameters this was zero, so another value was chosen as  $x_0$ .  $\Delta x$  was taken as 50% of  $x_0$ . The sensitivity runs were performed using precipitation input data of station STN0551 for the year 2018.

### 3.2.2. Uniform and static precipitation input

In modelling strategy 1 uniform and static precipitation input is used, meaning precipitation is constant over both space and time. This was done to test the model performance, do a first check of the boundary conditions and the water balance, and get a first rough estimation of the catchment behaviour in the model. This was done using two methods, the first was using complete static precipitation and the second using partial static precipitation, where only a differentiation was made between dry and wet season, but for both seasons the precipitation was static.

#### Complete static precipitation input

Runs were performed with spatial and temporal homogeneous precipitation of both 4 mm/d and 10 mm/d. For both test scenarios, first a run of 5 years was performed to get the initial conditions of the unsaturated and saturated groundwater content. Next, a 1 year run was performed to check the water balance, using the obtained initial conditions.

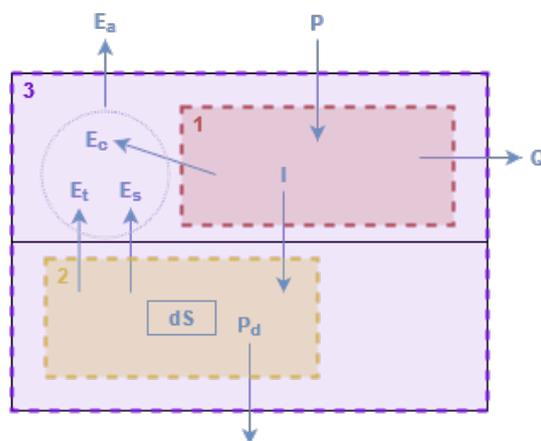


Figure 3.5: *Schematic illustration of the boundary conditions of the different water balances and their corresponding in- and outfluxes.*

With the output variables modelled by the CALEROS model, three different variations of the water balance can be checked, depending on the boundary conditions. These variations are schematically illustrated in Figure 3.5. The water balances corresponding to boxes 1, 2 and 3 are given in Equation 3.7, Equation 3.8 and Equation 3.9 respectively.

$$P = I + E_c + Q \quad (3.7)$$

$$\Delta S = I - E_s - E_t - P_d \quad (3.8)$$

$$P = E_a + Q + P_d + \Delta S \quad (3.9)$$

$P$  = precipitation,  $I$  = infiltration,  $E_c$  = canopy evaporation,  $Q$  = discharge,  $\Delta S$  = change in groundwater storage,  $E_s$  = soil evaporation,  $E_t$  = canopy transpiration,  $P_d$  = deep percolation, all in  $mm$

The results from Equation 3.9 were compared to the water balance estimations in subsection 2.4.6 to do a first tuning of the boundary conditions of the model.

### Partial static precipitation input

For the partial static precipitation input, a model run was performed using test precipitation data with a rainfall of  $0 \text{ mm/d}$  in the dry season (nov - apr) and of  $10 \text{ mm/d}$  during the wet season (may - oct) to check the model performance for a seasonal precipitation. This test run was also used to get initial conditions for the water content in the saturated and unsaturated zone. The resulting initial conditions were compared with the data from the telemetering station in Khao Yai National Park to check the plausibility.

With the partial static precipitation input, the recession coefficient of the discharge generated by the model was checked. Of the resulting discharge, the base flow was determined using the software discussed in subsection 2.3.1. Next the recession coefficient was calculated according to the method explained in subsection 2.3.4 and the resulting recession coefficients were compared to the ones obtained in subsection 2.4.4. Values for saturated hydraulic conductivity and soil thickness were adjusted in a trial-and-error process, as these parameters were expected to have the biggest impact on the recession coefficient.

### 3.2.3. Uniform and dynamic precipitation input

Modelling strategy 2 uses uniform but dynamic precipitation input, so constant in space but not in time. This was done using precipitation data from one rain gauge and assuming it to fall homogeneously over the study area. Two methods were used within this calibration strategy, which are event based calibration and Flow Duration Curve (FDC) calibration. After these two calibration strategies, the model was again run using observed precipitation data from a single rain gauge and the resulting modelled discharge was compared to the observed discharge at both stations.

#### Event based calibration

Event based calibration was performed to calibrate the model to event based discharge peaks, as this is most relevant for prediction of flash floods, which are also event based. To do this, isolated discharge peaks of  $50 \text{ m}^3/\text{s}$  and higher were identified from the observed discharge data (re-sampled to daily values) at both M43a and M89. This resulted in four discharge peak events, which can be seen in Figure 3.6

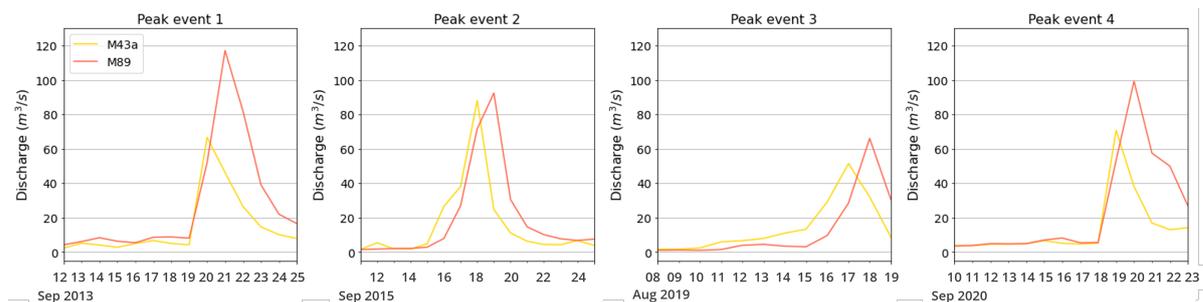


Figure 3.6: Identified isolated discharge peaks with a peak larger than  $50 \text{ m}^3/\text{s}$ .

For event 1 and 2 only precipitation data from rain gauge TLPC, downstream in the study area, was available. For peak events 3 and 4 also precipitation data from rain gauges STN0551 and STN00548 was available. However, gauge STN0551 did not show any significant precipitation prior to either of the two events and STN0548 did not show any significant precipitation peak prior to event 4. For that reason peak event 3 was chosen to use for calibration, in combination with precipitation data from gauge STN0548.

The model was run over a period spanning from a week prior to the peak event until the end of the event. The parameter set obtained after the recession check was adjusted in a trial-and-error process in order to properly simulate the timing, magnitude and propagation of the peaks using visual calibration. Visual calibration was applied, meaning modelled and observed discharge were compared by looking at the result and not by mathematical error equations. After tuning of the parameters, the model was run for a longer period of time in order to validate the optimised parameters set.

Multiple parameters were changed in order to capture the observed discharge with the model. From the sensitivity analysis it had been decided to tune the saturated hydraulic conductivity, initial unsaturated water content, cropfactor and layer thickness. Tuning of the parameters was an iterative process.

### Flow Duration Curve

FDC calibration was used to calibrate the modelled discharge to the overall distribution of the discharge over a year. An FDC is a plot showing the percent of time the discharge is equal to or exceeded by a certain value during a given period of time. In Figure 3.7 the FDC of both discharge stations in the year 2018 can be seen. The first 10% of the FDC shows the highest peaks, after which the decrease of the slope becomes more gradual. For this reason, calibration was only applied to the first 10% of the FDC, as this is where the flash flood peaks are captured.

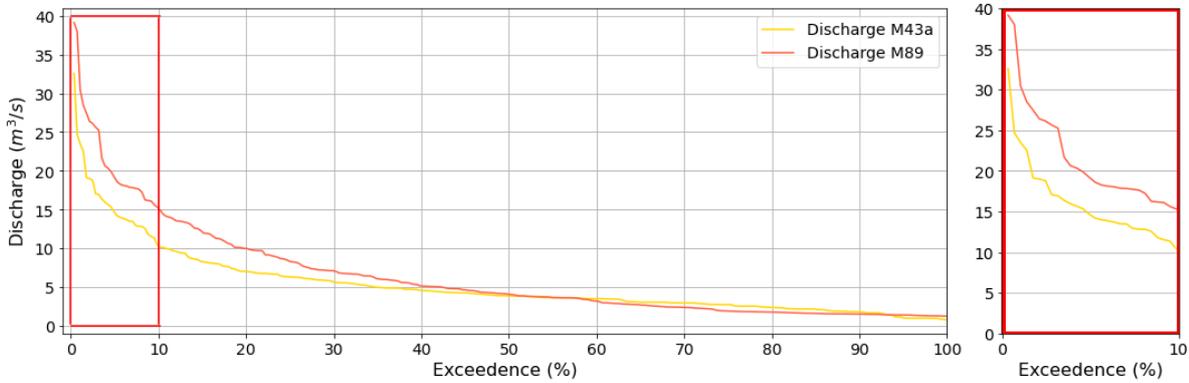


Figure 3.7: Flow Duration Curve of the discharge observed in 2018, with only the first 10% shown in the right plot.

The model was run for the year 2018, using precipitation data from rain gauge STN0551. Rain gauge STN0551 is located closest to discharge station M43a, which is most important to properly model due to a higher risk of flash flood. Additionally, the gauge data for the year 2018 shows to be most reliable according to the  $R^2$  values determined in section 2.2.

Initially the model was run using the parameter sets obtained after the recession check and event based calibration. Next, tuning of different parameters was performed in order to lower the error of the fit of the FDC of the modelled discharge compared to the FDC of the observed discharge. Three methods were used to assess the error between the observed and modelled discharge, which are the Root Mean Square Error (RMSE) (Equation 3.10), Nash-Sutcliffe Efficiency (NSE) (Equation 3.11) and Kling-Gupta Efficiency (KGE) (Equation 3.12).

$$N_{RMSE} = \sqrt{\frac{\sum_{i=1}^n (Q_{obs_i} - Q_{mod_i})^2}{n}} \quad (3.10)$$

$$N_{NSE} = 1 - \frac{\sum_{i=1}^n (Q_{obs_i} - Q_{mod_i})^2}{\sum_{i=1}^n (Q_{obs_i} - \bar{Q}_{obs})^2} \quad (3.11)$$

$$N_{KGE} = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{mod}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\bar{Q}_{mod}}{\bar{Q}_{obs}} - 1\right)^2} \quad (3.12)$$

$Q_{obs}$  = observed discharge,  $Q_{mod}$  = modelled discharge,  $\overline{Q_{obs}}$  = mean of observed discharge,  $r$  = Pearson coefficient (Equation 3.13),  $\sigma_{mod}$  = standard deviation of the modelled discharge,  $\sigma_{obs}$  = standard deviation of the observed discharge,  $\overline{Q_{mod}}$  = mean of modelled discharge

$$r = \frac{\sum_{i=1}^n (Q_{obs_i} - \overline{Q_{obs}})(Q_{mod_i} - \overline{Q_{mod}})}{\sqrt{\sum_{i=1}^n (Q_{obs_i} - \overline{Q_{obs}})^2} \sqrt{\sum_{i=1}^n (Q_{mod_i} - \overline{Q_{mod}})^2}} \quad (3.13)$$

For the RMSE an error closer to zero generally means a better fit of the modelled data to the observed data. However, it is dependant on the range of the data set, so when comparing two different years, the year with a higher RMSE value might still have a better fit. It is therefore not a very convenient method to compare runs with different observed data sets, but it is useful to compare different parameter sets for the same year. For both the NSE and the KGE, the closer the efficiency is to 1, the more accurate the model is. When the efficiency is below 0, the mean of the observed discharge would be a better indicator than the modelled discharge.

### Annual model run

From the different tuning methods, multiple optimal parameter sets were obtained. The model was run with all sets for the year 2018, using precipitation data from rain gauge STN0551. For each run the error was calculated using the three methods mentioned above to determine which parameter set gave the most optimal fit. The observed discharge in 2018 at both location can be seen in Figure 3.8.

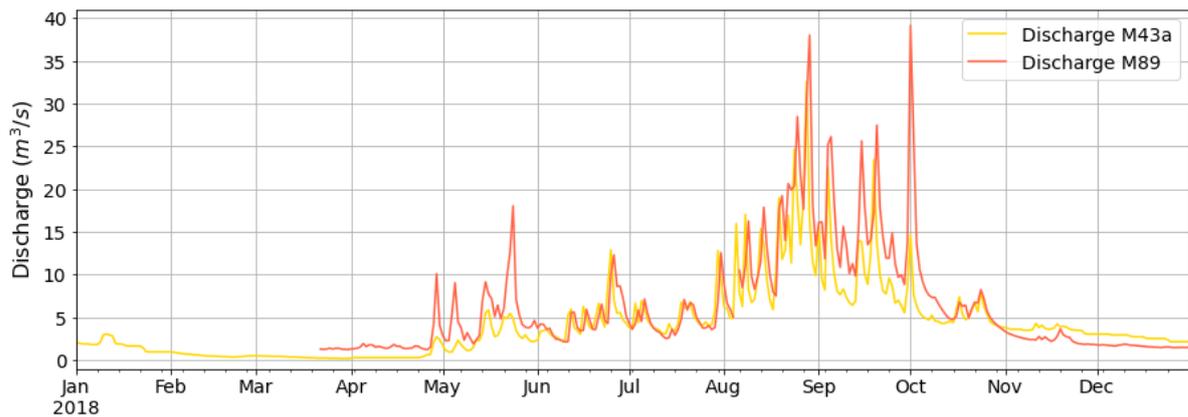


Figure 3.8: Observed discharge at the two discharge station for the year 2018, resampled to daily resolution.

### 3.2.4. Spatial and static precipitation input

Modelling strategy 3 uses spatial and static precipitation input, meaning the precipitation is constant over time, but not in space. For these runs, the study area was divided into four sub areas, as shown in Figure 3.9. The northern part comprises  $174 \text{ km}^2$ , the eastern part  $244 \text{ km}^2$ , the southern part  $135 \text{ km}^2$  and the western part  $149 \text{ km}^2$ . Four runs of two weeks were performed, where for each run only precipitation occurred in one of the four sub areas. This was done to analyse the effect of the location of a precipitation event on the resulting discharge. A constant precipitation of  $50 \text{ mm/day}$  was used to imitate a period of heavy precipitation. The model run was done using the parameter set with the lowest error obtained after calibration.

### 3.2.5. Spatial and dynamic precipitation input

Modelling strategy 4 includes spatial and dynamic precipitation, meaning the precipitation is heterogeneous in both space and time. First, the model was run using synthetic precipitation data, both at a steady location and with the precipitation moving in different directions across the catchment. Secondly, the model was run using interpolated rain gauge data. Again, the model run was done using the best performing parameter set after calibration.

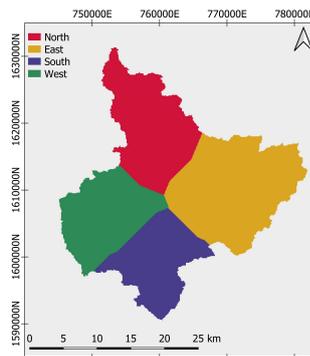


Figure 3.9: The four sub areas for the spatial and static precipitation input runs.

### Synthetic data runs

A precipitation event based on true precipitation data was created which can be seen in Figure 3.10. With this design storm different runs were performed. Similar to subsection 3.2.4 the model was run with the storm only occurring in one of the four sub areas as illustrated in Figure 3.9, to see the effect of the location of the storm. Additionally, runs were performed with the precipitation event moving in different directions across the catchment, to see the effect of the direction of the storm on the modelled discharge. This resulted in four runs, where the precipitation was moving from north to south, south to north, east to west and west to east. All four runs were done using two different precipitation inputs. For the first the precipitation moved across the catchment using polygons distribution and for the second Inverse Distance distribution was used (Figure 3.11), resulting in eight runs in total. This was done to analyse the effect of the direction of precipitation event.

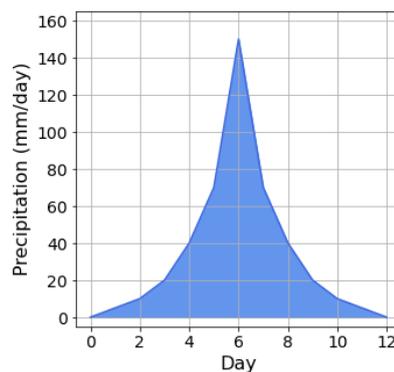


Figure 3.10: Design storm used for the spatial, dynamic precipitation modelling.

### Interpolated rain gauge data runs

The last run performed used interpolated rain gauge data of the year 2018. Distributed precipitation maps were created using precipitation data observed at rain gauges across the catchment. The Inverse Distance method was used to create these maps, as this method results in a relatively natural looking rainfall distribution. The resulting modelled hydrographs were compared to the results obtained from the annual model run using uniform data, in order to analyse the effect of spatial versus uniform precipitation input with regard to the accuracy of the modelled discharge.

### Catchment characteristics

From the modelled discharge, characteristics as discussed in section 2.3 were determined to check whether it corresponds to the observed characteristics of the catchment. For the interpolated rain gauge runs, the BFI and runoff coefficient were determined, according to the methods explained in section 2.3, and compared to the obtained results for 2018 in section 2.4.

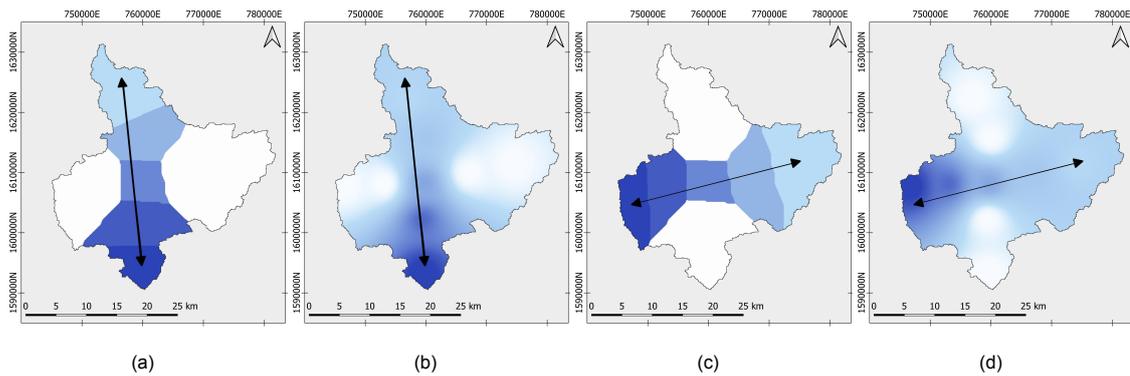


Figure 3.11: The precipitation event moving (a, b) between north and south and (c, d) between east and west, using distribution according to (a, c) polygons and (b, d) Inverse Distance.

### 3.3. Results modelling

In this section the results of the above explained calibration methods will be presented as well as those of the different model runs performed, following the four identified modelling strategies.

#### 3.3.1. Sensitivity Analysis

In Table 3.3 an overview of the resulting sensitivity classes for the different parameters can be found. In Appendix I an overview of the sensitivity indices can be found. For modelling of the discharge, the temperature and the cropfactor showed to be the parameters with the strongest correlation. Next, the soil parameters showed to be important, affecting factors such as saturated hydraulic conductivity and porosity, mainly changing the infiltration capacity and water holding capacity of the soil. Lastly, the initial unsaturated water content showed to be important as well as the initial saturated water content and saturated hydraulic conductivity, although slightly less significant.

	$Q$	$E_a$	$P_d$	$I_{eff}$	$\Delta S$
Temperature	IV	III	IV	IV	III
Clay content	III	II	IV	III	III
Silt content	III	II	IV	II	IV
Sand content	III	II	II	II	IV
Stone content	II	I	I	I	III
Organic matter content	III	I	III	II	III
Soil thickness	III	II	III	II	IV
$K_{sat}$ boundary condition	I	I	III	I	I
Unsaturated water content	III	II	IV	II	IV
Saturated water content	II	II	II	I	IV
Interception	I	I	I	I	I
Soil albedo	I	I	I	I	I
Canopy albedo	I	I	I	I	I
Canopy height	I	I	I	I	I
Cover fraction	I	I	I	II	I
Leave area index	I	I	II	II	I
Soil cropfactor	I	I	I	I	I
Cropfactor	IV	III	III	III	III
Interception capacity	I	I	I	I	I
Root depth	I	I	I	I	I
$K_{sat}$	II	II	II	II	III
Root cohesion	I	I	I	I	I
Surcharge	I	I	I	I	I

Table 3.3: Overview of determined the sensitivity class per parameter. Blue cells indicate a negative correlation, red cells indicate a positive correlation. White cells indicate a sensitivity index of zero.

Resulting from the sensitivity analysis, the main parameters on which tuning was performed are soil

thickness, saturated hydraulic conductivity, initial water content in the soil and crop factor. Temperature input was not changed, as it is a dynamic variable in the model which could be obtained through daily observations.

### 3.3.2. Uniform and static precipitation input

#### Complete static precipitation input

Both the 4 mm/hr run and the 10 mm/hr run, gave a closing water balance. However, for the initial parameter run no discharge was modelled for the 4 mm/day run and hardly any discharge was modelled for the 10 mm/day run. With an annual potential evaporation of 2060 mm, all precipitation of the 4 mm/day run is evaporated, explaining the lack of discharge. However for the 10 mm/day run it showed that almost half of the incoming water left the system through deep percolation (Figure 3.12). For this reason the boundary condition of the saturated hydraulic conductivity was lowered in an iterative process, where the resulting water balance was compared to the water balance in subsection 2.4.6. The best results were modelled for  $K_{sat_{bc}}$  being equal to zero, meaning no water leaving the system through deep percolation.

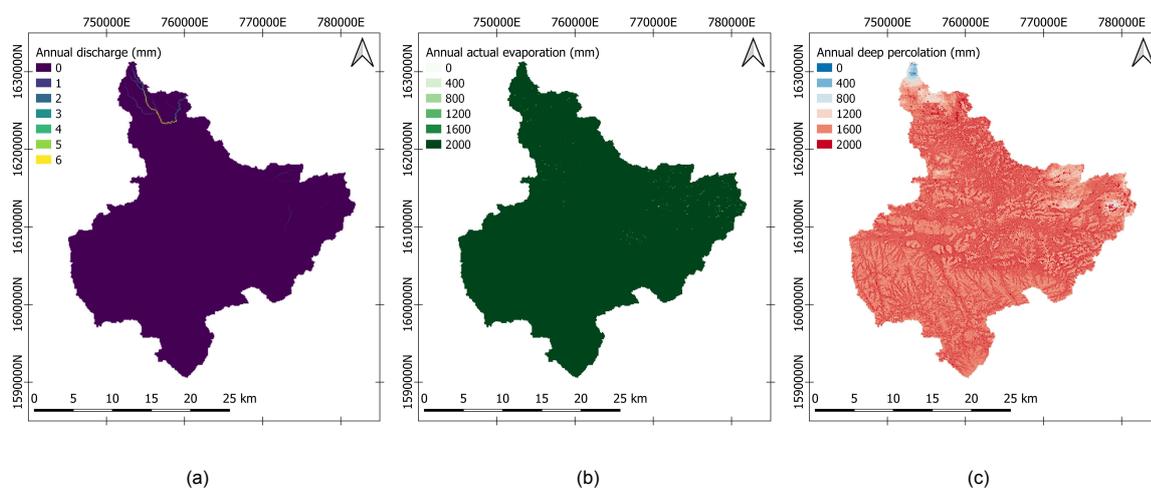


Figure 3.12: (a) Modelled total annual discharge, (b) modelled annual actual evaporation and (c) modelled annual deep percolation for the 10 mm/day run using the initial parameters.

#### Partial static precipitation input

Runs with the initial parameter set gave a recession coefficient at M43a of 13.6 days and at M89 of 12 days. For M89 this corresponds to the recession coefficients determined in subsection 2.4.4, however for M43a it was too low.

Lowering of the hydraulic saturated conductivity did not have the desired effect and decreased the recession coefficient, as it resulted in less water infiltrating into the ground and thus a lower groundwater storage and base flow and more direct runoff. However, increasing the layer thickness had the desired effect. As an increase of the layer thickness resulted in an increase of the recession coefficient at both location, the study area was divided into the two HRU (Figure 2.6a), and the layer thickness was changed for both sections individually. In an iterative process, the layer thickness was increased with 10% in the National Park and decreased by 10% in the provincial area, resulting in a recession coefficient of 26.5 days and 18.7 days at M43a and M89 respectively. The precipitation event, modelled discharge and corresponding base flow can be seen in Figure 3.13. The recession coefficient at M89 is still slightly too high compared to the ones determined in subsection 2.4.4, however further decreasing of the layer thickness would make the layer very narrow (< 1 meter), which is why it has been decided to do this. The obtained parameter set from the recession check will from now on be referred to as *Parameter set A*.

### 3.3.3. Uniform and dynamic precipitation input

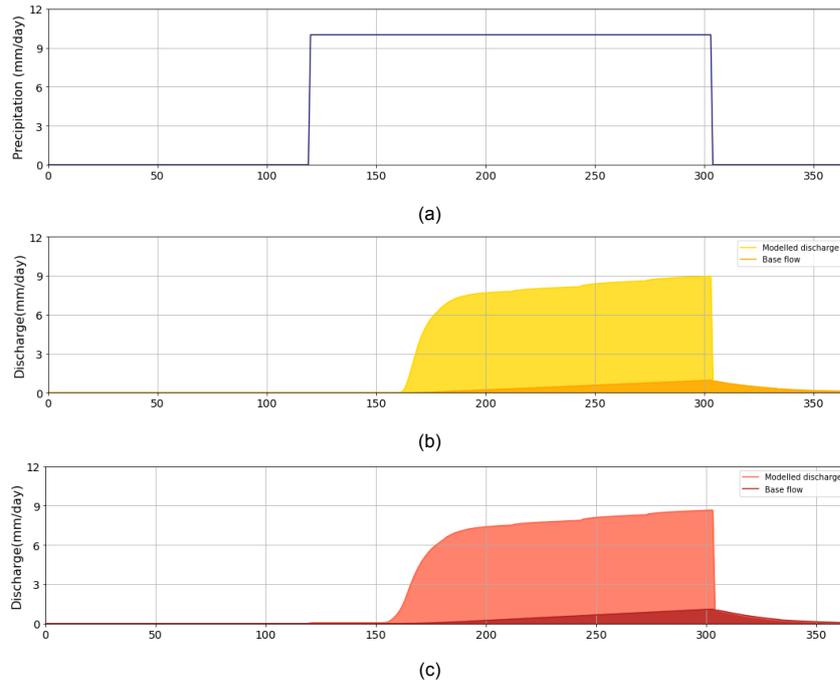


Figure 3.13: (a) Test precipitation used for the partial static precipitation run and the resulting modelled discharge and determined base flow for (b) M43a and (c) M89.

**Event based calibration**

Running the model with *Parameter set A* gave an adequate result of the magnitude of the discharge peak at station M43a, however the peak at station M89 was too high. Increasing the saturated hydraulic conductivity resulted in a higher infiltration and thus a lower runoff. As the magnitude of the modelled runoff at M43a was already satisfactory, only the saturated hydraulic conductivity in the provincial area was altered. Additionally, the value of the initial water content in the unsaturated zone was adjusted. This resulted in two possibly best parameter sets obtained through visual calibration, of which the results can be seen in Figure 3.14, and which will be called *Parameter Set B* and *Parameter Set C* from now on. The details of the parameter sets can be found in Appendix J. *Parameter Set B* resulted in an adequate modelling of the magnitude of the peaks, however the timing, onset and recession did not match the observed discharge. For *Parameter Set C* the discharge modelled at station M43a resulted in adequate results of the magnitude, but an inadequate timing, onset and recession. Discharge modelled at station M89 *Parameter set C* resulted in an adequate timing of the peak, but an inadequate magnitude and onset.

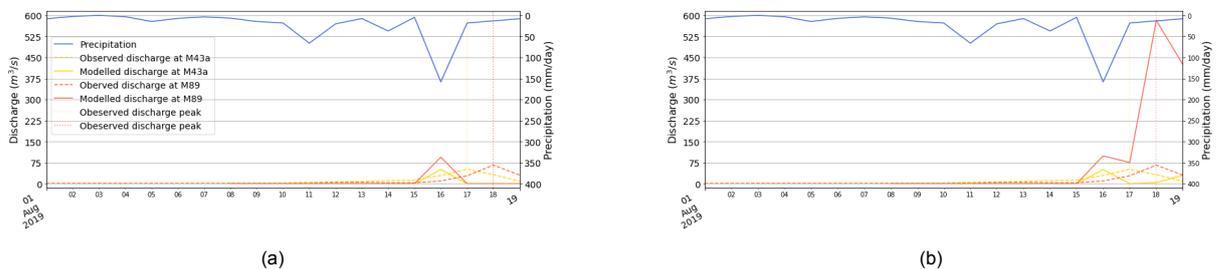


Figure 3.14: Modelled discharge after parameter tuning based on an isolated individual discharge event for (a) *Parameter set B* and (b) *Parameter set C*.

Changes in the saturated hydraulic conductivity were most effective in altering the magnitude of the modelled discharge, but also changing the initial unsaturated water content helped in improving the results. Changes in the crop factor and layer thickness did not result in any significant changes in the modelled discharge peaks.

For validation both parameter sets were run for the whole month of August. The results can be seen in Figure 3.15. For *Parameter set B*, the magnitude of the modelled peaks at M43a is similar to that of the observed peaks. The timing of the peaks however is either too early or too late. The modelled discharge at M89 shows both too high magnitudes and wrong timing of the peaks. For *Parameter set C*, both timing and magnitude are significantly off for the discharge modelled at M43a. At M89 the timing seems to be adequate for most peaks, although some false peaks are modelled. However, the discharge is significantly too high compared to the observed discharge peaks.

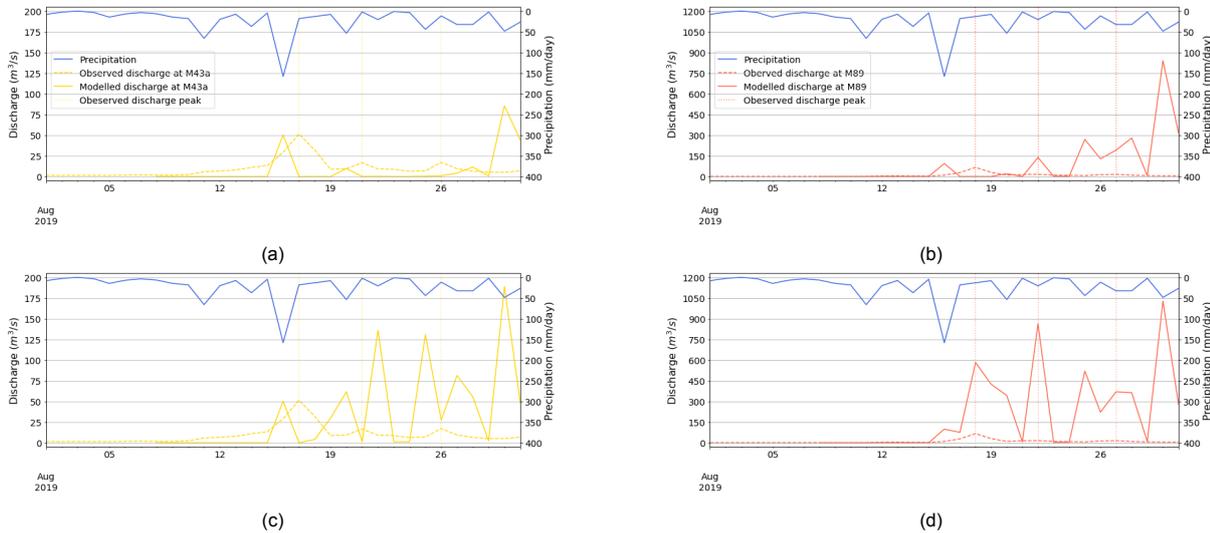


Figure 3.15: Results of the validation run for (a, b) Parameter Set B and (c, d) Parameter Set C, modelled at (a, c) discharge station M43a and (b, d) discharge station M89.

**Flow Duration Curve**

The FDC results of *Parameter sets A, B and C* for discharge modelled at station M43a and station 89 can be seen in Figure 3.16. The corresponding calculated errors can be found in Table 3.4. For the discharge modelled at M43a the calculated errors do not differ a lot between the three parameter sets. The errors calculated at M89 are significantly higher than those calculated at M43a. For all three errors, *Parameter set B* shows to be the best fit for modelling discharge at M89.

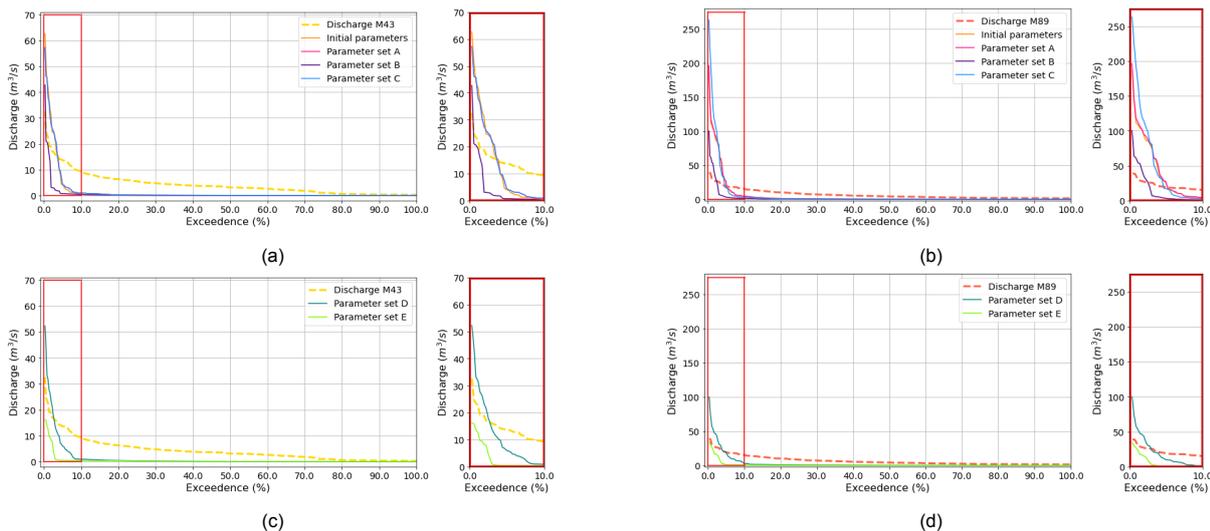


Figure 3.16: The Flow Duration Curves of the modelled discharge at (a, c) station M43a and (b, d) station M89, using (a, b) the initial parameter set as well as Parameter sets A-C and (c, d) Parameter sets D and E.

After further tuning of the parameters, two new parameter sets (*Parameter set D and E*) were obtained,

		<i>Parameter set A</i>	<i>Parameter set B</i>	<i>Parameter set C</i>	<i>Parameter set D</i>	<i>Parameter set E</i>
<b>M43a</b>	<b>RMSE</b>	10.80	11.12	10.81	<b>8.60</b>	11.71
	<b>NSE</b>	-3.70	-3.98	-3.71	<b>-1.98</b>	-4.52
	<b>KGE</b>	-1.32	-0.42	-1.33	<b>-0.87</b>	<b>-0.25</b>
<b>M89</b>	<b>RMSE</b>	47.77	19.37	67.14	16.67	<b>16.41</b>
	<b>NSE</b>	-56.46	-8.45	-112.49	-5.99	<b>-5.78</b>
	<b>KGE</b>	-5.77	-2.00	-8.78	-1.73	<b>-0.87</b>

Table 3.4: Errors calculated for the FDC of modelled discharge compared to observed discharge using the different parameter sets obtained.

with a lower error of the FDC fit. For both sets, the initial unsaturated water content was set to zero in the provincial area, while it was not changed for the National Park. Additionally, the layer thickness in both parameter sets was increased. A more detailed overview of the changes made for each parameter set can be found in Appendix J. The errors of the new parameter sets can be found in Table 3.4, the corresponding FDC plots for both stations are shown in Figure 3.16. When looking at the KGE *Parameter set E* shows to be the best fit, while for the other two errors *Parameter set D* would overall score better. However, for all parameter sets the NSE and KGE still show a negative value, indicating a poor fit. When looking at the FDC plots, both parameter sets shows a less steep slope than the previous sets, better corresponding to the slope of the observed discharge, although recession is still too quick. However, they both either still overestimate or under estimate the peak discharges.

#### Annual model run

The resulting modelled discharge of the runs for 2018 can be seen in Figure 3.17. In Figure 3.18 and Figure 3.19 the same results can be found, but then in to the months May, and August and September, which are the months where most peaks in the discharge are observed and modelled, and thus most relevant for checking the models capability of flash flood predictions. When comparing the modelled peaks with the observed peaks, it can be seen that the timing is overall modelled quite poorly. At station M89 none of the observed peaks are modelled at the right timing. Most peaks are modelled a day too early, but also many false peaks are modelled, making it difficult to determine. At station M43a some of the modelled peaks coincide with observed peaks, however, also still a large amount of discharge peaks is not modelled. Additionally, at a lot of false peaks are modelled. In May three peaks were observed, while nine peaks were modelled at both locations and for the months August and September a similar amount of peaks were observed as modelled, although no correspondence between the two could be found. What can be seen is that the modelled discharge often is a direct response to the input rainfall data. The different parameter sets do not show any difference in timing of the peaks.

The magnitude of the peaks is greatly overestimated in the month May for *Parameter sets A, C and D*, while *Parameter set E* hardly modelled any discharge for that month. In the months August and September, where observed discharge values are higher, the magnitude of the modelled discharge better corresponds to the magnitude of the observed discharge. Furthermore, hardly any base flow is modelled for either of the parameter sets.

In Table 3.5 the calculated errors of the different parameter sets can be found. When looking at the RMSE and NSE determined for the different parameterisations, *Parameter set E* shows to model the best fit. When looking at the KGE *Parameter set D* models the best fit. However, all errors are negative, indicating neither of the parameter sets to model a proper fit. The months August and September are the months where most intense precipitation events occur and thus the probability of flash floods is highest. Consequently accurate modelling of the discharge is most important for this period, which is why the error has also been determined regarding these months only. The results can be found in Table 3.6. *Parameter Set E* again shows to be the best fit.

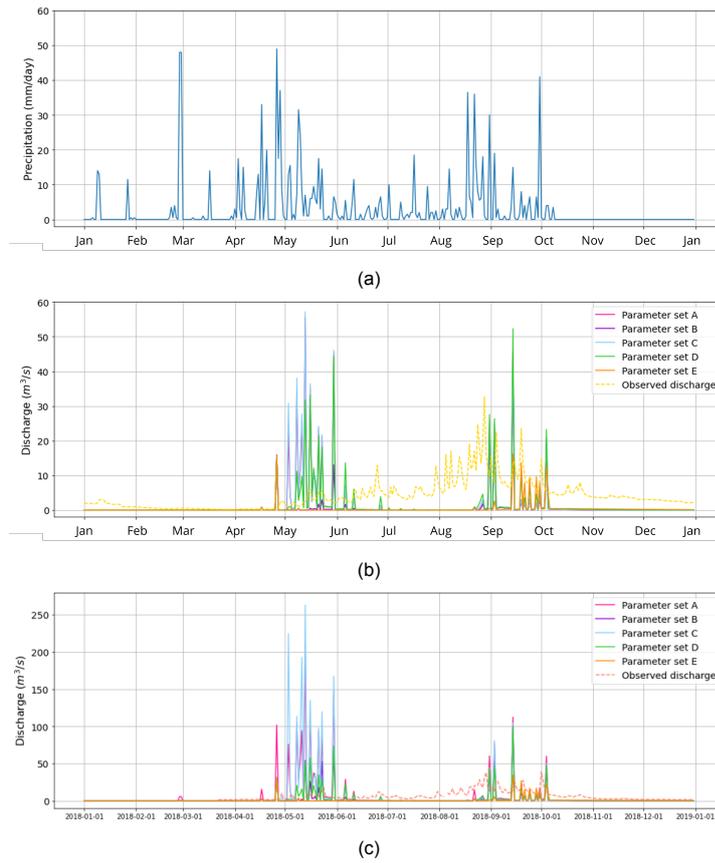


Figure 3.17: (a) Observed precipitation of the year 2018 at rain gauge STN0551 and the resulting modelled discharge for the five parameter sets at (a) station M43a and (b) station M89 when using this precipitation as input for the model.

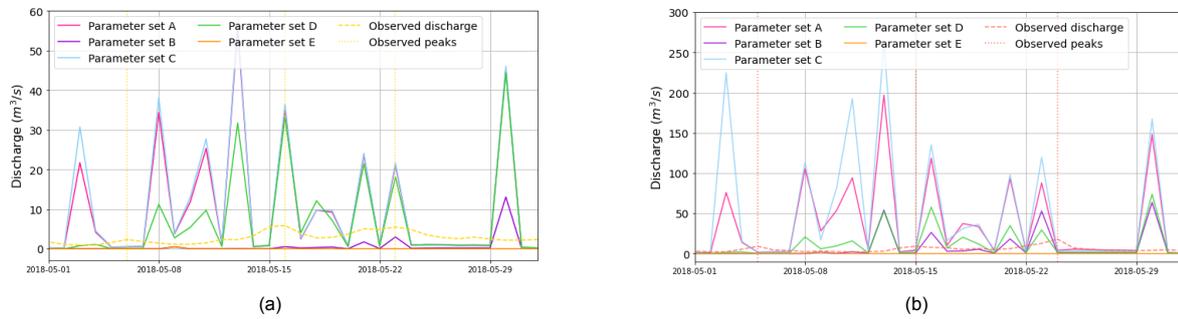


Figure 3.18: Observed and modelled discharge using uniform precipitation zoomed in to May 2018 for (a) M43a and (b) M89 using the obtained parameter sets after tuning.

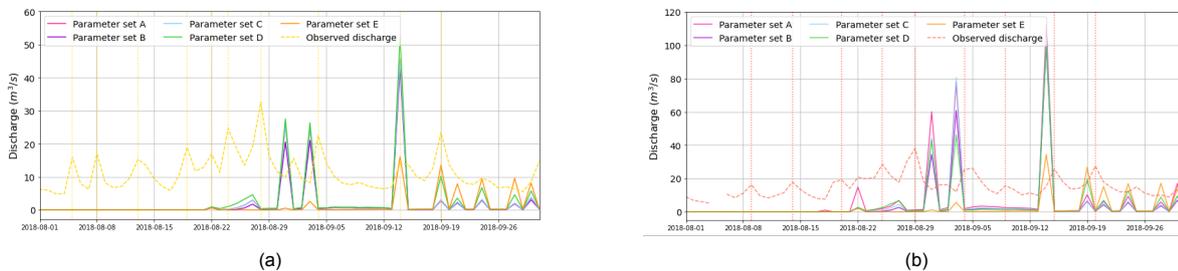


Figure 3.19: Observed and modelled discharge using uniform precipitation zoomed in to August and September 2018 for (a) M43a and (b) M89 using the obtained parameter sets after tuning.

		Parameter set A	Parameter set B	Parameter set C	Parameter set D	Parameter set E
M43a	RMSE	7.84	6.19	8.05	7.06	<b>5.82</b>
	NSE	-2.21	-1.00	-2.38	-1.60	<b>-0.76</b>
	KGE	-0.27	-0.37	-0.28	<b>-0.24</b>	-0.47
M89	RMSE	21.02	10.55	27.34	10.84	<b>7.86</b>
	NSE	-10.46	-1.89	-18.38	-2.05	<b>-0.60</b>
	KGE	-1.51	-0.28	-2.47	<b>-0.26</b>	-0.40

Table 3.5: Errors calculated for the annual model runs where discharge is modelled using the different parameter sets obtained after various calibration methods using uniform precipitation input.

		Parameter set A	Parameter set B	Parameter set C	Parameter set D	Parameter set E
M43a	RMSE	12.83	12.66	12.83	12.83	<b>11.79</b>
	NSE	-4.52	-4.37	-4.52	-4.52	<b>-3.66</b>
	KGE	-0.33	<b>-0.32</b>	-0.33	-0.35	-0.39
M89	RMSE	21.69	19.93	21.23	19.25	<b>15.67</b>
	NSE	-10.22	-8.47	-9.75	-7.84	<b>-4.86</b>
	KGE	-1.20	-0.83	-1.07	-0.78	<b>-0.31</b>

Table 3.6: Errors calculated for the period August and September of the annual model runs where discharge is modelled using the different parameter sets obtained after various calibration methods using uniform precipitation input.

### 3.3.4. Spatial and static precipitation input

The four runs executed with spatially varying, but temporal steady precipitation input resulted in modelled discharge as can be seen in Figure 3.20. The constant input precipitation results in a constant modelled discharge. It can be seen that the different locations result in different magnitudes of the discharge. It should be considered that the four parts are not equally large, however, while the southern part has the smallest area, it results in the highest levels of discharge. Precipitation in the northern and eastern part of the catchment does not result in any discharge at station M43a.

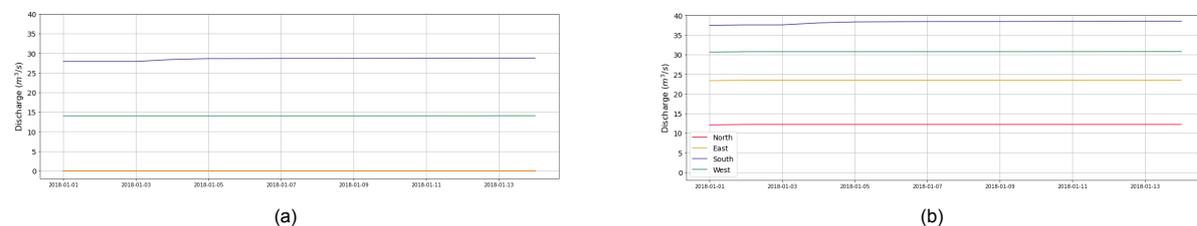


Figure 3.20: Modelled discharge at (a) M43a and at (b) M89 for the spatial, static precipitation input.

### 3.3.5. Spatial and dynamic precipitation input

#### Synthetic data runs

For the synthetic data runs two kinds can be distinguished, one where the precipitation event has a steady location and one where the precipitation event moves across the catchment. Both the results will be discussed below.

**Steady location runs** The results of the steady location precipitation run can be found in Figure 3.21. Just as in subsection 3.3.4, the largest peak in discharge occurs for precipitation falling in the south and the least for precipitation falling in the north. The discharge peaks increase with distance, so precipitation falling further away from the discharge station seems to result in higher peaks. Additionally, it shows that precipitation falling in the north and east does not result in any discharge at M43a.

**Moving runs** For the precipitation moving across the catchment two methods were used. One in which the precipitation pattern was obtained using polygons and one where it was obtained using Inverse Distance interpolation. The results can be found in Figure 3.22 and Figure 3.23 respectively.

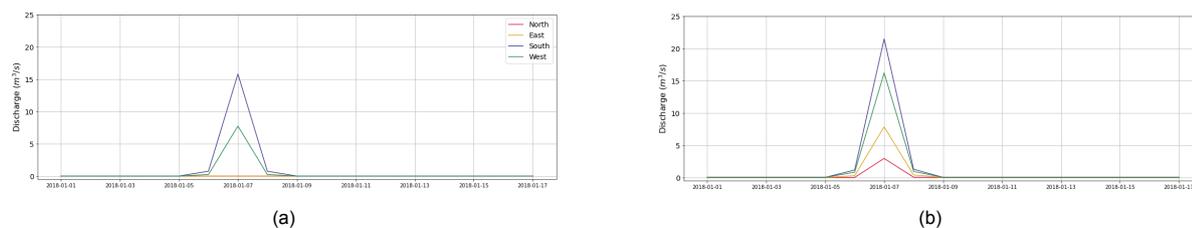


Figure 3.21: Modelled discharge at (a) M43a and at (b) M89 for the spatial, dynamic precipitation input, with precipitation only occurring at one sub area of the catchment.

It can be seen that movement of the precipitation event between north and south gives higher peaks than movement between east and west. Moreover, the timing of the peak is affected by the direction of the precipitation event. Each direction results in a different timing of the discharge peak. A symmetrical response can be seen regarding the direction of the precipitation event. Additionally, the distribution of the precipitation has an effect on the occurring discharge, as can be seen in the difference between peaks resulting from the polygon distribution and the Inverse Distance distribution. An overview of the modelled discharge using polygon distribution compared to Inverse Distance distribution can be found in Appendix K.

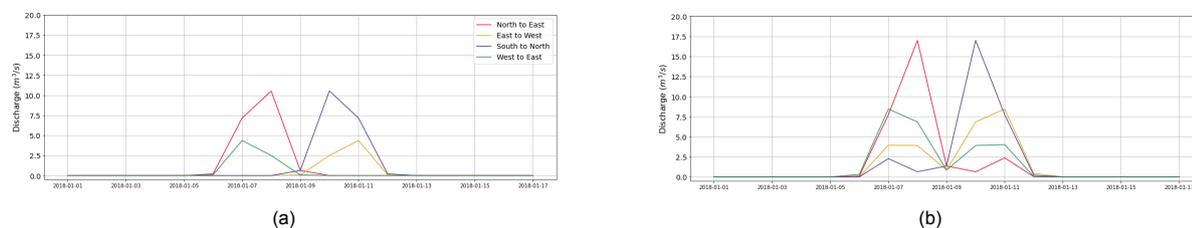


Figure 3.22: Modelled discharge at (a) M43a and at (b) M89 for the spatial, dynamic precipitation input, with precipitation moving across the catchment using polygon distribution.

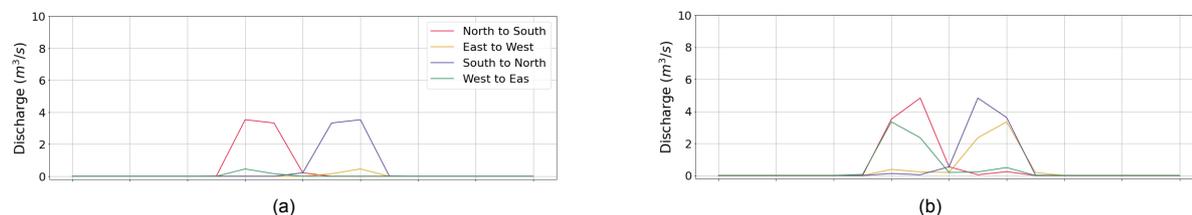
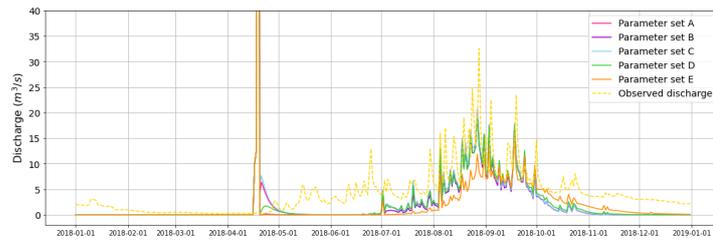


Figure 3.23: Modelled discharge at (a) M43a and at (b) M89 for the spatial, dynamic precipitation input, with precipitation moving across the catchment using Inverse Distance distribution.

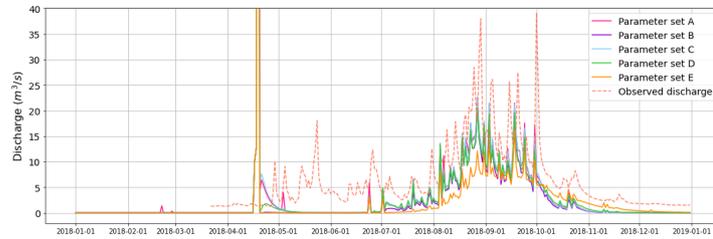
### Interpolated rain gauge data runs

The resulting modelled discharge for the different parameter sets using interpolated rain gauge data can be seen in Figure 3.24. The y-axis has been limited to better visualise the modelled discharge during the wet season, however this results in the incomplete visualisation of the large discharge peak in April. In Appendix K the complete peak is included. Besides this large discharge peak in April, the modelled discharge quite well represents the observed discharge. The base flow is modelled and also the pattern of the peaks is captured. However, when zooming in to the months August and September in Figure 3.25 it can be seen that the timing of the peaks is still off in most cases, although the magnitude and pattern of the discharge is decent.

When looking at the calculated errors, it can be seen that those for the complete year are higher than the errors obtained from uniform precipitation modelling. However, when only looking at the errors for the period August and September, this is significantly better than for the uniform runs. *Parameter Sets A, B and C* even show positive errors for the NSE. Where *Parameter Set E* showed to be the best fit for uniform modelling, it is the worst fit for the distributed modelling.

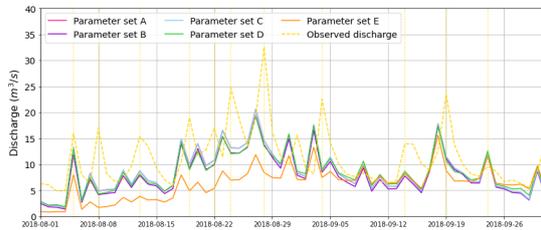


(a)

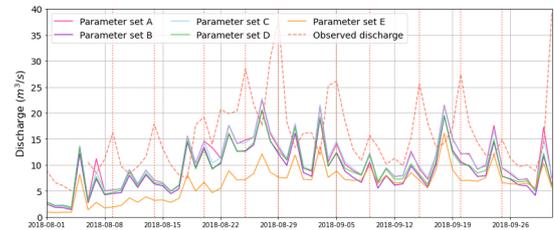


(b)

Figure 3.24: Modelled discharge of 2018 for the five parameter sets at (a) station M43a and (b) station M89 when using distributed precipitation obtained from rain gauge data as input for the model.



(a)



(b)

Figure 3.25: Modelled discharge using distributed precipitation interpolated from rain gauge data zoomed in to August and September 2018 for (a) M43a and (b) M89 using the obtained parameter sets.

		Parameter set A	Parameter set B	Parameter set C	Parameter set D	Parameter set E
M43a	RMSE	13.73	14.37	13.78	<b>13.70</b>	13.84
	NSE	-8.82	-9.76	-8.89	<b>-8.78</b>	-8.99
	KGE	-1.37	-1.51	-1.38	<b>-1.36</b>	<b>-1.34</b>
M89	RMSE	<b>14.71</b>	14.79	14.73	14.73	14.99
	NSE	<b>-4.61</b>	-4.67	-4.63	-4.63	-4.83
	KGE	-0.71	-0.71	-0.71	<b>-0.70</b>	<b>-0.70</b>

Table 3.7: Errors calculated for the annual model runs where discharge is modelled using the different parameter sets obtained after various calibration methods using distributed precipitation input.

		Parameter set A	Parameter set B	Parameter set C	Parameter set D	Parameter set E
M43a	RMSE	5.39	5.43	<b>5.38</b>	5.58	7.35
	NSE	0.02	0.01	<b>0.03</b>	-0.04	-0.81
	KGE	-0.04	<b>-0.03</b>	<b>-0.03</b>	-0.05	-0.17
M89	RMSE	7.64	8.38	<b>7.52</b>	8.15	10.47
	NSE	-0.39	-0.67	<b>-0.35</b>	-0.58	-1.62
	KGE	<b>-0.06</b>	-0.12	-0.07	-0.11	-0.25

Table 3.8: Errors calculated for the period August and September of the annual model runs where discharge is modelled using the different parameter sets obtained after various calibration methods using distributed precipitation input.

### Catchment characteristics

For the determination of the catchment characteristics, the large discharge peak modelled in April was corrected to get more realistic results.

**Base Flow Index** The results of the determined BFI for the different parameter sets can be found in Table 3.9. It shows that the modelled discharge has a lower BFI than the one determined for the observed discharge. Additionally, the difference in BFI between the upstream and downstream station in the study area is not captured in the modelled discharge.

	Observed	Parameter Set A	Parameter Set B	Parameter Set C	Parameter Set D	Parameter Set E
<b>M43a</b>	0.67	0.46	0.45	0.46	0.48	0.50
<b>M89</b>	0.73	0.46	0.45	0.46	0.48	0.48

Table 3.9: Base Flow Index values determined from modelled discharge.

**Runoff coefficient** As distributed precipitation is used, the total annual precipitation amount is determined by taking the average precipitation amount of the upstream catchment of both discharge stations as illustrated in Figure 2.6b, which results in 1511 mm for discharge station M43a and 929 mm for discharge station M89. The runoff coefficients determined for the modelled discharge from the different parameter sets can be found in Table 3.10. They correspond quite well with the ones obtained from the observed data.

	Observed	Parameter Set A	Parameter Set B	Parameter Set C	Parameter Set D	Parameter Set E
<b>M43a</b>	0.25	0.15	0.15	0.15	0.15	0.11
<b>M89</b>	0.05	0.06	0.05	0.06	0.06	0.04

Table 3.10: Runoff coefficients determined from modelled discharge.

## 3.4. Discussion

### Limitations and uncertainties

All data used as input for the model comes with uncertainties. The distribution maps obtained for various of the parameters all come with specific uncertainties. The soil parameters were scaled based on soil samples collected at one location in the National Park, however this scaling was applied to the entire catchment, although different ratios of the different soil types might be present at different locations in the catchment. For multiple parameters a constant was used, while in reality all parameters will be heterogeneous over the area. Additionally, assumptions were made to get to the different parameters which might not be true to reality. Other big uncertainty factors are the initial saturated and unsaturated water content used as input for the model. Both were obtained through a combination of model runs and observed data. For the saturated water content the observed data consisted of multiple ground water well throughout the study area, whereas for the unsaturated water content it consisted of only one telemetering station installed in the National Park, which is very minimal for an area of more than 700 km<sup>2</sup>. Also the saturated hydraulic conductivity of the bedrock layer below the soil layer comes with high uncertainty, as it is complicated to obtain through field observations. For this research it has been obtained through multiple runs, resulting in a saturated hydraulic conductivity of the bedrock of zero, meaning no deep percolation occurs. This is however unlikely to be true in reality, as a large part of Thailand is underlain by deep permeable soils, facilitating the presence of deep percolation (Chappell et al., 2007)

For the conducted sensitivity analysis, linearity between the input parameters and output variables was assumed, which is not always accurate. Additionally, as a local sensitivity analysis was performed, only the individual affect of input parameters was examined, while for some parameters the effect might be intensified when jointly altered. From the sensitivity analysis, the most important parameters were found to be temperature and cropfactor. The temperature was assumed to be uniform over the study area and monthly means were used. No changes in temperature were made during calibration of the model, as temperature is a dynamic and meteorological characteristic that can be observed and not a static characteristic of the catchment. Furthermore, the sensitivity analysis only checked the effect of

the different input parameters on the annual totals of the output variables. Effects on event scale and aspects such as timing and single peak magnitudes were not considered, while this would be beneficial to look at for flash flood modelling.

Furthermore, the test data and synthetic data used did not always represent real precipitation. For the uniform precipitation runs, daily precipitation of 4 and 10 *mm* was used over the time span of a year and seasonal precipitation of 0 *mm/day* in the dry season and 10 *mm/day* in the wet season. This is not a realistic precipitation pattern, and solely used to check the model performance and get a first understanding of the catchment and its boundary conditions, not to get realistic discharge values. The same applies to the precipitation used for the spatial and static runs with a steady location. A precipitation event of 50 *mm/day* occurring for two weeks is not realistic, however does help to check the models response to distributed precipitation as well as to get an initial idea of the catchments response to local precipitation.

For the runs based on observed precipitation data, a big limitation can be found in the lack of data. As for most gauges the precipitation data is limited, it is not possible to first perform model runs over a longer period of time for calibration and later run the model over the remaining years for validation. Models generally have a spin up period of a few years in which they need to adjust and stabilise, which could improve the performance of distributed hydrological models as the one used in this research (Wang et al., 2016). With the limited precipitation records in the study area this is not possible. This limitation in precipitation data also results in the FDC calibration being performed for one year only, as well as the event calibration, possibly causing biased results. Calibration over multiple years would be better. Additionally, the uncertainty of the rain gauge data as discussed before in section 2.5 plays a role. Interpolating the observed rainfall over the study area further increases this uncertainty.

Furthermore, for some parameter sets spatial calibration has been applied, making a differentiation between the the HRU in the study area. This has been done by dividing the area into two parts, resulting in a harsh division between the two areas, while in reality the change in characteristics between the two areas would be more gradual. As this division interrupts the connectivity of the soil layer, it could have affect on the modelled soil hydrology around this boundary, and result in a reduced accuracy around of the models performance around this area.

Lastly, a lot of information is lost by changing the temporal resolution of the precipitation and discharge data from hourly to daily. By doing this, the exact timing of peaks is lost and can thus not be properly captured by the model. Additionally, the exact timing of discharge peaks is lost as well, making validation less reliable.

## Results

From the sensitivity analysis it was found that the most important parameters affecting the modelled discharge were temperature and cropfactor, followed by soil content, layer thickness, soil water content and saturated hydraulic conductivity. However, when changing the cropfactor during event calibration, little effect was observed. A reason for this could be that the sensitivity analysis was performed looking at annual totals, while for the event calibration only two weeks were considered. During calibration, the biggest effects were observed when changing the saturated hydraulic conductivity, as this is the main component for determining the infiltration capacity of the soil.

The uniform and static precipitation runs were performed to get a first understanding of the system. As mentioned above, the precipitation input used does not give a realistic representation of the precipitation occurring in the study area, so the recession coefficient determined from the seasonal precipitation input is solely a first indication on how the catchment would respond to precipitation, using the specific parameterisation. From this initial recession check, it could be concluded that distributed calibration is necessary to capture the hydrological response of the catchment.

From the uniform and dynamic precipitation runs, it becomes clear that uniform precipitation input is not suitable when properly trying to capture the hydrological response of the study area. For the event calibration, both obtained parameter sets show an incorrect timing, following the precipitation pattern. Due to the uniform precipitation, the resulting discharge is a direct response to this, not showing any delay. Additionally, it can be seen that when running for a longer period of time, the modelled discharge peaks are significantly overestimated, especially at the downstream discharge station, which also is a

result of the uniform precipitation. The volume of the precipitation is overestimated as the input file gives uniform precipitation over the whole catchment while in reality precipitation would only cover a fraction of the total area. This validation shows that even when the magnitude seems acceptable for the single event calibration, this has been obtained for the wrong hydrological reasons. Both the event calibration and FDC calibration show the difficulty of properly capturing the base flow using uniform input as well as properly capturing the magnitude of the discharge peaks.

The lack of base flow modelling and overestimation of the peak flow is confirmed when running the model for a year using uniform rain gauge data. Furthermore, this run shows that the different parameterisations do not have any effect on the modelled timing of the peaks. They only affect the magnitude of peak discharge, but not the onset, timing or recession, making uniform input data very incompetent for adequate modelling of the discharge. The modelled discharges using uniform precipitation are a very poor fit to explain the observed discharge.

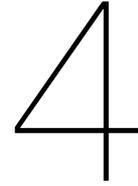
Calibrating the model using uniform precipitation data is very difficult, if not impossible. Some parameter sets give more satisfactory results than others, however none could be used for useful flash flood predictions when combined with uniform precipitation, as representation of the observed discharge was very poor for all. Due to time constraints of this research, certain parameter values have been accepted as the best options for a specific method, although there could potentially be more optimal parameter sets. However, it is not expected that different parameter sets would significantly increase the capability of the model to adequately model the discharge using uniform precipitation. Additionally, some of the obtained parameter sets might not be physically realistic, however necessary to get adequate results. An example of this is the initial unsaturated water content of 0 *mm*, which is not realistic for the study area, as observations from the telemetering station in the study area showed a moisture content of roughly 30% during the dry season. As it is a physically-based model, having physically impossible input parameters should not be necessary and thus indicates errors in other input data.

When looking at spatially distributed precipitation runs the importance of using distributed precipitation over uniform precipitation as input for the system is confirmed. The spatial runs with precipitation occurring at a steady location do not show any effect on the timing of the peak, however this can be explained by the daily resolution of the model. The duration of the discharge peak to move from upstream to downstream is generally less than a day, as determined in subsection 2.4.2, so using a daily resolution does not capture this difference in timing. On the other hand, when looking at the moving runs it clearly shows the effect of distributed precipitation on the timing of a discharge peak. The four different directions of the precipitation event show four different timings of the peak. Additionally, the difference between the modelled discharge using polygon based rainfall input and Inverse Distance rainfall input again shows the effect of distributed precipitation. Although the event moving across the system is the same, the distribution over the catchment is different, thus resulting in different hydrographs.

When looking at the steady location runs it is interesting to see that precipitation in the biggest sub area does not result in the highest discharge. The highest discharge is modelled for precipitation occurring in the most upstream area, even though this is the smallest area. A reason for this could be the strong precipitation-discharge relations in this area, as was determined in section 2.5. Runoff at M43a is only modelled when precipitation either falls in the southern or the western part. This is as expected, as the northern and eastern part are located downstream of station M43a and thus unlikely to significantly affect the discharge at that location. Also, the discharge modelled at station M89 is higher than that modelled at station M43a, which corresponds to observed discharge.

Running the model using interpolated rain gauge data especially confirms the importance of using spatially distributed precipitation data. Base flow is modelled during the wet season, which was lacking for the uniform runs. However, overall errors are higher than those for the uniform runs, which can be explained by the large discharge peak modelled in April. Nevertheless, when calculating the error for only the wet season, errors get significantly lower and some parameter sets even show positive errors for the determined NSE, clearly showing the added value of distributed precipitation. The differences in error between the five parameter sets are very limited however. It shows the rainfall pattern is most important for modelling of the hydrological response of the catchment, and although parameterisation has some effect this is not as important as the rainfall patterns. The precipitation in the study area determines its hydrological response, and not the soil or vegetation properties, making it very important to have accurate precipitation input.

The determined BFI using the modelled discharge gives lower values than the ones determined from the observed data and additionally the differentiation between the upstream and downstream area is not represented. A reason for the lower BFI could be the calibration of the Eckhardt parameters, but also the lacking base flow throughout the year. Although base flow is modelled during the wet season, the rest of the year, discharge is almost equal to zero. The calculated annual runoff coefficient corresponds quite well to those calculated using the observed data. Even the differentiation between the upstream and downstream catchment is captured. The annual precipitation obtained using the distributed precipitation gives a slightly higher annual precipitation for the upstream part than for the downstream part, although differences are minimal. When converting the modelled discharge to *mm/day* this results in higher discharge values upstream, and with that a larger runoff coefficient.



# Radar data analysis

This chapter gives a preliminary analysis of real precipitation radar data, which is ultimately used as input for the model. The analysis was done comparing the radar imagery with distributed precipitation maps obtained from the various rain gauges in the study area in order to determine what the additional value of using radar data would be.

## 4.1. Methodology Radar Data

A period of two weeks was selected to compare the radar data to the rain gauge observations. This period was selected based on the occurrence of precipitation and peak discharge events. The event was chosen from the options of isolated discharge events identified in subsection 3.2.3. It was decided to look at the event of 08-08-2019 until 23-08-2019, which is the same period as used for the event calibration. Additionally, for the year 2019 most rain gauge data was available, resulting in a more complete distributed rainfall pattern. The radar data has a spatial resolution of 600x600 meter and a temporal resolution of one hour.

### 4.1.1. Radar pixel comparison

To get an estimation of the reliability of the rain gauge data compared to the radar data, a point comparison was performed. Such comparison provides a first check on the accuracy of the rain gauge and radar data. The rain gauge observations for the selected period were compared to the data values of the radar pixels corresponding to the rain gauge locations. This was first done using the hourly resolution, after which it was repeated for the daily resolution.

### 4.1.2. Radar image comparison

From the available rain gauges in the study area, distributed precipitation maps were made using the Inverse Distance interpolation as explained in subsection 3.2.5. The resulting precipitation maps were compared to radar imagery to analyse the expected effect of using radar data.

For the initial comparison, the temporal resolution of the radar images was resampled from hourly to daily resolution, to correspond to the resolution of the model. For each day, the difference between the precipitation maps obtained from rain gauge interpolation and the radar maps was determined. The resulting difference maps were analysed by determining the (absolute) minimum difference, the (absolute) maximum difference, the range. Additionally, the standard deviation of the difference map as well as the absolute difference map was determined. These factors were calculated to determine the spatial variability of the difference map and analyse whether a pattern could be observed.

Furthermore, the NSE of each day was calculated, assuming the radar data to be the observed data and the rain gauge data to be the modelled data. Calculation of the NSE was done to determine whether interpolated rain gauge maps would be an adequate alternative for discharge modelling, would radar data no be available.

## 4.2. Results Radar Data

### 4.2.1. Radar pixel comparison

In Figure 4.1 the results of the hourly observations for both gauge and radar data at the different locations in the study area is provided. At the locations of gauge STN1475, STN1481 and TLPC the timing of the precipitation peaks observed by the gauge and the radar correspond quite well, just as the magnitude, although rain gauge observations seem to measure slightly higher rainfall amounts. At gauge STN0548, during the first five days quite some precipitation is observed by the rain gauge whereas only little precipitation is observed by the radar. After those five days gauge and radar peaks correspond quite well, although again peaks are higher from gauge observations. At gauge STN1477 the correspondence between radar and gauge observation is weaker, with multiple peaks either only measured by the gauge or only measured by the radar. The same holds for gauge STN1482.

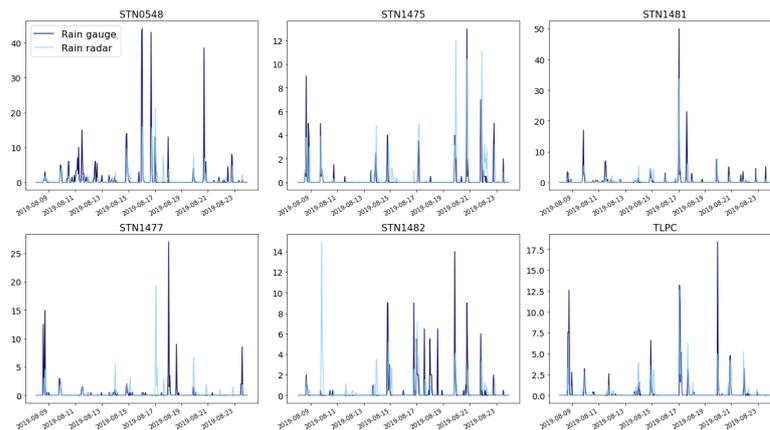


Figure 4.1: Hourly precipitation observations from multiple rain gauges in the study area and the radar pixels corresponding to those locations.

When looking at the daily precipitation amounts (Figure 4.2), it can be seen that precipitation patterns correspond quite well. Again, correspondence is best for precipitation measured at gauges STN1475, STN1481 and TLPC, but also at the other gauges similarities in the patterns can be observed. Only measurements at gauge STN1482 show relatively big differences between gauge and radar. At gauge STN1477 a clear delay in the observation of the precipitation peak can be seen, with rain gauge data showing the peak a day later than radar data.

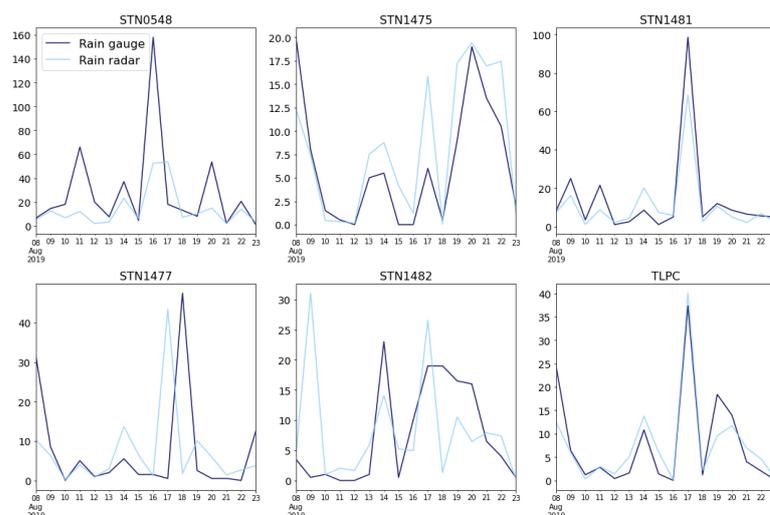


Figure 4.2: Daily precipitation observations from multiple rain gauges in the study area and the radar pixels corresponding to those locations.

### 4.2.2. Radar image comparison

In Appendix L an overview can be found of all daily precipitation maps obtained with interpolation over the rain gauges, precipitation maps obtained from radar images and the difference between the two. Additionally, an overview of the minima, maxima, means and standard deviations of the different maps can be found, as well as the calculates NSE.

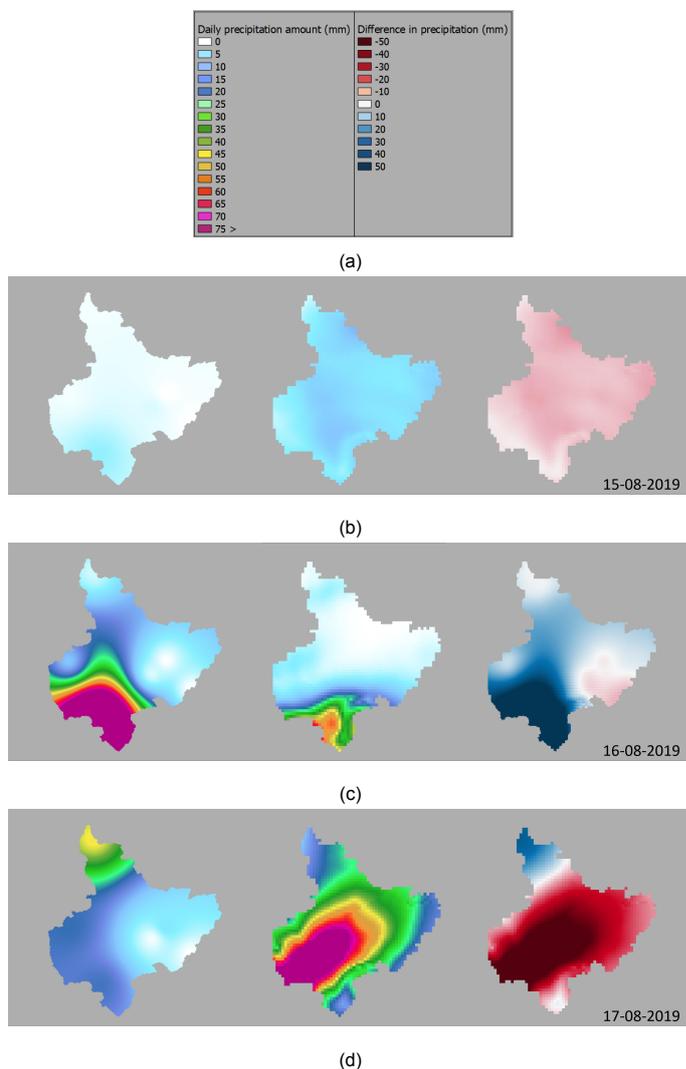


Figure 4.3: Overview of Inverse Distance of rain gauge data, Radar data and difference between rain gauge data and radar data (left to right), on (a) August 15<sup>th</sup>, (b) August 16<sup>th</sup> and (c) August 17<sup>th</sup> (c) 2018. The legend is provided in (d).

From a first look at the maps in Appendix L, it can be seen that for most days the difference between the two types of precipitation maps is quite clearly visible. In Figure 4.3 the precipitation and difference maps of three days can be found. The 15<sup>th</sup> of August is a day with relatively little precipitation and with that a relatively low error and difference between the precipitation maps. For the 16<sup>th</sup> of August, the NSE is has the lowest value and the maximum difference between the maps is largest, additionally on that day the highest amount of precipitation is measured by the rain gauges. On the 17<sup>th</sup> of August, the radar data observes the highest amount of precipitation, and additionally the largest range in the difference between the precipitation maps is observed for this day. For all three maps, the difference between the interpolation map and the radar maps shows different results. For 15<sup>th</sup> of August the rain gauges show an underestimation of precipitation throughout the whole area, although limited, while the 16<sup>th</sup> of August shows a large overestimation for most of the area and the 17<sup>th</sup> of August shows a large underestimation for the majority of the catchment.

In the difference maps, positive values indicate an overestimation of the precipitation by the interpolated

rain gauge maps, while negative values indicate an underestimation. In Figure 4.4 a plot is shown indicating in what percentage of the area the precipitation amount is over- and underestimate when using interpolated rainfall data. When looking to the overall results, overestimation of the precipitation occurs more (65%) than underestimation (35%), however when looking to the separate days, it is quite variable for each day. When looking at the boxplot in Figure L.2 the spatial variability in differences between rain gauge maps and radar maps is shown more evidently. On many days outliers can be identified, and minima and maxima range span wide ranges from positive to negative values. The difference between the two types of precipitation maps varies a lot. From values in Table L.1 it can be concluded that especially for days with high observed precipitation the range is very wide.

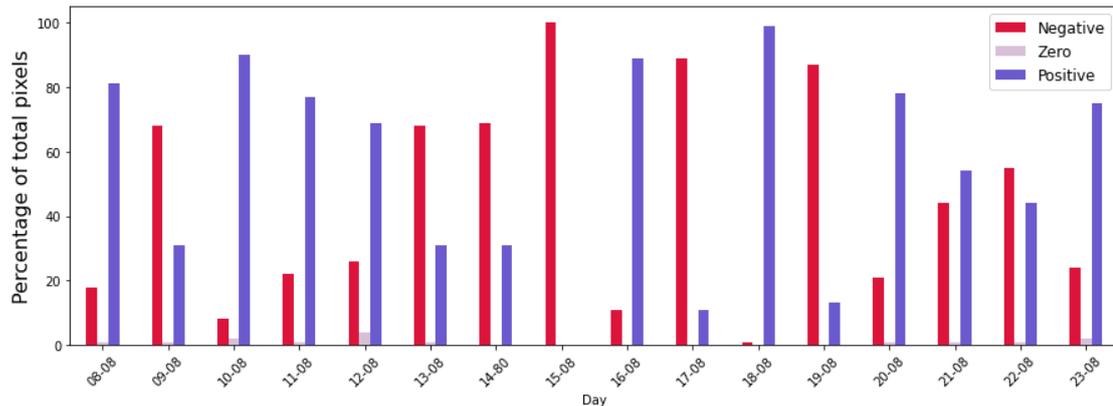


Figure 4.4: Percentage of overestimated (positive) and underestimated (negative) precipitation values for each day when using interpolated rain gauge data opposed to radar data.

### 4.3. Discussion

#### Limitation and uncertainties

Both the radar imagery and the rain gauge data come with uncertainties. This makes comparisons between the two subject to even larger uncertainties. Additionally, it makes it difficult to draw proper conclusions, as it is uncertain whether differences between the two are a result of errors in the rain gauge data or in the radar data. Radar data should give accurate representation of the distribution of the precipitation across the catchment, but precipitation amounts could be off, as they still need to be tested. Furthermore, the comparison between radar and gauge data has only been conducted for a period of two weeks, a relatively short time period, potentially giving biased results. The chosen period is during the wet period, which might give different results than comparisons made during dry or pre monsoon season. Although precipitation during the wet season is most important to capture properly, as this is the period in which flash floods can occur, data in the drier periods are important to properly capture, as this determines the initial conditions of the model.

#### Results

When looking at radar pixel to gauge comparison, the patterns as well as the magnitudes of precipitation look quite similar, for daily resolution. The hourly resolution observations show bigger variety (**elaborate a little on this**). For some locations, the correspondence between rain gauge values and radar pixel values is weaker than for others, but overall, a clear correlation can be seen in both magnitude and pattern. When looking at distributed maps however, differences are significantly larger. Especially on days with high amounts of precipitation, the difference between precipitation measured by the rain gauges and precipitation observed by the rain radar is large. This is inconvenient as days with high precipitation are most important to capture correctly, as these are the days with increased possibility of flash floods.

There seems to be hardly any correlations between the interpolated precipitation pattern and radar observed precipitation. For some days, the patterns show similarities, but for most days this is not the case (**geef wat voorbeelden van welke dagen. waar zou het verschil aan kunnen liggen?**). As the radar pixel to rain gauge comparison shows pretty similar results, the differences between interpolated

rain gauge data and radar data are not a result of poor quality of the rain gauges but rather of the scarcity of them. More gauges would likely increase the accuracy of the interpolated images, although numbers would need to increase significantly to properly capture the rainfall patterns within the entire catchment of over 700  $km^2$ .

It is inconvenient to use the interpolated rain gauge data as indication of the real precipitation as for most days it shows little correspondence with the real rainfall patterns throughout the area. As the catchment is very sensitive to the heterogeneity of the precipitation, misrepresentation of the course of the precipitation event would lead to a misrepresentation of the hydrological response. Especially as it has been shown in section 3.3 how big the effect of different rainfall patterns is to modelling of the timing of the peak.



# Conclusion and Recommendations

In this chapter the sub questions divined in the introduction will be answered, after which an answer to the research question will be formulated. Additionally, recommendations for further research will be provided.

## 5.1. Conclusion

The aim of this research was to answer the research question: *"How does distributed precipitation data improve the accuracy of flash flood modelling?"* To get to the answer of this question, three sub questions were investigated.

### **Sub question 1: *How does the catchment respond to precipitation and vary throughout the area?***

Based on the morphology, elevation and land use in the study area, the catchment could be divided into two parts. The Southern part of the catchment is located in the National Park, where elevations are higher, slopes are steeper and the land cover mainly consists of forest. The Northern part of the catchment is significantly flatter, elevation is lower and the main land cover classes are croplands and build up areas. The difference in these geomorphological features between the two areas is also reflected in the hydrological response.

Using observed discharge and precipitation data, it was found that the two different parts in the study area have indeed a different hydrological response to rainfall. Base flow values are lower in the National Park, and corresponding to this the annual runoff values are significantly higher. The response to rainfall is overall faster in the National Park, with higher levels of direct flow, making the area more susceptible to flash floods.

Additionally, from the hydrological catchment characterisation it could be concluded that also within the two parts, the hydrological response is very heterogeneous. At both the upstream and downstream discharge station precipitation events with similar characteristics were found to have different responses in terms of lag time and runoff coefficient, showing the heterogeneity of the precipitation in the study area. Based on this and the radar analysis it is clear that spatially distributed rainfall is key to understand and model river discharge.

### **Sub question 2: *To what extent is the model capable of recreating the hydrological response of the study area and how is this affected by the spatial resolution of the input data?***

The different modelling strategies used to recreate the hydrological response of the study area give a clear insight on the effect of using spatially distributed precipitation as input for the model. In the catchment characterisation, a large spatial variation in precipitation patterns was observed as well as the resulting heterogeneity in the catchments response. This indicates that in order to adequately model the hydrological response of the catchment, uniform precipitation input is not be sufficient. This was confirmed by the calibration strategies used, which were all based on uniform precipitation data. None of the used calibration methods lead to a representative modelling of the observed discharge.

Both timing and magnitude of the modelled discharge peaks were inconclusive and additionally the model was not capable of modelling the base flow in the area.

When looking at the precipitation maps obtained from the interpolated rain gauge data, the spatial variability and localised nature of events becomes more apparent. As precipitation patterns are very heterogeneous, it is as expected that using data of a single rain gauge does not result in a proper modelling of the hydrological response of the catchment. When using the interpolated rain gauge map as input for the model, the importance of precipitation patterns to the hydrological response of the study area becomes more evident. Although modelled discharges are still not an accurate representation of reality, the overall discharge pattern is captured more adequately. It follows the trend of the observed discharge and is capable of capturing some of the base flow in the catchment. The catchment characteristics determined from the modelled discharge roughly capture the hydrological response of the area. The synthetic model runs emphasise the importance of properly capturing the movement and patterns of precipitation in this area, as this greatly influences not only the shape and magnitude of the hydrograph but the timing of the discharge peak as well.

Differences in parameterisation of the model seem only to have an effect on the magnitude of the modelled discharge and not so much on the trend of the hydrograph. This is shown by the catchment characteristics determined from the modelled discharge, where differences in parameterisation have a minimal effect on the determination of the characteristics.

Therefore it can be concluded that the model is capable of capturing and recreating the hydrological response of the study area, but accurate spatially distributed precipitation data is key in obtaining this.

### **Sub question 3: *What will be the effect of using radar data as input for the model?***

When comparing precipitation maps obtained from spatial interpolation of rain gauge data to precipitation maps from the weather radar, big differences in observed precipitation can be seen. Although the observed precipitation by the rain gauges seems to correspond to observed precipitation by the rain radar at that same location, interpolation of these observations does not produce rainfall patterns matching reality.

The difference in observed rainfall patterns of the two methods is not constant, showing great variability over both space and time. Especially on days with high amounts of precipitation observed in part of the catchment differences between the precipitation maps are large, as variations throughout the area are high. As concluded before, the spatial pattern of precipitation is of high importance for satisfactory modelling of the catchments hydrological response. This is important for days with high precipitation amounts, as these are days on which probability of flash floods is high, but also more average days, as they provide a starting point for calibration of the model and are important for obtaining initial conditions of the catchment.

The interpolated rainfall maps show both over- and underestimations of the precipitation in the study area, misinterpreting the spatial pattern of precipitation throughout the study area. Assuming the precipitation observed from the radar data is relatively accurate, this would significantly increase the accuracy of the output of the model, as it better captures the rainfall patterns.

### **Final conclusion**

Using distributed precipitation data is crucial when modelling discharge for the study area of this research. Precipitation patterns in the area are very heterogeneous, meaning uniform data will always lead to wrong interpretation of the study area. Uniform precipitation data results in over-estimations of the discharge magnitude and wrong timing of its peaks, generating very inaccurate representations of the discharge, while distributed data better captures the overall response of the study area.

Distributed precipitation data could be obtained by interpolating precipitation measured at rain gauges, however the density of rain gauges data is too sparse, resulting in low accuracy precipitation maps. This is validated when comparing interpolated rainfall maps with radar imagery. Using radar data as input for the model instead of rain gauge data would offer a good solution to increase the accuracy of the model. Although the accuracy of the rainfall depths measured by the rain radar in Thailand still needs to be tested, its capability of accurately capturing the spatial variation of the rainfall will be a significant improvement for the prediction of the discharge in the area.

In conclusion, using distributed data is a crucial factor for proper modelling of the discharge in the catchment and using rain radar data is a promising method to obtain this data and increase the accuracy of flash flood modelling.

## 5.2. Recommendations

Currently, the model runs on a daily resolution. By using a daily resolution instead of finer resolution, of for example an hour or quarter of an hour, a great deal of information is lost. All observed precipitation and discharge data has a resolution of 15 minutes or one hour, capturing quite detailed information on the timing and shape of the resulting hyetographs and hydrographs. When resampling the data to daily resolution, this detail is lost, leading to less accurately modelled discharge. In addition to information loss, daily resolution is not convenient for forecasting of flash floods and landslides, as these events happen within a few hours after a precipitation event. However, running the model for an hourly resolution costs a lot of computational power, which is also not desired. As during the majority of the year, precipitation amounts are low with minimal chance of flash floods, running the model on daily resolution would be sufficient during these periods, keeping the initial conditions of the catchment updated but not requiring too much computational power. During the wet season, temporal resolution should be increased to hourly, providing for an appropriate modelling of the timing of flash floods and thus ensuring the possibility of timely warning. As flash floods occur within 6 hours, daily resolution would not be sufficient to put warnings in place in time, when flash flood risks would occur.

Furthermore, currently constant and static temperature data is used in the model. From the sensitivity analysis it was found that the modelled discharge is very sensitive to the input temperature. No tests have been conducted using distributed temperature maps, but testing this would be recommended to determine its added value. The effect of changing temperatures over the course of a day should be investigated. For this research, monthly data obtained over multiple years was used as temperature input. Future research should use observed temperature values. The exact resolution of the temperature input should be determined by doing some additional tests.

Moreover, many of the input parameters used are subject to high uncertainties. Although it has been shown that for the modelling of discharge parameterisation only has a limited effect, it is still recommended to do some additional field measurements to improve the input maps used. Differences in parameter sets used showed to influence the magnitude of the modelled discharge, which is important for evaluating whether a flash flood might occur. Additionally, certain parameters might not be relevant for better modelling of discharge, but could play an important role in modelling the probability of slope instability and thus the probability of landslides occurring. As the region of interest is substantial, additional sampling is advisable. While the whole area would not be practical, additional sampling in the national park could be feasible. Since this is the area is most susceptible to flash floods and landslides, proper conceptualisation of this area is more critical than of the downstream area.

Additionally, the current model consists of only one soil layer. Adding more layers to the model could improve results by better capturing the hydrological soil processes and with that an improved conceptualisation of the study area.

Lastly, a more extensive sensitivity analysis and parameter tuning is recommended. Currently, a sensitivity analysis has been done solely based on total annual values of the different fluxes and states in the water balance. However, a sensitivity analysis looking at event characteristics could be of additional value. Factors such as BFI and runoff coefficient could be used, ensuring that calibration is not only performed regarding discharge amounts but also considers the hydrological response of the catchment. Added to this, calibration of the model has been done using uniform precipitation data, whereas calibrating to distributed precipitation could result in better capturing the catchments response. For the spatial calibration performed in this research, only two HRU were identified. Possibly, more could be identified when looking more closely at certain characteristics of the catchment, better capturing the heterogeneous nature of the study area.



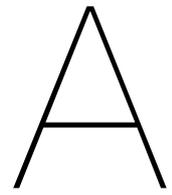
# Bibliography

- Aghakouchak, A., & Habib, E. (2010). Application of a conceptual hydrologic model in teaching hydrologic processes. *International Journal of Engineering Education*, 26(4 (S1)), 963–973.
- Allen, R. G., Pereira, L. S., Raes, D., Smith, M., et al. (1998). Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56. *Fao, Rome*, 300(9), D05109.
- An, N., Hemmati, S., & Cui, Y.-J. (2017). Assessment of the methods for determining net radiation at different time-scales of meteorological variables. *Journal of Rock Mechanics and Geotechnical Engineering*, 9(2), 239–246.
- Bangkok Post. (2020). Pak chong floods worst in 15 years. <https://www.bangkokpost.com/thailand/general/1999923/pak-chong-floods-worst-in-15-years>
- Beck, H. E., Van Dijk, A. I., Miralles, D. G., De Jeu, R. A., Bruijnzeel, L., McVicar, T. R., & Schellekens, J. (2013). Global patterns in base flow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research*, 49(12), 7843–7863.
- Bidorn, B., Chanyotha, S., Kish, S. A., Donoghue, J. F., Bidorn, K., & Mama, R. (2015). The effects of thailand's great flood of 2011 on river sediment discharge in the upper chao phraya river basin, thailand. *International Journal of Sediment Research*, 30(4), 328–337.
- Bogaard, T. A., & Greco, R. (2016). Landslide hydrology: From hydrology to pore pressure. *Wiley Interdisciplinary Reviews: Water*, 3(3), 439–459.
- Borga, M., Anagnostou, E., Blöschl, G., & Creutin, J.-D. (2011). Flash flood forecasting, warning and risk management: The hydrate project. *Environmental Science & Policy*, 14(7), 834–844.
- Budimir, M., Uprety, D. R., Ares, M., & Svensson, A. (2021). How do you build an effective early warning system? <https://www.preventionweb.net/news/how-do-you-build-effective-early-warning-system> (accessed: "11.04.2022")
- Campbell, R. (2017). Introduction to the flash flood guidance (ffg) system for southeast asia. *Hydrological Research Center*. [https://wmoomm.sharepoint.com/sites/wmocpdb/eve\\_activityarea/Forms/AllItems.aspx?id=%5C%2Fsites%5C%2Fwmoocpdb%5C%2Feve%5C%5Factivityarea%5C%2FHydrology%5C%20and%5C%20Water%5C%20Resources%5C%20Programme%5C%20%5C%28HWRP%5C%29%5C%5F1a994a92%5C%2D1373%5C%2De911%5C%2Da965%5C%2D000d3a396ff4%5C%2FHRC01%5C%5FOverview%5C%5FRG%5C%5F11202017%5C%2Epdf&parent=%5C%2Fsites%5C%2Fwmoocpdb%5C%2Feve%5C%5Factivityarea%5C%2FHydrology%5C%20and%5C%20Water%5C%20Resources%5C%20Programme%5C%20%5C%28HWRP%5C%29%5C%5F1a994a92%5C%2D1373%5C%2De911%5C%2Da965%5C%2D000d3a396ff4&p=true](https://wmoomm.sharepoint.com/sites/wmocpdb/eve_activityarea/Forms/AllItems.aspx?id=%5C%2Fsites%5C%2Fwmoocpdb%5C%2Feve%5C%5Factivityarea%5C%2FHydrology%5C%20and%5C%20Water%5C%20Resources%5C%20Programme%5C%20%5C%28HWRP%5C%29%5C%5F1a994a92%5C%2D1373%5C%2De911%5C%2Da965%5C%2D000d3a396ff4%5C%2FHRC01%5C%5FOverview%5C%5FRG%5C%5F11202017%5C%2Epdf&parent=%5C%2Fsites%5C%2Fwmoocpdb%5C%2Feve%5C%5Factivityarea%5C%2FHydrology%5C%20and%5C%20Water%5C%20Resources%5C%20Programme%5C%20%5C%28HWRP%5C%29%5C%5F1a994a92%5C%2D1373%5C%2De911%5C%2Da965%5C%2D000d3a396ff4&p=true) (accessed: "17.01.2022")
- Chappell, N. A., Sherlock, M., Bidin, K., Macdonald, R., Najman, Y., & Davies, G. (2007). Runoff processes in southeast asia: Role of soil, regolith, and rock type. *Forest environments in the Mekong River basin*, 3–23.
- Chen, A., Giese, M., & Chen, D. (2020). Flood impact on mainland southeast asia between 1985 and 2018—the role of tropical cyclones. *Journal of Flood Risk Management*, 13(2), e12598.
- Cheng, G. (2021). Asia's floods: Once a blessing, now a curse? <https://kontinentalist.com/stories/climate-change-has-made-asias-flooding-a-natural-disaster> (accessed: "17.01.2022")
- Connors, D. (2019). What causes landslides and mudslides? <https://earthsky.org/human-world/what-causes-landslides/> (accessed: "20.01.2022")
- De Roo, A., Wesseling, C., & Ritsema, C. (1996). Lisem: A single-event physically based hydrological and soil erosion model for drainage basins. i: Theory, input and output. *Hydrological processes*, 10(8), 1107–1117.
- Department of Mineral Resources. (n.d.). Landslides in thailand. [http://www.dmr.go.th/main.php?filename=landslide\\_En](http://www.dmr.go.th/main.php?filename=landslide_En) (accessed: "20.01.2022")
- Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A review on hydrological models. *Aquatic Procedia*, 4, 1001–1007.

- DisasterAWARE. (2011). 9.5 million people experience flooding in southeast asia. [https://disasteraware.com/weather\\_wall/9-5-million-people-experience-flooding-in-southeast-asia/](https://disasteraware.com/weather_wall/9-5-million-people-experience-flooding-in-southeast-asia/) (accessed: "11.01.2022")
- Dobos, E. (2003). Albedo.
- Eckhardt, K. (2005). How to construct recursive digital filters for baseflow separation. *Hydrological Processes: An International Journal*, 19(2), 507–515.
- Feiken, H. (2014). *Dealing with biases: Three geo-archaeological approaches to the hidden landscapes of italy* (Vol. 26). Barkhuis.
- Floodlist. (n.d.). Asia. <https://floodlist.com/asia> (accessed: "12.01.2022")
- Floodlist. (2020). Thailand – thousands affected by floods. *Floodlist News*. <https://floodlist.com/asia/thailand-floods-october-2020>
- Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161–2181.
- Fuster, B., Sánchez-Zapero, J., Camacho, F., García-Santos, V., Verger, A., Lacaze, R., Weiss, M., Baret, F., & Smets, B. (2020). Quality assessment of proba-v lai, fapar and fcover collection 300 m products of copernicus global land service. *Remote Sensing*, 12(6), 1017.
- Goel, M. K. (2011). Recession coefficient. In V. P. Singh, P. Singh, & U. K. Haritashya (Eds.), *Encyclopedia of snow, ice and glaciers* (pp. 922–922). Springer Netherlands. [https://doi.org/10.1007/978-90-481-2642-2\\_436](https://doi.org/10.1007/978-90-481-2642-2_436)
- Habert, J., Ricci, S., Le Pape, E., Thual, O., Piacentini, A., Goutal, N., Jonville, G., & Rochoux, M. (2016). Reduction of the uncertainties in the water level-discharge relation of a 1d hydraulic model in the context of operational flood forecasting. *Journal of Hydrology*, 532, 52–64.
- Hall, F. R. (1968). Base-flow recessions—a review. *Water resources research*, 4(5), 973–983.
- Hapuarachchi, H., Wang, Q., & Pagano, T. (2011). A review of advances in flash flood forecasting. *Hydrological Processes*, 25(18), 2771–2784.
- Hirota, T. (2001). Estimation of seasonal and annual evaporation using agrometeorological data from the thai meteorological department by the heat budget models. *Journal of the Meteorological Society of Japan. Ser. II*, 79(1B), 365–371.
- Ikeda, M., & Palakhamarn, T. (2020). Economic damage from natural hazards and local disaster management plans in japan and thailand.
- IUSS Working Group WRB. (2015). World reference base for soil resources 2014, update 2015; international soil classification system for naming soils and creating legends for soil maps. *World soil resources report no. 106*. FAO, Rome.
- Jain, M. K. (2011). Recession of discharge. In V. P. Singh, P. Singh, & U. K. Haritashya (Eds.), *Encyclopedia of snow, ice and glaciers* (pp. 922–924). Springer Netherlands. [https://doi.org/10.1007/978-90-481-2642-2\\_437](https://doi.org/10.1007/978-90-481-2642-2_437)
- Jha, R., Herath, S., & Musiaka, K. (2000). Application of iis distributed hydrological model (iisdhm) in nakhon sawan catchment, thailand. *Journal of Hydroscience and Hydraulic Engineering*, 18(1), 21–28.
- Kása, I., Gelybó, G., Horel, Á., Bakacsi, Z., Tóth, E., Koós, S., Dencsó, M., Deelstra, J., Molnár, S., & Farkas, C. (2017). Evaluation of three semi-distributed hydrological models in simulating discharge from a small forest and arable dominated catchment. *Biologia*, 72(9), 1002–1009.
- Kumar, D., & Bhattacharjya, R. K. (2020). Evaluating two gis-based semi-distributed hydrological models in the bhagirathi-alkhnanda river catchment in india. *Water Policy*, 22(6), 991–1014.
- Land Development Department. (2011a). General climatic condition. [https://www.idd.go.th/web\\_eng56/Soil\\_Resource/Physial\\_Environment\\_of\\_Thailand/General\\_Climatic\\_condition.html](https://www.idd.go.th/web_eng56/Soil_Resource/Physial_Environment_of_Thailand/General_Climatic_condition.html) (accessed: "14.09.2021")
- Land Development Department. (2011b). Rainfall. [https://www.idd.go.th/web\\_eng56/Soil\\_Resource/Physial\\_Environment\\_of\\_Thailand/Rainfall.html](https://www.idd.go.th/web_eng56/Soil_Resource/Physial_Environment_of_Thailand/Rainfall.html) (accessed: "15.09.2021")
- Land Development Department. (2011c). Season. [https://www.idd.go.th/web\\_eng56/Soil\\_Resource/Physial\\_Environment\\_of\\_Thailand/Season.html](https://www.idd.go.th/web_eng56/Soil_Resource/Physial_Environment_of_Thailand/Season.html) (accessed: "15.09.2021")
- Land Development Department. (2011d). Surface temperature. [https://www.idd.go.th/web\\_eng56/Soil\\_Resource/Physial\\_Environment\\_of\\_Thailand/Surface\\_Temperature.html](https://www.idd.go.th/web_eng56/Soil_Resource/Physial_Environment_of_Thailand/Surface_Temperature.html) (accessed: "15.09.2021")
- Lenhart, T., Eckhardt, K., Fohrer, N., & Frede, H.-G. (2002). Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth, Parts A/B/C*, 27(9-10), 645–654.

- Liechti, K., Panziera, L., Germann, U., & Zappa, M. (2013). The potential of radar-based ensemble forecasts for flash-flood early warning in the southern swiss alps. *Hydrology and Earth System Sciences*, 17(10), 3853–3869.
- Lu, J., Sun, G., McNulty, S. G., & Amatya, D. M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern united states 1. *JAWRA Journal of the American Water Resources Association*, 41(3), 621–633.
- Mapiam, P. P., Methaprayun, M., Bogaard, T., Schoups, G., & Ten Veldhuis, M.-C. (2022). Citizen rain gauge improves hourly radar rainfall bias correction using a two-step kalman filter. *Hydrology and Earth System Sciences*, 26(3), 775–794.
- Marino, P., Peres, D. J., Cancelliere, A., Greco, R., & Bogaard, T. A. (2020). Soil moisture information can improve shallow landslide forecasting using the hydrometeorological threshold approach. *Landslides*, 17(9).
- Marra, F., Nikolopoulos, E. I., Creutin, J. D., & Borga, M. (2014). Radar rainfall estimation for the identification of debris-flow occurrence thresholds. *Journal of Hydrology*, 519, 1607–1619.
- Masaki, Y., Ishigooka, Y., Kuwagata, T., Goto, S., Sawano, S., & Hasegawa, T. (2011). Expected changes in future agro-climatological conditions in northeast thailand and their differences between general circulation models. *Theoretical and Applied Climatology*, 106(3), 383–401.
- Merzdorf, J. (2020). Climate change could trigger more landslides in high mountain asia. <https://climate.nasa.gov/news/2951/climate-change-could-trigger-more-landslides-in-high-mountain-asia/> (accessed: "01.09.2022")
- Moore, R. J., Cole, S. J., & Bell, V. A. (2006). Issues in flood forecasting: Ungauged basins. *Frontiers in Flood Research*, (305), 103.
- Mwakalila, S., Feyen, J., & Wyseure, G. (2002). The influence of physical catchment properties on baseflow in semi-arid environments. *Journal of Arid Environments*, 52(2), 245–258.
- NASA JPL. (2013). *Nasa shuttle radar topography mission global 1 arc second* ([Data set]). NASA EOSDIS Land Processes DAAC. <https://doi.org/https://doi.org/10.5067/MEaSURES/SRTM/SRTMGL1.003>
- Padiyedath, S. G., Kawamura, A., Amaguchi, H., & Azhikodan, G. (2017). Baseflow estimation for tropical wet and dry climate region using recursive digital filters. *Geographical Modelling*, 73(5), 1\_9–1\_16.
- Pinker, R. T., Thompson, O. E., & Eck, T. F. (1980). The albedo of a tropical evergreen forest. *Quarterly Journal of the Royal Meteorological Society*, 106(449), 551–558.
- Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). Soilgrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *Soil*, 7(1), 217–240.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C. E., et al. (2021). Mapping global forest canopy height through integration of gedi and landsat data. *Remote Sensing of Environment*, 253, 112165. <https://doi.org/https://doi.org/10.1016/j.rse.2020.112165>
- Price, K. (2011). Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review. *Progress in physical geography*, 35(4), 465–492.
- Prudden, R., Adams, S., Kangin, D., Robinson, N., Ravuri, S., Mohamed, S., & Arribas, A. (2020). A review of radar-based nowcasting of precipitation and applicable machine learning techniques. *arXiv preprint arXiv:2005.04988*.
- Rao, L., Sun, G., Ford, C., & Vose, J. (2011). Modeling potential evapotranspiration of two forested watersheds in the southern appalachians. *Transactions of the ASABE*, 54(6), 2067–2078.
- Sapač, K., Vidmar, A., Bezak, N., & Rusjan, S. (2020). Lag times as indicators of hydrological mechanisms responsible for no3-n flushing in a forested headwater catchment. *Water*, 12(4), 1092.
- Schmidt-Thomé, P., Tatong, T., Kunthasap, P., & Wathanapridc, S. (2018). Community based landslide risk mitigation in thailand. *Episodes*, 41(4), 225–233.
- Searcy, J. K., & Hardison, C. H. (1960). *Double-mass curves*. US Government Printing Office.
- Shah, S. M. H., Mustafa, Z., Teo, F. Y., Imam, M. A. H., Yusof, K. W., & Al-Qadami, E. H. H. (2020). A review of the flood hazard and risk management in the south asian region, particularly pakistan. *Scientific African*, e00651.
- Silva, J. S., & Rego, F. C. (2003). Root distribution of a mediterranean shrubland in portugal. *Plant and Soil*, 255(2), 529–540.

- Singh, K., & Kumar, V. (2021). Rainfall thresholds triggering landslides: A review. *Sustainable Environment and Infrastructure*, 455–464.
- Thai Meteorological Department. (2015). The climate of thailand. <https://www.tmd.go.th/en/downloads.php> (accessed: "18.10.2021")
- The Nation. (2020). Khao yai closed this weekend to clear storm damage. *The Nation Thailand*. <https://www.nationthailand.com/in-focus/30395921>
- Torti, J. (2012). Floods in southeast asia: A health priority. *Journal of Global Health*, 2(2).
- Trelo-ges, V., & Sriboonlue, V. (2002). Study on physical and hydraulic properties of major soil series in northeast, thailand. *Warasan Wichakan Kaset*.
- UN/ISDR. (2006). *EWC III Third International Conference On Early Warning: Developing Early Warning Systems: A Checklist* (tech. rep.). UN Inter-Agency Secretariat of the International Strategy for Disaster Reduction.
- UNISDR, C. et al. (2015a). The human cost of natural disasters: A global perspective.
- UNISDR, C. et al. (2015b). The human cost of weather related disasters: 1995-2015.
- United Nations. (n.d.-a). Early warning systems. <https://www.un.org/en/climatechange/climate-solutions/early-warning-systems> (accessed: "01.07.2021")
- United Nations. (n.d.-b). What is climate change? <https://www.un.org/en/climatechange/what-is-climate-change> (accessed: "01.09.2022")
- Wang, W., Lu, H., Yang, D., Sothea, K., Jiao, Y., Gao, B., Peng, X., & Pang, Z. (2016). Modelling hydrologic processes in the mekong river basin using a distributed model driven by satellite precipitation and rain gauge observations. *PLoS one*, 11(3), e0152229.
- Wieczorek, G. F. (1996). Landslides: Investigation and mitigation. chapter 4-landslide triggering mechanisms. *Transportation Research Board Special Report*, (247).
- Xu, X., Sun, C., Huang, G., & Mohanty, B. P. (2016). Global sensitivity analysis and calibration of parameters for a physically-based agro-hydrological model. *Environmental Modelling & Software*, 83, 88–102.
- Yang, W., Xiao, C., Zhang, Z., & Liang, X. (2021). Can the two-parameter recursive digital filter baseflow separation method really be calibrated by the conductivity mass balance method? *Hydrology and Earth System Sciences*, 25(4), 1747–1760.
- Zhao, C., & Lu, Z. (2018). Remote sensing of landslides—a review. *Remote Sensing*, 10(2), 279.
- Zheng, C., Jia, L., Hu, G., & Lu, J. (2019). Earth observations-based evapotranspiration in northeastern thailand. *Remote Sensing*, 11(2), 138.



## Station Locations

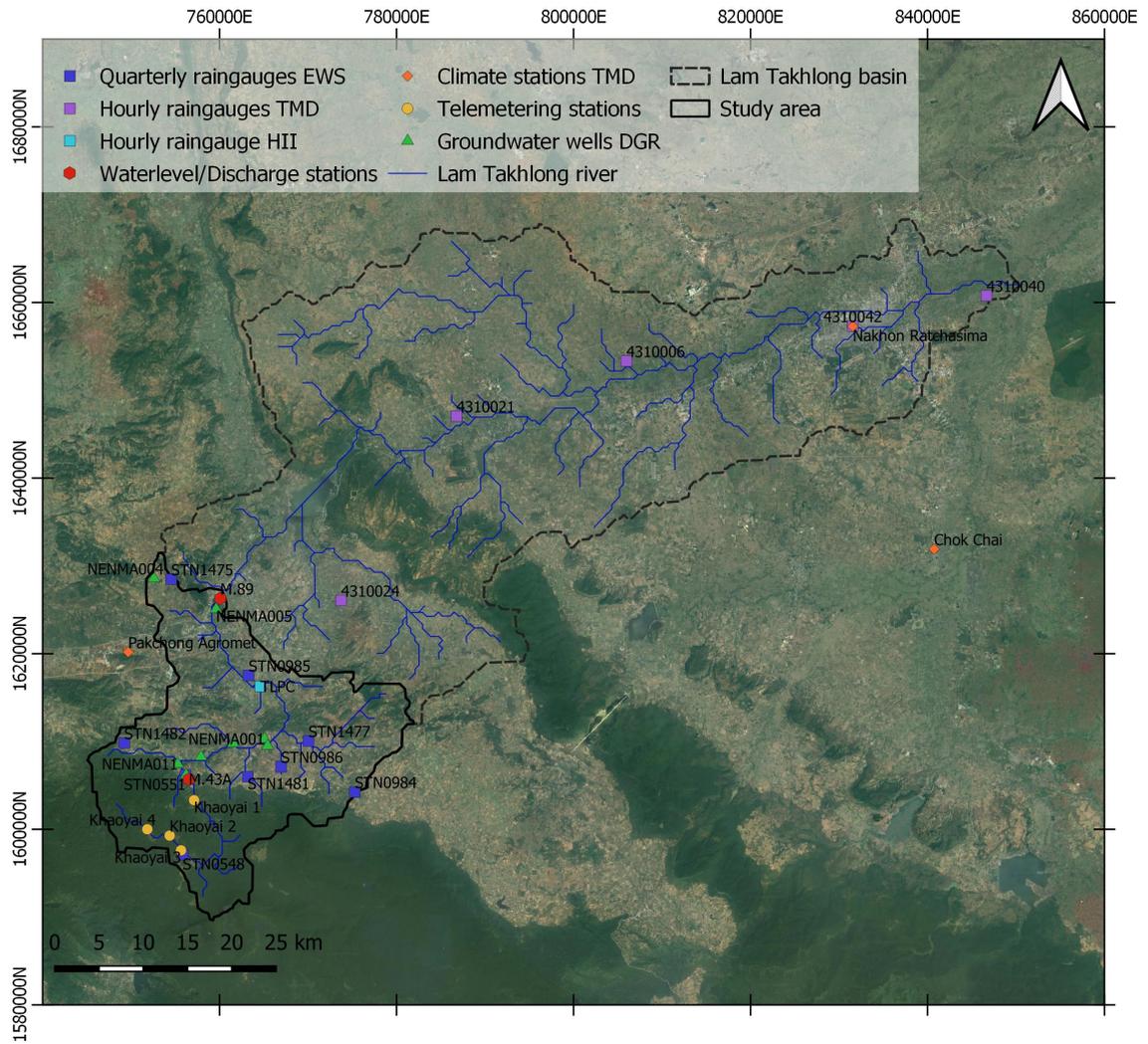


Figure A.1: Locations of stations in and around the study area.

# B

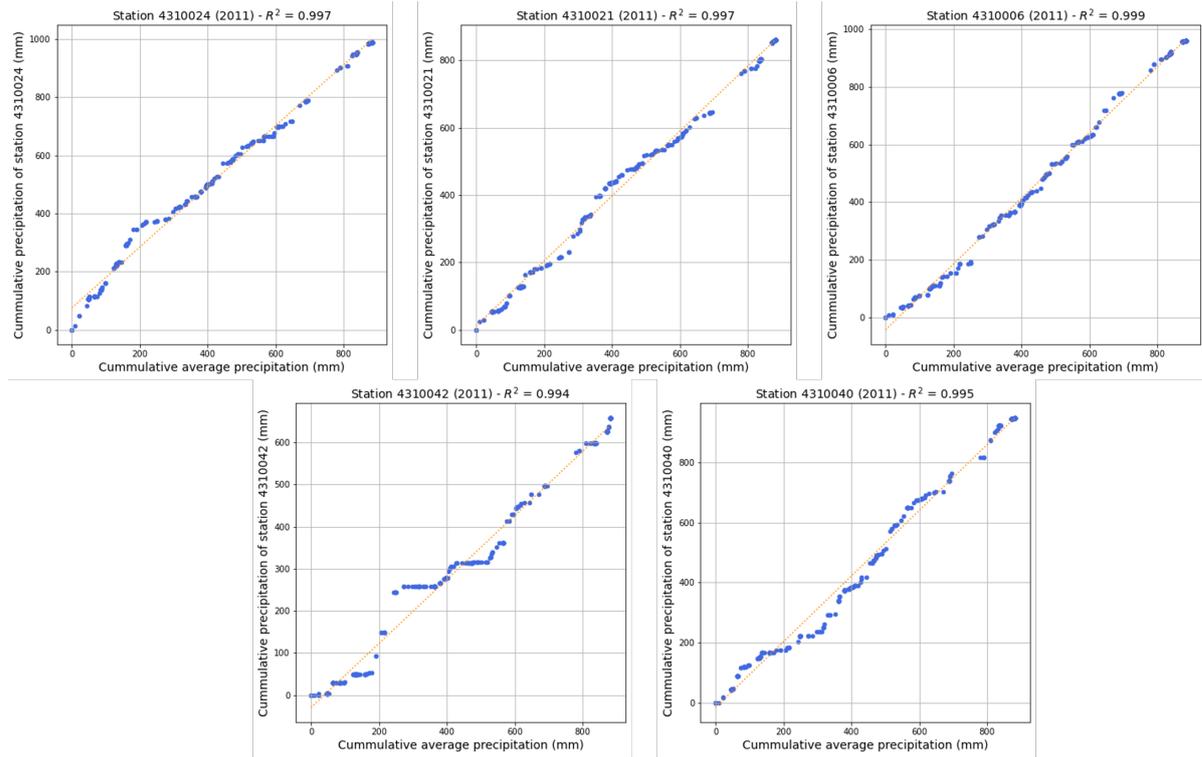
## Data consistency

### B.1. Missing Data

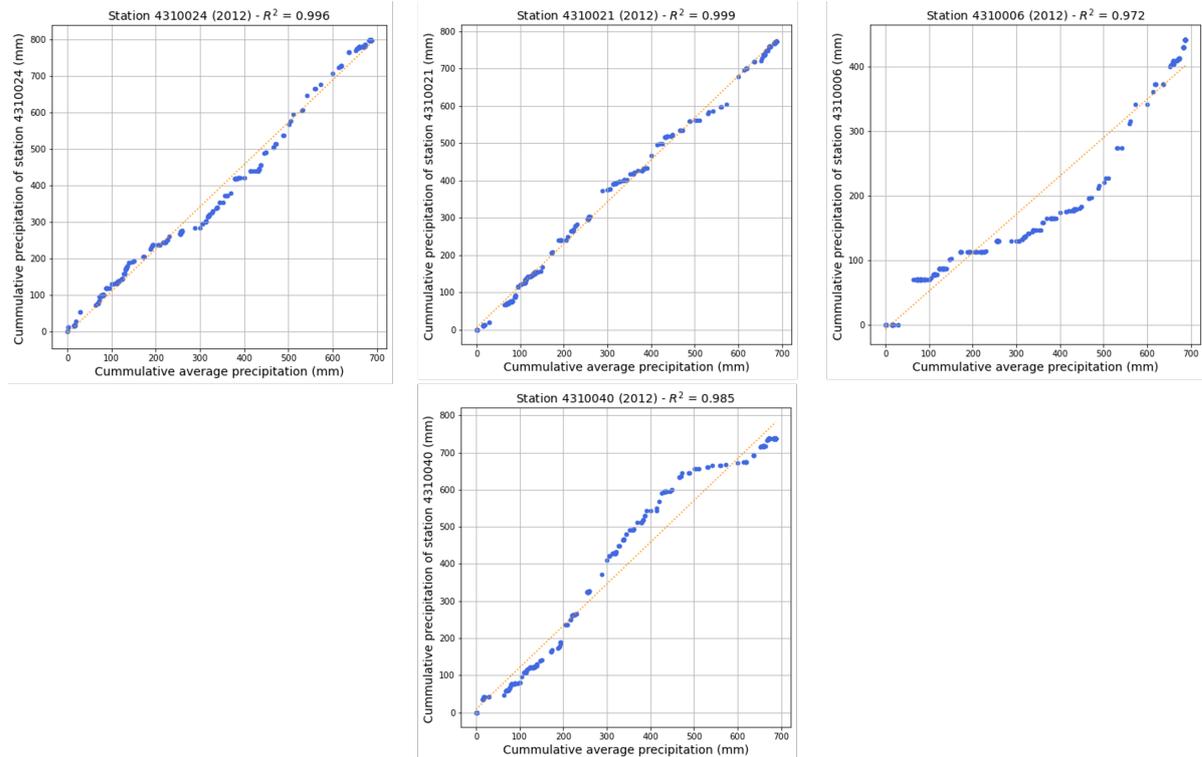
	Discharge		Precipitation																	
	RID		HII	TMD						EWS										
	M89	M43a		TLPC	431002	4431002	1431000	6431004	2431004	0	STN0548	STN0551	STN0984	STN0985	STN0986	STN0987	STN1475	STN1477	STN1481	STN1482
2011	41.6%	0.0%	100.0%	0.4%	0.5%	0.1%	49.9%	0.5%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2012	5.5%	3.0%	100.0%	0.4%	0.2%	43.6%	30.6%	23.7%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2013	14.1%	0.0%	1.6%	3.2%	1.6%	16.6%	86.2%	31.2%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2014	0.4%	0.0%	2.5%	7.7%	8.0%	53.7%	63.6%	8.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2015	0.0%	0.0%	3.4%	61.4%	63.6%	66.0%	59.8%	61.4%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2016	34.8%	16.7%	0.6%	5.1%	14.0%	47.3%	90.5%	11.9%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2017	62.2%	0.0%	0.3%	5.8%	39.2%	11.9%	100.0%	5.8%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2018	22.8%	0.0%	17.6%	56.4%	95.1%	56.3%	100.0%	66.3%	2.8%	5.2%	7.1%	2.2%	2.2%	4.6%	0.5%	12.0%	4.6%	0.7%		
2019	0.0%	0.0%	1.3%	26.6%	12.0%	45.8%	100.0%	29.6%	20.4%	14.3%	19.4%	41.5%	51.8%	13.2%	24.8%	13.8%	22.5%	13.1%		
2020	0.0%	0.0%	0.6%	100.0%	100.0%	100.0%	100.0%	100.0%	8.4%	32.1%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
2021	33.6%	75.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	5.4%	9.8%	0.3%	4.7%	15.0%	0.2%	100.0%	2.8%	2.2%	3.8%		

Table B.1: Missing data per station per year. Green values indicate 0 - 20% of the data is missing, yellow means 20 - 40% is missing, for orange 40 - 60% is missing, light red implies 60 - 80% is missing and for dark red 80 - 100% is missing.

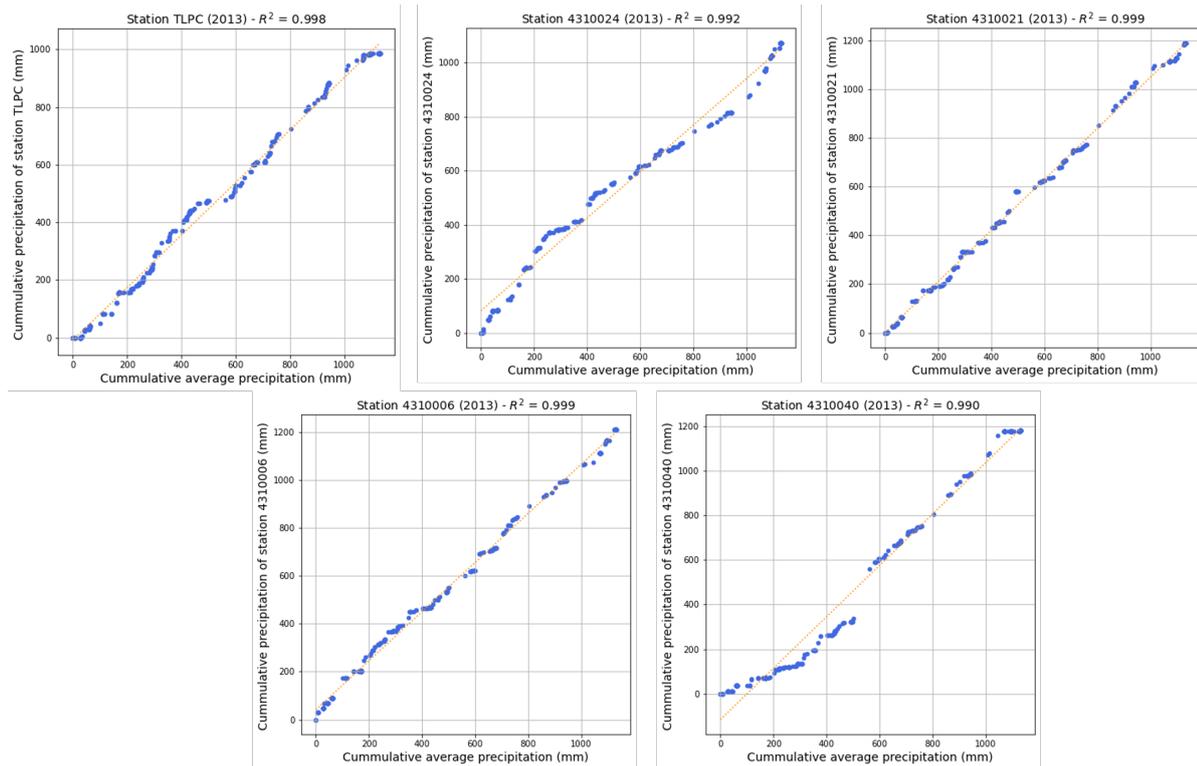
## B.2. Double Mass Curves



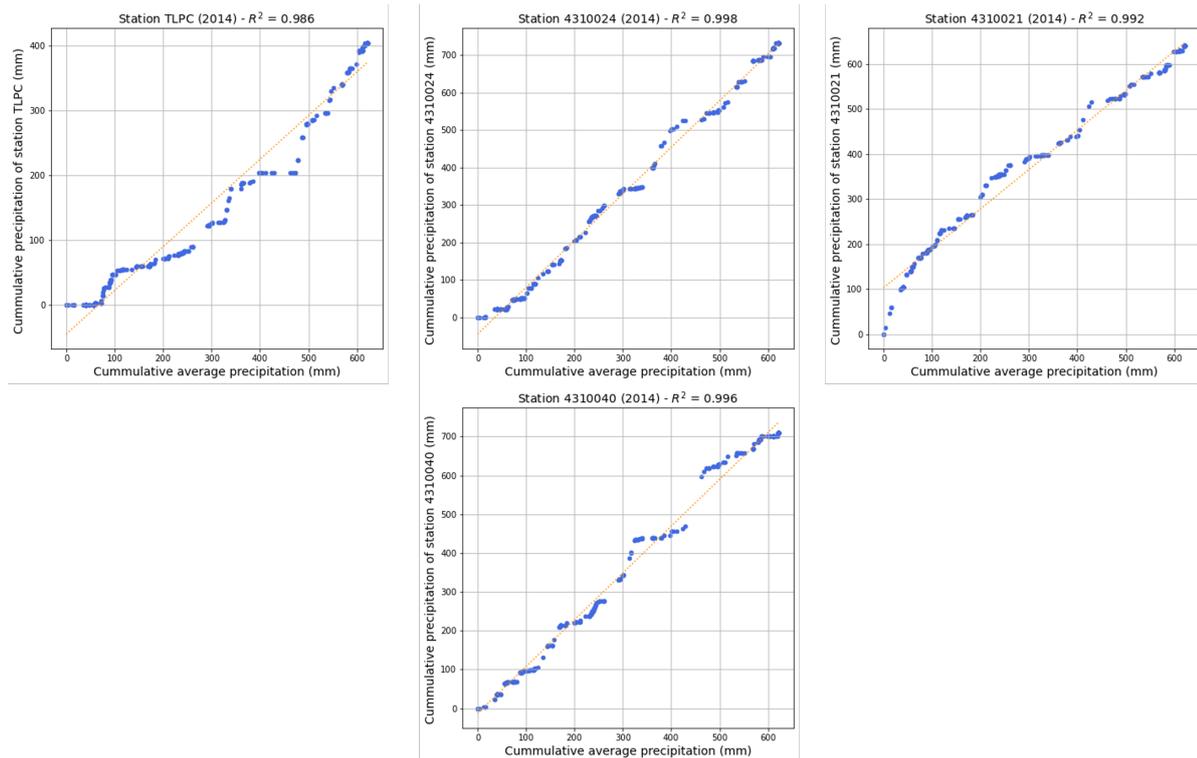
(a) 2011



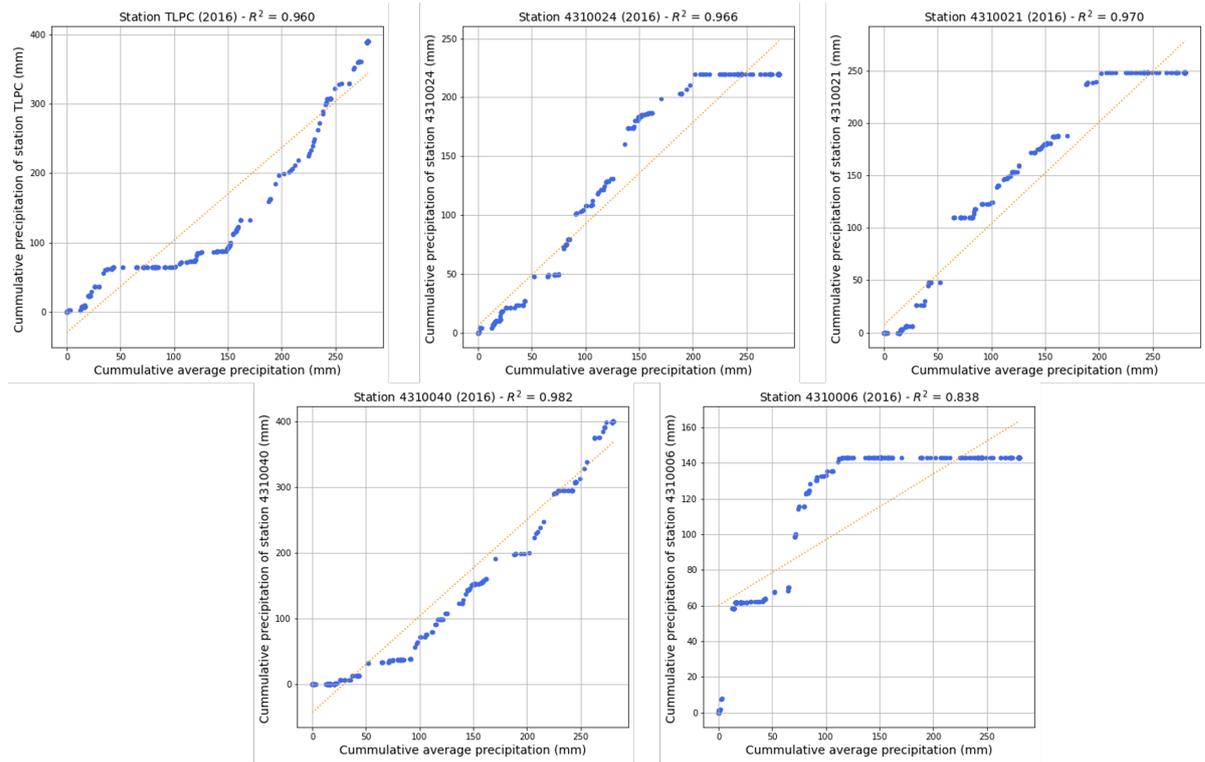
(b) 2012



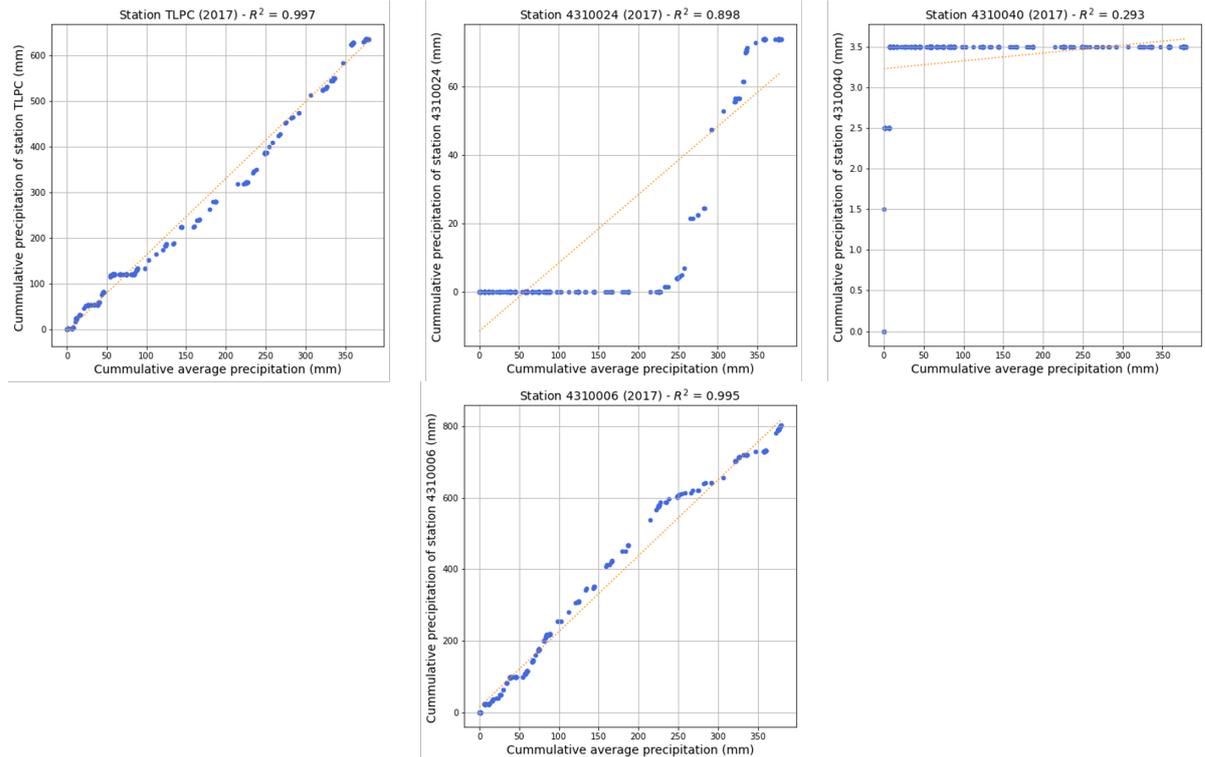
(c) 2013



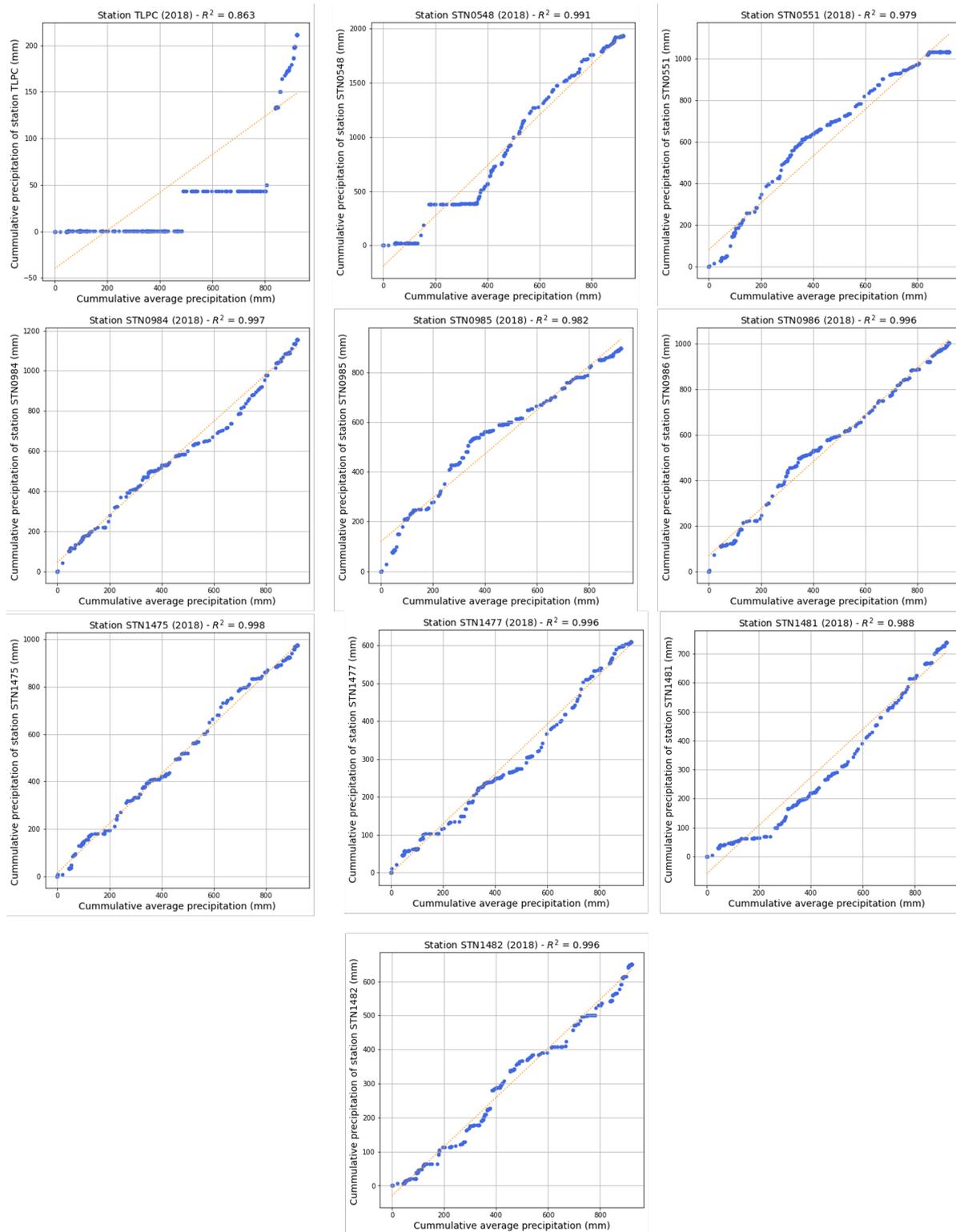
(d) 2014



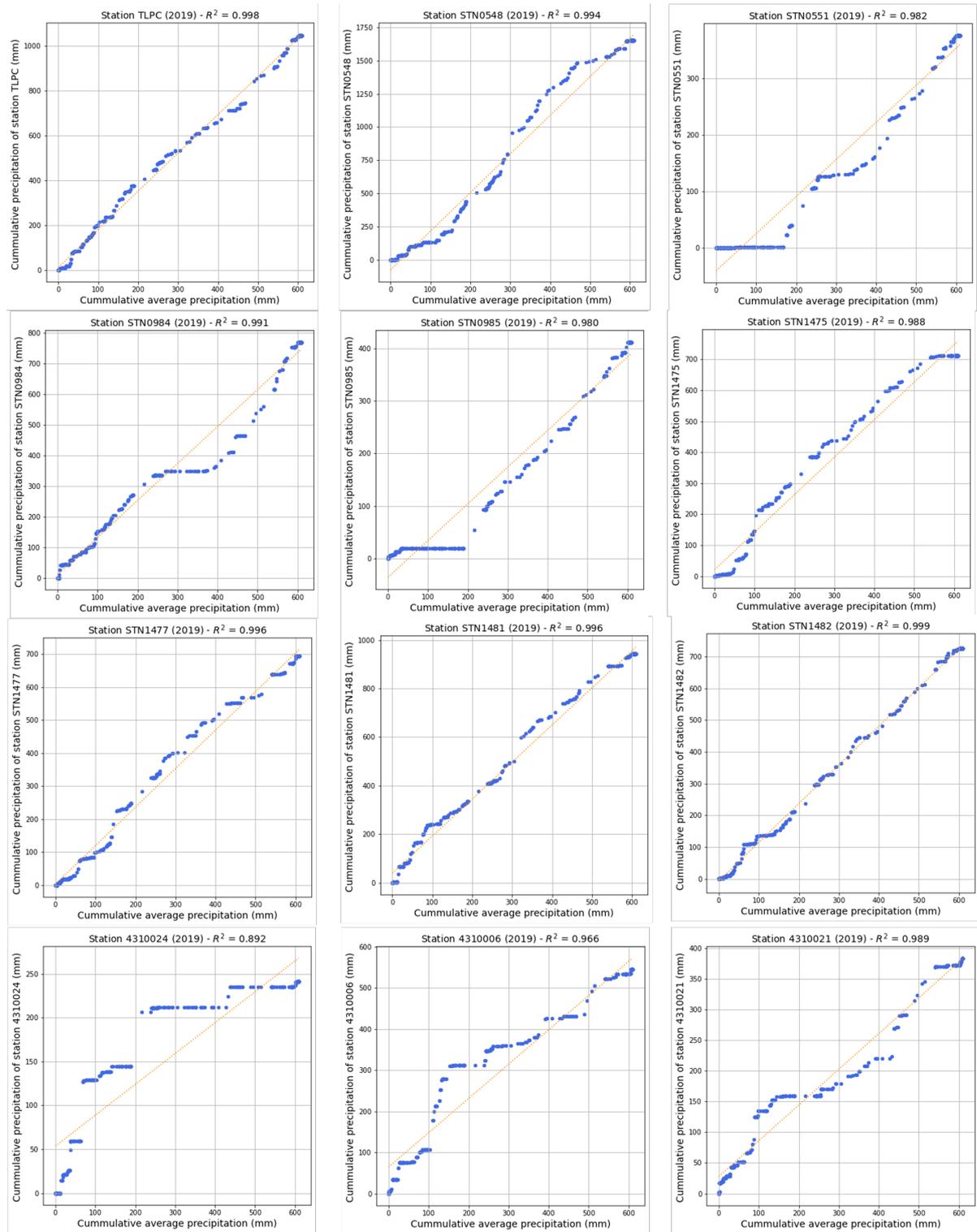
(e) 2016



(f) 2017



(g) 2018



(h) 2019

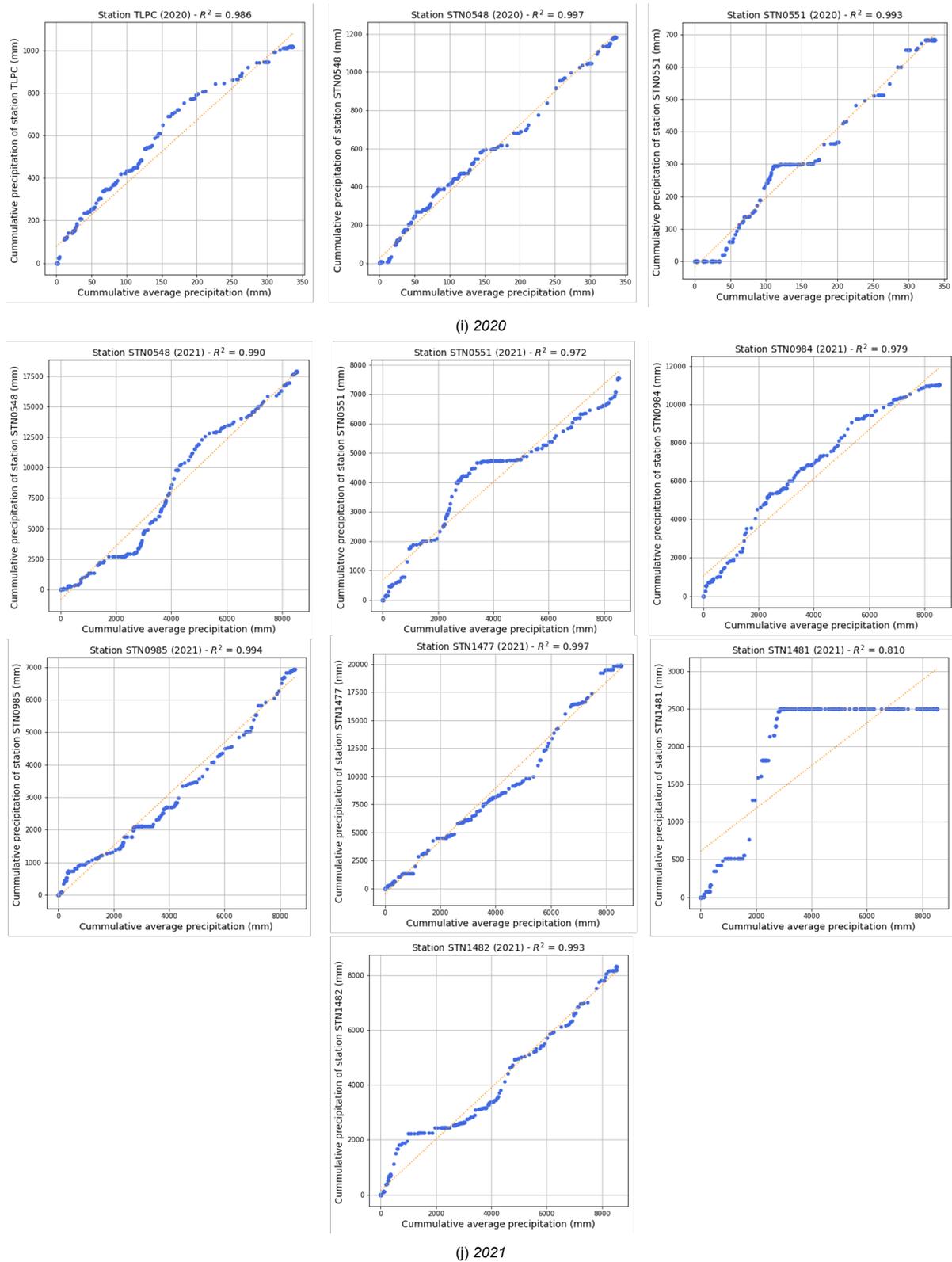


Figure B.1: Double Mass Curves of the rain gauges for the different years.



# C

## Base Flow

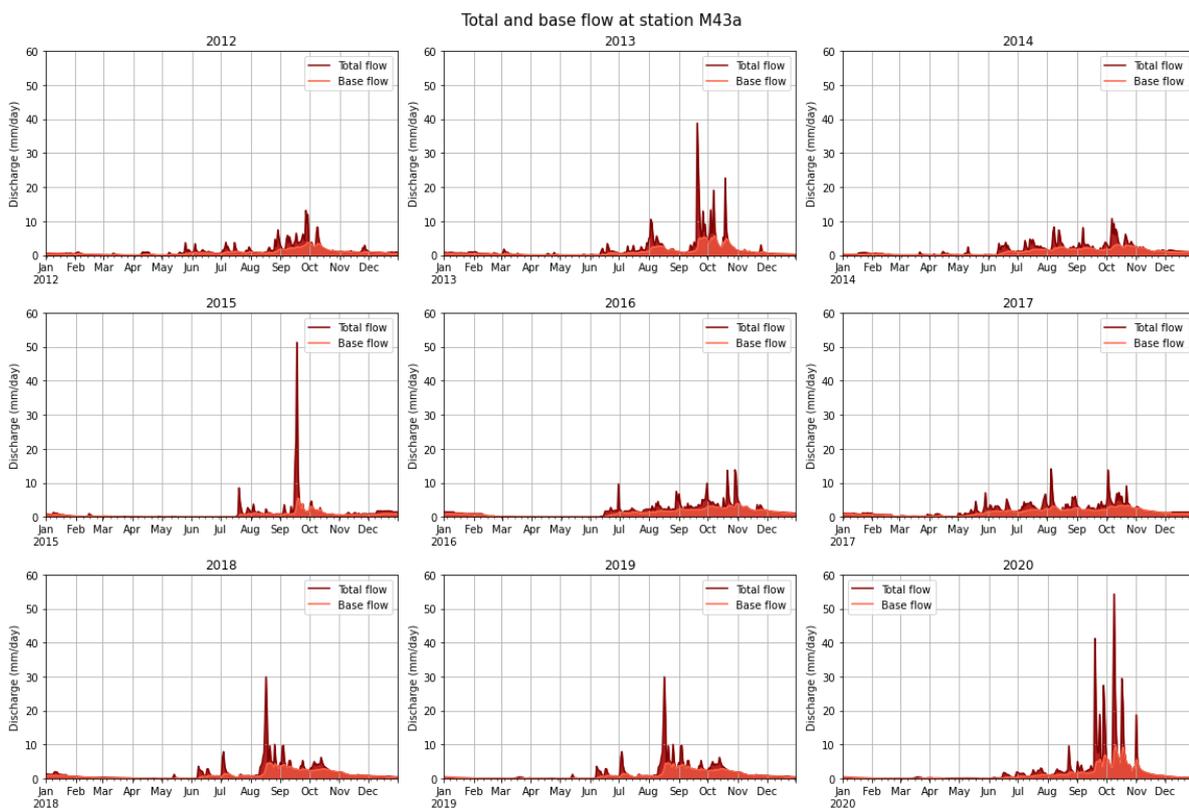


Figure C.1: Total flow measured at discharge station M43a and the corresponding baseflow as determined with the Eckhardt Algorithm where  $\alpha$  is 0.999 and  $BF I_{max}$  is 0.7 for the years 2012 until 2020.

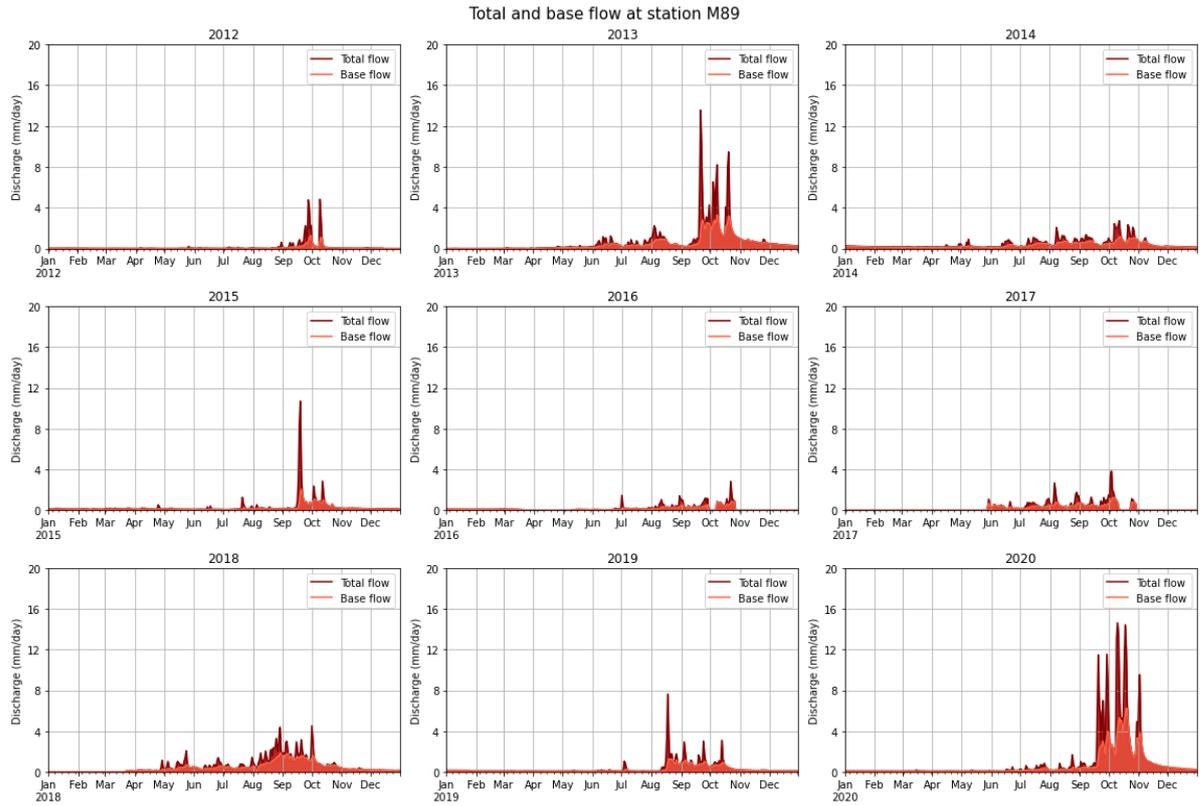
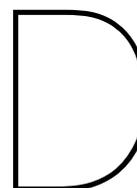


Figure C.2: Total flow measured at discharge station M89 and the corresponding baseflow as determined with the Eckhardt Algorithm where  $\alpha$  is 0.999 and  $BFI_{max}$  is 0.8 for the years 2012 until 2020.



# Lag Time and Runoff Coefficient

## D.1. Event overview

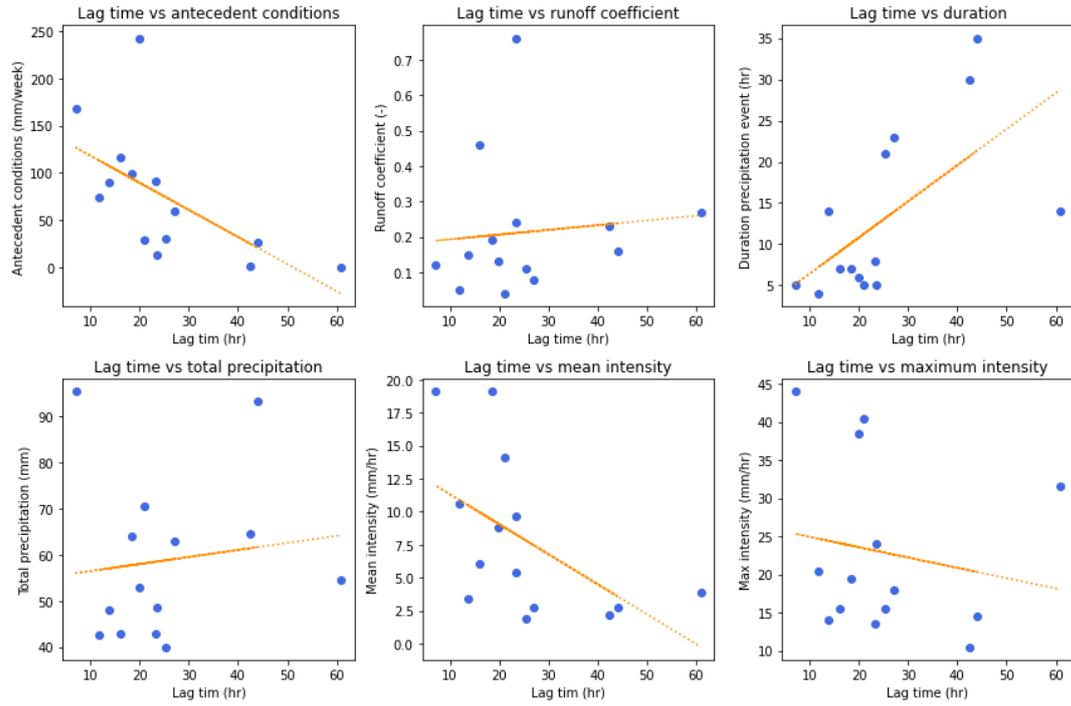
		Precipitation						Discharge		Lag Time (hr)	Runoff coefficient (-)
		Start Event	Duration (hr)	Total precipitation (mm)	Mean intensity (mm/hr)	Max intensity (mm/hr)	Antecedent conditions (mm/week)	Start Peak	Total discharge (mm)		
STN0551	Event 1	02-07-19 17:00	30	64.5	2.2	10.5	0.5	3-07-19 11:00	14.92	42.4	0.23
	Event 2	23-09-19 3:00	21	40.0	1.9	15.5	30	23-09-19 15:00	4.38	25.4	0.11
	Event 3	27-09-20 17:00	5	48.5	9.7	24	12.5	27-09-20 21:00	36.93	23.4	0.76
STN0548	Event 4	20-06-18 21:00	4	42.5	10.6	20.5	73.5	21-06-18 3:00	2.01	11.8	0.05
	Event 5	02-07-18 17:00	5	70.5	14.1	40.5	29.0	2-07-18 17:00	2.62	21.0	0.04
	Event 6	19-07-18 20:00	23	63.0	2.7	18	59.0	20-07-18 4:00	4.86	27.1	0.08
	Event 7	18-08-18 14:00	7	64.0	9.1	19.5	99.5	18-08-18 17:00	12.17	18.5	0.19
	Event 8	27-08-18 14:00	7	43.0	6.1	15.5	116.5	27-08-18 17:00	19.62	16.0	0.46
	Event 9	03-09-18 17:00	8	43.0	5.4	13.5	91	4-09-18 0:00	10.35	23.3	0.24
	Event 10	02-07-19 14:00	35	93.5	2.7	14.5	26.5	3-07-19 11:00	14.92	44.0	0.16
	Event 11	16-08-19 0:00	5	95.5	19.1	44	167.5	16-08-19 3:00	11.85	7.1	0.12
	Event 12	20-08-19 16:00	6	53	8.8	38.5	242.5	20-08-19 20:00	6.61	19.9	0.12
	Event 13	25-08-19 16:00	10	48	4.2	14	89.5	25-08-19 19:00	7.21	13.7	0.15
	Event 14	8-10-20 2:00	65	194.5	3.0	12	42.5	8-10-20 11:00	90.59	-2.7	0.47
	Event 15	29-10-20 18:00	14	54.5	3.9	31.5	0	31-10-20 19:00	14.5	60.9	0.27

Table D.1: Information on the precipitation events used to determine the lag time for direct runoff peaks at discharge station M43a.

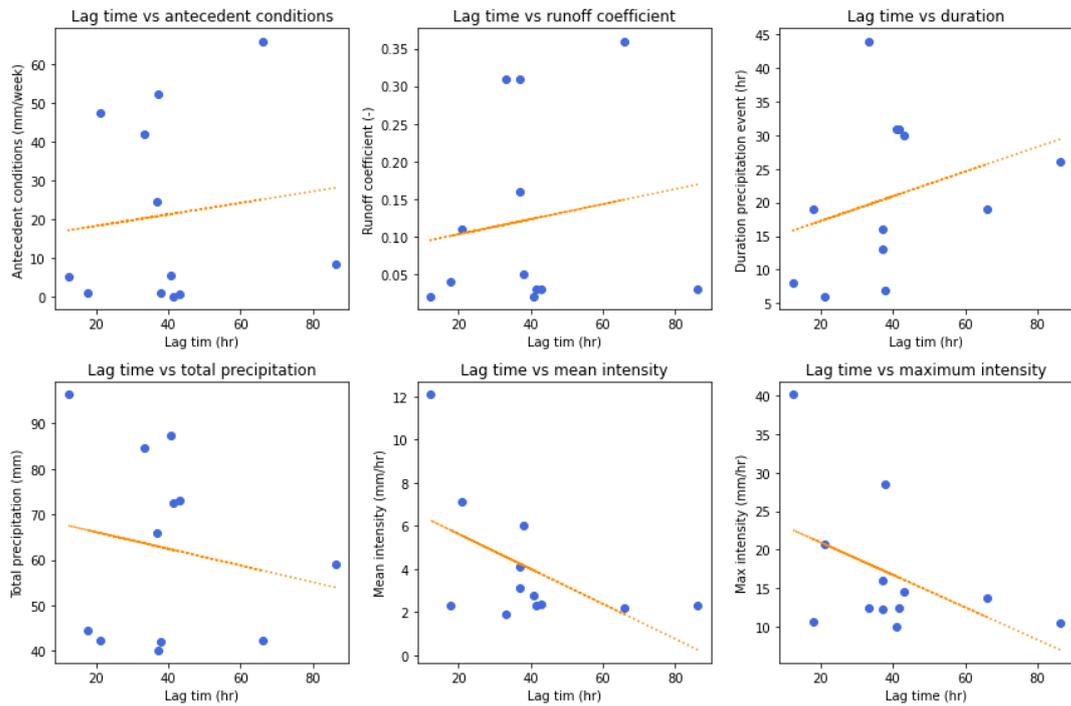
		Precipitation						Discharge		Lag Time (hr)	Runoff coefficient (-)
		Start Event	Duration (hr)	Total precipitation (mm)	Mean intensity (mm/hr)	Max intensity (mm/hr)	Antecedent conditions (mm/week)	Start Peak	Total discharge (mm)		
STN1475	Event 1	16-07-18 17:00	26	59.0	2.3	10.5	8.5	17-07-18 13:00	1.73	86.3	0.03
	Event 2	07-08-18 19:00	7	42.0	6.0	28.5	1	8-08-18 10:00	2.08	38.0	0.05
	Event 3	02-07-19 16:00	31	87.5	2.8	10	5.5	2-07-19 19:00	1.91	40.8	0.02
STN0985	Event 4	02-07-19 16:00	30	73	2.4	14.5	0.5	2-07-19 19:00	1.91	43.1	0.03
TLPC	Event 5	19-09-13 14:00	44	84.6	1.9	12.4	42	19-09-13 18:00	26.10	33.3	0.31
	Event 6	16-10-13 16:00	19	42.4	2.2	13.8	6.6	16-10-13 22:00	15.15	66.1	0.36
	Event 7	02-10-17 17:00	13	40.2	3.1	16	52.4	2-10-17 17:00	6.55	37.1	0.16
	Event 8	30-09-18 19:00	6	42.4	7.1	20.8	47.4	30-09-18 23:00	4.65	21.0	0.11
	Event 9	02-07-19 18:00	31	72.6	2.3	12.4	0	2-07-19 19:00	1.91	41.5	0.03
	Event 10	19-09-19 9:00	8	96.4	12.1	40.2	5	19-09-19 11:00	1.54	12.3	0.02
	Event 11	18-09-20 17:00	16	65.8	4.1	12.2	24.6	18-09-20 17:00	20.19	37.0	0.31
	Event 12	29-10-20 2:00	19	44.6	2.3	10.6	1	29-10-20 21:00	1.77	17.8	0.04

Table D.2: Information on the precipitation events used to determine the lag time for direct runoff peaks at discharge station M89.

## D.2. Lag time and runoff correlations



(a)



(b)

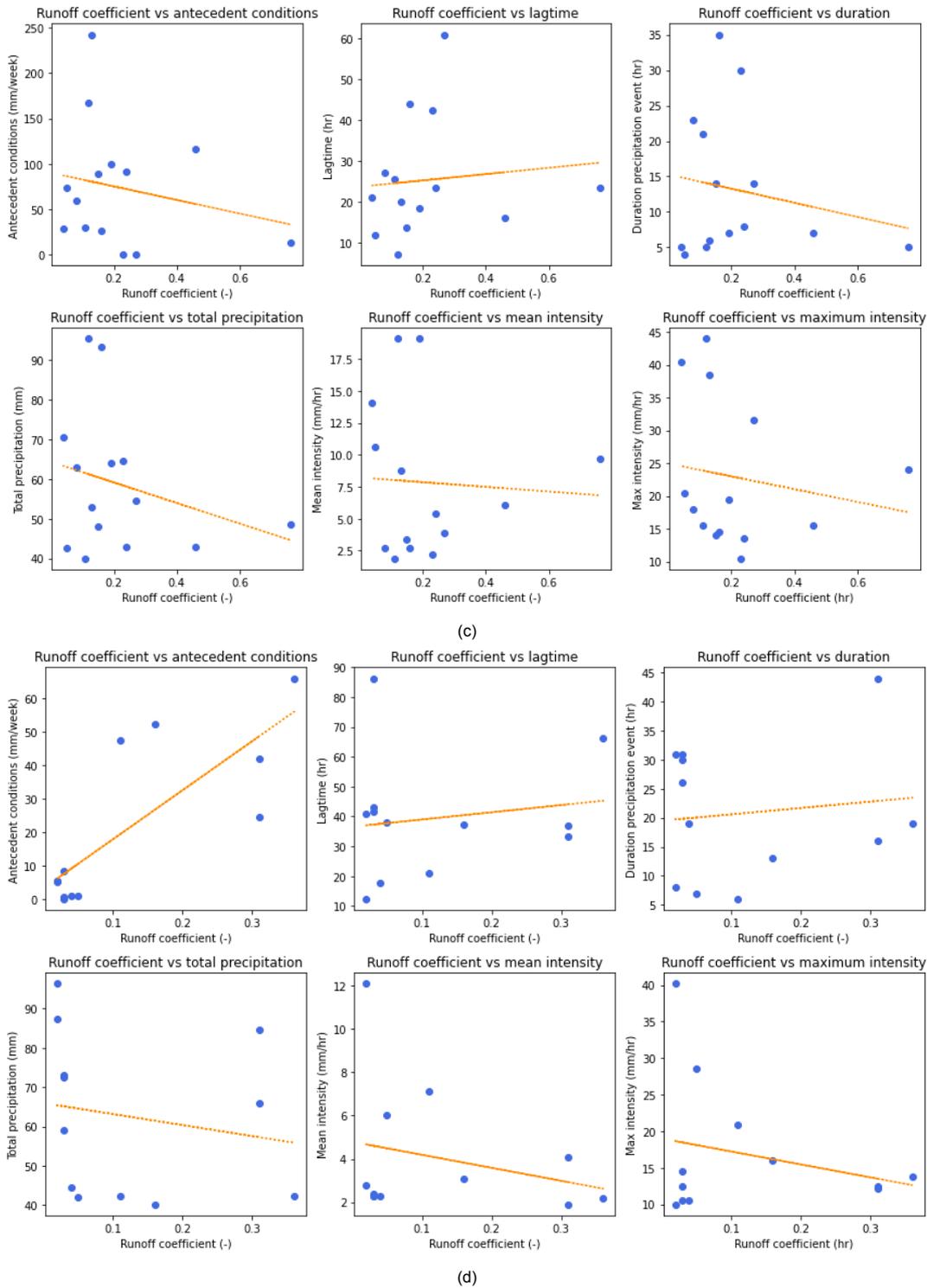
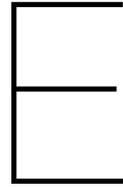


Figure D.1: Correlation between (a, b) lag time, (c, d) runoff coefficient and various characteristics of the precipitation event for discharge measured at (a, c) station M43a and (b, d) station M89.





## Pattern Description

Below the spatial rainfall patterns of the precipitation events occurring in 2018 and 2019 shown in Table D.1 and Table D.2 will be described. This will be done based of hourly distributed rainfall maps obtained from applying inverse distance to the rain gauge data.

### **20/06/18 - 21/06/18 (Event 4 - M43a)**

Intense precipitation occurs only in the south of the area. Both the start and end of the event are sudden, with a quick increase and decrease of the intensity. The rain event lasts for two to three hours. In the rest of the study area only light precipitation (<1.5 mm/hr) occurs.

Calculated runoff coefficient: 0.05

Calculated lag time: 11.8 hours

### **02/07/18 - 02/07/18 (Event 5 - M43a)**

Intense precipitation mainly occurs in the south, with a spread to the middle of the area, where precipitation is less intense. In the southern part the precipitation there is a sudden peak, which lasts for about an hour, after which some less intense rainfall continues for two more hours. In the middle the precipitation stops after an hour.

Calculated runoff coefficient: 0.04

Calculated lag time: 21.0 hours

### **16/07/18 - 17/07/18 (Event 1 - M89)**

Light precipitation (< 3 mm/hr) throughout the whole area for about a day and a half. Intensities reach the highest point in the north, where a peak of around 10 mm/hour continues for 2 hours, a day after the start of the precipitation.

Calculated runoff coefficient: 0.03

Calculated lag time: 86.3 hours

### **19/07/18 - 20/07/18 (Event 6 - M43a)**

Precipitation occurs in the southern part of the study area, although intensities are not very high. The event shows some hours without rainfall in between, after which rain starts again.

Calculated runoff coefficient: 0.08

Calculated lag time: 27.1 hours

### **07/08/18 - 08/08/18 (Event 2 - M89)**

Heavy precipitation starts in the north of the study area, showing a rapid increase and a slightly slower decrease. After an hour, precipitation occurs throughout the whole area, with overall intensities around 5 mm/hr. Higher intensities occur in the north and south and later on a high intensity peak occurs in the eastern part. Precipitation leaves towards the south.

Calculated runoff coefficient: 0.05

Calculated lag time: 38.0 hours

**18/08/18 - 18/08/18 (Event 7 - M43a)**

The event starts with low intensity rainfall in the south, after which intensity increases. The south experiences around 4 hours of precipitation intensities higher than 10 mm/hr. The precipitation spread across the whole study region, with larger intensities occurring throughout the area. Calculated runoff coefficient: 0.19

Calculated lag time: 18.5 hours

**27/08/18 - 27/08/18 (Event 8 - M43a)**

Precipitation occurs throughout the whole area, with the highest intensities in the south and middle.

Calculated runoff coefficient: 0.46

Calculated lag time: 16.0 hours

**03/09/18 - 04/09/18 (Event 4 - M43a)**

The precipitation starts in the west, after which it spreads to cover the whole area. Over a period of 6 hours, higher intensity peaks occur on multiple location in the study area.

Calculated runoff coefficient: 0.24

Calculated lag time: 23.3 hours

**30/09/18 - 30/09/18 (Event 8 - M89)**

Precipitation occurs throughout the whole area, however high intensity precipitation only occurs in the middle of the study region, with a rapid increase and decrease of the peak.

Calculated runoff coefficient: 0.11

Calculated lag time: 21.0 hours

**02/07/19 - 03/07/19 (Event 1 & 10 - M43a, event 3, 4 & 7 - M89)**

High precipitation peaks occur throughout the whole area.

Calculated runoff coefficient: 0.23, 0.16, 0.02, 0.03, 0.16

Calculated lag time: 42.4, 44.0, 40.8, 43.1, 37.1 hours

**16/08/19 - 16/08/19 (Event 11 - M43a)**

Precipitation only occurs in the south of the area, where there is both a rapid increase and rapid decrease of the precipitation intensity. The high intensity lasts about 2 hours.

Calculated runoff coefficient: 0.12

Calculated lag time: 7.1 hours

**20/08/19 - 20/08/19 (Event 12 - M43a)**

High intensity precipitation starts in the south, after which also some precipitation occurs in the north and west, although of lower intensities.

Calculated runoff coefficient: 0.13

Calculated lag time: 19.9 hours

**25/08/2019 - 25/08/19 (Event 13 - M43a)**

The rainfall event begins in the south after which the majority of the area is covered with precipitation. The precipitation seems to leave the study region in the north east.

Calculated runoff coefficient: 0.15

Calculated lag time: 13.7 hours

**19/09/19 - 19/09/19 (Event 10 - M89)**

Precipitation starts in the west after which large peaks begin in the north, north east and east of the study area.

Calculated runoff coefficient: 0.02

Calculated lag time: 12.3 hours

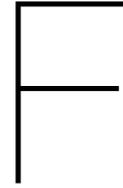
**23/09/19 - 23/09/19 (Event 2 - M43a)**

Low intensity rainfall starts in the northern part of the area, after which it spreads with decreasing intensities. After, high intensity rainfall begins in the east of the area, and spreads towards the middle and west.

Calculated runoff coefficient: 0.11

Calculated lag time: 25.4 hours





# Model Description

## F.1. Data Input

### Meteorology

The meteorological input consists of precipitation and temperature data, both with a daily temporal resolution. The precipitation data includes information on the amount of precipitation (in  $kg/m^2$ ), the maximum intensity of the rain event (in  $kg/m^2/s$ ), the duration of the event ( $s$ ) and the timing of the peak ( $s$ ). The temperature data gives information of the temperature in  $K$ .

In the study area, there are multiple rain gauges providing information on precipitation amount and intensity (see section 2.2), either with a quarterly or hourly resolution. These data sets have been re-sampled to daily precipitation to fit the temporal resolution of the model. The hourly data sets have been used to find the maximum intensity of the rain event for each day. With the maximum intensity and the total daily precipitation amount, an estimation of the rainfall duration was made. The timing of the rainfall peak was assumed to be halfway each event.

For the temperature data, climatological information from the TMD weather station Pakchong Agromet was used. This weather station is located near the study area (Figure 2.5) and provides information on monthly minimum, maximum and mean temperature obtained in the period between 2006 and 2021. In this research the monthly temperature means have been used as input and being constant for each year. In Table F.1 these monthly means can be seen, together with the average monthly minimum and maximum measured at the climatological station.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<b>Minimum</b>	18.0	19.6	21.5	22.9	23.7	24.1	23.8	23.5	22.7	22.0	20.7	18.5
<b>Maximum</b>	29.9	32.1	33.5	33.8	32.9	32.2	31.2	30.7	30.3	30.1	29.8	29.0
<b>Mean</b>	23.5	25.4	27.0	27.6	27.6	27.5	26.9	26.5	25.8	25.3	24.7	23.3

Table F.1: Minimum, maximum and mean monthly temperatures ( $^{\circ}C$ ) from Pakchong Agromet weather station determined over the years 2006-2021.

As there is only one climatological station near the study area, for this research the temperature is assumed to be spatially homogeneous over the entire catchment. For the precipitation, both uniform and spatially distributed data has been used.

### Topography

The topography of the catchment is characterized by a DEM, with a spatial resolution of 100 m (Figure 2.2a). The data was obtained from the NASA SRTM (NASA JPL, 2013) with an initial resolution of 30m. This resolution was lowered for computational convenience.

## Soil

The CALEROS model assumes a soil distribution which is steady state. Information is required on the content of silt, sand, clay, organic matter and coarse fragments. Soil data has been obtained from the ISRIC World Soil Information database (Poggio et al., 2021). This database provides soil information for the depth intervals of 0-5 cm, 5-15 cm, 15-30 cm, 30-60 cm, 60-100 cm and 100-200 cm. The SoilGrids of the ISRIC database used for this research include the bulk density of the fine particles (sand, silt, clay) in  $cg/cm^3$  and sand, silt and clay content in  $g/kg$ . As the desired input for the soil maps in the CALEROS model is in  $kg/m^2$ , a combination of the bulk density, particle content and soil depth was used to calculate this, according to Equation F.1:

$$M_{soil} = \rho_{bulk} * 10 * d * c_{particle} \quad (F.1)$$

$M_{soil}$  = soil mass ( $kg/m^2$ ),  $\rho_{bulk}$  = bulk density ( $cg/cm^3$ ),  $d$  = soil thickness of the data set ( $m$ ),  $c_{particle}$  = particle content ( $g/kg$ ).

When looking at these maps, the percentual soil content of clay, sand and silt is 46.2%, 25.0% and 28.8% respectively. According to the results of (Treløges and Sriboonlue, 2002), who studied soil properties in two different soil series in northeast Thailand, clay content is 9.2% - 17.8%, sand content is 46.2% - 66.6% and silt content is 24.2% - 36.0%. Later information obtained from soil samples collected in the study region at telemetering station Khaoyai 1 (Figure 2.5) confirm that the soil distributions is closer to the one found in literature than the one derived from the SoilGrids, with a sand content of 64.4%, a silt and clay content of 17.1% and a gravel content of 18.5%.

As these values deviate a lot from the soil contents according to the ISRIC datasets, the maps have been scaled to better correspond to the literature. This resulted in a mean clay, sand and silt content of 13%, 62.36% and 24.76% respectively, which also resulted in a saturated conductivity corresponding better to the values found in literature. The content of coarse fragments and organic matter has been set to zero, as no data was found on this.

**Soil parameters** From the soil maps, different soil parameters are calculated by the model using pedotransfer functions. These soil parameters include soil thickness, the bulk density of the soil ( $\rho_b$ ), the saturated hydraulic conductivity ( $k_{sat}$ ) and the porosity ( $\theta_{sat}$ ).

## Surface hydrology

In the surface hydrology the initial ponded water mass in the catchment is specified. For this research it is set to zero, as there are no significant lakes or other water bodies in the study area.

## Soil hydrology

The soil hydrology includes information on the water content in the saturated as well as the unsaturated zone. Also, here the saturated hydraulic conductivity of the ground layer below the modelling layer is specified. The water content in the saturated zone is set to zero for this research, as the layer thickness corresponding to the parameters used in the model has a maximum of 2.24 m, while groundwater wells throughout the study area (Figure 2.5) show the groundwater table to be at significantly larger depths (see Table F.2). Additionally, test runs resulted in initial conditions for the saturated water mass of zero.

The water content in the unsaturated zone is quite a big uncertainty within the model. Multiple-year test runs have been performed (subsection 3.2.2) to obtain initial values for the unsaturated zone. The results were compared with the field observations from the telemetering station installed in the National Park, giving information on the volumetric water content at different depths. As the observation stations have only been installed in the last half year, the stations mainly provide information on the moisture content during the dry season, which is around 30%.

Well	001	002	003	004	005	006	007	008	008	009	010	011	012	013
Min	8.9	6.0	3.5	8.5	6.9	4.7	19.0	10.0	3.3	4.0	4.4	3.5	4.5	6.2
Max	87.7	13.6	44.5	15.6	9.3	8.6	31.7	18.5	11.1	7.8	10.6	9.4	9.4	11.3
Mean	25.4	11.3	11.1	12.1	7.9	6.8	26.8	15.5	8.0	5.5	8.8	7.1	7.6	9.8

Table F.2: *Minimum, maximum and mean groundwater levels (in meters from the ground surface) observed by the groundwater stations in the study area.*

Besides the water content in the subsurface, also the saturated conductivity of the boundary conditions needs to be specified, which is done as a fraction of the hydraulic conductivity of the top layer. This "leakage factor" is a big uncertainty in the model, and therefore obtained through calibration and multiple test runs.

### Vegetation stand

In the section on vegetation stand, multiple aspects of the vegetation are specified including; initial interception, soil albedo, canopy albedo, canopy height, cover fraction, LAI, soil cropfactor, cropfactor, interception capacity, root depth, saturated hydraulic conductivity, root cohesion and surcharge. For some vegetation parameters constants are used, while others have a spatially distributed input.

**Initial interception** In this research the initial interception has been set to zero. As the model runs are started during a moment in the year with little to no precipitation. Additionally, temperatures in the area are high and as the maximum interception capacity is not significant, it can be assumed that any interception as a result of previous events has been evaporated.

**Soil albedo and canopy albedo** For the soil and surface albedo, constant values have been used as no distributed data was available. Pinker et al., 1980, have investigated the albedo of a tropical evergreen forest at the Khorat Plateau, located around 200 km northeast of the study area. The average albedo obtained from their findings was 0.14, which is therefore used as the canopy albedo in this research.

Dry, sandy soils generally have an albedo ranging from 0.25 to 0.45 (An et al., 2017; Dobos, 2003). As the study region has a relatively wet soil for majority of the year, a slightly lower albedo of 0.20 has been used in this research.

**Canopy height** Potapov et al., 2021 developed a global forest canopy height map with the GLAD with a spatial resolution of 30 x 30 meters. Information from this data set has been used as input for the model, after lowering the resolution to 100 x 100 m, to fit the spatial resolution of the model.

**Cover fraction and LAI** The data for both the cover fraction and the LAI was obtained from data sets from the Copernicus Global Land Service (Fuster et al., 2020). Both data sets had a spatial resolution of 300 x 300 m, which has been changed to a resolution of 100 x 100 m to match the spatial resolution of the model.

**Soil cropfactor and cropfactor** The soil cropfactor and cropfactor are two components affecting the evapotranspiration potential of the system. The cropfactor reflects the potential of vegetation transpiration, while the soil cropfactor reflects that of the bare soil evaporation.

For the soil cropfactor and the cropfactor constant values were used. For the cropfactor a value of 1.00 has been used, based on the general crop factors determined by the FAO (Allen et al., 1998). The soil cropfactor is very depending on the water availability, and thus the weather. When the topsoil is wet, so for example after a precipitation event, the soil cropfactor is maximal (1.00), while when the soil surface is dry, the soil cropfactor can reach as low as zero, meaning no more water can be evaporated from the soil (Allen et al., 1998).

**Interception capacity** The interception capacity was calculated based on the LAI, according to Equation F.2 (De Roo et al., 1996).

$$S_{max} = 0.935 + 0.498 * A_{leaveindex} - 0.00575 * A_{leaveindex}^2 \quad (F.2)$$

$S_{max}$  = interception capacity (mm),  $A_{leaveindex}$  = LAI.

**Root depth** Root depth has been determined based on the land cover types in the study area (see Figure 2.2b). For water bodies, herbaceous wetlands and build up areas the root depth was assumed to be zero. Croplands, shrub lands (Silva and Rego, 2003) and herbaceous vegetation were assumed to have a root depth of 30 cm. For forest covers the root depth was set to be equal to the soil thickness.

**Hydraulic saturated conductivity** The hydraulic saturated conductivity of the soil is determined by the model using pedotranfer-functions, based on the soil content of the different soil classes in the system.

**Root cohesion and surcharge** Similar as to the hydraulic saturated conductivity, both root cohesion and surcharge are determined by the model as well.

## F.2. Data output

There are a large number of variables that can be written as output in this model. In this chapter the output sets most relevant to this research will be discusses. An overview of all possible output variable can be found in Appendix H.

### Evaporative fluxes

**Potential evaporation** The potential evaporation in the model is determined using the Hamon equation. The Hamon equation is a convenient method to use when only temperature data is available. It is often used in hydrological studies as limited data is required and it has shown to be closely correlated with the actual evaporation (Lu et al., 2005; Rao et al., 2011).

The potential evaporation estimated with this method is based on air temperature and day length, according to Equation F.3:

$$E_{hamon} = 0.1651 * L_d * \rho_{sat} * k \quad (F.3)$$

$E_{hamon}$  = potential evaporation (mm/day),  $L_d$  = amount of sunshine hours from sunrise to sunset (fraction of 12 hours),  $\rho_{sat}$  = saturated vapour density at the daily mean temperature ( $g/m^3$ ),  $k$  = the calibration factor (-).

In the model the calibration factor is set to zero.  $e_{sat}$  and  $\rho_{sat}$  are calculated using Equation F.4 and Equation F.5 respectively.

$$e_{sat} = 6.108 * \exp(17.27 * T / (T + 237.3)) \quad (F.4)$$

$$\rho_{sat} = 216.7 * e_{sat} / (T + 273.3) \quad (F.5)$$

$T$  = temperature (°C).

The potential evaporation determined by the model using the Hamon equation was checked by calculating the monthly potential evaporation using the Penman-Monteith equation, with data obtained from the climatological TMD station Pakchong Agromat (Figure 2.5). The potential evaporation determined with Penman-Monteith gave higher values as can be seen in Figure F.1, so the calibration factor,  $k$ , in the model was set to 1.4 to better match the calculated potential evaporation.

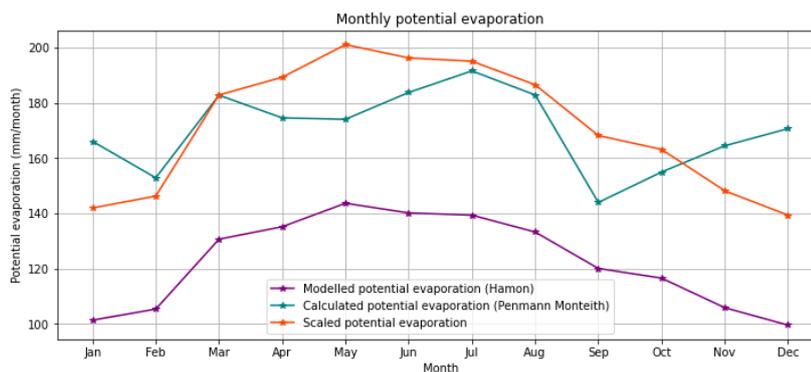


Figure F.1: The potential evaporation as determined by the model using Hamon, as determined using the Penman-Monteith equation and as determined using Hamon, after scaling it to the Penman-Monteith results.

**Actual evaporation** The actual evaporation is determined based on the combination of canopy or interception evaporation, soil evaporation and canopy transpiration. Canopy evaporation is equal to the interception stored on the canopy. The soil evaporation and canopy transpiration are determined by multiplying the potential evaporation with the soil cropfactor and cropfactor respectively. After this, a scale factor is applied to limit the maximum evaporation rate. This limiting rate can be e.g. the atmospheric demand or the allowable evaporation from the soil.

### Vertical groundwater fluxes

In the unsaturated zone of the soil, water flows vertically through percolation as a result of gravitational forces. This flow is controlled by the saturated hydraulic conductivity of the soil and the bedrock layer. The saturated hydraulic conductivity ( $k_{sat}$ , in  $m/d$ ) is estimated on the basis of soil properties using the relationship of Brutsaert:

$$k_{sat} = \frac{(\theta_{sat} - \theta_{res}) * \lambda_{bc})^2 * \psi_{sat_{bc}}^{-2}}{(\lambda_{bc} + 1) * (\lambda_{bc})^{-1}} \quad (F.6)$$

$\theta_{sat}$  = saturated volumetric water content ( $m^3/m^3$ ),  $\theta_{res}$  = residual volumetric water content ( $m^3/m^3$ ),  $\psi_{sat_{bc}}$  = air entry value of the Brooks & Corey soil water retention curve (m),  $\lambda_{bc}$  = pore size distribution of the Brooks & Corey soil water retention curve (-),  $a$  = optional constant (default value is 3.01124)

A higher value of  $k_{sat}$  allows more water to infiltrate into the subsurface and thus reduces the runoff in the system.

The saturated hydraulic conductivity of the bedrock layer,  $k_{sat_{bc}}$ , is a fraction of the saturated hydraulic conductivity of the soil. When set to zero, no water leaves the system through deep percolation, which means all precipitation either leaves the system as evaporation, runoff or stays in the system as additional groundwater storage.

### Lateral fluxes

The main lateral fluxes in this model are groundwater flow and runoff.

**Runoff** The runoff in this model is determined using the PCRaster function *pcr.accuthresholdflux*. This function describes the accumulation of water masses in a drainage network and limits its transport by thresholds. Transport of the water will only occur when a particular threshold of losses has been attained. When the threshold is not reached, the water will be stored. For overland flow, the threshold is equal to the storage capacity of the soil, so that flow will only develop once the soil is saturated.

**Groundwater flow** Lateral groundwater flow occurs across the saturated zone due to differences in total head. The lateral flow is dependant on the saturated hydraulic conductivity of the soil, the water level and the width of a cell. It is calculated using Darcy's law given in Equation F.7.

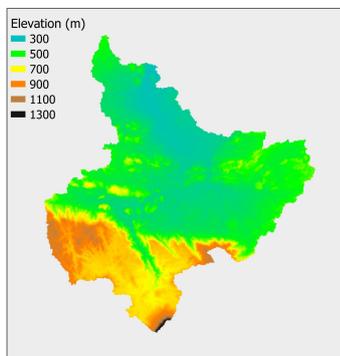
$$Q_{gw} = -k_{sat} \frac{\Delta h}{\Delta x} \quad (\text{F.7})$$

$Q_{gw}$  = Lateral groundwater flow ( $m/d$ ),  $k_{sat}$  = saturated hydraulic conductivity ( $m^3/d$ ),  $\Delta h$  = head difference ( $m$ ),  $\Delta x$  = distance ( $m$ )

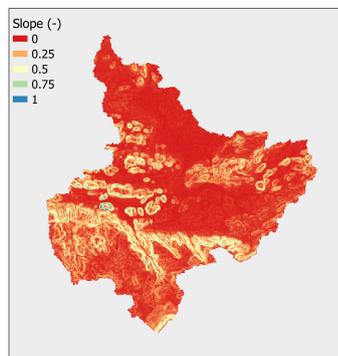
# G

## Initial input variables

### Topography

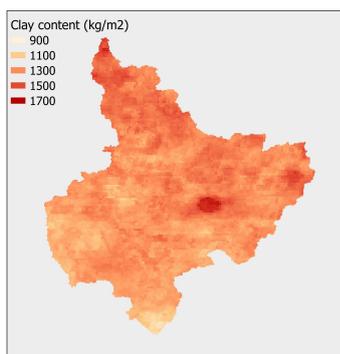


(a) DEM

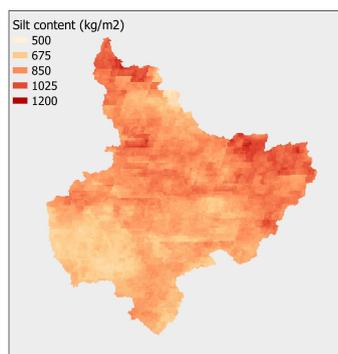


(b) Slope

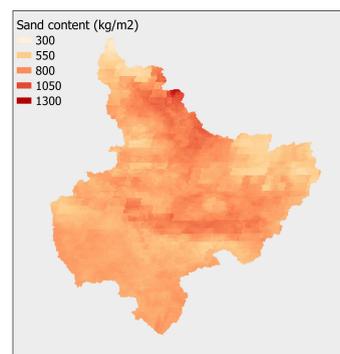
### Soil



(a) Clay mass



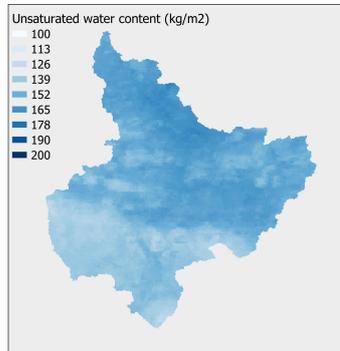
(b) Silt mass



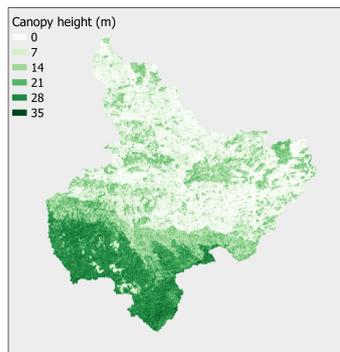
(c) Sand mass

Stone mass: 0 kg/m<sup>2</sup>

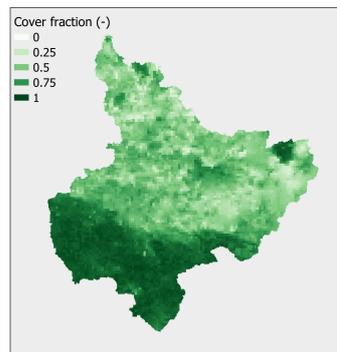
Organic matter mass: 0 kg/m<sup>2</sup>

**Surface hydrology**Ponded water mass:  $0 \text{ kg/m}^2$ **Soil hydrology**

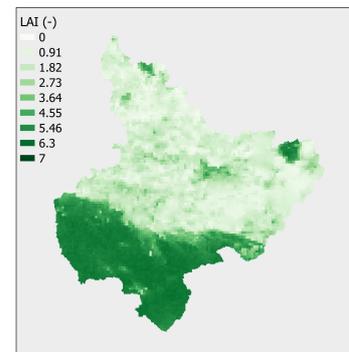
(a) Unsaturated water mass

Saturated water mass:  $0 \text{ kg/m}^2$ **Vegetation stand**

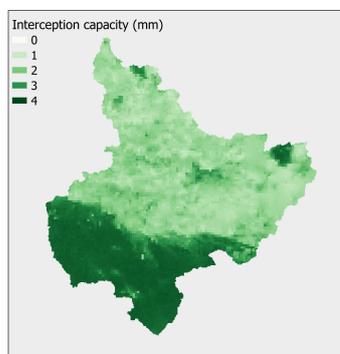
(a) Canopy height



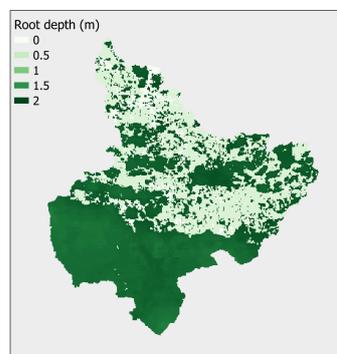
(b) Cover fraction



(c) Leave Area Index



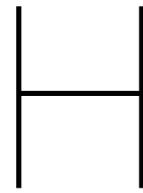
(d) Interception capacity



(e) Root depth

Initial interception: 0 *mm*  
Soil albedo: 0.20  
Canopy albedo: 0.15  
Soil cropfactor: 0.20  
Cropfactor: 1.0  
Root cohesion: 0 *kPa*  
Surcharge: 0 *kPa*





Model variable list

	Variable name	Unit	Description
Soil properties	Soil_void_volume	$m^3/m^2$	Total void volume
	Soil_thickness	$m^3/m^2$	Total soil thickness
	Soil_density	$kg/m^3$	Soil bulk density
	Soil_theta_sat	$m^3/m^3$	Volumetric moisture content at saturation
	Soil_theta_res	$m^3/m^3$	Volumetric moisture content at residual moisture content
	Soil_psi_sat_bc	$m$	Air entry value
	Soil_lambda_bc	-	Pore size index
	Soil_sorptivity	$m$	Sorptivity at the wetting front
	Soil_ksat	$m/day$	Saturated hydraulic conductivity
	Soil_surface_stoniness	$m^2/m^2$	Fraction of soil surface stoniness
	Soil_tan_phi	-	Tangent of the angle of internal friction
	Soil_cohesion	$kPa$	Drained soil cohesion
	Soil_viscosity_ratio	$s$	Ratio of the plastic viscosity to the yield stress
Soil_shear_zone_thickness	$m$	Thickness of the shear surface	
Vegetation stand	Interception	$kg/m^2$	Interception stored on the canopy and litter
	Albedo	-	Overall albedo of the vegetation stand
	Soil_albedo	-	Soil albedo
	Canopy_albedo	-	Canopy albedo
	Canopy_height	$m$	Average height of the canopy
	Cover_fraction	$m^2/m^2$	Cover fraction - equivalent of the gap fraction
	Lai	$m^2/m^2$	One-sided leaf area index over the vegetation stand
	Cropfactor	-	Crop factor: directly based on input
	Soil_cropfactor_soil	-	Crop factor of the soil surface: directly based on input
	Interception_capacity	$kg/m^2$	Interception capacity over the stand
	Root_depth	$m$	Average root depth over the vegetation stand
	Ksat	$m/day$	Saturated hydraulic conductivity averaged over the vegetation stand
	Root_cohesion	$kPa$	Overall root cohesion for the stand
	Surcharge	$kPa$	Surcharge due to the vegetation present in the stand
	Soil_evaporation	$kg/m^2/day$	Total evaporation from the soil at the stand level
Canopy_evaporation	$kg/m^2/day$	Total evaporation from the canopy and litter of the vegetation stand	
Canopy_transpiration	$kg/m^2/day$	Total transpiration for the vegetation stand	
Throughfall	$kg/m^2/day$	Throughfall: net precipitation passing the canopy and litter	
Meteo	Precipitation_amount	$kg/m^2$	
	Air_temperature	$K$	
	Potential_evaporation	$kg/m^2/day$	
Surface hydrology	Ponded_water_mass	$kg/m^2$	Water ponded at surface
	Infiltration	$kg/m^2/day$	Infiltration into the soil
	Surface_water_evaporation	$kg/m^2/day$	Surface evaporation from ponded water
	Exfiltration	$kg/m^2/day$	Groundwater exfiltration to the surface
	Mass_flux	$kg/s$	Runoff
Soil hydrology	Mass_balance_error	$kg/m^2$	Mass balance error in the total water stored in the soil layer
	Unsaturated_water_mass	$kg/m^2$	Water in the unsaturated zone
	Saturated_water_mass	$kg/m^2$	Water in the saturated zone
	Base_elevation	$m$	Height of the layer base
	Surface_elevation	$m$	Height of the layer top
	Saturated_thickness	$m$	Height of the groundwater table above the layer base
	Watertable_depth	$m$	Depth of the groundwater below the surface
	Volumetric_moisture_content	$m^3/m^3$	Volumetric moisture content
	Total_head	$m$	Total head at the base of the soil layer
	Pressure_head	$m$	Pressure head above the base of the soil layer
	Soil_infiltration	$m^3/m^2/day$	Infiltration into the soil
	Soil_exfiltration	$m^3/m^2/day$	Groundwater exfiltration to the surface
	Percolation	$m^3/m^2/day$	Internal percolation: exchange between the saturated and the unsaturated
Deep_percolation	$m^3/m^2/day$	Deep percolation constituting the total lower boundary flux	
Soil_moisture_evaporation	$m^3/m^2/day$	Soil evaporation	
Actual_evapotranspiration	$m^3/m^2/day$	Actual evapotranspiration	
Mass wasting	Pore_pressure	$kPa$	Pore pressure at the potential shear plane
	Safety_factor	-	Safety factor for the soil mass

# Sensitivity Analysis

	Q	Ea	Pd	leff	dS	Mean
Temperature	-4.26	0.56	-1.53	1.22	-0.82	1.68
Clay content	-0.87	-0.08	-2.99	0.25	0.72	0.98
Silt content	-0.53	-0.07	-1.28	0.15	1.04	0.61
Sand content	0.50	-0.05	-0.07	-0.15	1.16	0.39
Stone content	0.06	-0.01	0.02	-0.02	0.34	0.09
Organic matter content	-0.87	-0.03	-0.50	0.13	0.49	0.40
Soil thickness	-0.52	-0.14	0.81	0.15	1.61	0.65
$K_{sat}$ boundary condition	0.00	0.00	0.52	0.00	0.00	0.11
Unsaturated water content	0.42	0.12	7.38	-0.12	-1.60	1.93
Saturated water content	0.08	0.05	0.06	-0.03	-1.11	0.27
Interception	0.00	0.00	0.00	0.00	0.00	0.00
Soil albedo	0.00	0.00	0.00	0.00	0.00	0.00
Canopy albedo	0.00	0.00	0.00	0.00	0.00	0.00
Canopy height	0.00	0.00	0.00	0.00	0.00	0.00
Cover fraction	0.02	0.00	0.00	-0.09	0.00	0.02
Leaf area index	0.01	0.00	-0.10	-0.06	0.00	0.03
Soil cropfactor	0.00	0.00	0.00	0.00	0.00	0.00
Cropfactor	-1.35	0.37	-0.45	0.95	-0.43	0.71
Interception capacity	0.01	0.00	0.00	-0.04	0.00	0.01
Root depth	0.00	0.00	0.00	0.00	0.00	0.00
$K_{sat}$	0.18	0.10	0.20	-0.05	-0.31	0.17
Root cohesion	0.00	0.00	0.00	0.00	0.00	0.00
Surcharge	0.00	0.00	0.00	0.00	0.00	0.00

Table I.1: Overview of the sensitivity indices for different input parameters and output variables. The mean is determined from the absolute values.





# Parameterisation

## Parameter set A - Recession check

Layer thickness: x 1.10 in National Park  
x 0.90 in provincial area

## Parameter set B - Event calibration

Layer thickness: x 1.10 in National Park  
x 0.90 in provincial area  
 $K_{sat}$ : x 1.00 in National Park  
x 3.00 in provincial area  
Unsat water: x 0.00

## Parameter set C - Event calibration

Layer thickness: x 1.10 in National Park  
x 0.90 in provincial area  
 $K_{sat}$ : x 1.00 in National Park  
x 2.40 in provincial area

## Parameter set D - FDC calibration

Layer thickness: x 1.32 in National Park  
x 1.35 in provincial area  
 $K_{sat}$ : x 1.00 in National Park  
x 3.00 in provincial area  
Unsat water: x 1.00 in National Park  
x 0.00 in provincial area

## Parameter set E - FDC calibration

Layer thickness: x 2.20 in National Park  
x 1.80 in provincial area  
 $K_{sat}$ : x 1.00 in National Park  
x 3.00 in provincial area  
Unsat water: x 1.00 in National Park  
x 0.00 in provincial area





# Synthetic spatial and dynamic precipitation input

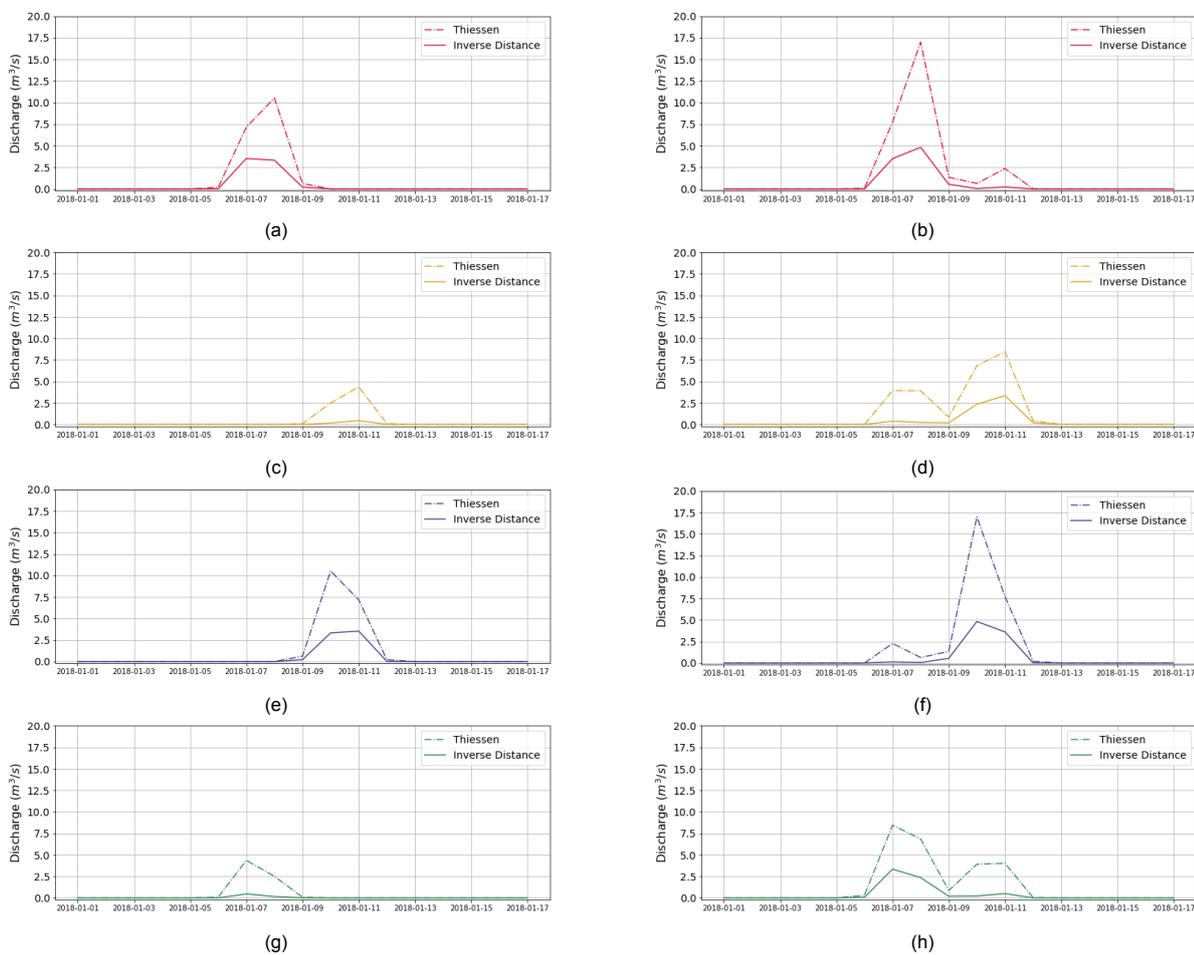
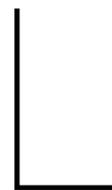


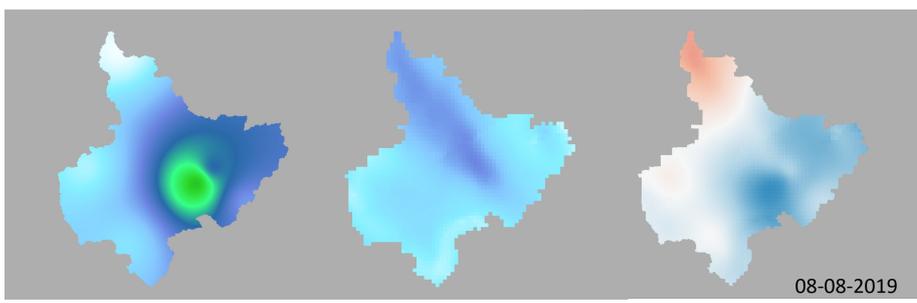
Figure K.1: Modelled discharge for the synthetic precipitation input using both Thiessen and Inverse distance interpolation for the movement of the precipitation event in different directions. The direction of the modelled discharges in (a) and (b) are from North to South, in (c) and (d) from East to West, in (e) and (f) from South to North and in (g) and (h) from West to East. (a), (c), (e) and (g) are the discharges at station M43a and (b), (d), (f) and (h) at station M89.



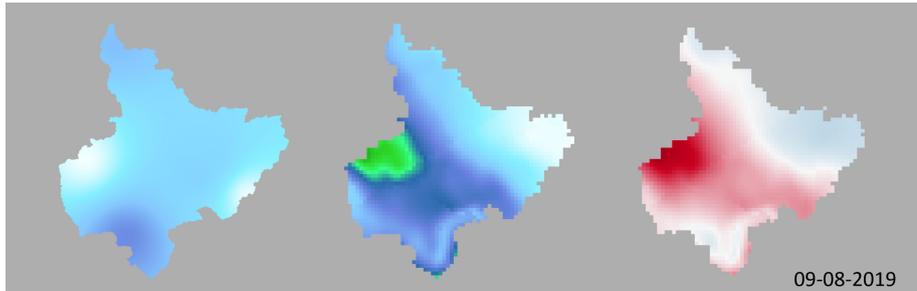


# Radar data comparison

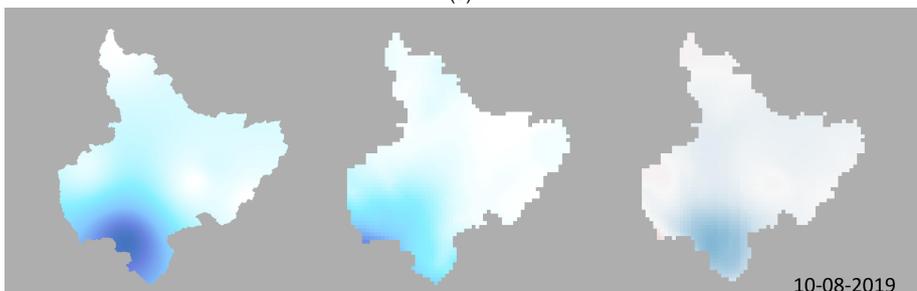
## L.1. Data images



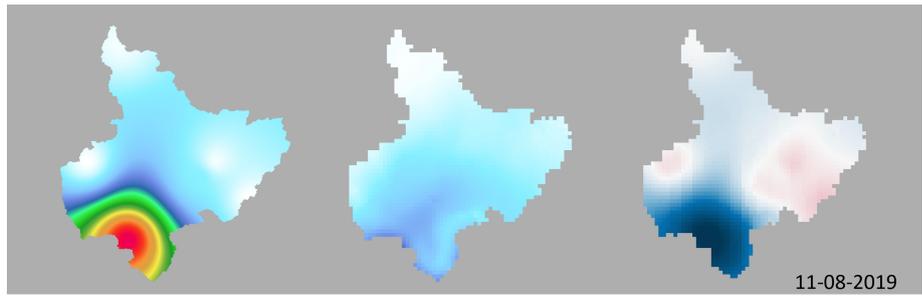
(a)



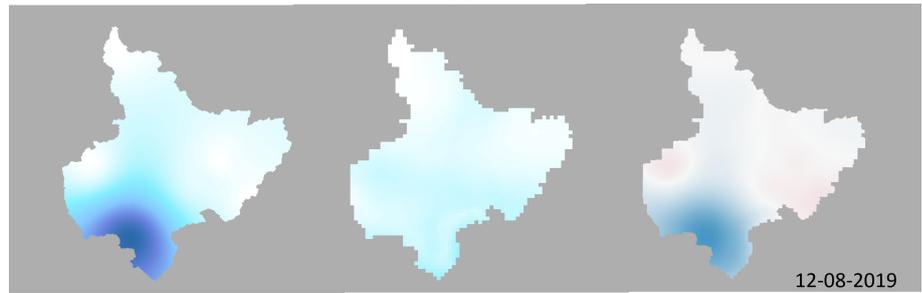
(b)



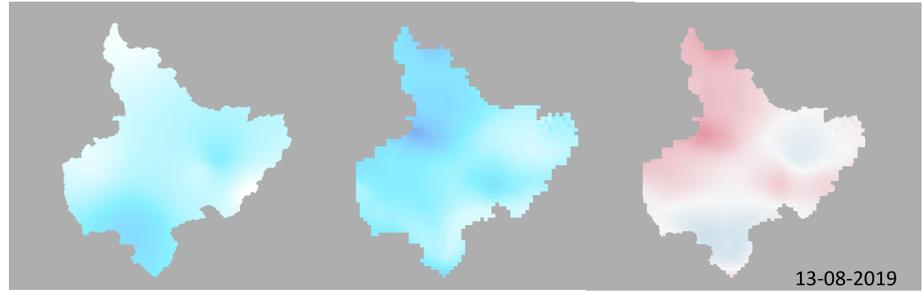
(c)



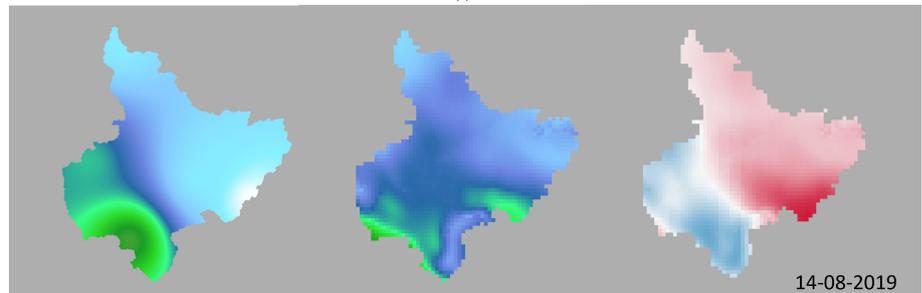
(d)



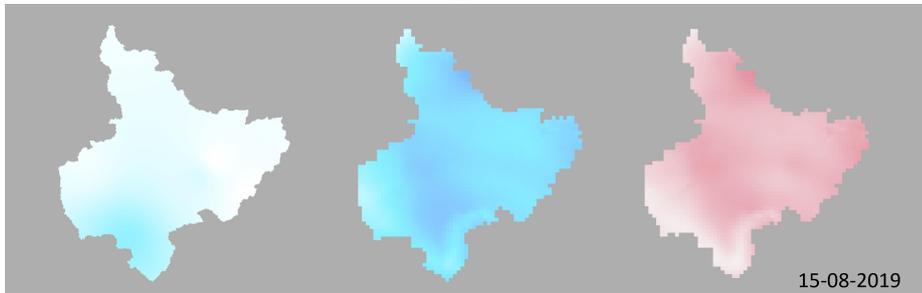
(e)



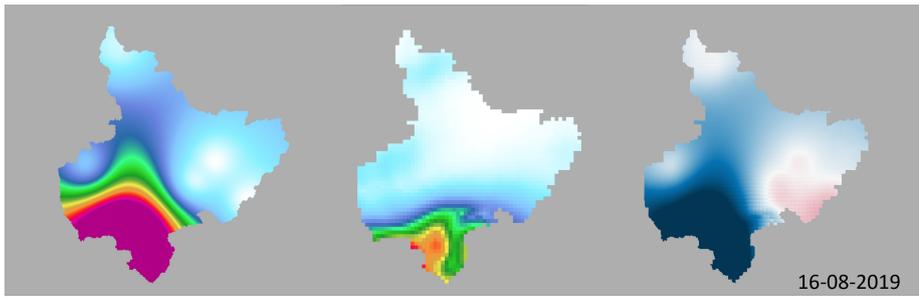
(f)



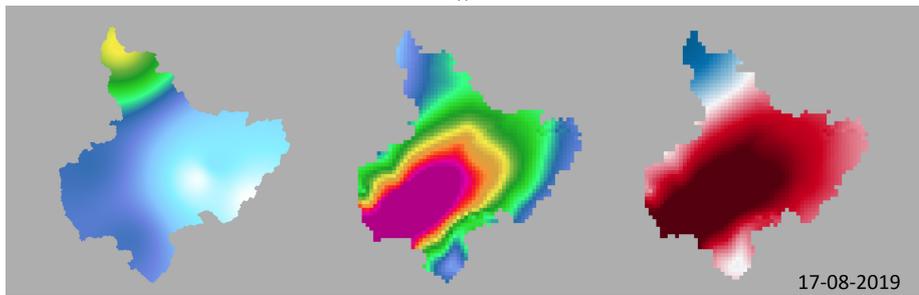
(g)



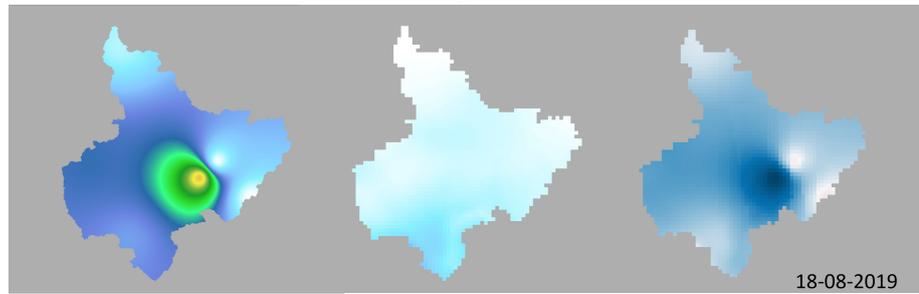
(h)



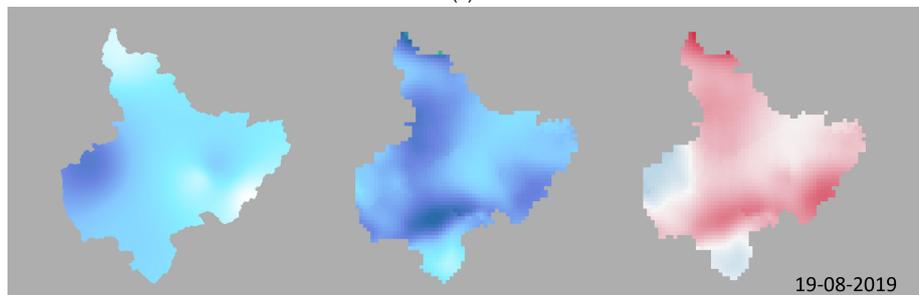
(i)



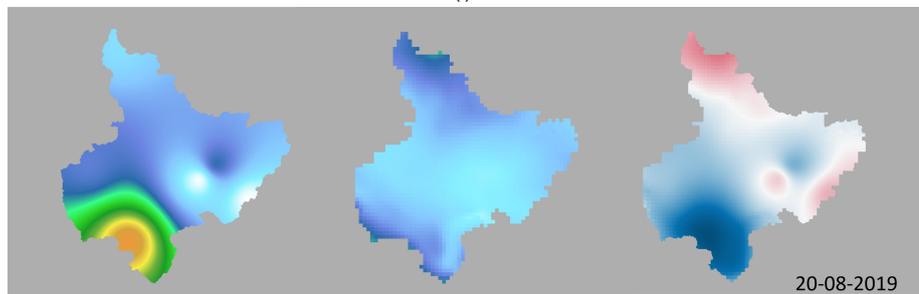
(j)



(k)



(l)



(m)

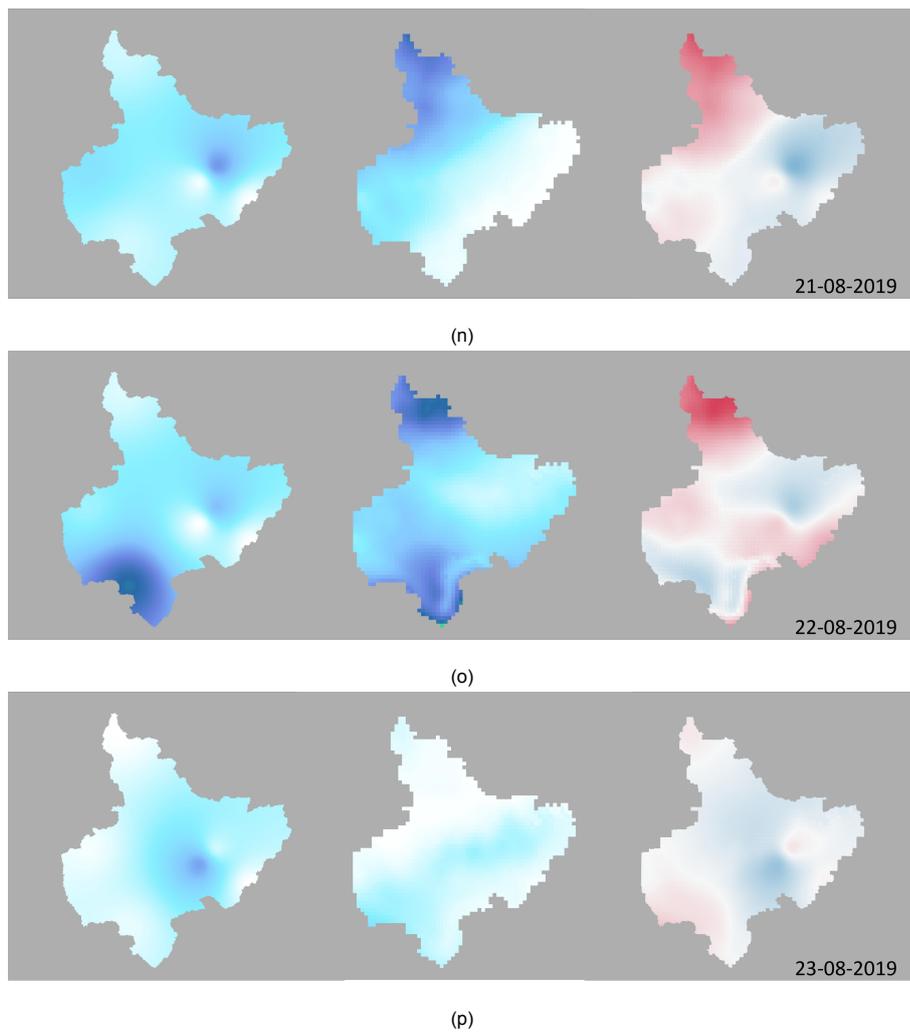


Figure L.1: Overview of (f.i.t.r.) the interpolated rain gauge map, the rain radar map and the difference map between rain gauge data and radar data of the different days.

### L.2. Error results

		8-8	9-8	10-8	11-8	12-8	13-8	14-8	15-8	16-8	17-8	18-8	19-8	20-8	21-8	22-8	23-8
Radar data	Min	2.36	0.43	0.03	0.27	0.03	1.33	7.67	1.86	0.01	10.28	0.01	3.07	3.73	0.04	1.91	0.08
	Max	14.80	33.01	14.29	16.57	4.06	11.57	34.14	10.96	64.17	120.47	10.48	22.57	22.55	18.83	23.81	6.72
	Mean	8.02	12.24	1.80	5.00	1.31	4.67	16.54	6.70	8.45	45.01	2.25	11.68	9.61	4.96	8.10	1.61
	Std	3.17	6.86	2.34	3.23	0.73	1.91	4.26	1.54	13.02	26.18	1.58	3.25	3.68	4.51	4.51	1.16
Rain gauge data	Min	0.50	0.50	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00
	Max	31.50	14.50	18.00	66.00	20.00	7.50	37.00	4.50	158.00	45.50	47.50	16.50	53.50	13.50	20.50	12.50
	Mean	13.79	7.42	3.74	13.12	3.82	3.00	13.78	1.27	30.98	13.51	15.92	6.96	17.36	4.60	6.73	3.21
	Std	6.83	2.75	4.36	16.16	4.94	1.54	9.01	1.03	39.11	9.67	7.37	3.29	11.50	2.07	4.35	2.25
Difference	Min	-12.75	-32.44	-2.67	-4.97	-2.13	-9.49	-22.54	-10.89	-7.83	-105.24	-1.79	-19.92	-14.26	-16.75	-18.45	-5.01
	Max	12.70	6.08	12.05	54.43	17.19	4.40	14.86	-0.49	116.80	33.43	45.16	7.11	40.27	12.27	7.82	9.88
	Range	25.45	38.52	14.71	59.40	19.31	13.89	37.39	10.39	124.63	138.67	46.95	27.03	54.53	29.02	26.27	14.89
	Std	7.31	7.81	2.73	13.95	4.61	2.94	7.65	1.77	28.51	28.67	7.16	4.51	11.15	5.16	5.10	2.36
Absolute difference	Max	12.75	32.44	12.05	54.43	17.19	9.49	22.54	10.89	116.80	105.24	45.16	19.92	40.27	16.75	18.45	9.88
	Mean	7.61	6.57	2.15	9.20	2.91	2.68	6.86	5.53	23.33	35.03	13.78	5.62	9.69	3.78	3.71	2.20
	Std	5.43	6.43	2.69	13.46	4.40	2.03	4.31	1.77	27.94	24.48	7.14	3.39	9.61	3.49	3.67	1.83
Error	NSE	-7.51	-0.26	-0.33	-17.13	-49.97	-1.99	-2.14	-12.39	-169.64	-1.60	-19.07	-2.98	-6.81	-0.27	-0.31	-4.90

Table L.1: Overview of some specifics of the different maps obtained to determine the correlation between radar data and interpolated rain gauge data.

### L.3. Difference map evaluation

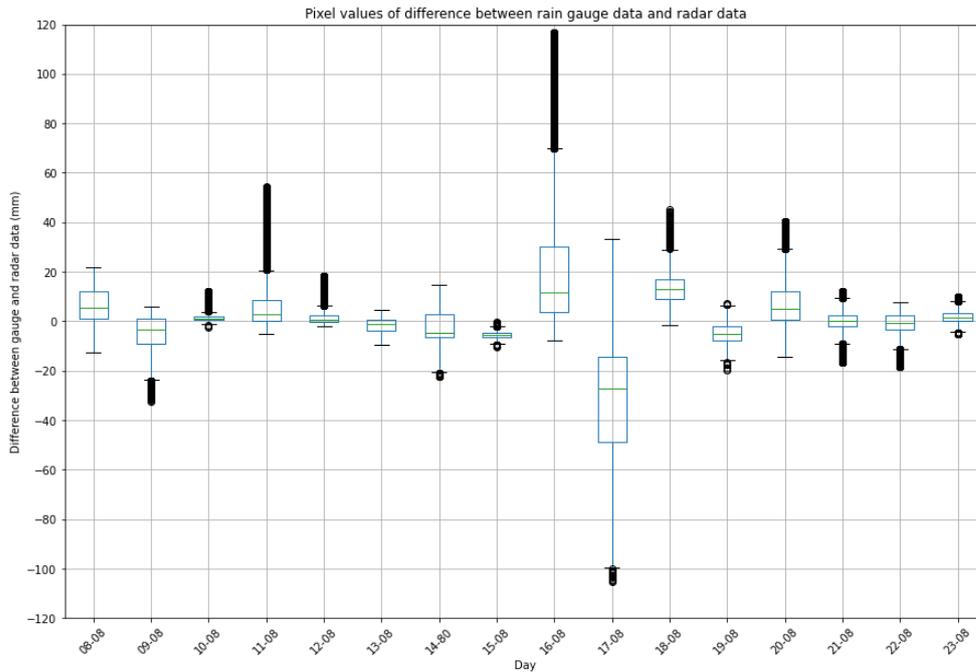


Figure L.2: Boxplots of the pixel values of the difference maps for each day.